

PART II: DETAILED DESCRIPTIONS OF STEPS

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.1: Understanding the ATM Problem and Operational Concept****a) Purpose.**

The purpose of Activity 1.1 is to ensure a good understanding of the ATM Problem and the operational concept that is proposed to address it.

b) Description.

The initial activity of Step 1 is to ensure that there is a good understanding of the operational concept that is being validated and of the ATM problem it is intended to address. This activity is very important because the overall objective of all validation exercises must be to determine how well an operational concept or element of an operational concept meets the need for which it was developed.

A statement, describing the ATM problem, should be provided by the development team. This statement may have been generated as the first stage in the requirements capture process. The description of the problem should provide information such as:

- Type of problem, (e.g., safety, limitations in capacity), indicating performance requirements or targets;
- Components of the ATM system involved, communication limitations;
- Geographic location and timeframe should also be documented.

The development team should also provide a description of the ATM operational concept. This statement should describe how the particular components of the ATM system relevant to the ATM problem are configured and operated. It should provide information on the issues presented below:

- *Maturity of the ATM Concept.* The level of maturity will help define the type, rigour and scope of the validation process. The maturity of the concept will have an impact in later steps and activities related to aspects such as: identification validation aims (activity 1.3) and validation objectives (activity 1.4), selection of the validation technique and associated validation platform (activities 1.5 and 2.1) ²¹.
- *Actors* that are involved. In the ATM context these are usually air traffic controllers, pilots, airline dispatchers, airport managers, etc.
- *Actors' tasks and human responsibilities.* Each actor may have specific tasks and responsibilities, though overlap is possible. In the latter case it is necessary to describe which actor has a responsibility in what circumstances.
- *Enablers.* Actors use systems based on enabling technologies to help them perform their tasks. It needs to be clear how they use these systems and their rôle (e.g., provision of information or advising the human in decision-making). The enablers may also include new procedures, human factor developments or institutional arrangements, which, when exploited potentially deliver improved performance in terms of safety, capacity, economics or other ATM 2000+ Strategy objectives. The validation exercise design should test out the enablers to determine their maturity, reliability and effectiveness.
- *Events and the drivers for events.* Events may occur (e.g., a thunderstorm or

²¹ As a concept matures from idea to product, the validation process should mature also. Validation in early stages of a concept's lifespan can take on a less rigorous approach while in later phases it needs to adopt a high fidelity approach as the concept becomes more clearly defined and closer to implementation. (Ref. GTG - D0.4.1: Validation Methodology and Strategy, Issue 1.0, page 17 "Identify the level of readiness").

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES

Activity 1.1: Understanding the ATM Problem and Operational Concept

emergency) or be triggered by another event and/or an action of an actor (e.g., the issue of a clearance by a controller triggers the pilot to manoeuvre the aircraft).

- *Processes* that take place and their relation to each other. An example of a process is the set of steps a controller follows in order to issue a clearance to an aircraft. The outputs from one process can act as inputs to other processes, or trigger other processes.
- *Organisation of airspace and supporting systems*. The general airspace structure and properties should be defined.
- *Information flows* can be recognised at different levels: between processes and between actors. There may also be information to and from external processes/actors not being part of the scope.
- *Procedures* that are applied in order to execute the Operational Concept describing sequences of actions for given operational events.

Initially the contents of the operational concept may be relatively vague; many issues and assumptions need clarification. As the operational concept passes through the initial phase of the validation life-cycle, more information from early validation exercises will allow these assumptions to be confirmed or adjusted.

Therefore, for each validation exercise, the current version of the operational concept will need to be confirmed. This means that the operational concept is checked for changes that need to be made following experience in the previous development phases. Moreover, it may need to be adjusted to specific environmental circumstances of the validation exercise, such as the geographical location and the traffic mix.

To ensure an unbiased validation, it is recommended that definition and development of the operational concept, and the performance of the validation exercise, are the responsibility of different actors. Therefore, any need to refine/modify the concept identified by the validation team should be referred back to and agreed with the developers of the concept.

Where the advanced Operational Concept is presented as an Operational Improvement, the baseline concept needs to be defined and understood, along with the changes to the baseline that comprise the Operational Improvement.

c) **Criteria for Application.**

This activity should always be applied.

d) **Inputs.**

- Descriptions of the ATM Problem and the ATM Concept²².

e) **Outputs.**

- Good, uniform understanding.

²² Appendix-7 provides examples and good practices on Operational Concept Template.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.2: Identification of Stakeholders****a) Purpose.**

The purpose of Activity 1.2 is to ensure that all parties relevant to the validation of the ATM concept are known so that they can provide information and develop confidence in the proposed ATM concept in meeting the operational need.

b) Description.

ATM is a complex system involving many agencies, services and countries. If any change in the operational concept is to be successfully brought into operation, it is critical that all the actors who will be involved in its implementation and its use are fully committed to its success. Validation exercises are frequently used as one of the means to provide these actors with the necessary confidence in the concept.

These actors, or stakeholders, will be different according to the nature of the operational concept and its intended rôle. If the stakeholders are identified at an early stage in the validation exercise, and their needs considered at critical points, the eventual outcome is more likely to be acceptable and convincing to all, thereby minimising the problems for the eventual operational implementation of the concept.

The exercise framework highlights key points where stakeholder involvement will be valuable.

The stakeholders provide five principal viewpoints for the validation

- Airline or other airspace user.
- Airports.
- Pilot.
- Air traffic controller.
- Air traffic service provider.

The ATM Problem and performance shortfalls will be understood and judged differently according to these viewpoints, so it is important that these are all taken into account in setting the aims of the validation exercise.

c) Criteria for Application.

This activity should always be applied.

d) Inputs.

- Information from the Programme Manager and previous similar validation exercises.

e) Outputs.

- A list of nominated individuals and organisations who are informed of their responsibilities with regard to the validation exercise.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.3: Identification of Validation Aims****a) Purpose.**

The purpose of this activity is to clarify what is to be achieved from the validation exercise.

b) Description.

The validation aim is an unambiguous, qualitative definition of what is to be achieved through the conduct of the validation exercise. In the context of ATM validation, it is generally to provide information that demonstrates the feasibility of the ATM operational concept and that the concept provides a solution to the specific ATM problem.

There are many different stakeholders, the actors in the ATM system (such as air traffic service providers, pilots and airport operators) and other parties involved (e.g., people living in the neighbourhood of an airport). These actors may all have different viewpoints on the problem and different views of what will represent a satisfactory solution.

Therefore, before starting any validation exercise, the aim has to be clearly understood, agreed and stated unambiguously. The main contributors to this process will be the Programme Manager, developers and the validation team. It may be necessary to involve other stakeholders, if their views are critical to the acceptance of the operational concept. Agreement needs to be reached on the nature of the ATM problem and what factors will demonstrate that the problem has been addressed. The validation aim, which may be expressed as a high-level question, will define this clearly and unambiguously. An unambiguous statement of the validation aim will reduce the risk that the validation exercise will not deliver what was expected.

As results of the exercises may need to be measured against failure condition, it is strongly recommended to analyse which targets/objectives are expected to be met in order to establish at this stage the Failure Criteria (failure threshold) for the validation exercise.

c) Criteria for Application.

This activity should always be applied.

d) Inputs.

- Outline aims for the exercise provided by the Programme Manager for clarification by the Validation Team.

e) Outputs.

- Description of the Validation Aim of the Exercise (including Failure Criteria) as clarified by the Validation Team²³.

²³ Appendix-7 provides examples and good practices on Validation Aims.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.4: Identification of Validation Objectives****a) Purpose.**

The purpose of this activity is to convert the Validation Aim, couched in qualitative terms, into quantitative objectives that can be measured in a validation exercise²⁴.

b) Description.

A validation objective is the formulation of the validation aim in terms of measurable factors. Within the 5th FP, one or more of the following eight high-level ATM2000+ strategic objectives provide the basis for identifying the measurable factors in ATM validation exercises:

- Safety.
- Economics.
- Capacity.
- Environment.
- National security and defence requirements.
- Uniformity.
- Quality.
- Human involvement and commitment.

These ATM strategic objectives indicate target directions in the further evolution of the European ATM system. They only give a direction of improvement for the factor (such as, that capacity should be increased), without stating specific numbers. Quantification of these objectives will be different for specific geographic areas and timeframes. For capacity and safety, the Performance Review Commission (PRC) and the Safety Review Commission (SRC) will set these targets. It is generally up to the local parties in ATM development to endeavour to meet the targets set for their area by introducing new ATM operational concepts (which need to be validated).

For each validation exercise, the high-level objectives that are associated with the validation aim should be related to these quantified targets. Related to these strategic objectives are questions that contribute to the validation aim. A question related to "safety" could be for example "will the new operational concept reduce safety?".

Since it usually proves difficult to answer these questions immediately, lower level objectives need to be derived. The lowest level objectives need to be parameters that can be measured using a known technique. This may mean that more than one detailing step is required. The decomposition should continue until the lowest level objectives are measurable and related to elementary ATM items. This is when feasible metrics and indicators can be identified.

The relationship between high-level objectives, low-level objectives and metrics is illustrated in Figure 10. In some cases, there may not be a one-to-one relationship between the levels.

²⁴ It is important to consider how to address unexpected occurrences during exercises runs, leading to objectives different to the originally planned.

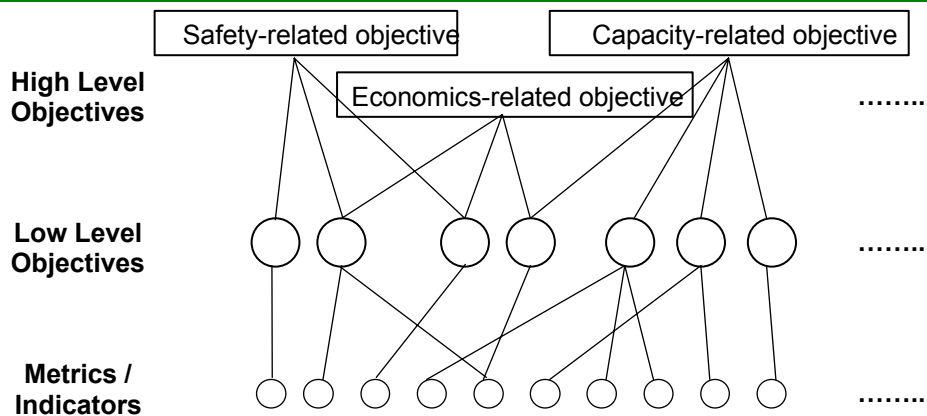
STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.4: Identification of Validation Objectives**

Figure 10 Relationship between Higher Level Objectives, Lower Level Objectives and Metrics/Indicators

The lowest-level objectives and associated metrics and indicators are defined in an iterative process. On the one hand it may be difficult/impossible to find a feasible metric/indicator for a specific objectives (for example, because the validation platform does not provide it) and on the other hand, metrics and indicators that are feasible may require a re-definition of the objectives. If the process of iteration indicates that the objectives may need to be changed, the validation aim needs to be critically examined to ensure that it is reasonable and that any changes do not reduce the value of exercise.

In addition to strategic and lower level objectives there will be subsidiary objectives in many validation exercises. These objectives are not related to the strategic objectives and usually result from specific requirements by the customer. They may well depend on the validation technique being employed, i.e., the node on the validation route map (see Figure 3). For example, a real-time simulation may have the subsidiary objective of increasing confidence in the user community that a new tool will easily integrate with existing processes. An operational exercise may have the subsidiary objective of providing training for the future users of a new tool.

c) Criteria for Application.

This activity should always be applied.

d) Inputs.

- Validation Aim.
- ATM 2000+ Strategy Objectives.
- Performance Targets.

e) Outputs.²⁵

- High-level objectives.
- Low-level objectives.

²⁵ Appendix-7 provides examples and good practices on specification of Validation Objectives.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES

Activity 1.5: Establish Validation Platform Requirements

a) Purpose.

The purpose of this activity is to decide objectively the requirements that should be placed on a validation platform that: (a) are sufficient to identify the validation technique to be used and (b) can specify the precise configuration of the validation platform to be used in the validation exercise.

b) Description.

Identification of the validation aim, seen in the light of the ATM 2000+ strategic objectives, is a prerequisite to establishing the requirements of the validation platform. In addition, an understanding of the ATM problem and the operational concept directly affect decisions made at this point.

The statement of requirements for the validation platform will identify options for the validation techniques that may be used for the exercise. Normally only one technique is used within an exercise, but there may be circumstances where the use of two or more techniques (and therefore platforms) will provide the most cost-effective solution.

The validation platform requirements are determined in six sub-steps.

- Sub-activity 1.5.1 – Identify scope of ATM to be incorporated in the platform.
- Sub-activity 1.5.2 – Decide on fidelity/resolution requirements.
- Sub-activity 1.5.3 – Define Geographic requirements.
- Sub-activity 1.5.4 – Define time based requirements.
- Sub-activity 1.5.5 – Balance requirements with constraints.
- Sub-activity 1.5.6 – Consider options for validation platform.

The contents of these activities are summarised in Figure 11 and are detailed below in the six following sub-sections.

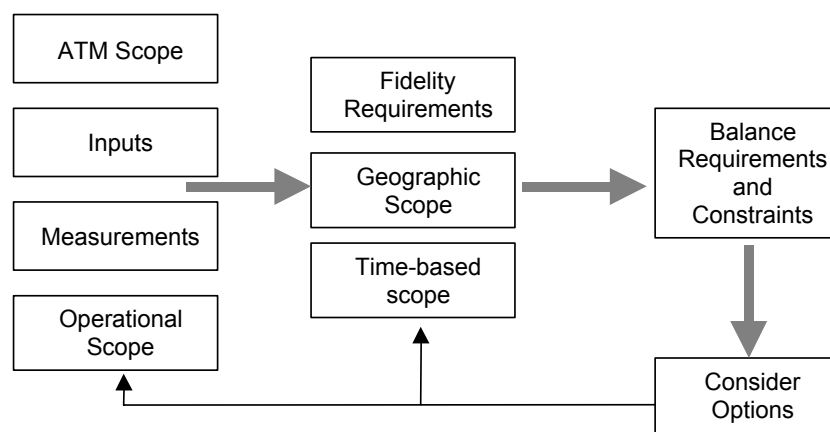


Figure 11 Determination of Validation Platform Requirements

c) Criteria for Application.

