

Safety Risk Management Methodologies (SRM) STPA – System-Theoretic Process Analysis



This document was developed by the Safety Management Panel (SMP). It is intended to support safety experts in the application of risk management methodologies. Any comments to this material should be forwarded to safetymanagement@icao.int.

CONTENTS

1.	Description	3
	1.a) Purpose of the methodology	3
	1.b) Theoretical basis	3
	1.c) Method overview	4
	1.d) Key terms and definitions used in this document	6
	1.e) Tools available	7
2 . L	Jser factors	8
	2.a) Applications	8
	2.b) Users	8
	2.c) Evaluation of complexity	8
	2.d) Availability of training	9
3. 0	Quality and consistency	10
	3.a) Consistency/differences from SMM concepts, terms, and definitions	10
	3.b) Validity and reliability of outputs	10
	3.c) Overall pros and cons	11
	3.d) Team assessment of usability	11
4. <i>A</i>	Additional information	12
	4.a) Abbreviations	12
	4.b) Literature, references	12
Арр	pendices	13
	Appendix 1 – Comparison of assumptions on accident causation (Leveson)	13
	Appendix 2 – Examples	14

1. DESCRIPTION

1.a) Purpose of the methodology

System-Theoretic Process Analysis (STPA) is a proactive and holistic hazard analysis and risk management tool. It derives from an accident causality model called System-Theoretic Accident Model and Processes (STAMP). Unlike other methods (e.g., Fault/Event Tree Analyses, Hazard and Operability (HAZOP), etc.), STPA attempts to identify hazards and risks across an entire system and not to focus on risks at the task level, which are typically associated with the human operator(s).

STPA was specifically designed to better address interactions between system components, rather than only component failures. In the STAMP model, safety is seen as an *emergent* property of work systems, and not as a property of their components. Indeed, in complex systems such as the ones extensively used in aviation, components that are deemed safe or reliable in isolation may nevertheless interact in ways that erode safety and may ultimately lead to an incident or accident.

STPA shouldn't be confused with another STAMP-based tool called Causal Analysis based on System Theory (CAST). Although they both use the same systemic model from STAMP, STPA and CAST have different purposes. CAST is a reactive incident/accident analysis technique. In other words, STPA seeks to identify and analyse the potential causes of negative events before they occur and to eliminate or control hazards and risks, whereas CAST is more suited to investigate a particular event that already happened.

1.b) Theoretical basis

Influenced by seminal work on risk management conducted by Jens Rasmussen among others, MIT professor Nancy Leveson launched STAMP in 2004. A 2011 book complemented her initial theoretical proposition with two analysis techniques intended for practitioners (i.e., STPA and CAST) and includes examples and comparisons with other approaches (e.g., domino theory, probabilistic assessments, Swiss Cheese Model). From a STAMP perspective, accidents "are the result of a complex process that results in the system behaviour violating the safety constraints" (Leveson, 2011, p. 92). See also appendix 1 for a summary of the assumptions and tenets in Leveson's perspective to engineer a safer world.

STAMP is deeply rooted in two complementary theoretical frameworks that were developed after World War II to deal with increasingly complex systems: systems thinking and control theory. More precisely:

- Systems thinking doesn't only look at component failures at the 'sharp end' (i.e., frontline work), and goes beyond linear, direct causality relationships between system components. It takes a broader perspective by also considering non-technical aspects, system objectives and constraints, emerging and potentially unpredictable interactions involving people and machines, etc. Systems thinking is primarily concerned with emergent properties that arise when all technical and social (i.e., human) components of a system interact, thereby extending our understanding of both normal operations and accident causation in dynamic, complex systems.
- Control theory looks at how interrelated components of a system interact (or not) between themselves and with their environment, whether they can settle into a state of equilibrium, and whether they can be thrown out of equilibrium. At the core of control theory, feedback loops of information and control play a central role to keep the components in a state of dynamic equilibrium. As an example of rudimentary control loops, consider how a thermostat senses an

ambient temperature and controls an associated heating system to achieve and maintain a target temperature in a room.

