# Safety management – a multi-level control problem

**Björn Wahlström[a], Carl Rollenhagen[b]**
[a]Vattenfall SMI, AX-22920 Brändö, Finland
[b]Vattenfall and KTH, SE-162 87 Stockholm, Sweden

**Abstract:** Safety management is a crucial activity in maintaining acceptable safety levels of large hazardous industrial facilities. Risk analysis and safety engineering are two important activities of safety management by which safe designs of such facilities can be achieved. A continued safety during the operation of the facilities relies furthermore on successful and efficient experience feedback and management of change. Activities in safety management build on a control metaphor by which control loops built into the technical, peoples and organisational systems ensure safety of the facilities. In this paper we take a closer look on concepts of control theory to investigate their relationships with safety management. A conclusion of the paper is that the control metaphor provides useful insights in suggesting requirements to be placed on safety management. The paper draws on experience from the Vattenfall Safety Management Institute (SMI), which started its operation in 2006.

**Keywords:** Safety management, Control structures, Modelling safety, Nuclear power.

## 1. INTRODUCTION

Safety of large scale industrial activities such as nuclear power, production of chemicals and off-shore oil production, has through some devastating accidents become a growing concern in society. Safety management and safety culture within organisations that design, build, operate and maintain such facilities are today seen as crucial components in a continued safety. Unfortunately however, in spite of available guidance on how to implement safety management and to enhance safety culture, there are still controversies involved in what these two concepts actually mean and how they should be applied. This paper reflects on safety management and safety culture using a conceptual frame of control engineering.

A common approach in ensuring safety is governed by a control metaphor, which is applied at several hierarchical levels [1]. The society exercises control of companies that operate hazardous facilities through laws and regulations. The companies implement policies and management systems to ensure that plants are designed, built, operated and maintained in a safe manner. The controls include both feedback and feed forward paths by which outcomes are monitored and correcting actions are initiated when deficiencies are detected.

A generalised control task assumes a system S that can be manipulated with inputs u and that gives observed outputs y. In addition it is assumed that the system S can be characterised with a state variable x, which has the property that by giving its state $x_0$ at time $t_0$ and the system input u(t) from $t_0$ onwards, unambiguously define future outputs y(t) [2]. This means that the state $x_0$ can be seen as integrating the history of past inputs u(t) for $t<t_0$. The following four conditions are then necessary for successful control

- a system model, i.e. some way to make predictions (deterministic or probabilistic) what certain control actions may give as an outcome,
- observability, i.e. one has to be able to determine the state of the system from outputs that can be measured,
- controllability, i.e. one has to be able to control the state of the system with inputs that can be manipulated.
- a preference relation, i.e. some way to separate between good and bad outcomes.

## 2. CONTROL OF SAFETY

Control of safety can applying a generalised control task be interpreted as two tasks, firstly that the system is kept in a safe region of the state space (state control) and secondly if the system has entered an unsafe region there exists an input, which transfers it to safe state (transition control). Control of safety would therefore suggest a need for distinguishing between safe and unsafe regions of the state space of the system, as well as

control inputs that will maintain system state within or transfer it to a safe region of the state space. In the following we look at the four necessary conditions for successful control.

### 2.1. The system model

The first important step in selecting a specific system model is concerned with the intent of the modelling effort. Which variables and influence mechanisms should be included and which can be left without consideration? A system model is assumed to have an internal structure of subsystems, which give micro-explanations of macro-behaviour that can be observed through system inputs and outputs. The subsystems may also have their own internal structure.

The next step in building a system model is to make an assessment of the system state, which is reflected in the selected model. Modelled subsystems provide a clue to state components, but an aggregation that is connected to the intent of modelling may be necessary. There are several possibilities to construct models of industrial facilities, which all are connected to some specific purpose to understand or to control. When the purpose is to control safety, the model has to be built on subsystems, influence mechanisms and variables that have an influence on safety. Conditions defined in the safety analysis report for a facility can for example provide one set of necessary conditions for safety, to define a safe region in the state space of the facility.

### 2.2. Observability and controllability

Observability and controllability are strictly speaking connected to a specific model of the system. Both concepts are defined trough the state of the system, the space of inputs and the space of outputs. A system is observable if its state can be determined from observed outputs and it is controllable if its state can be controlled from its inputs. Observability implies therefore that we can say if the system is in a safe state or not. Similarly we should also be able to say if the system is in an unsafe state from where controls can be initiated to transfer it to a safe region of the state space.

Observability is a weaker condition than measurability, which requires a suitable scale (nominal, ordinal, interval, ratio) on which observed variables can be given a value. If the system state is measurable it means that it can be obtained using a measuring instrument at a specific time. If it is only observable, a state estimator is needed, which collects information over time through measurable components of system outputs.

In control of safety the controllability requirement implies that there are controls that keep the system within a safe region of the state space or controls that transfer it from an unsafe to a safe state. This means that state components should be possible to manipulate in desired directions through system inputs either directly or indirectly. The system model gives information on how these controls should be constructed.