This activity should always be applied, though the level of detail produced in terms of requirements will vary according to the maturity of the ATM Concept being validated.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****d) Inputs.**

- Descriptions of the ATM Problem, the ATM Concept and the validation aims and objectives of the exercise.
- There is likely to be iteration between this activity and activity 1.6.

e) Outputs.

- A statement of requirements for the validation platform.
- Initial selection of the validation technique to be used.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES

Activity 1.5: Establish Validation Platform Requirements

Sub-Activity 1.5.1: Identify Scope of ATM to be Incorporated in the Platform

a) Description.

The inputs to this sub-activity are the statements of the ATM Problem and ATM Concept.

First of all the ATM Concept needs to be broken down into the lowest level components which can be mapped on to the invariant processes from the description. The same process must then be followed for the ATM Problem, paying particular attention to the relationships between the Concept and the Problem to ensure that all components are covered. Figure 12 illustrates the outcome of this process.

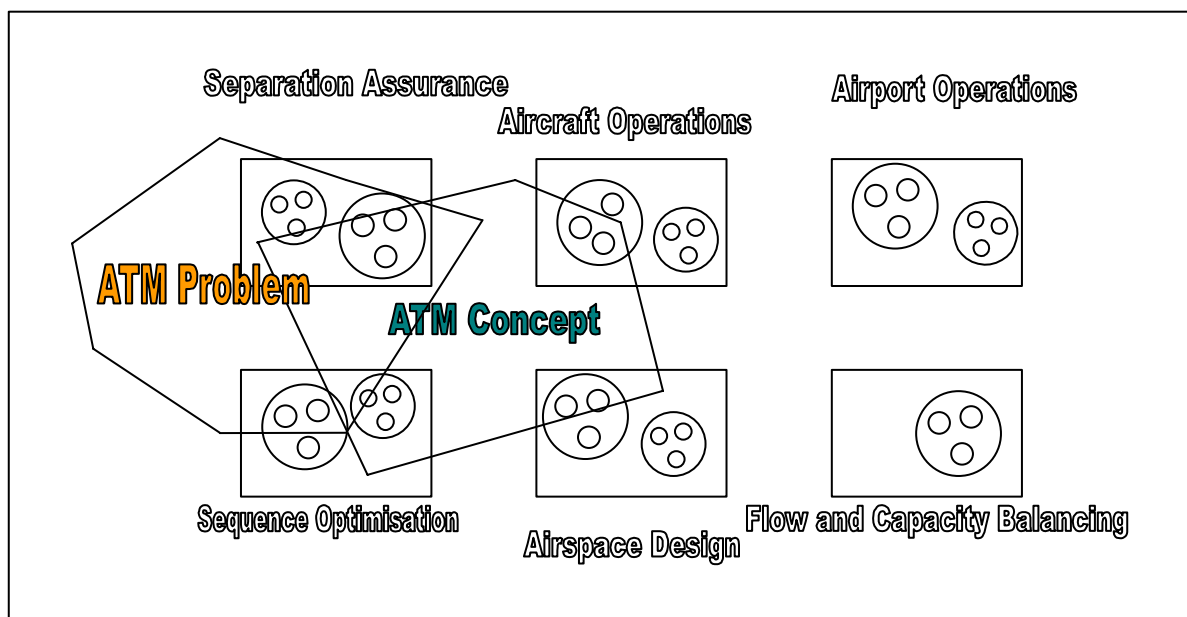


Figure 12 Mapping Concept and Problem to Invariant Processes

Figure 12 shows a system in which six of the Invariant Processes have been broken down into their major components and these have been further broken down into sub-components. The figure illustrates the parts of the complete ATM process that form a particular ATM operational concept: these are significant parts of the Separation Assurance and Sequence Optimisation Invariant Processes and smaller elements of aircraft operations and airspace design. It incorporates components and sub-components from four of the Invariant Processes. This concept has been developed to address an ATM problem also mapped on to the Invariant Processes, these involving one major component of the Separation Assurance and two sub-components of the Sequence Optimisation Invariant Processes. This particular breakdown has no overlap between the components representing the problem and those representing the operational concept. It is quite possible that there could be an overlap, for example where the problem entails controller workload and the operational concept involves the use of new procedures for the controller.

The output from this sub-activity is the set of components forming the concept and the problem mapped in terms of the Invariant Processes.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.1: Identify Scope of ATM to be Incorporated in the Platform****Identify Inputs.**

The inputs to this step are a statement of the aims of the experiment and the validation scenario.

The inputs to the platform need to be established. The exercise aim is normally to validate the concept under a range of circumstances, as represented by the validation scenario(s). It must be possible to input the required data to represent these scenarios. There may also be a requirement to assess the concept in terms of its sensitivity to certain factors. The factors may need to be varied in the course of an exercise, in which case it must be possible to vary the related inputs.

All of the ATM components identified in the previous step need to be examined to establish if any particular inputs are required to configure them at initialisation or if they are read in periodically during the simulation run to determine the course of the exercise.

Finally the inputs identified need to be compared with the format in which it can be provided. If this format does not match input requirements, an additional process within the platform is needed in order to convert them to a suitable format. A simple example of this is that if a model uses metres to represent height, but the controller inputs data as Flight Level, the value will need to be converted. An example of a more complex requirement would be when wind needs to be represented as a function of height and a process is needed to convert height to the value for wind at the height of interest.

The output from this step is a list of additional processes to be included in the platform that are required to provide the input data in the correct form.

Identify Measurement Requirements.

The input to this step is the description of the measurements to be taken during the validation exercise.

The measurements to be taken should be reviewed. Each one needs to be compared with the components identified as part of the ATM operational concept and Problem. If the measurement cannot be directly measured from these components, an additional system to facilitate the made will be required. For example, if workload is to be measured, an additional system such as an ISA (Instantaneous Self Assessment) panel and supporting software will need to be included in the statement of requirements for the platform.

At this stage it may be useful to consider the requirements for recording input data and also the tools required for analysing the measurements.

The output from this step is a list of processes required to measure and record data from the validation exercise.

Define Operational Scope.

The input to this step is the statement of the aims of the validation exercise defined in Activity 1.3.

Finally, the components to be incorporated in the platform need to be aligned with the Validation Aims, identified in Activity 1.3. If the aim of the validation is to assess the impact of the system on more than one phase of flight, for example on gate-to-gate operations, it may be necessary to include further Invariant Processes to provide the overall ATM context of the validation. These aspects would only be represented at a coarse level of

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES

Activity 1.5: Establish Validation Platform Requirements

Sub-Activity 1.5.1: Identify Scope of ATM to be Incorporated in the Platform

granularity. The ATM context indicated by the Validation Aims should be reviewed to ensure that it is reasonable and does not increase the cost of the validation for limited additional benefit.

The output is a list of any additional invariant processes required to represent the intended operational scope for the validation exercises.

List Components to be Incorporated in Platform.

The ATM components to be represented on the platform are the sum of all the components required to represent the following:

- The ATM operational concept,
- the ATM problem;
- the operational scope;
- the inputs and any related processes to ensure these inputs are in the correct form;
- any additional systems required for the measurements to be taken.

This is illustrated in Figure 13.

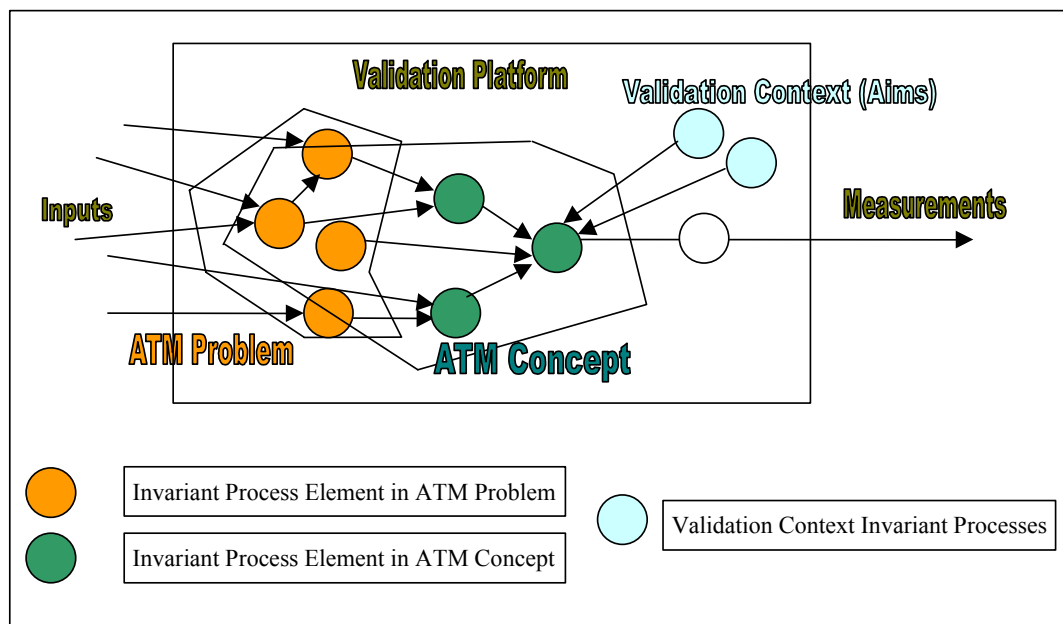


Figure 13 Components to be Incorporated in Platform

b) Criteria for Application.

This step should always be followed.

c) Inputs.

- Descriptions of the ATM Problem and the ATM Concept.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.1: *Identify Scope of ATM to be Incorporated in the Platform***

- The inputs to the model.
- The measurements to be taken.
- The Operational Context for the exercise.

d) Outputs.

- A list of components that will be represented on the ATM Platform.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.2: *Decide on Fidelity/Resolution Requirements*****a) Description.**

The ATM components identified in Sub-activity 1.5.1 can be represented in the validation platform in a number of ways. The principal categories are listed below:

- Real systems: real aircraft, controllers or pilots may be required;
- Pseudo roles: for some purposes the role of the pilot can be adequately represented by a pseudo-pilot;
- Models: a model will allow the system being represented to interact dynamically with the other systems represented on the platform;
- Input files: are suitable to represent an initial situation or changes which only occur on a limited basis;
- Assumptions: it is possible to make an assumption on the effect of a given factor given experience from exercises or operations and thereby simplify the experiment.

The individual responsible for deciding or determining the validation platform requirements needs to decide on the method to be used to represent each of the components that will form part of the validation platform. The decision on representation will depend on the fidelity requirements, the dynamics of the scenario and the availability of models and resources.

When considering the fidelity required, the following issues are relevant:

- Level of understanding of the system to be represented: if the system is not well understood, it will not be possible to develop a representative model. In this case either a real actor is required or a simple high level model would be more suitable, depending on the relevance of the system to the results of the exercise;
- Participation of actors: if actors are key to the validation of the concept and are participating in the exercise, it is important that the platform and environment is sufficiently representative for them to work effectively;
- Confidence requirements: if it is important that a high level of confidence is achieved, the fidelity will need to be high;
- Impact that a particular system has on the measured results will also indicate how true a representation is required. If the system only indirectly relates to the results, a lower level of fidelity may be sufficient.

An example of varying levels of fidelity can be seen in the representation of an aircraft within validation exercises.

- Commercial aircraft in operational use may be involved, particularly if the exercise does not involve active control of the aircraft using exercise procedures;
- Research aircraft may be used in restricted airspace under full control of a pre-operational system;
- Manned high fidelity model (qualified pilot), for example a non-dynamic flight simulator. If the concept is dependent on the pilot being able to operate the new system while flying the aircraft then a qualified pilot will be required;
- Manned low fidelity model (pseudo pilot), where the controller is the focus of the

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.2: *Decide on Fidelity/Resolution Requirements***

system a pseudo-pilot may be sufficient;

- Unmanned low fidelity model is used in a fast time simulation where the interaction between the pilot and the controller are well understood and characterised.

When Invariant Processes and components are introduced merely to provide the operational scope of the validation, for example, if the impact of en-route tools is to be assessed in the context of gate-to-gate operations, a relatively low fidelity representation of Airport Operations would be sufficient.

b) Criteria for Application.

This step should always be followed.

c) Inputs.

- A description of the ATM Concept.
- The list of ATM components resulting from Activity 1.5.1.

d) Outputs.

- The required fidelity level.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.3: Define Geographic Requirements****a) Description.**

The geographic location and extent of the validation exercise has an impact on the requirements for the validation platform. The geographic scope will be determined by the physical extent of the ATM Problem, operational concept and operational scope. The range of geographic scope is follows:

- Local, covering an area of a few square metres. This would be the case with Controller Working Positions or the layout of aircraft control panels;
- Involving only one phase of flight, for example terminal control;
- Limited number of Sectors. The impact of tools on the workload of a controller will need to be assessed on aircraft using one or two sectors;
- Multi-sector. A complex re-sectorisation will require a simulation covering all the sectors which will be affected;
- Multiple ACC. This geographic scope will need to be considered when the system being assessed will change co-ordination procedures between different ACCs.
- Global perspective. This scope should be used when the future ATM systems envisaged make use of airborne generated data and this data will be used across the whole flight. The data requirements for long-haul flights may therefore need to be considered.

b) Criteria for Application.

This step should always be followed.

c) Inputs.

- Description of the ATM Problem and the ATM Concept.

d) Outputs.

- The geographical location to be represented.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.4: Define Time-Based Requirements****a) Description.**

The validation needs to be assessed in the correct time frame (future date) and correct elapsed time. The time frame to be used will be related to the expected/planned Operational date for the concept and its expected operational life. The time frame is needed to ensure that the models and systems being used are representative of the systems to be in operation at that date.