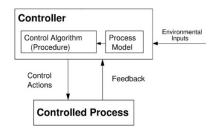


Figure 1 – Example of a simple control loop (Leveson, Thomas)

1.c) Method overview

An STPA analysis typically contains four steps.

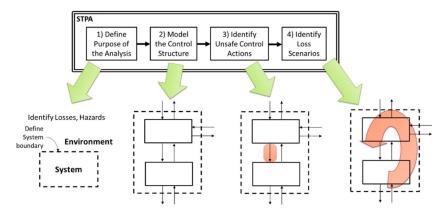


Figure 2 – Overview of the basic STPA method (Leveson, Thomas)

- 1. Define the purpose of the analysis: this step aims to identify all kinds of losses to be prevented, as well as both the hazards and the constraints present at system level (i.e., rather than at the level of an individual component). The scope of the analysis can be defined by answering questions such as: will the losses to be considered and analysed only concern safety matters, or will it also include other pertinent considerations such as security, organizational performance, customer satisfaction, etc.? What is the system to be analysed? What are its boundaries versus other systems and its environment? Who are its stakeholders?
- 2. Model the system (also called a Control Structure), using feedback control loops. This Control Structure captures functional relationships and interactions. Various elements and components of the system may be represented. However, the focus in this second step is primarily on what functions, actions, constraints, behaviours, etc. take place within the system rather than laying down each of its individual components. It usually begins at a very abstract level and is refined and detailed through various iterations. There are only few but important conventions when drawing a Control Structure:
 - The vertical placement of the various elements represents the hierarchy of control, from high-level controllers at the top to the lowest-level entities at the bottom,

- Downward arrows represent Control Actions (i.e., commands), and
- Upward arrows represent feedback.
- 3. Analyse the Control Structure to identify Unsafe Control Actions: once the Control Structure is built in sufficient detail (e.g., typically with three or four levels as a start), its Control Actions must be analysed to examine how they could lead to the losses defined in step 1. This third step focuses on the existence and timing of each Control Action, since they can only create a hazard and become unsafe in one of four ways:
 - Providing the Control Action leads to hazard(s), or
 - Not providing the Control Action leads to a hazard, or
 - Providing a potentially safe Control Action too early, too late, or in the wrong order leads to a hazard, or
 - The Control Action lasts too long or is stopped too soon.

While the graphic representation of the Control Structure will likely continue to evolve throughout the STPA analysis, step 3 generally calls for a methodical approach where the four ways for Control Actions to become a hazard are systematically considered, and the answers are recorded (e.g., by using a simple spreadsheet or similar tool). Answering questions such as "in which condition(s) can providing this Control Action lead to a hazard?" or "what if that Control Action is stopped too soon?" provides insights on that particular Control Action, on its feedback loop(s), but also on system behaviours, constraints, etc.

- **4. Identify Loss Scenarios**: done through analysis of the reasons why Control Actions can become unsafe (as found in step 3). This identification effort must consider two types of Loss Scenarios, and answer the following:
 - Why would Unsafe Control Actions occur?
 - Why would Control Actions be improperly executed or not executed, leading to hazards?