The elapsed time over which the validation platform should exercise the ATM Concept being validated, will be determined by the following issues:

- Operational cycle of the Problem and operational concept relating to how long it takes to see the concept to start up, to execute and the outcome of events to be observed;
- Duration of events of interest;
- Rarity of events. If a concept is to be validated using Monte Carlo type techniques it is important that the simulation is run for a sufficient number of the rare events to be observed;
- If a number of measurements are required, then the validation needs to be run for a long enough period for the measurements to be made. This may require some thought if the measurements are event based rather than time based.

b) Criteria for Application.

This step should always be followed.

c) Inputs.

- A description of the ATM Concept and its In Service Date.

d) Outputs.

- One or more future dates.
- Duration in terms of minutes, days or months.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.5: *Balance Requirements with Constraints*****a) Description.**

The output from the first four sub-activities of Step 1 described above makes up the requirements for the validation platform. These requirements need to be considered in the light of the constraints on the validation. The constraints are in terms of cost, resources and feasibility.

The requirements should be assessed to determine the type of validation technique needed. This will give an indication of the costs and resources required. The costs should be compared with the budget allowed for the validation exercise. If only limited budget is available it may not be possible to have a full large-scale real-time exercise, even though that appears to be the ideal option. If there is insufficient budget the platform requirements should be reviewed in detail to establish whether a lower-cost alternative could provide satisfactory results.

Similarly the requirements need to be reviewed against their feasibility. For example if there are no valid fast time models available to represent all the components identified in the requirements when fast-time assessment seems most appropriate, an alternative approach should be considered. Where real systems need to be involved in the validation exercise, safety issues need to be considered.

The project developing the concept frequently has a view on the type of validation technique that should be applied at a particular point in the programme. If this is the case, the list of requirements developed through the objective process described above needs to be compared with the proposed validation technique and any major differences resolved with the development project manager.

Consideration of the requirements against previous experience will indicate the validation technique that will be suitable for the validation exercise.

Details of the fundamental validation techniques are given in Appendix 3 under Activity 1.5.5 - Validation Techniques.

b) Criteria for Application.

This step should always be applied, unless it is clear from previous experience that the requirements can be met easily.

c) Inputs.

- Initial list of platform requirements.

d) Outputs.

- Refined list of platform requirements.
- Initial view of the validation technique for the exercise.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.5: Establish Validation Platform Requirements****Sub-Activity 1.5.6: Consider Options for Validation Platform****a) Description.**

The list of requirements then needs to be reviewed in the light of the available validation platforms. A list of platforms is provided in Appendix 3 under Activity 1.5.6 - Validation Platforms. There are two options possible at the conclusion of this assessment. These are that the list of requirements is:

- Feasible for a single Validation Platform;
- Not feasible within a single Validation Platform.

It is not a problem if the required validation exercise is not possible on a single platform. A cost-effective validation activity may combine the results from exercises on more than one type of platform. There are therefore three options available if the set of requirements identified cannot be met by a single platform. These are to:

- Reduce scope of exercise - this may require a number of iterations and the manager should ensure that the reduction does not compromise the value of the exercise and should investigate ways for the validation programme for the concept eventually can achieve all the validation objectives;
- Use more than one type of platform;
- Adjust the platform, so that the requirements are met.

This sub-activity should be considered as a first step in an iterative process related to the selection of the platform and validation technique, which will not be concluded until later in Step 2.

b) Criteria for Application.

This step should always be applied unless it is clear that the requirements can be easily met.

c) Inputs.

- List of requirements and platforms that are available.

d) Outputs.

- Statement on feasibility of the proposed validation exercise.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.6: Identification of Metrics and Indicators****a) Purpose.**

The purpose of this activity is to convert the low level objectives into metrics and indicators that can be measured on a validation platform (examples in Appendix-3).

b) Description.

As stated previously, the high-level objectives cannot normally be measured directly in a validation exercise. These objectives, therefore, are broken down into low-level objectives as described above in order to identify system parameters related, directly or indirectly, to the high-level objectives of interest. The measurements of these system parameters are known as metrics and indicators.

A metric is a parameter that can either be measured directly, or be calculated from several measurements, and that expresses a significant quality of a system. A 'system', in the context of ATM, may also include the human operators. An indicator is a metric that is only indirectly related to the objective(s) of interest. Consequently a change in the value of an indicator does not necessarily lead to a definite conclusion that the related objective has been achieved. An example of this is controller workload in relation to airspace capacity; a reduction in workload is an indication of the potential for capacity to increase. The measurement of workload, however, is not a direct measurement of capacity.

As many different metrics as possible are needed to provide an unambiguous answer. The factors determining the selection of metrics and indicators depend amongst others on the selection of the validation platform. Metrics will usually be specific for each validation and will depend on the types of data to be collected. Data types applicable to ATM validation exercises and examples of metrics used in Fast-time and Real-time exercises are provided in Appendix 3 under Activity 1.6 - Data Types Applicable to ATM Validation Exercises. Also requirements for specific metric/indicators may imply requirements on the validation platform. Thus the selection of metrics and indicators (and the validation platform) will be determined through an iterative process.

c) Criteria for Application.

This step must be applied for fast-time and real time techniques. It may be useful in literature searches and judgemental techniques.

d) Inputs.

- Low-level objectives.

e) Outputs.

- Metrics and indicators to be measured in the validation exercise.²⁶
- There is likely to be iteration between this activity and activity 1.5.

²⁶ Appendix-7 provides examples and good practices on specification of derived metrics and indicators.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.7: Identification of Hypotheses****a) Purpose.**

The purpose of this activity is to convert the low-level objectives into a rigorous mathematical framework that can be tested using statistical techniques.

b) Description.

In the context of the validation exercise, a hypothesis is a proposition about operational concepts couched in statistical terms, constrained by the metrics that are available. Further information on the format of statistical hypotheses can be found in and statistical text books.

Hypotheses arise from a direct consideration of the low-level objectives, which are directly linked to the metrics and indicators that the hypothesis will use. Comparisons of the values of metrics and indicators in different operational concepts will decide whether a hypothesis is accepted or rejected.

The hypothesis must relate to the success of the operational concept in addressing the problem; otherwise it is of little use.

c) Criteria for Application.

This step is appropriate to fast time and real-time exercises.

d) Inputs.

- Low-level objectives.
- Metrics and indicators.

e) Outputs.

- A set of hypotheses for the exercise.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.8: Definition of High-Level Experimental Design****a) Purpose.**

The purpose of this activity is to set the framework for defining the individual measured runs to be performed in the exercise.

b) Description.

The high-level experimental design is a broad statement of the experimental approach to be used in the validation exercise. This will determine a number of issues, (such as whether a baseline system will be used), that are required for the detailed validation exercise design.

In general, every component of the ATM system can operate at various “settings”, affecting the resultant quality of service of the ATM system (such as capacity, efficiency, safety, etc.). Experimentation with a new ATM operational concept allows the experimenter to adjust the settings of the ATM system components in a systematic manner and to learn which factors have the greatest impact on the resultant quality of service. Using this information, the system can be optimised.

An experimental design represents a similar systematic and controlled approach to data collection, in order to extract the maximum amount of unbiased information (i.e. to support valid, reliable and usable conclusions) regarding the factors affecting the ATM operational concept studied from as few (costly) exercise runs (or observations) as possible. This allows a validation exercise to be planned, that supports the aim and that can be executed without significant risks with regard to, for example, effort and time. Consequently the experimental design is a crucial phase of a validation exercise.

The high-level experimental design specifies in outline (amongst others):

- Experimental design type. There are several standard experimental design types that can be chosen. They determine different ways of obtaining the optimum settings for the different factors that affect the ATM processes. This includes, for example, what factors are varied and how that is done, such as independently or in combination. Also the high level experimental design establishes the type of analysis:
 - Relative vs. absolute (e.g. is a baseline concept used or not?);
 - Quantitative vs. qualitative analysis.
- An estimate of the number of exercise runs that will be needed. At this stage it is possible to have an overview of the size of the validation exercise (thus having an idea of the costs) and to identify potentially insoluble problems, especially related to obtaining sufficient representative participants.

c) Criteria for Application.

This step should always be conducted for fast-time and real-time exercises. It will not always be applicable to literature search and judgemental techniques.

d) Inputs.

- Validation aims and objectives.
- Metrics, indicators and hypotheses.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.8: Definition of High-Level Experimental Design****e) Outputs.**

- Experimental design type.
- Approximate number of necessary exercises runs.
- High-level experimental design requirements.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES

Activity 1.9: Determine Operational and Statistical Significance

a) Purpose.

The aim of this activity is to obtain operationally significant results, with the required level of statistical significance, using the minimum number of exercise runs.

b) Description.

Operationally-significant changes are those that have a significant impact on ATM operations and are measured by looking at the high-level objectives. Statistically-significant changes relate purely to the behaviour of the measures involved, independently of their ultimate context, and are measured in terms of statistical confidence. A statistically-significant change does not necessarily imply an operationally-significant change.

For example, establishing with statistical confidence that an activity may be performed significantly faster may have no bearing on significantly improving ATM operations because the time differences are too small to affect the high-level objectives noticeably, due to their being swamped by the behaviour of other parts of the process. Therefore, an understanding of how the high-level objectives are related to the metrics is a vital factor in deciding what level of statistical significance is required.

Statistical significance associated with a hypothesis is usually expressed as a percentage chance that the null hypothesis is correct. This is referred to as confidence, and indicates to what extent the measurements represent a systematic change rather than occurring purely by chance.

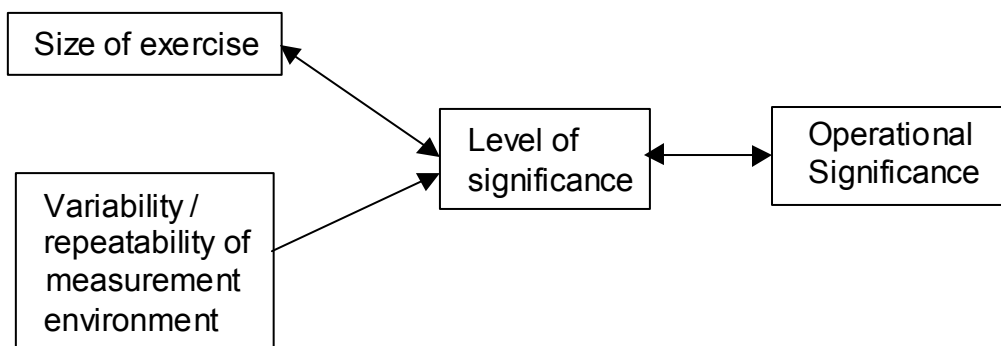


Figure 14 Relationship between Size and Significance of Exercise

Figure 14 illustrates how the level of significance relates to the size of the exercise and underlying variability in the measurements. All other things being equal, as the size of the exercise increases, then the significance of the results will increase. However, this is not a linear relationship; for example, halving the probability of error may imply quadrupling the sample size. If it is assumed that the magnitude of the variability is reasonably understood (from past experience) then the size of the exercise can be established once a level of significance is set.

The target level of statistical significance required is not governed by strict rules. In combination with the fidelity of the model used, it will be a factor in the level of confidence in the outcome of the validation. Generally, the more mature the ATM operational concept that is assessed, the more certain the results need to be, and thus the more confidence the performer wants to have in the (key) measurements.

STEP 1: IDENTIFICATION OF VALIDATION AIMS, OBJECTIVES AND HYPOTHESES**Activity 1.9: Determine Operational and Statistical Significance**

In the ATM validation community, a 95% target statistical significance level is usually accepted as sufficient. The main reason is that, in practice, 95% seems to be achievable within sensible costs, and that allowing a 1-in-20 chance of error is generally the default for a wide range of statistical applications. However, it cannot necessarily be selected on the basis of cost; safety issues may explicitly drive the level of confidence required. If for some reason 95% cannot be achieved, then establishing a difference with a 90% significance level may be worth reporting this result, but it is better to describe this as a trend that has been observed.

The statistical significance achieved is usually derived from the distribution of the comparison values (for example each difference between a metric in the baseline concept and the same metric in the new concept). There are many different tests to determine the statistical significance from this distribution. The type of test will depend on several issues.

- The nature of the measure (e.g. whether it is continuous or discrete, absolute or ranked. See Appendix 3 Activity 1.6 - Data Types Applicable to ATM Validation Exercises);
- What assumptions have to be made (e.g. whether or not the standard deviation is known);
- How the null hypothesis is expressed (e.g. does it relate to a change in location, variation or extreme value).

For some types of validation, the rigorous statistical framework described above may be impossible to apply. In such cases it may be necessary to rely upon the judgement of experts in to support a less rigorous application of the statistical tools.

c) Criteria for Application.

This step should always be applied when hypotheses and metrics/indicators have been identified.

d) Inputs.

- An understanding of the high-level design requirements of the exercise.
- The critical values of system parameters that will make an operational impact on ATM performance.

e) Outputs.

- An agreed target value of statistical significance to be achieved in the exercises and targets for operational significance.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.1: Selection of the Technique and Facility and Detailed Experimental Design****a) Purpose.**

The purpose of this activity is to provide the information needed for detailed planning of the exercise runs and the preparation of the Measurement and Analysis Specification (MAS).

b) Description.

The first step in planning and preparing the validation exercise is to make a final selection of the validation technique and facility that will be used in the validation and to draw up a Detailed Experimental Design. This step is carried out in three substeps, as follows:

- Sub-activity 2.1.1 - Final Selection of Validation Technique.
- Sub-activity 2.1.2 - Selection of Validation Platform.
- Sub-activity 2.1.3 - Detailed Experimental Design.

The contents of these activities are detailed in the three following sub-sections.

c) Inputs.

- Validation aims, objectives and hypotheses.
- Selected metrics.

d) Outputs.

- Decision on platform to be used for the validation.
- Detailed Experimental Design.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.1: Selection of the Technique and Facility and Detailed Experimental Design****Sub-Activity 2.1.1: *Final Selection of Validation Technique*****a) Description.**

The platform requirements specification and the high-line outline scope of the experimental design will be used to make a final selection of the validation technique to be used. This selection will also take into account platform availability. The choice will be between the different techniques presented in Appendix 3 Sub-Activity 1.5.5 –Balance Requirements with Constraints.

b) Criteria for Application.

This sub-activity will always be applied.

c) Inputs.

- Platform requirements.
- High-level experimental design.

d) Outputs.

- Selection of technique.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.1: Selection of the Technique and Facility and Detailed Experimental Design****Sub-Activity 2.1.2: Selection of Validation Platform****a) Description.**

The validation platform requirements are compared against a list of available platforms. A selection is then made of the most suitable available platform that meets the requirements. A suggested list of platforms, indicating their capabilities and possible generic metrics, is provided in Appendix 3 under Sub-Activity 1.5.6 – Validation Platforms.

b) Criteria for Application.

This will always be performed for fast time and real time simulations, as well as shadow mode and operational trials.

c) Inputs.

- Validation Platform requirements.
- List of platforms that are available.

d) Outputs.

- Selection of Platform.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.1: Selection of the Technique and Facility and Detailed Experimental Design****Sub-Activity 2.1.3: Detailed Experimental Design****a) Description.**

The aim of experimental design is to provide a structure for the specification of the exercise runs, to ensure that the results from the runs provide robust evidence to confirm or deny the hypotheses developed for the validation exercise. Robustness can be compromised both if the measurements made are not clearly related to the factors of interest and if the number of measurements is insufficient and representative of the future use of the system. These aspects must be addressed correctly in the experimental design so that clean and robust measurements are used to reach accurate and sound conclusions.

The method used for detailed experimental design will vary according to the type of technique to be used in the exercise and the resources available, including the capability of the validation platform, the availability of experienced participants, financial resources and available time. A brief summary of the detailed experimental design sub-activity will be presented here; further details are provided in Appendix 4 under Sub-Activity 2.1.3.

The inputs to this sub-activity are illustrated in Figure 15.

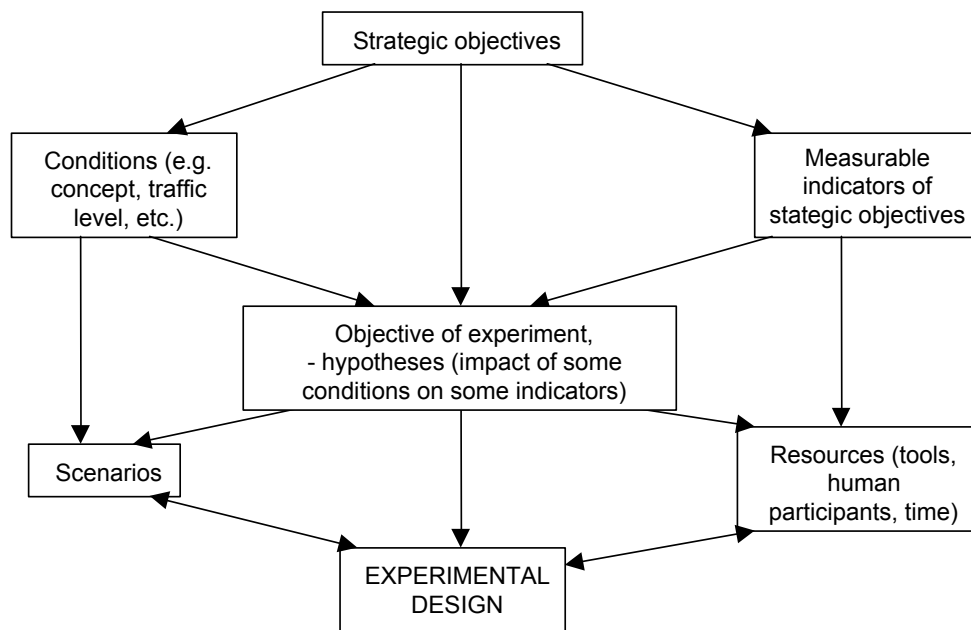


Figure 15 Inputs to Experimental Design

Two questions that must be addressed in this sub-activity are how to select scenarios and resources and how to best use them to meet objectives. The choices to be made include the validation technique, the platform, the number of scenarios (encompassing geographic area, traffic level, etc.), the number of participants involved and any other aspect that may have an impact on the achievement of the objectives in terms of budget and time. These choices differ depending on the technique used. The factors that need to be specified for each type of validation technique can be summarised as follows:

- Judgemental techniques: the selection of the experts (including number, experience and duration of involvement) and the design of the material to be used by the experts.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.1: Selection of the Technique and Facility and Detailed Experimental Design****Sub-Activity 2.1.3: Detailed Experimental Design**

- Fast-time simulation: the time allocated, the modelling tool that will be used, the number of geographic areas to be simulated, the number and type of traffic samples, the reference baseline used and the workload calibration (if workload is one of the variable to be measured).
- Real-time simulations: the size of the exercise as a function of the number of participants; the number of runs; the number, size and type of the geographic areas to be simulated; the number of traffic samples; the type of participants and the manner in which they are to be prepared; the reference baseline used; and the content and degree of realism of the scenarios.
- Shadow mode: the manner in which the participants are to be prepared; the length, time and conditions of the trials; the type of traffic under shadow-mode control; and the possible installation of new equipment in aircraft.
- Operational trials: the scale of the trial in terms of space, the scale of the trial in terms of the flights concerned, the conditions and length of the trial, the preparation of participants.

As stated above, guidelines and recommendations for making each of these selections is provided in Appendix 4 under Sub-Activity 2.1.3.

b) Criteria for Application.

This sub-activity always needs to be considered, with the exception of literature studies, for which it is not applicable.

c) Inputs.

- Descriptions of the ATM Problem and ATM Concept.
- Validation aims, objectives and hypotheses.

d) Outputs.

- Specifications for the exercise runs to be performed. These specifications will vary depending on the validation technique that will be used:
 - For judgemental techniques, the specifications will consist of the list of experts to be consulted, and the questions and concepts to be discussed.
 - For fast-time techniques and real-time simulations, the specifications will be of the input information to the validation platform. For example: traffic samples, environment characteristics (e.g. airspace organisation), precise configuration (e.g. components) of the validation platform, organisation of the exercise(s), identification of participants, measurements to be taken, etc.
 - For shadow mode and operational techniques, the specifications will be for timetable and nominated operational staff and, for example, if aircraft are the focus of the particular validation, flight plans.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.2: Preparation of Outline Plan (including Measurements and booking of Facilities)****a) Purpose.**

The purpose of this activity is to summarise the exercise runs planned and to draft the type of measurements that will be made (the final decision on the measurement to be made will take place on activity 2.5.1).

b) Description.

This activity is normally applied to real time exercises (real time simulations, shadow mode trials and operational trials) for which significant resource commitments and planning activities are required.

c) Inputs.

- Validation objectives, hypotheses and metrics.
- Detailed experimental design.
- Selected validation platform.

d) Outputs.

- Outline plan.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****a) Purpose.**

The purpose of this activity is to produce a specification for a scenario that exercises the ATM concept as required by the validation exercise objectives in an operationally realistic manner.

b) Description.

The third step in planning and preparing the validation exercise is to produce the scenario specification. This step is carried out in six substeps, as follows:

- Sub-activity 2.3.1 - collate and review inputs to scenario definition;
- Sub-activity 2.3.2 - develop baseline and advanced operational concept scenarios;
- Sub-activity 2.3.3 - develop traffic samples;
- Sub-activity 2.3.4 – list scenario assumptions;
- Sub-activity 2.3.5 – review scenario definition;
- Sub-activity 2.3.6 - get approval for use of scenario.

The contents of these activities are detailed in the three following sub-sections.

c) Criteria for Application.

This activity will always be performed in order to specify the scenario to be considered.

d) Inputs.

- Description of the ATM Problem and the ATM concept.
- Validation aims, objectives, hypotheses and metrics.
- Experimental design.
- Platform requirements document.

e) Outputs.

- Scenario specification.
- Traffic samples (if required).

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.1: Collate and Review Inputs to Scenario Definition****a) Description.**

The first step to producing the scenario specification is to collate and review the inputs to the scenario definition (example in CARE/ASAS/ISDEFE/02-030, 'Initial Validation Framework & Scenario Report', Annex B, pages B1-B4). In collating and reviewing the inputs to the scenario definition the validation team should:

- Restate the validation aims;
- Consider Experimental Design Specification and Metrics;
- Review the Platform Requirements Document.

The reason these issues should be reviewed at this stage are now described:

- *Validation Aims* - It is always important before starting any task to remind a team of the requirements it should be addressing. This process of clarifying the aims and developing the objectives must necessarily have captured the intended Operational Improvements and any Performance Targets that the ATM Concept being validated should achieve. The scenario definition document should restate the validation aims and high level objectives.
- *Experimental design and metrics* - The experimental design specification defined in activity 1.8 and modified in activity 2.1 provides information needed for defining the scenario. First and foremost it will indicate whether the baseline operational concept needs to be explicitly represented in the validation exercise.
- The metrics identified in activity 1.6 needs to be reviewed and considered when developing the scenario to ensure that the measurements of the metrics will be sensitive to changes in the input parameters to the scenario (examples in Appendix-3).
- The statistical significance requirements, the experimental design and the statistical tests to be applied will affect the number of independent measurements required. This will determine the number of traffic samples required.
- *Platform requirements* – The level of detail and flexibility that can be captured in the scenario will be determined by the platform to be used. The analysis undertaken when developing the platform requirements in activity 1.5 can be re-used when developing the scenario definition and will have given initial ideas on the likely characteristics of the scenario.

b) Criteria for Application.

This sub-activity will always be applied.

c) Inputs.

- Validation aims.
- Experimental design specification and metrics.
- Platform requirements.

d) Outputs.

- Clear understanding of the requirements for the scenario.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.2: *Develop Baseline and Advanced Operational Concept Scenarios*****a) Description.**

The detailed analysis of the ATM Concept from activity 1.1 and the Platform Requirement Specification from activity 1.5 should be to hand when performing this sub-activity. The scenario comprises three main elements:

- ATM Concept, (different for the Baseline concept and the Advanced concept being validated);
- ATM Environment comprising geographical scope and Timeframe (the same for both concepts);
- Events (the same for both concepts).

Each of these elements needs to be considered in turn and this is described below:

Where the experimental design specifies a baseline, two concepts will need to be considered in this sub-activity, the Advanced ATM concept being assessed and the baseline ATM concept against which it is being compared. It is important that there is agreement on the baseline ATM concept. The baseline concept may be represented by current day operations or it may represent an intermediate step towards the advanced ATM concept. The concept is the fundamental element of the scenario and it needs to be adequately represented and exercised in the validation exercise. Where a baseline concept is required for the experimental design, this needs to be defined to the same level as the advanced ATM concept. In Sub-activity 1.5.1 the ATM problem and concept were analysed to establish the Invariant Processes that needed to be represented. This needs to be repeated for the baseline concept. At this point all the constituent elements (components & sub-components) of the Invariant Processes involved must be identified and the nature of their "change" from the baseline concept to the future concept must be made explicit. These "changes" are the operational improvements of the advanced ATM concept that must be demonstrated when the scenario is run. Likewise, the enablers of the operational improvements of the advanced ATM concept (technologies, human factors, procedures, etc that it is thought will enable an operational improvement to be achieved) should be clearly identified and their known characteristics and sensitivities recognised. This understanding should then be used to ensure that the scenario is devised such that the enablers are exercised thoroughly.

The term "Geographic Area" covers a number of distinct elements:

- The actual physical size and location – i.e. a plan view of the area within Europe to be covered;
- The aircraft routes to be used – i.e. where the aircraft are coming from and going to;
- The terrain – i.e. the 3D aspect of the geography;
- The weather (or meteo).

These elements are discussed in Appendix 4 under Sub-Activity 2.3.2 Geographic area.

When deciding the timeframes, within which the operational concept under investigation will be validated, there are three clear and distinct references that can be made:

- Baseline Date – i.e. the date against which the future performance of the operational concept will be compared;

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.2: *Develop Baseline and Advanced Operational Concept Scenarios***

- In Service Date – i.e. the date on which the operation concept may be expected to start operation within the European ATM system;
- Evaluation Dates – i.e. dates in the future on which the performance of the operational concept will be evaluated against a known target – e.g. demand on that date.

These elements are discussed in Appendix 4 under 2.3.2 Timeframe.

All proposed advanced ATM concepts should be able to perform effectively in normal operations. It is also important that the performance, when there are unusual events, possibly safety critical, should be considered. Types of events that might be considered here include:

- Ground or Airborne Equipment failures;
- Extreme weather conditions, storms, cross winds, fog;
- Aerial photography and surveying;
- Military missions;
- Emergency flights.

Early on in the development lifecycle, the types of contingencies may be considered using judgmental techniques in parallel to the simulation exercises. However, it is important at some stage in the development of the concept that they are considered in as realistic a situation as possible. Of course, where ATM concepts are developed specifically to improve operational performance for specific events, it is essential that these are fully represented in the validation exercise scenarios. In this activity, a decision needs to be made on whether the scenarios are to include events or represent purely normal operations. If events are to be represented, the types of event should be selected and scripts must then be prepared for the controllers, pilots/pseudo-pilots or simulation so that the events can be triggered correctly.

b) Criteria for Application.

This sub-activity will always be applied.

c) Inputs.

- Detailed analysis of ATM concept.
- Platform requirements specification.

d) Outputs.

- The key parts of the operational concept (HMI, datalink, pilot performance) that need to be exercised in the validation exercises.
- The size and location of geographic area for the validation exercise.
- The dates to be used to set the traffic levels.
- A list of any events to be represented in the scenario to script the scenario.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.3: *Develop Traffic Samples*****a) Description.**

All the information needed to generate traffic samples should now be available. The next stage is to develop the samples. This activity will also make use of information from other activities in the MAEVA five-step process. For example, the number of traffic samples needed will have been determined by the experimental design in activity 2.1 and the length of the traffic sample will have been determined in the platform requirement. Development of the traffic samples is as follows:

- Base traffic sample (the real life samples used);
- Timeframe;
- Generate traffic sample.

These activities are described below.