1.d) Key terms and definitions used in this document

Term	Explanation
Complicated system	Built from a significant number of elements, with well-defined function(s) to perform, and governed by clear and understandable laws. Examples: a jet engine, the hydraulic system of an airliner, a steam locomotive, etc.
	(Adapted from Grabowski & Strzalka, 2008)
	"A complicated system is () mechanistic in that all the parts, components, and their interactions are knowable. However, the structure and interactions in a complicated system may be difficult to understand, and the system may have multiple functions. Experts with appropriate qualifications can understand and analyse these systems with a high degree of accuracy. An understanding of linkages and interactions of system components is developed linearly, where an understanding of one element leads to an understanding of the next element. The impact of one element on another can be reasonably predicted. The relationship between cause and effect is linear. However, a single cause may have multiple effects, and a single effect may be the result of several possible causes." (ICAO website - Human Performance page)
Complex system	Built of many elements, which may or may not be identical, cooperating together according to rules which may or may not be well defined, but that can also change with time. These elements are connected in a network pattern (e.g., one-to-many and many-to-one network thinking) and can interact to create nonlinear and emergent behaviours. Example: an airliner, a nuclear reactor, etc.
	(Adapted from Grabowski & Strzalka, 2008; Sturmberg, et al., 2016)
	"A complex system is dynamic. The whole of the system is greater than the sum of its parts and components. Interactions between parts and components are diverse and nonlinear because everything is connected to, and depending on, something else. The behaviours of these systems may change in unpredictable ways. Analysis - such as causal analysis - of complex systems is not an effective method to improve system performance because the behaviour of the system cannot be predicted from examining the behaviour of its separate parts, and the system cannot be understood by only looking at one component or from one perspective.
	Complex systems often behave unpredictably or randomly due to the diversity of interactions within the system, the unpredictable nature of system components (such as humans or weather), and/or changing influences within the system. For example, an individual may change behaviour, adapting to internal influences, such as health or personal mood, as well as to external influences, such as environment or equipment. A complex system may exhibit unpredictable behaviour even though its performance may be defined by strict policies and procedures. Unlike simple or complicated systems, complex systems have the unique ability to learn and adapt, which can be attributed to the human component of the system." (ICAO website - Human Performance page)
Control Action	In general, an action taken to control some process and to enforce constraints on the behaviour of the controlled process. Impact is determined by feedback. (STPA Handbook)
Control Structure	"A hierarchical control structure is a system model that is composed of feedback control loops. An effective control structure will enforce constraints on the behaviour of the overall system."
Controller constraint	(STPA Handbook) "Specifies the controller behaviours that need to be satisfied to prevent Unsafe Control
	Actions." (STPA Handbook)

Term	Explanation
Hazard	"System state or set of conditions that, together with a particular set of worst-case
	environmental conditions, will lead to a loss."
	(STPA Handbook)
	"A condition or an object with the potential to cause or contribute to an aircraft incident or accident."
	(ICAO Annex 19)
Loss	"A loss involves something of value to stakeholders. ()"
	(STPA Handbook)
Loss Scenario	"Describes the causal factors that can lead to the Unsafe Control Actions and to hazards."
	(STPA Handbook)
Simple system	Consists of a relatively small number of elements that act together according to laws that
	are well-defined and easy to understand (for instance, a pendulum, a clock, most
	mechanical systems, etc.). However, simple systems do not necessarily remain so and,
	under some conditions (e.g., disturbances, chaos), may behave in complex ways.
	(Adapted from Grabowski & Strzalka, 2008)
	"A simple system is mechanistic, comprised of a defined number of components or parts which interact to accomplish one or a small number of known goals or functions. The
	behaviours of the system are fully predictable and do not change over time. Therefore,
	when the system is degraded or fails, cause is easily identifiable. The relationship between
	cause and effect is linear and clear. As such, these systems are straightforward to maintain
	and repair to ensure that they consistently meet their pre-identified performance standards."
	(ICAO website - Human Performance page)
Socio-technical system	Involves complex interactions between humans, machines (hardware and/or software),
	and their environment. Examples: an airline, a hospital, a powerplant.
System	"Set of components that act together as a whole to achieve some common goal, objective,
·	or end. A system may contain subsystems and may also be part of a larger system."
	(STPA Handbook)
System-level constraint	"Specifies system conditions or behaviours that need to be satisfied to prevent hazards (and ultimately prevent losses)."
	(STPA Handbook)
Unsafe Control Action	"Control action that, in a particular context and worst-case environment, will lead to a hazard."
	(STPA Handbook)

Table 1 - Key terms and definitions

1.e) Tools available

Although they're not required, half a dozen STAMP-based software tools have been developed to support STPA users. Some are open-source and available for free, whereas others require a license.