Traffic samples employed in simulator based Validation Exercises (fast-time and real-time) are mainly developed from actual, real-life traffic plans and/or recordings. In these cases, it is vitally important that the selected traffic samples – referred to herein as the base traffic samples – are carefully selected. Daily traffic is defined by two main criteria – the traffic pattern and the events that occurred on the selected day. In turn, the traffic pattern will be defined by the prevailing weather and the time of year, day of the week and time of the day. These three criteria – events, weather and time – are discussed in Appendix 4 under 2.3.3 Base Traffic sample.

It is recommended that two or three base traffic samples are generated. When preparing traffic samples for a validation exercise to be conducted on a real-time simulator, it must be remembered that participants should be trained on different traffic samples than those to be used in the measured runs. Thus, an additional set of traffic samples must be generated to cover this.

In order to develop traffic samples representative of the timeframes chosen in sub-activity 2.3.2 above from the base traffic samples; it is necessary to predict the traffic growth that will occur. Most organisations involved in ATM research have developed their own methodology for growing traffic; however, these tend to be country specific and often contain sensitive information that cannot be released. It is desirable to have a common assumption over all the projects in the VMP of the expected traffic growth. However, as the growth patterns will vary across Europe according to economic, geographic and capacity constraints, this ideal will be difficult to achieve. All models predicting the traffic growth over a number of years consider the following issues:

- Will the growth on the various origin destination pairs differ?
- What changes will there be to aircraft type and mix?
- What factors affecting growth need to be taken into account (alternative means of transport, economic factors, and demographic factors)?
- Are there any physical limitations to growth that need to be taken into account (airport capacity)?
- Are there any trends in growth?

The answers to the questions above are the assumptions used for the calculation of the

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.3: *Develop Traffic Samples***

growth factors applied to the base traffic samples. These assumptions should be recorded. The application of the growth model will result in the calculation of a set of annual growth figures. The overall figures estimated by the growth model to be applied to the traffic samples should also be recorded. Where there is no strong reason for using the local growth model, the STATFOR methodology is recommended as it is currently widely used across Europe.

The traffic samples will be developed using tools specific to the validation platform being used for the exercise. Operational experts should support the individual or team responsible for developing the traffic samples. Traffic samples generated should take into account any differential growth over the airspace organisation and limitations in forecast runway capacity. Detailed guidance on traffic sample generation cannot be given here, as it would need to be specific to the tools available for the data preparation and the target validation platform.

b) Criteria for Application.

This sub-activity will always be applied for fast-time and real-time validation methods.

c) Inputs.

- Size and location of geographic area selected.
- Base traffic samples for that region (from radar recordings).
- Dates for future traffic flows.
- Growth model for traffic.
- Any events to be incorporated.

d) Outputs.

- Traffic samples grown or modified in accordance with the scenario specification.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.4: *List Scenario Assumptions*****a) Description.**

Frequently, it is not possible to represent everything in the scenario to the desired level of fidelity or for the feasible range of values. In these cases the scenario designer has to make assumptions on the fidelity of representation that is deemed adequate and on the 'typical' values to use for these parameters. It greatly supports the integration and consolidation of validation exercise results if these assumptions are documented and available when the results are being analysed.

It is recommended that assumptions are included as part of the validation scenario specification document.

b) Criteria for Application.

This sub-activity will always be applied when assumptions have been made.

c) Inputs.

- Assumptions made in previous steps in activity 2.3.

d) Outputs.

- A statement of the scenario assumptions.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.5: Review Scenario Definition****a) Description.**

The scenario definition produced by following the sub-activities so far needs to be reviewed to ensure that it meets the requirements identified in sub-activity 2.3.1. The scenario definition is the set of documentation developed so far and the input files and scripts for use with during the exercise. A visualisation tool for the traffic samples will assist those involved in this activity.

The following issues should be addressed:

- *Correctness*: does the proposed scenario description address the stated level of requirements and objectives of the exercise?
- *Completeness*: does the proposed scenario content cover all necessary areas to the required level of detail?
- *Consistency*: are the contents of the proposed scenario coherently related?
- *Traceability*: do the proposed scenario areas identify the correlation between the solutions and problems they're addressing?
- *Usability*: is the defined scenario fit to be employed within a given validation process and purpose?
- With development team - will it test out the areas of interest in the advanced concept?
- With operational staff - is the scenario operationally reasonable?

b) Criteria for Application.

This sub-activity will always be applied.

c) Inputs.

- The scenario specification.
- Traffic samples.
- Tools to review the traffic samples.
- Operational expertise.

d) Outputs.

- Approved scenario and traffic samples.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.3: Scenario Specification****Sub-Activity 2.3.6: *Get Approval for Use of Scenario*****a) Description.**

Before the scenario and the associated traffic samples are used in the validation exercises or for training, final approval should be given by the customer for the validation exercise. Again a visualisation tool will help in the identification of the characteristics of the main traffic flows. If the scenario development process is going to take significant time and effort, it is clearly important to keep the customer informed of significant decisions and assumptions so that effort is not wasted.

b) Criteria for Application.

This sub-activity will always be applied.

c) Inputs.

- The scenario.

d) Outputs.

- Approval for use of the scenario.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.4: Production of the Overall Validation Management Plan****a) Purpose.**

The purpose of this activity is to formalise the planning and ensure that all parties are aware of the timing and scope of the activities in the validation exercise.

b) Description.

Management planning is a vital cornerstone of any project, and is core to the successful conduct of a validation exercise. The validation project management plan should be drafted in such a way as to give details about:

- The aim of the validation exercise: This is the high-level aim of validation exercise identified in Step 1;
- The activities to be undertaken, such as:
 - The tasks that need to be performed; in principle a description of all the steps described in this document (as applied to the technique selected);
 - The deliverables that will be prepared.
- How the activities are to be undertaken:
 - The approach applied and the method(s), technique(s) and tool(s) used. This includes the development of the high level experimental design, preparation for the exercise and an analysis plan;
 - Procedures to ensure quality of the work (including the skills/resources required for the exercise validation team – e.g. familiarisation with new tools/airspace/routes/stacks).
- When the activities will be performed:
 - Planning of tasks and meetings in time (visualised in a time line chart or equivalent). This will include the schedule of exercise runs;
 - Key decision points;
 - The milestones in the project, such as, the acceptance of the validation platform, the execution of the exercise and the delivery of the final report.
- Responsibilities in the exercise:
 - The individuals that will undertake the exercise, these people will perform the work and will supervise and observe the exercise runs;
 - Input and/or co-operation expected from the customer;
 - Input and/or co-operation expected from third parties; this may concern air traffic controllers, pseudo-pilots, etc.

c) Criteria for Application.

This activity should always be performed, although in a small exercise, other plans may be included in the document (for example the Measurement and Analysis Specification).

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.4: Production of the Overall Validation Management Plan****d) Inputs.**

- Outline Plan.
- Platform to be used.
- Detailed Experimental Design.
- Training requirements for the Validation Team.

e) Outputs.

- Validation Management Plan²⁷.

²⁷ GENOVA Project proposed a generic table of content for a Validation Plan (ref GENOVA - Validation Strategy and Plan, 25.06.1997, GEN/IS/WPR/4/FINAL 0.2, section 5 “Generic Validation Plan”).

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****a) Purpose.**

The purpose of this activity is to ensure that the platform and/or facilities for the measured runs are suitable for the measurements that need to be taken.

b) Description.

From this point on the general validation methodology will diverge for each validation exercise. The detail of the steps to be performed is significantly different for each of the validation techniques. Three main steps can nevertheless be identified. These are:

- Sub-activity 2.5.1 - Specification of the MAS.
- Sub-activity 2.5.2 - Preparation of the Platform or facility.
- Sub-activity 2.5.3 - Pre-exercise testing.

These steps will only be described in detail for fast-time and real-time techniques, in two sub-sections below. The preparation for literature search and judgemental techniques is described here.

Literature Search.

The aim of the preparation stage for a literature survey is to ensure that all the relevant documents have been identified and are obtained. The person conducting the literature study also needs to be clear about the information to be recorded and how it is to be analysed.

There may be very many documents that have been written, depending on the particular subject being studied. Selecting the right documents from this potentially long list is therefore important, both with regard to their suitability and the level of detail that can be achieved within certain (budgetary/time) constraints. Suitable means need to be established to ensure that the contents of the documents selected can in some way contribute to testing the hypotheses of the exercise. It is possible that an individual document may only contribute a small amount, but this may be vital in order to develop the best overall view of previous work. In general, the most suitable documents are those documents related to studies similar to the present validation exercise.

The number of documents to be analysed will be determined by the level of confidence required from the study, the availability of suitable documents and the timescales and budget available for the exercise. A small number of high level documents may be analysed quickly, but the level of confidence in the conclusions will be low. In contrast, if there are a large number of detailed documents that are directly relevant to the aims of the validation exercise, the level of confidence in the conclusions will be higher. If many potentially suitable documents are found, the selection constraints may be tightened to reduce the number to be read (and stay within budget/time constraints). The final selection of documents will depend ultimately on the availability of suitable documents.

The best way to find the most relevant documents is to check reference lists of known documents and to ask experts for their opinion. In addition the library and/or Internet may be searched.

A statement should be made in the validation management plan on what information is to be gathered and how it is to be analysed. This will include stating the criteria for determining the confidence level in data from the different sources.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Judgemental techniques.**

The use of judgemental techniques in validation is to elicit and draw together the views of experts on the operational concept. The preparation phase will therefore involve deciding the focus for the discussions, selecting the experts, briefing them on the operational concept and the approach to be taken, and preparing the facilities for the exercise.

The application of judgemental techniques will involve the collection of subjective views; it is in the careful selection of experts that the problems of excessive subjectivity or bias will be mitigated. The selection of representative experts from the stakeholders identified in Step 1 is a valuable first step in ensuring the best results. Where possible, more than one representative from each stakeholder function/organisation will further reduce subjectivity and bias. The numbers finally selected will depend on cost, availability and the type of judgemental technique selected; involving too many experts may prevent any conclusions being reached.

There should be a statement in the validation plan on the focus for discussions, what will be recorded and how the experts' subjective opinions will be analysed. The recorded data will be qualitative and criteria should be provided on how opinions should be collated and presented for any feedback sessions.

Briefing materials describing the aims of the exercise, the ATM problem and the operational concept must be produced and sent out to the participants at least six weeks before. The briefing material may also provide information on the conduct of the exercise and any particular scenarios to be considered.

The preparation of the facility to be used may involve training the validation team members who will be facilitating and leading the panel(s). It is very important that discussions are controlled sensitively to ensure they remain productive and reach conclusions if appropriate. Some approaches may require questionnaires or structured forms to be prepared so that the views and discussion is recorded efficiently in relation to the validation aims and objectives. Facilities such as video cameras or tape recorders may be required and need to be set up in such a way that they do not interfere with the proceedings.

Finally a pilot exercise may be required using a limited number of experts or pseudo-experts to shake out any problems with the application of the technique.

c) Criteria for Application.

This step should always be performed; although for literature search and straight forward validation exercises it may not be a significant task.

d) Inputs.

- Detailed Experimental Design.
- Scenario Specification.

e) Outputs.

- Measurement and Analysis Specification.
- Configured and tested Validation Platform.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE

Activity 2.5: Preparation (shakedown trials) for the Exercise Runs

Sub-Activity 2.5.1: Specification of the Measurement and Analysis Specification (MAS)

a) Description.

Each validation technique is suited to validating different stages in the development cycle of ATM concepts. Consequently there are different requirements for each validation exercise depending on the technique used. The description of this sub-activity will focus exclusively on fast-time and real-time techniques.

Fast-time techniques.

An MAS is required for each validation exercise to support the planning and execution of the measured runs. It draws together all the relevant information on the conduct of the measured runs and the subsequent analysis of the results. In this way any gaps or inconsistencies in these aspects of the validation exercise can be identified and addressed. The issues that should be taken into account for fast-time techniques are:

- The validation aims;
- measurements;
- recommended analysis methods.

These issues are now described:

- The validation aims to be detailed within the validation aims section of the MAS will be those identified in step 1. The validation aims identified in Activity 1.3 should be stated along with the high-level validation objectives identified in Activity 1.4. These objectives provide the basis for identifying the measurements for the validation exercise. Including aims and objectives within the MAS ensures that the document can be read in isolation. Examples of Validation Aims are provided in Appendix 4 under Sub-Activity 2.5.1 Validation Aims.
- The measurements section of the MAS will describe the measurements to be used to assess each low-level validation objective. At this stage, the final decisions on which measurements will be taken (drafted in previous activities of step-2; i.e. activity 2.2) should be taken. A number of measurements could contribute to more than one objective. The low-level and subsidiary validation objectives, identified in activity 1.4 should be stated under the heading of the related high-level validation objective or research aim. The hypotheses developed in activity 1.7 should also be noted along with a description of the measurements that are to be recorded. Examples of low-level objectives are given in Appendix 4 under Sub-Activity 2.5.1 Measurements.
- The experimental design defined in activity 1.8 and refined in activity 2.1 will impact on the analysis methods used. The operational and statistical significance levels determined in activity 1.9 will be stated here, with an explanation of the reasons why these levels were chosen. An example of the operational and statistical significance levels that will be stated is given in Appendix 4 under Sub-Activity 2.5.1 Analytical Methods. The selection of analytical approach section of the MAS will describe the structure of the analysis process. It is assumed that the team conducting the trial has experience of data analysis methods. There are three principle phases of the data analysis:
 - *Exploration* of the data;
 - *inference* about the objectives;

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.1: Specification of the Measurement and Analysis Specification (MAS)**

- *synthesis* of conclusions concerning the objectives.

The three principle phases are described in Appendix 4 under Sub-Activity 2.5.1 Selection of Analytical Approach.

Real-time techniques.

The issues described for fast-time techniques also apply to real-time techniques. In addition to these, two further issues that should be considered are:

- The requirements placed on controllers;
- training issues;

These issues are now described:

- The controller requirements section of the MAS should contain a description of the number of controllers required and the sectors on which they are valid. The requirements placed on the controllers will differ depending on:
 - The validation aims and objectives;
 - the type of trial;
 - the length of the trial;
 - the number of sectors.

The qualities that should be considered when requesting controllers are detailed in Appendix 4 under Sub-Activity 2.5.1 Controller Requirements.