For more information and download: http://psas.scripts.mit.edu/home/stamp-tools/

2. User Factors

2.a) Applications

STPA can be used for hazard identification and risk management in any type of aviation organization, and for any system. Although STPA can be used for both strategic and tactical risk management, it is arguably slightly better suited for the former. Put differently, STPA helps identifying hazards by creating detailed models of systems and by uncovering their sometimes-intricate interactions at any stage of a system's life cycle. However, STPA is not a rigid framework that precisely guides how to assess, quantify, or mitigate risks. In this sense, STPA complements the existing SMS framework as specified in Appendix 2 of Annex 19 – Safety Management.

However, (very) small organizations may encounter some specific challenges typically related to the demands already placed on each staff member by their normal workload, but also related to the fact that a significant amount of the activities of the organization may be contracted (i.e., to some extent, a number of key processes are likely to be both out of their sight and out of their control). In such cases, it would be advisable that STPA analyses are jointly conducted by several organizations pooling their resources for a broader and sharper systemic perspective, and in turn achieve maximum effectiveness and efficiency. Obviously, this is also valid for much larger organizations, for instance when considering how internal departments cooperate to deliver a product or service.

Note – papers presented in conferences related to the STAMP model indicate that, in aviation, STPA was mostly applied in either purely technical systems (i.e., engineering, design) or larger socio-technical systems (i.e., encompassing an entire organization), and only very rarely in purely social systems (i.e., no technology involved).

2.b) Users

STPA assists in modelling, designing, operating, and upgrading entire systems, primarily through numerous iterations and refinements. It is therefore not the type of 'tactical' risk management tool that can be sketched during an in-flight emergency or on the corner of a desk in the middle of a dynamic situation. STPA is more of a 'strategic' risk management tool, to be found in team meetings and working groups involving many if not all departments of one or more organizations (e.g., civil aviation authority, safety committee, safety review board, etc.). Implicating many participants doesn't mean they all need to be actively involved at every step of the development or review of an STPA analysis.

However, the comprehensiveness of STPA analyses doesn't preclude from using Control Structures (or simpler/partial versions) for training or communication purposes. Feedback loops usually are easy to explain and understand. Depending on the audience and key message(s) to be conveyed, it is not always necessary to show an entire Control Structure in all its detail. The granularity of STPA analyses can therefore be adjusted to suit particular needs.

2.c) Evaluation of Complexity

Although the method itself requires some initial training and effort, it is not complex to learn and apply, and abundant resources are available to support learning (many of which are free). As stated earlier, feedback loops usually are easy to explain and understand although they may become more complicated, for instance when the need arises to detail any actuator or sensor included in a feedback loop. Overall,

the complexity of the system to be analysed and the depth of the analysis are the principal determinants in the complexity of an STPA analysis. Organizations are therefore encouraged to start with simple or complicated systems before analysing complex systems too (see 1.d above for a distinction between the three). Put differently, STPA allows to model various types of systems and system behaviours in ways that other methods cannot replicate (i.e., at least not easily or not as clearly and comprehensively), including complex, nonlinear, and emergent behaviours. This inevitably impacts the complexity of the analysis but also comes with a better grasp of the system.

Moreover, the STPA Handbook pre-emptively lists mistakes that are commonly made during STPA analyses. This should support most organizations from the early stages of implementation.

To summarize, STPA is an appropriate method to model systems for the purpose of identifying hazards and managing risks, regardless of the level of complexity of the system. This ability also means that, in contrast with other risk management methods that became popular in the late 20th century, the STPA method does not require significant if not counterproductive workarounds and simplifications that are needed when using linear causality to model complex systems.