- While it is possible to make requests there are usually only a limited number of available controllers who cannot be guaranteed to meet the requirements. When the controllers and their qualities are known, the trial design should be re-examined in order to assess the impact on the analysis and to identify any steps that could be taken to ensure that the aims of the trial are not comprised. An example of how controller requirements may appear within the MAS is given in Appendix 4 under Sub-Activity 2.5.1 Controller requirements.
- A detailed description of the training issues should be given in the training issues section of the MAS. The training issues will vary depending on:
 - Whether controllers meet the requirements set out in the controller requirements section of the MAS;
 - the language the training is to be conducted in;
 - the understanding of the concept by the controller;
 - the scale of the difference between current day systems and the advanced system;
 - the level of knowledge of the current and advanced systems required;
 - familiarity with the simulation environment;
 - the aims and objectives of the trial;
 - It is likely that the training will not be sufficient for controllers to have learnt everything there is to know about a system before measurements begin to be taken.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.1: Specification of the Measurement and Analysis Specification (MAS)**

It is further expected that the controllers will become familiar with the traffic samples, although the level of familiarity can be minimised by careful experimental design. Learning effects should therefore be considered during the analysis. The amount of time dedicated specifically to training during the trial should be minimised. As far as possible, training should be undertaken a few weeks prior to the trial, with some time spent during the trial to refresh the controllers' memory. For example in a two-week trial approximately 2 days should be dedicated to reinforcing the previous training;

- In a trial where a significant amount of training is necessary it is prudent to produce a training plan. The first stage in developing a training plan is to consider the objectives of the training. It is unlikely that the controllers will be familiar with everything before the start of the validation exercise; however, it is useful to define exactly what knowledge each controller should have. This will enable decisions to be made on the best way in which to train the controllers and the extent to which the training has achieved its aim will also be clear. Example training objectives are given in Appendix 4 under Sub-Activity 2.5.1 Training Issues. Various training methods and examples are also given in Appendix 4 under Sub-Activity 2.5.1 Training Issues.

b) Criteria for Application.

This step should always be performed.

c) Inputs.

- Detailed Experimental Design.
- Validation Hypotheses and Metrics/Indicators.

d) Outputs.

- A Measurement and Analysis Specification.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.2: Preparation of the Platform or Facility****a) Description.**

The description of this sub-activity will focus exclusively on fast-time and real-time techniques. An identification of validation techniques characteristics is provided in Appendix 3 under Sub-Activity 1.5.5 - Balance Requirements with Constraints. A suggested list of platforms, indicating their capabilities and possible generic metrics, is provided in Appendix 3 under Sub-Activity 1.5.6 – Validation Platforms.

Fast-time techniques.

The aim of preparing the platform is to configure it to represent the ATM problem, as expressed through the scenarios, and the operational concept. The elements of the platform that need to be considered are shown in Figure 16. The core of the platform consists of the simulator itself, but it cannot be considered in isolation from the input files, and the requirements of analysis and data presentation.

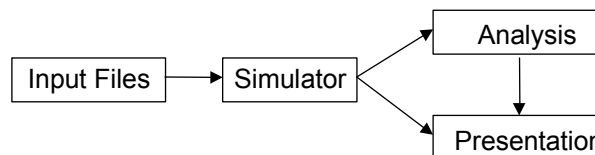


Figure 16 Elements of the Simulation Platform

These elements are now considered:

- A set of input data files will be derived from the scenario description. It is important to know how much work this will involve. Typically this may include traffic sample extraction and growth, or the creation of an airport ground layout.
- The preparation of the simulator will be driven by the scope of required modelling. This will be driven by the validation objectives. For example, it is unlikely that an en-route study would need aircraft ground movement to be modelled in detail. Again, it will be important to know to what extent new software is required to reflect new functions relating to the ATM concept, and what the implied time scales are.
- As analysis requirements have been stated, these need to be considered in the light of currently available analysis tools, as time may be needed to develop or acquire new tools. The customer may have particular requirements, and the way in which results are to be presented may be influential too.
- Presentation requirements may have an impact on both the analysis tools and the choice of simulator. They should take into account both the testing/validation stage of the simulation, where a particular form of output associated with the functioning of the operational concept may be of use, and the analyses themselves. Different types of presentation may be required for different parties.

Real-time techniques.

The platform(s)/tool(s) must resemble the real-life situation as closely as possible; they may therefore need to be adapted by adding functionality specific to the operational concept considered. If for example a free flight concept is examined, a CDTI (Cockpit Display of Traffic Information) needs to be implemented in the cockpit of the aircraft involved. Other

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.2: Preparation of the Platform or Facility**

preparations may include for example the airspace structure and measurement data storage.

Also preparations need to be made for the measurements and for collecting feedback from the exercise participants. This may comprise of the selection and set-up of measuring equipment like ISA boxes (Instantaneous Self Assessment) and the development of questionnaires. Observations are recorded using observers during the measured runs, and the validation staff, who will be acting as observers, will need to be trained.

b) Criteria for Application.

This step will normally be performed; its scope will depend on the complexity of the platform.

c) Inputs.

- Unconfigured Platform.
- Scenario description.
- Detailed Experimental Design.

d) Outputs.

- Configured platform, including input files and additional functionality provided for the model.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.3: Pre-exercise Testing****a) Description.**

The description of this sub-activity will focus exclusively on fast-time and real-time techniques.

Fast-time techniques.

Once the platform has been prepared, it must be tested and its suitability to support the validation exercise demonstrated, so that it is acceptable to all parties, including the customer. Some iterative development may be necessary. This process familiarises the customer with the simulation platform and builds confidence. A visual representation of the simulation may be very powerful at this stage.

The fidelity of the simulation depends upon the extent to which it represents the ATM system to be validated, and its real world context. To an extent this is subjective, but any tools that allow controllers to assess whether or not the correct procedures are being modelled should be used at this stage. This may result in suggested changes to simulation parameters (e.g. arrival separation) or to the underlying behaviour of the simulator. It is also possible that additional analyses, not specified in the original validation plan, are seen to be useful.

This step is only complete when all parties are satisfied with the model developed.

Real-time techniques.

The real-time platform is tested through a shakedown trial. This is a trial of up to one week, it may involve the use of operational participants. There will be no plans to make any measurements for analysis.

The objectives of the shakedown trial are to make sure that the platform satisfies the following criteria:

- Form a representative model of the real world to be validated;
- Provide a robust platform for measurements, recently developed software can be unstable and liable to fail;
- Deliver the best possible output. When measuring equipment is used it needs to be calibrated. Questionnaires need to be reviewed.

Exercise participants need to be trained since they are usually not acquainted with the new procedures and/or supporting tools. They may need to develop a different working method. By using specific training runs, the participants are focused on the differences compared to the situation they are used to. This way the learning effect during the measured runs is minimised.

Only after acceptance by the customer and the training of the participants can the execution of the validation exercise begin.

b) Criteria for Application.

This step should always be performed. Its scope will depend on the complexity of the platform and the modifications required for the validation exercise.

STEP 2: VALIDATION DESIGN – PLAN AND PREPARE THE VALIDATION EXERCISE**Activity 2.5: Preparation (shakedown trials) for the Exercise Runs****Sub-Activity 2.5.3: *Pre-exercise Testing*****c) Inputs.**

- Platform configured to represent the scenario (including the ATM Concept).

d) Outputs.

- Tested validation platform ready for the validation exercise runs to be performed.

STEP 3: CONDUCT OF VALIDATION EXERCISE RUNS**a) Purpose.**

The purpose of this step is to execute the validation exercise, as specified in the detailed experimental design, on the validation platform in order to obtain a set of measurements (including unexpected occurrences), replies and other data that will be analysed in Step 4.

b) Description.

At the end of Step 2, everything has been left ready for performance of the individual validation exercise run(s). No generic guidelines can be given in this regard that would be applicable to all validation techniques, so that this section will provide two types of general information in separate subsections below.

Training for participants will be required before undertaking the real-time exercise. This might include familiarisation of new airspace including standing agreements, routes changes, stack changes, new working practice and familiarisation of new tools.

First, information will be provided regarding different types of data which may be collected during the run(s), because the type of data determines the type of analysis that will be able to be performed in step 4. In addition, a brief overview of data collection methods will be given.

Second, a brief series of guidelines for conducting the validation exercise will be provided for each of the four overall types of validation techniques.

c) Criteria for Application.

This step must be performed for every validation exercise.

d) Inputs.

- Detailed Experimental Design.
- Validation Scenarios.
- Validation Management Plan.
- Configured and tested Validation Platform.

e) Outputs.

- “Raw” data, in a form suitable for analysis, obtained from the measurements and questionnaires used during the validation exercise run(s).
- Logs of the system parameters recording the status of validation platform.

STEP 3: CONDUCT OF VALIDATION EXERCISE RUNS

Types of Data and Data Collection Methods

This section provides an overview of data types and data collection methods. More detailed information on data types with examples for fast-time and real-time validation exercises is presented in Appendix 3 under Activity 1.6 Data Types Applicable to ATM Validation Exercises. More detailed information on data collection methods is presented in Appendix 5 under Step 3 Categorisation of Data Collection Methods. Examples of data collection methods for fast-time, real-time, shadow-mode and operational trial techniques are also given in Appendix 5 under Step 3 Examples of Data Collection Methods.

The following data types are, in general, applicable to ATM validation exercises:

- Qualitative/quantitative.
- Objective/subjective.
- Intrusive/non-intrusive.
- Binary/not binary.
- Nominal/ordinal/interval/ratio-level.

Data can be collected during validation exercise runs by either automated means, humans (the participants in the exercise or the validation team), or some combination of both. Data collection methods can be broken down into the following generic categories, illustrated in Figure 17 below:

- Data measurable by the automated/simulated system.
- Data gathered through observation of the exercise.
- Physical data measurements from the participants in the exercise.
- Data and/or opinions provided by the participants.

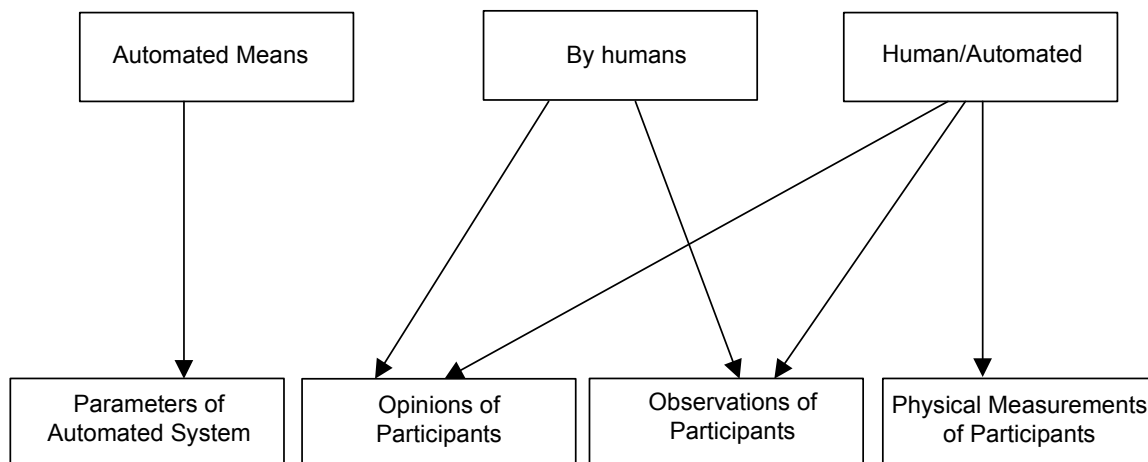


Figure 17 Generic Data Collection Methods

STEP 3: CONDUCT OF VALIDATION EXERCISE RUNS

Guidelines for Executing Validation Exercises

This subsection provides general guidelines for executing exercises using each of the four overall types of validation techniques. As has been stated above, no detailed guidelines can be provided for execution of validation exercises because of the enormous variation between techniques as applied to individual exercise runs.

Literature Study.

The performance of a literature study will involve reading all or parts of the documents identified and making structured notes based on the analysis plans.

Executive summaries of the selected documents will provide a background for the data obtained from the documents. This will set the context for individuals who have not read the document, and can provide important information for the performer of the validation exercise as well. This allows the analyst to weigh data obtained from the different documents describing validation exercises with similar aims and scope. It is also important to record any assumptions either in the documents surveyed or by the exercise performer.

The output data from the literature study are a list of qualitative statements, and possibly numerical results, from relevant documents providing the "evidence material" for testing the hypotheses.

Judgemental Techniques.

In judgemental techniques, Steps 3 and 4 are frequently part of an iterative process, that is, the output of the one is used as an input of the other.

The exercise runs will involve eliciting and recording the opinions from the expert panel members; these will be recorded in accordance with the plans. This may be in the form of a meeting, at which the validation team facilitates the discussions or it could be purely paper-based, circulating specific information and questionnaires e.g. by post or e-mail to the experts. The results may then be collated and analysed and used to support further discussion or the review of conclusions.

The output data consist of a set of recordings of subjective views. These will normally be text based, but may be supported by video and audio recordings.

Fast-Time Techniques.

Before the execution of the validation exercise run(s) starts, the validation team will have a clear plan in place. This will provide a detailed specification of the configuration of the validation tool and the inputs representing the scenario.

Execution of a fast-time validation exercise runs therefore comprises executing the process or running the tool for each planned scenario. In order to avoid bias in the results, each scenario should be run a multiple number of times (10-20 is typical) with traffic representative of the validation scenario. This avoids one particular ordering of the traffic sample having too great an effect. Detailed advice cannot be provided for this step, as it will be very dependent on the particular platform being used.

Where the validation platform is a detailed ATM simulator, the process can be automated. Batch scripts can be created to allow the fast-time simulator to run with little or no operator intervention. Simulations can therefore be run overnight or over a weekend.

The "raw" results from a fast-time validation exercise could range from a small list of calculated data from a spreadsheet to large data sets of measured data. The data are numerical and objective.

STEP 3: CONDUCT OF VALIDATION EXERCISE RUNS

Guidelines for Executing Validation Exercises

Real-Time Techniques.

When the application of a real-time technique reaches Step 3, a detailed plan will be available setting out the precise configuration for each measured run. There will also be planned sessions for debriefing and for visitors. The planning phase should also have identified the actions to be taken in case of equipment failure.

Each exercise goes through a similar process. It starts with briefing, to give the participants an idea of what to expect from the exercise runs and to know what they are expected to do. When this is complete, the measured runs start and the data is collected. The data collected may include logs recorded digitally, notes from observers, recordings from videos, and questionnaires completed after the measured run finishes. Periodically there should be debriefing sessions to collect feedback from the (measured) participants.

There may also be data provided through the visits of invited expert stakeholders whose views on the operation of the operational concept also need to be recorded.

The "raw" results of a real-time validation exercise are varied. There will be different types of data:

- Objective measured data obtained from:
 - A prototype implementation of the operational concept;
 - the ATM data processing system;
 - monitoring equipment, e.g. eye movement tracking, heart rate monitor.
- Subjective measured data (e.g. from the ISA panel and the records from the exercise observers, video recordings);
- Subjective opinions of the participants recorded in questionnaires and gathered in the debriefing sessions.

STEP 4: ANALYSIS OF THE RESULTS**Activity 4.1: Performance of the Analysis Methods Specified in the MAS****a) Purpose.**

The purpose of this activity is to perform the methods of analysis as specified in the Measurement and Analysis Specification (MAS) on the measurements generated in Step 3, in preparation for the proposal of conclusions and recommendations in Activity 4.2²⁸.

b) Description.

Whatever the validation technique used, the analysis techniques available can be divided into different categories that can be applied to different types of data:

- **Qualitative Analysis:** Qualitative techniques aim to extract the meaning and conclusions from a set of non-numerical data, but without necessarily giving a specific numerical result. These techniques look at the ranking of factors or opinions, and the overall characteristics of the data, rather than the specific values. The techniques may involve ranking answers to questions, investigating the similarity between responses, following a line of logical reasoning and conducting a content analysis on a set of subjective responses. It is possible to generate numerical data, such as counts, through qualitative analysis and then subject this to quantitative analysis.
- **Quantitative Analysis:** Quantitative techniques intend to reach a specific numerical conclusion, often with an associated level of confidence. There are many quantitative statistical tests, and an attempt will not be made here to describe them all or to state which are most accurate. Further detail can be found in many statistics books (e.g. [Statistic for Management] and references therein). The quantitative analysis can be sub-divided into descriptive and inferential techniques:
 - Descriptive techniques are concerned with making concise and intuitive descriptions of the data to aid interpretation, such as by making graphs and histograms;
 - Inferential techniques aim to derive a general result from a limited set of measurements and then testing hypotheses. This leads to statements such as "with 95% confidence, the number of flights per day from a specific runway can be increased by more than 5% by implementing a new procedure".

Inferential techniques fall into two types, parametric and non-parametric: Parametric analysis techniques assume some distribution of results, while non-parametric studies suppose no assumptions on underlying distribution of data. Hence the latter are less powerful.

Further information on data types and the use of analytical methods can be found in Appendix 6 under Activity 4.1 Types of Data Analysis Methods.

When the descriptive and inferential statistics have been performed, and any qualitative analysis on the subjective data has been done, the generality of the results along with any assumptions and constraints on the exercise data needs to be considered. Also the statistical significance needs to be calculated for all results where possible, and be reported together with the results.

Careful use of appropriate techniques, including a clear view of the assumptions made, will enable the most value to be obtained from the data. Some dangers are:

²⁸ A set of suitable statistics reference books are included in Appendix 2.

STEP 4: ANALYSIS OF THE RESULTS

Activity 4.1: Performance of the Analysis Methods Specified in the MAS

- **Implicit assumptions:** Some analysis techniques will assume the nature of the distribution from which the data is chosen (in particular, they may assume a Normal probability distribution), and if such assumptions are not true, the analysis may be fundamentally flawed.

Inappropriate generalisation: Any analysis will only be as reliable as the data used. Though the analysis should aim to generalise the specific results to a wider context, this should only be done when there is sufficient volume and variety of independent data.

- **Over-reliance on statistical significance:** Statistical significance is only related to the number of measurements and error model, the overall confidence should take into account the validation exercise fidelity and non-independent measurements and equipment inaccuracies.
- **Misleading presentation:** Descriptive statistics can easily be misinterpreted. Non-linear, broken, or non-zeroed axes should only be used where strictly necessary. Error or confidence bars should be added where possible, such that it is clear to the reader where data is sparse and inconclusive.

The relevance of these analytical techniques for each of the four overall types of validation techniques is as follows:

Literature Studies.

The literature study provides only qualitative data for analysis. The data must therefore be analysed using qualitative techniques. The analysis will typically start with ordering the list of documents by importance, which means by relevance of the conclusions/results in that document with regard to the present validation exercise. Relevance depends on the scope of the reference document, the subject studied and the approach used. Generally the more similarity between the reference document and the own validation exercise, the more relevant it is. It is however up to the experience of the exercise performer to judge this.

The documents with highest relevance will determine whether a hypothesis is accepted or rejected. The less relevant documents affect the level of confidence in the hypothesis. In particular, when there are significant differences between higher and lower importance sources, the level of confidence in the trend identified will be lower than when they all point in the same direction.

Judgemental Techniques.

Judgemental techniques initially only generate qualitative data for analysis, qualitative techniques must therefore be applied to draw together views and synthesise intermediate conclusions. A key element of the analysis with judgemental techniques is its close feedback with the previous step where the data is gathered. The qualitative techniques applied at this stage therefore need to be efficient so those summaries can be produced quickly.

Fast-time Techniques.

At this stage in the exercise, the analyst should follow the detailed analysis plan that was produced in the second step. The types of analysis that are feasible will have been determined by the original experimental design that has been selected for the measured runs.

STEP 4: ANALYSIS OF THE RESULTS

Activity 4.1: Performance of the Analysis Methods Specified in the MAS

Some general steps need to be followed:

- *Check the data* – The first thing to do after a calculation/simulation run has been executed is to check if it has run correctly. Some output files may have been corrupted for some reason, or an incorrect input parameter may have been used, despite verification of the input files. If this is the case, the cause must be traced and corrected. After that the specific calculation/simulation run will be repeated.
- *Perform the analysis* – The data collected in a fast-time exercise is numerical and therefore amenable to quantitative statistical techniques and the use of statistical measures and hypothesis testing should be considered.
- *Consider assumptions* – The quality of the assumptions in the modelling should be considered as part of the analysis. This is particularly important with fast-time validation exercises because there are no human actors involved who can provide immediate feedback if the simulation is not realistic. Also the affect of any limitations imposed by the platform in the scenarios or operational concept modelled need to be considered.

Real-time Techniques.

The analysis of data obtained using real-time techniques is very similar to the analysis of data from fast-time techniques described in the preceding sub-section, although the type and amount of data may be different. Therefore other analysis techniques may be required. For example the subjective data can be used to support the quantitative analysis conducted on the numerical measurements.

Some specific issues related to real-time techniques are:

- Unusable measurements that may be the result of:
 - Inaccurate measuring equipment;
 - Equipment faults or technical problems;
 - Incorrect input files/parameters;
 - Poorly controlled conditions.
- Since ATM actors are involved, the amount of subjective data relative to objective data is probably higher. Moreover these subjective data are usually from a small number of individuals; it should be considered whether they are representing the views of the whole ATM community. Both will affect the conclusions that can be drawn, and especially the level of confidence applied to its operational significance.

c) Criteria for Application.

This activity must be performed for every validation exercise, regardless of the technique used.

d) Inputs.

- Measurements/Data collected in Step 3.
- Measurement and Analysis Specifications.

e) Outputs.

- Analysed data.

STEP 4: ANALYSIS OF THE RESULTS**Activity 4.2: Proposal of Conclusions and Recommendations****a) Purpose.**

The purpose of this activity is to draw up initial conclusions and recommendations for finalisation, reporting and dissemination in Step 5 on the basis of the analysed data from Activity 4.1.

b) Description.

This activity will allow the robust conclusions of the exercise to be deduced and presented, along with recommendations for the underlying development project, and for future validation activities. The significance of the data, both statistical and operational, plays an important rôle in deriving these conclusions and recommendations.

c) Criteria for Application.

This step should be performed for every validation exercise.

d) Inputs.

- Analysed data from Activity 4.1.

e) Outputs.

- Proposed conclusions and recommendations (for refinement and dissemination in Step 5).

STEP 5: DEVELOP AND REPORT CONCLUSIONS AND RECOMMENDATIONS**Activity 5.1: Develop Conclusions and Recommendations****a) Purpose.**

The aim of this activity is to place the results in the correct operational context, taking into account the interests of the customer of the validation exercise. This will allow robust and convincing conclusions to be drawn.

b) Description.

The analysis of results should be used to derive conclusions from the bottom up. Thus the hypotheses will be used to derive answers to the lower level objectives; these in turn will address the higher level objectives. This process will continue through all the objectives until the validation aim is fully addressed. In all cases the level of confidence in the conclusions should be provided, though not in statistical terms. If there is insufficient positive evidence for the anticipated conclusions, this should not be seen as a failure of the validation exercise, but as support for a recommendation for further development or further validation exercises, or for work on the operational concept to be cancelled.

The customer and key stakeholders need to be fully involved through the development of the conclusions. Their operational expertise is required to ensure that the conclusions and recommendations are set in the correct operational context. They can also provide advice on the relative importance of the conclusions and recommendations to operational staff.

c) Criteria for Application.

This step should be performed for every validation exercise.

d) Inputs.

- Performance Requirements identified in Step 1.
- Initial conclusions and recommendations defined in Step 4.

e) Outputs.

- Final conclusions and recommendations from the validation exercise.

STEP 5: DEVELOP AND REPORT CONCLUSIONS AND RECOMMENDATIONS**Activity 5.2: Prepare Report****a) Purpose.**

Validation is a continuous parallel process throughout the development life cycle of a new AMT concept. As such, a validation exercise should not be seen as an isolated activity. The report on the validation exercise should make clear, practical recommendations to improve the underlying development project and keep it on track to achieve the strategic objectives. The report should also recommend the areas of focus for future validation activities.

b) Description.**Intended Audience.**

As in the preparation of any report, it is important to determine the expected audience of the report before the first word is put on paper. A report destined for higher management will have a different scope - focused more on the conclusions and recommendations - than a report that is meant for technicians or scientific staff, who may have more interest in the detailed results, exercise design and the analytical techniques applied. These considerations may require the production of more than one report.

During the preparation of the report, care needs to be taken to consider the perception and sensitivity of the report's results and conclusions to these different audiences. Different communities will interpret the results in different ways. For example, the controllers may consider the results of a specific workload indicator to be of major impact to their working methods, whereas other parties may be more concerned with the technical capabilities of a new tool.

Scope of Report.

The report should state and discuss the strategic objectives, lower level objectives and metrics or indicators that were targeted in the exercise. The exercise methodology should be stated, or described, as well as the analytical techniques applied. It should discuss the level of success in the measurement of the factors, along with the conclusions as to whether the underlying project is on target to meet the strategic aims. If relevant, the report should also discuss the subsidiary objectives of the exercise and the success in their achievement.

The level of detail presenting these subjects can be limited; the main objective is to provide the reader with a background of the validation exercise. The results need to be placed in the right context with regard to the life-cycle development of the ATM concept and provide evidence in order for the reader to be confident in the conclusions.

An executive summary should be provided. This should present the aims of the validation exercise, be relevant to the important high level management decision making and provide the conclusions in a well argued manner.

Present Results and Conclusions.

Depending on the primary audience, the results should be listed mainly in appendices. In the main report only high-level results should be presented, preferably graphically or in tabular form. The results should be clearly related to the conclusions and recommendations that are derived from them. As such they provide an audit trail to these conclusions and recommendations.

The conclusions and recommendations should be presented concisely and in their operational context.

STEP 5: DEVELOP AND REPORT CONCLUSIONS AND RECOMMENDATIONS**Activity 5.2: Prepare Report****Review and Dissemination.**

The report should be reviewed in the context of the priorities of potential readers, and consideration given as to whether the report results could be misinterpreted or quoted out of context by others thereby suggesting unintended conclusions.

A selection of reviewers should be used to bring to bear a wide range of experience, and to give both internal and external opinions on the report structure and content. By having a wide reviewing audience the different perceptions can be covered and the sensitivity of the conclusions can be minimised.

The report's distribution should be considered early in the report preparation phase. There may be security or commercial reasons that force a limited distribution, but where possible the intention should be to disseminate the report conclusions and recommendations to all those who could be affected by the findings, along with standard information centres. It may be suitable to release a high-level synopsis, or abstract, of the report to a wide audience and to make the full report available on request. In addition, presentations to appropriate audiences should be considered to coincide with the report release.

c) Criteria for Application.

This step should be performed for every validation exercise.

d) Inputs.

- Final conclusions and recommendations defined in Activity 5.1.

e) Outputs.

- Final report on the validation exercise.

STEP 5: DEVELOP AND REPORT CONCLUSIONS AND RECOMMENDATIONS**Activity 5.3: Make Presentations and Demonstrations****a) Purpose.**

The purpose of this activity is to provide presentations and demonstrations of the results of the validation exercise to the stakeholders and the general user community.

b) Description.

It is important to provide presentations of the results to the stakeholders and user community. The considerations described above for report preparation need to be taken into account in the preparation of presentations as well. It is very important with presentations that material be provided in a visual form. Summaries of results in the form of histograms or graphs can demonstrate the potential benefits effectively to even a non-technical audience. Systems and processes are better described using block diagrams than with bulleted text. Conclusions need to be stated briefly so that the impact is not lost in a mass of words.

A demonstration is a very effective way of building confidence in the conclusions from the validation and in the feasibility of the operational concept. As the aim of the demonstration should be to build confidence, it is very important that demonstrations to non-technical stakeholders should only be considered when the implementation of the operational concept is very reliable and reasonably comprehensive. Demonstrations are most easily arranged as part of the schedule of measured runs, however, the implementation needs to be sufficiently mature at this point in the programme to provide a robust demonstration. If there are doubts about reliability, more limited demonstrations of the operational concept may be given as part of the presentation when the final results are available.

c) Criteria for Application.