2.d) Availability of training

An MIT website dedicated to STAMP offers a free STPA Handbook available in several languages: http://psas.scripts.mit.edu/home/materials/

Moreover, several online tutorials on the basics of STAMP/STPA/CAST are also available for free: http://psas.scripts.mit.edu/home/mit-stamp-workshop-tutorials/

MIT also hosts free, week-long STAMP workshops every year since 2012. These training events address STAMP, STPA, CAST, and serve as a platform for professionals from all industries to share their experience and ideas. Introductory material and theoretical courses on the framework itself can be watched online and are recommended before attending. Workshops are generally held on-campus in Cambridge, MA (USA). For more information, including past presentations: http://psas.scripts.mit.edu/home/stamp-workshops/

Other training events, workshops, and conferences are sometimes organized in various parts of the world (e.g., Europe, Japan). http://psas.scripts.mit.edu/home/other-stamp-meetings/

Finally, the number of publications on STAMP and particularly on STPA sharply increased since the mid-2010s. Although some publications are behind a paywall, many others can be easily accessed. The MIT maintains a PDF file organized by industry, topic, etc., including links: http://sunnyday.mit.edu/theses/STAMP-publications-sorted-new.pdf

3. Quality and Consistency

3.a) Consistency/Differences from SMM concepts, terms, and definitions

The STPA Handbook defines several key terms differently from ICAO publications, but without any significant incompatibility (see 1.d).

Broadly speaking, the STAMP model and its STPA and CAST tools are consistent with concepts found in the SMM and other ICAO publications such as the Human Performance section of the ICAO website, ICAO Doc 10151, or the Safety Management Implementation (SMI) Website. Aviation personnel who have recently embarked on their journey to learn about Safety Management Systems and others who discover STAMP, STPA or CAST for the first time may perceive differences in STAMP's approach to human error, to accident causation (for instance, see 5.a for Leveson's perspective on this topic), to probabilistic risk assessments, etc. Different models and paradigms coexist in safety science for decades. The question is not to define which is right or wrong but, considering their complementarity, to identify which is the most suitable and effective to address the specific need(s) of an organisation.

It is worth noting that a number of publications explored modifications, combinations, and extensions to the methodology in an effort to better embed STPA into an overall risk management framework, and in risk assessments in particular. Similar initiatives have appeared in the area of security management.

3.b) Validity and reliability of outputs

Many publications, including peer-reviewed, scientific papers, have confirmed the value of using STPA to identify hazards, support risk management activities, and analyse complex sociotechnical systems in high-risk industries such as aviation. An increasing number of well-established aerospace firms have made experiments and adopted the STAMP model, but generally do not publicly communicate about it. Although the validity and reliability of outputs are not questioned, the current absence of built-in methods or tools to quantify risks within a system is sometimes perceived as a hindrance.

3.c) Overall pros and cons

Pros & strengths

- Underpinned by systems theory.
- Considers the loss of control(s) across the whole sociotechnical system in a more comprehensive manner than traditional models (e.g., Fault Tree Analyses, BowTie, etc.) and therefore doesn't only look for risks at front-line level.
- Considers hazards residing throughout the system but also their interactions.
- Shown to be more effective at the identification and analysis of early warnings than conventional approaches.
- Likely to be more efficient than traditional models (i.e., richer results obtained from comparatively less resources), without excluding linear causality which remains relevant for simpler systems and components/elements.
- Numerous resources, including training, documentation, and IT tools, are freely available and regularly updated.

Cons & limitations

- Requires at least an understanding of systems thinking principles.
- Requires some training and practice.
- Doesn't seek or help to quantify risks or losses (quantification deemed often too unreliable in complex sociotechnical systems to be useful, although it may remain valuable for purely technical, engineering aspects of the system)
- While effective at analysing complex systems, it inevitably requires proportional investments to achieve its goal, which may prove problematic for the smallest organisations using STPA in isolation.
- STPA being steeped in engineering, the role of front-line workers adapting to real-time conditions and operational variability may be difficult to model with precision.
- Does not address the dynamic, non-linear behaviours of complex systems.

Table 2 – Perceived strengths and limitations of STPA

3.d) Team assessment of usability

As one of the few methodologies to have embraced the systems thinking school of thought on 'human error' and safety management, STPA appears useful to thoroughly analyse complex sociotechnical systems at any stage of their lifecycle. In addition to being recognized by academics, STPA is well supported by an "open source" platform of practitioners and users steered by the main author and contributors from MIT (i.e., Nancy Leveson, John Thomas, etc.). Experience with STPA is growing across high-risk industries and indicates that it is a valuable complement to traditional analysis techniques, if not an alternative. While STPA may not be an optimal choice for tactical risk assessments during high tempo operations or in (very) small organisations, this methodology grounded in engineering principles fosters a better refined understanding of complex systems.