This step should be performed for every validation exercise.

d) Inputs.

- The final report generated in Activity 5.2.

e) Outputs.

- Presentations and demonstrations.

EXAMPLE: A REAL-TIME VALIDATION**STEP 1: Identify Aims, Objectives and Hypotheses**

1. Problem.

The high level ATM Problem to be addressed in the validation exercise described here is that of airspace capacity limits imposed by controller workload. Given that controller workload can be reduced by providing tools, there must be a migration path from today's paper based operational concept, to one where all information is held electronically. The particular problem to be addressed on this project is the transition from a paper based controller interface system to a fully electronic interface. This should be achieved without compromising capacity or safety.

2. Operational Concept.

The Operational Concept for this project required the minimum change from that of the paper based operations. The concept was therefore based on the use of planner and tactical rôles with electronic system support replacing the functions of paper strips.

3. Stakeholders.

The stakeholders in the trial were:

- Senior Management (who need to make the investment decisions regarding the future of the project);
- The controller community (who would ultimately have to use the system);
- And the design team (who would use the results of the trial to influence the future design).

4. Aims.

The aims of the trial were to:

- Explore the impact of the system on the tactical and planner rôles;
- Observe and measure how the system was used to aid further development.

5. High-Level Objectives.

High-level objectives of the trial included:

- Assessing the impact of the new system on workload, compared to the initial system configuration;
- Assessing the usability of the prototype.

6. Low-Level Objectives.

Low-level objectives for the trial included:

- Determine whether the overall level of planner workload is changed by the introduction of the prototype;
- Determine whether the overall level of tactical workload is changed by the introduction of the prototype;
- Identify key contributors to changes in level or type of workload;
- Identify any changes in working practices for individual tactical and planner rôles, and in the team, that affect workload;
- Assess the acceptability of the system;

EXAMPLE: A REAL-TIME VALIDATION

- Assess the learnability of the system;
- Assess the ease of interaction with the system;
- Identify any features for improvement or re-design.

7. Platform Requirements and Options.

As the trial was one in a series of validation exercises for the operational concept. In all of these the same real-time validation platform had been used, so there was no explicit stage in this exercise where an explicit decision was made on the platform. However, as the operational concept is mature and its success is critically dependent on the ease with which the controllers use the electronic support tools replacing the paper strips, it can be seen that a real-time simulation would provide a suitable platform for meeting the validation aims.

8. High-Level Experimental Design.

Typically, the real-time trials used matched designs and employed non-parametric statistics. If the system to be simulated is immature and subject to change throughout the trial, then there would only be plans to use descriptive statistics, and the experiment designer may choose not to simulate a baseline. If it is to be a fully measured trial, then consideration is given to how many operational concepts are to be simulated. Although it may be desirable to test as many variants of the operational concept as possible, there is a trade-off between the number that can be validated whilst still having sufficient information on each one to support reliable results.

As the operational concept is mature, it was decided to use a matched design with the data analysed using non-parametric inferential statistical tools.

9. Hypotheses Observations and Metrics.

The hypotheses below provide two sets of examples of the hypotheses used in the trial and the measurements used to justify them, or otherwise:

Null hypothesis, H_0 – Tactical workload is the same in each of the organisations;

Alternative hypothesis, H_1 – Tactical workload is different in at least one pair of organisations.

The measurements for these hypotheses included:

- ISA;
- Number of aircraft on frequency;
- Number and duration of R/T calls;
- Number of altitude, heading and speed instructions;
- Number of aircraft requiring tactical intervention.

Another pair of hypotheses used in the trial were:

Null hypothesis, H_0 – The usability is the same in each of the organisations.

Alternative hypothesis, H_1 – The usability is different in at least one of the organisations.

The measurements for these hypotheses included:

- Eye Movement Tracking;
- Software Usability Measurement Inventory (a usability questionnaire that produces quantitative results);

EXAMPLE: A REAL-TIME VALIDATION

- Observations;
- Number of times the different tools were used.

STEP 2: Plans and Preparations**10. Final Technique Selection and Detailed Experimental Design.**

It was agreed to use real-time simulation using the existing platform.

In the trial, two organisations were simulated together with a baseline operational concept representing the initial paper based operational concept. A fully matched design was used with the three operational concepts, in order to identify whether any one of the three differed from the others in terms of the metrics employed. Consideration was also given as to whether all controllers would cycle through the all measured positions; for example, whether retired controllers should be kept on the feeds. Another decision to be made, was the length of measured runs that were required, and the number of runs needed, in order for results to be valid.

The timetable generated for the two week (10 working day) trial included:

- Two days of training time;
- 27 measured runs (a set of 9 'matched triples');
- three 'development runs', where new variants of the software could be run for subjective feedback only;
- debrief time.

11. Outline Plans.

An Outline Trial Design was produced that gave a high-level statement of the aims of the trial together with: some examples of measurements; a description of the airspace; the timing of the trial; and the analysis techniques to be used. This was discussed and agreed with the customer.

An Initial Project Plan was also produced which detailed staffing levels and outline timings for the trial preparation and analysis. This was approved by the Head of Section.

12. Scenario Development.

The measured airspace used was chosen as it presented a two-sector airspace configuration that could be simulated reasonably realistically with surrounding feed sectors. Another reason for using the airspace was because it presented very different types of sectors to those simulated earlier in the programme. This airspace was chosen partly to fit within the facilities available on the platform, but had there been a need to use a larger airspace configuration (and therefore more machines) the platform could have been extended to accommodate this.

The number of traffic samples produced, eight, was what was needed to fulfil the experimental design. Since the electronic based system is a future system, a range of samples were produced starting with current day traffic levels, and ending with predicted traffic levels when the system is planned to go out of service (circa 2010). The idea behind this range is to assess whether it will be able to function with the predicted traffic levels, but also to enable exploration of how the new system supports controllers in varying traffic levels. For example, it is of interest to see which tools are used when the controller is quiet, and which ones cease to be used as traffic levels increase, together with the impact that this has on their 'picture'.

EXAMPLE: A REAL-TIME VALIDATION**13. Select and Book Facilities.**

The simulator time is booked with reference to the rest of the simulation programme and the likely availability of controllers. In addition to the main trial, extra time in the simulator needs to be booked for shakedown trials, pilot trials (in-house large-scale tests) and traffic sample testing. This was fitted in around the on-going software development and testing.

14. Detailed Plans (Including Analysis Plan).

A number of detailed plans are produced. Chronologically, the first was the Engineering and Test Plan. This specifies the configuration that needs to be implemented in the validation platform including:

- Which machines will be used and which sector they will simulate;
- The video recording equipment needed;
- Telephone facilities required between sectors;
- The R/T configuration to be used;
- When the airspace will be tested;
- When traffic samples will be tested;
- The requirement for the 'history computer' recording ISA, R/T and telephones;
- The pilot, shakedown and main trial dates.

A Detailed Project Plan was also produced which built on the initial project plan. This specified system delivery and test dates, and for the post-trial period identifies which analysts will undertake which tasks and what the timings need to be for report delivery.

The Measurement and Analysis Specification was then produced which detailed the whole analysis in terms of aims, objectives, hypotheses, metrics and analysis methodology.

The Data Logging Specification was produced which detailed which system interactions needed to be recorded, and in what format. This was passed to the software team to implement.

A Report Plan which detailed the structure of the final report, an outline of the contents of each section, and the timing of all the drafts and reviews was also produced and agreed by the Head of Department.

It is worth noting that all of these plans, though agreed in principle, were subject to change throughout the preparation and analysis phases. Where possible, and beneficial, they were updated to reflect changes. However, particularly in the latter stages of analysis, novel issues were examined as they arose without retrospectively putting them in to the analysis plan. It is essential to keep this sort of flexibility, and these sorts of decisions were logged in the documentation that describes the output of the analysis.

15. Prepare Facilities.

A number of different teams prepare the facilities in accordance with the timetable in the project plan. These include the software engineers who develop the prototype, the hardware engineers who ensure that the correct machines are working on the validation platform (including R/T, videos, etc.), the data preparation and mapping teams, and the validation team.

EXAMPLE: A REAL-TIME VALIDATION

The detailed work on the part of the validation team included:

- Developing training material for the controllers (on the prototype and the trial environment). This included computer based training, briefings covering measurements and operational procedures to be followed, together with a number of 'mentored runs' where the mentors from the development team or the trials team used a checklist of functionality to ensure that the controllers were familiar with all the system functions;
- Preparing briefing material for other analysts in the trial, such as a briefing for the observers;
- Writing and testing scripts to extract the measurements required from the data logs;
- Preparing template spreadsheets to generate descriptive and comparative statistics.

16. Test Facilities.

Airspace maps and traffic samples were tested prior to the shakedown trial. The shakedown itself also provides an ideal opportunity for operational controllers to review the traffic samples in terms of realism. Our samples were based on traffic volume, but as at present there is no automatic tool for measuring other traffic characteristics that affect controller workload, the controllers' input was also used to assess the difficulty of the samples.

The software was tested throughout its development. This includes small-scale tests where one or two engineers or designers focused on different aspects of the system. Additionally, large-scale tests were carried out with an in-house person on every sector testing the software in staffed runs.

Data logging is tested as soon as it is implemented, typically 3 weeks before a shakedown. This consists of an analyst making number of inputs (and writing them down) for example on the prototype system or on the R/T, and then checking the subsequent logs. This is further tested by extracting the logs from the manned runs and 'sanity checking' them. Generally, this process is highly iterative as any changes in the software can make unexpected changes to the data logging. The extraction scripts and analysis templates are tested using this data. This testing can take a lower priority since it is vital that the correct events are logged during the trial, but errors in extraction scripts can be corrected after the trial.

The training is tested during the shakedown and evaluated through a debrief and a short questionnaire.

STEP 3: Execute Exercise

The range of measurements recorded in the trial were:

- Data logging;
- Controller questionnaire
- Group debriefs;
- One-to-one debriefs;
- Observations;
- Videos;
- Talk-through runs (where the controllers provided a monologue of the tools that they are using and why – 'thinking aloud').

EXAMPLE: A REAL-TIME VALIDATION

Though the trial timetable is set before the trial, there are a number of things that can happen to upset this. For example in this particular trial: a controller was sick, a controller attended a job interview, there was a fire alarm, and a power surge that crashed the R/T system for a morning. All these things had to be accommodated in the timetable by re-scheduling matched runs, questionnaires and debrief slots. For example, the controller who was for half a day to attend a job interview, was notionally allocated an exercise run where, had he not been away, he would have been operating a feed sector, rather than a measured sector. His replacement by another controller in this rôle did not therefore significantly affect the 'matching' of exercises. In the case of the R/T failure, the morning was re-assigned to questionnaires and debriefs.

At the end of each day, the data log files were checked, firstly to ensure that all the expected files were present and, where time permitted, to allow a quick sanity check of the data being produced. Additionally, the observation notes were reviewed to ensure that the data generated would be of sufficient quality; it is very easy to write a lot of comments that do not give the reader a full picture.

Another important consideration is to aim to have the controller questionnaire completed before the group debrief takes place. The debrief should be used to address the initial aims of the trial together with issues that have become apparent during the trial. Using the questionnaires as input to this provides the opportunity to focus on the areas where there is a range of opinion within the group, or to substantiate other responses.

STEP 4: Analyse Results

The analysis phase started with each analyst in the validation team following their part of the plan. Each worked independently on the different analyses required:

- The analysis of controller tasks;
- usability;
- the quantitative analysis of workload etc; and the subjective analysis of workload etc.

There were a series of meetings held either weekly or fortnightly in which all the analysts shared their main findings as they emerged. This helped to integrate the whole analysis and ensured that relevant questions and issues were aired among all in the analysis team. Additionally, daily meetings were held in which the project leader and the two main quantitative analysts reviewed the day's results. These meetings and reviews helped to catch all inconsistencies in the data, and to raise a number of new avenues of enquiry. Thus, though the analysis plan was followed, the flexibility of the analysis team ensured that it was augmented. Examples of the new lines of enquiry include, the questions 'Do we have any quantitative information on how the use of electronic strips changed over time?', or 'All the controllers had the full flight plan open, observation suggests that it was not used, can we find any objective evidence to suggest whether it was used or not?'

The videos were used primarily to check the consistency of the data. Where there were some suspicious data logging results, an appropriate exercise was reviewed to ascertain whether the logging was correct or not. Additionally, where one controller or one exercise was found to be noticeably different from the rest, the videos were used to supplement the subjective assessment of why they differed.

EXAMPLE: A REAL-TIME VALIDATION**STEP 5: Disseminate**

The target audience for the findings of the trial was identified very early on in the process. It consisted predominantly of the group who would be using of the system, but the report was meant to be intelligible to the wider Air Traffic Control community.

The reporting to the customer (from the user group) took three distinct stages. Firstly there was a brief presentation of initial results about 3 weeks after the trial was completed. This was to provide an early view of the likely headlines in the report, but was of course caveated by the fact that the results were from an immature analysis and may have been subject to some change.

Once the report was in an early draft form, and the analysis had matured, about 7 weeks after the trial, the customer was given a more detailed presentation of the results. This served two purposes. Firstly to ensure that the customer had an early view of the results rather than wait for the necessarily long report review procedure to be complete. And secondly, to allow full discussion of the results to identify any misunderstandings prior to report release.

Finally, release of the report served as the primary vehicle for dissemination of results. It was widely distributed and additionally copies were made available in the library. Using this as the primary vehicle, and making the main author the contact for questions, ensured consistency in presentation of the messages contained within the report.

As stakeholders, the design team was involved in the analysis team's weekly meetings. The aim of this was to ensure that all those who were present for the trial developed a consistent set of conclusions and the process worked well. If the design team were not present for the trial, then this type of involvement is still recommended as it will engender ownership of the results.

In this trial, the final stakeholders, the controllers, were not involved to any extent in the discussion of the results. This was because their views had been adequately represented in the collection of subjective data in questionnaires and debriefs.