4. Additional information

4.a) Abbreviations

Abbreviations	Meaning	Notes
CAST	Causal Analysis based on Systems Theory	
MIT	Massachusetts Institute of Technology	
SMM	Safety Management Manual	ICAO Doc. 9859
STAMP	System-Theoretic Accident Model and Processes	
STPA	System-Theoretic Process Analysis	

Table 3 - Abbreviations

4.b) Literature, references

Allison, C.K., Revell, K.M., Sears, R., Stanton, N.A. (2017). *Systems Theoretic Accident Model and Process (STAMP)* safety modelling applied to an aircraft rapid decompression event. Safety Science. Volume 98.

Castilho, D.S., Urbina, L.M.S., de Andrade, D. (2018). *STPA for continuous controls: a flight-testing study of aircraft crosswind takeoffs*. Safety Science. Volume 108.

Dallat, C., Salmon, P., Goode, N. (2019). Risky systems versus risky people: To what extent do risk assessment methods consider the systems approach to accident causation? A review of the literature. Safety Science. Volume 119.

Grabowski, F., Strzalka, D. (2008). *Simple, complicated, and complex systems – The brief introduction*. Rzeszow University of Technology. Department of Distributed Systems.

Hulme, A., Stanton, N.A., Walker, G.H., Waterson, P., Salmon, P.M. (2019). What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. Safety Science. Volume 117.

Le Coze, J.C. (2022). The 'new view' of human error. Origins, ambiguities, successes and critiques. Safety Science. Volume 154.

Leveson, N. (2002). System Safety Engineering: Back to the Future. Massachusetts Institute of Technology.

Leveson, N. (2004). *A new accident model for engineering safer systems*. Safety Science. Volume 42. Issue 4. Retrieved from: http://sunnyday.mit.edu/accidents/safetyscience-single.pdf

Leveson, N. (2010). Applying systems thinking to analyze and learn from events. Safety Science.

Leveson, N. (2011). Engineering a Safer World: Systems Thinking Applied to Safety. The MIT Press. Cambridge, MA.

Leveson, N. (2020). An Improved Design Process for Complex, Control-Based Systems Using STPA in a Conceptual Architecture. Massachusetts Institute of Technology. Retrieved from: http://sunnyday.mit.edu/conceptual-architecture-final.pdf

Leveson, N., Dulac, N., Marais, K., Carroll, J. (2009). *Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems*. Organization Studies. Volume 30.

Leveson, N., Thomas, J.P. (2018). STAMP Handbook. Retrieved from: http://psas.scripts.mit.edu/home/materials/

Patriarca, R., Chatzimichailidou, M., Karanikas, N., Di Gravio, G. (2022). *The past and present of System-Theoretic Accident Model and Processes (STAMP) and its associated techniques: A scoping review.* Safety Science. Volume 146.

Sturmberg, J.P., Martin, C.M., Katerndahl, D.A. (2016). *It is complicated! – Misunderstanding the complexities of 'complex'*. Journal of Evaluation in Clinical Practice.

APPENDICES

Appendix 1 – Comparison of assumptions on accident causation (Leveson)

In her 2011 book (p. 57), Leveson contrasts different assumptions on normal work and accident causation. Her 'new assumptions' can be considered as core tenets of the STAMP framework. They're reproduced here to assist the reader in better understanding the resolute turn taken in STAMP towards a more holistic approach of systems thinking and risk management. STAMP does not negate or ignore the progress made in the '80s and '90s nor the concepts that emerged at the time (e.g., mainly from Reason, Perrow, Turner, and many others) but pushes several arguments farther, partly thanks to the insights gained with linear models of accident causation.

Old assumption (in linear models)	New assumption (systems thinking)
Safety is increased by increasing system or component	High reliability is neither necessary nor sufficient for
reliability; if components do not fail, then accidents will	safety.
not occur.	
Accidents are cause by chains of directly related	Accidents are complex processes involving the entire
events. We can understand accidents and assess risk by	sociotechnical system. Traditional event-chain models
looking at the chains of event leading to the loss.	cannot describe this process adequately.
Probabilistic risk analysis based on event chain is the	Risk and safety may be best understood and
best way to assess and communicate safety and risk information.	communicated in ways other than probabilistic risk analysis.
Most accidents are caused by operator error.	Operator error is a product of the environment in
Rewarding safe behaviour and punishing unsafe	which it occurs. To reduce operator "error" we must
behaviour will eliminate or reduce accidents	change the environment in which the operator works.
significantly.	
Highly reliable software is safe.	Highly reliable software is not necessarily safe.
	Increasing software reliability will have only minimal
	impact on safety.
Major accidents occur from the chance simultaneous	Systems will migrate toward states of higher risk (i.e.,
occurrence of random events.	drift). Such migration is predictable and can be
	prevented by appropriate system design or detected
	during operations using leading indicators of increasing
	risk.
Assigning blame is necessary to learn from and prevent	Blame is the enemy of safety. Focus should be on
accidents or incidents	understanding how the system behaviour as a whole
	contributed to the loss and not on who or what to
	blame for it.

Table 4 – Assumptions on accident causation (Leveson, 2011)

Appendix 2 - Examples of control structures

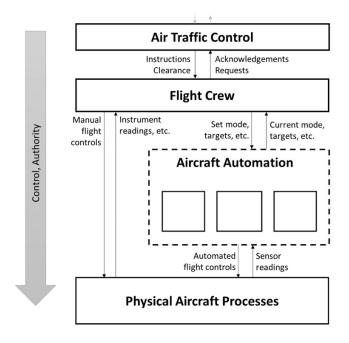


Figure 3 – Simple example of a control structure in aviation (Leveson, Thomas)

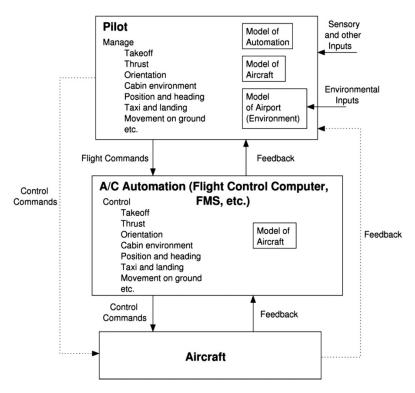


Figure 4 – Simple example of a control structure at the level of an aircraft (Leveson, Thomas)

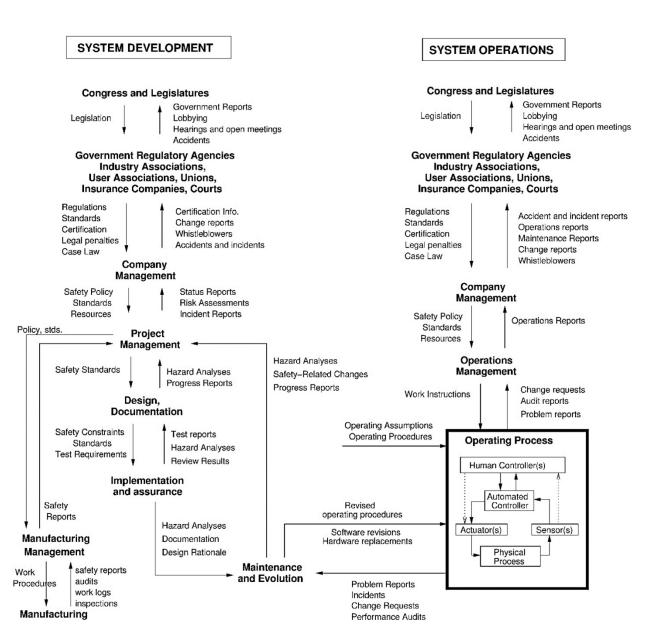


Figure 5 – High-level control structures depicting system development and operations (Leveson, Thomas)