### 2.3. The preference relation

The preference relation gives a way to calculate the value of certain control inputs in a time interval from $t_0$ to $t_1$ in which the system moves from an initial state $x_0$ to an end state $x_1$. In control of safety the path of state transition is important, because it should as far as possible be maintained in safe regions of the state space. Other component of the preference relation may be related to the efforts of calculating and applying necessary inputs to the system.

In control of safety a simple preference relation may be constructed by to associating very large costs to state transitions that take momentary excursions outside safe regions of system space. Uncertainties contained in model predictions and state estimates may make it necessary to require margins in setting safety targets. The time integral of such margins could then provide measurement of safety in state transitions. An extension from deterministic to probabilistic system models can be handled using the concepts of probability, probability distributions and expected values [3].

### 2.4. Models of how safety is constructed

The challenge in constructing safety is to ensure reliability with non-reliable components. Safety engineering has with time and increasing experience built models and strategies for ensuring safety [4]. Some of the safety strategies have been integrated in guidelines for system design and others in how to monitor and

improve safety during operation. Among them are the principles of a graded approach to safety and defence in depth. These principles are embedded in the requirement that multiple independent barriers should be erected against unwanted system excursions. Protection of the barriers can then be achieved by introducing the single failure criterion and applications of redundancy, separation and diversity.

An important lesson has been that technical systems are operated by humans that are working in an organisation. This means that the system to be controlled consists of three very different subsystems that should be modelled based on their own assumptions. These systems have sometimes been termed man, technology and organisation (MTO) or plant, people and processes (PPP). A large industrial facility can therefore be considered as three subsystems all with their own internal controls and with some system level controls to ensure coordination between the three subsystems.

## 3. CONTROL STRUCTURES IN USE

In modelling, designing, operating and maintaining specific controls one may separate between five simple control structures, which in the control of large industrial facilities are used and combined in various ways. Since these simple controls have their own characteristics, it is important to understand how they can be applied and combined. Controls are implemented through some control actor, which reads outputs of the controlled system and applies control inputs that are calculated according to some control algorithm. The control actor may be an automatic system, a single person, an organisational unit or some combination of them.

### 3.1. Open loop control

Open loop control is the simplest control structure. It is a control in which no direct feedback from the system state is used to determine the control to be used. When this control structure is used in practice, it is based on earlier experience, i.e. a model of cause and consequence that is applied to select the controls to be used in specific situations. This control structure is simple to implement, but has the drawback that it requires a large set of pre-calculated control inputs for situations that may occur.

This type of control structure was often applied for the control of emergency situations before the TMI accident. The rationale for this control structure is that such situations require accurate and rapid responses, which if calculated in advance are likely to perform better than ad hoc responses. Some feedback from the system state, can with this control structure be introduced, using a set of if-then rules.

### 3.2. Closed loop control

Closed loop control relies on a continuous feedback of system outputs for calculating control inputs. It therefore provides larger possibilities of adapting to given situations. Closed loop controls assume that an agreed set point can be defined, i.e. a norm for what can be considered as a target state of the system. When the target state is given, the control can be calculated based on the difference between actual and target state.

Closed loop control gives improvements over open loop control, but it still have drawbacks. One is connected to the existence of large differences in time constants in the system to be controlled. In this case the controller should be carefully tuned not to let short term behaviour of the system influence controls that are aimed at long term behaviour. A similar problem is connected to non-linearities in system behaviour, because they would in most cases make it necessary to change control parameters depending on the state of the system.

### 3.3. Adaptive control

Adaptive control provides a control structure, which adds a second control loop that is aimed at adjusting specific control parameters to make the overall control performance adapted to situational needs. In the simplest case this may take place with a tuner from system state to change control parameters in a closed control loop. Such schemes can be used when the system model is accurately known and the necessary parameter adjustments can be calculated in advance.

A more intricate adaptive control structure can be called model reference adaptive control. That means that a system model is calculated continuously based on inputs and output to the system. This model is then used to

synthesize the best controller to be applied. This control structure has its advantages in noisy environments, where the controlled system is exposed to variations in its environment. This control structure is nearly related to the structures of double loop learning that have been proposed [5].

### 3.4. Learning control

Learning control takes a step towards the possibility to change not only local control parameters, but also the whole control structure. The difficulty in this connection is to determine on what grounds one control structure should be regarded as better than another. If we have a good model of the system and its controls, we may use simulation and simply with trial and error select the most appropriate control structure. When more refined system models are introduced to include new situations and operational modes, they may be possible to suggest better control structures to capitalize on the improvements. Innovations may also provide an impetus to search for better control structures.

Observed flaws or problems in used controls may also initiate a search for improvements. Typical problems involve dysfunctional feedback, inaccurate models, flawed algorithms, failed control agents, etc. A classification of archetypes of control flaws leading to accidents has been proposed [6]. It is however important that changes in existing control structures are analysed carefully before implementation, to avoid introducing of new problems in trying to correcting old ones.

### 3.5. Hierarchical control

A common extension of single control loops is to build several controls into a hierarchical structure, where controls on a higher level are used to influence set points, control parameters and preference functions for controls on lower levels [7]. The benefit of building controls in a hierarchical structure is that control design can be partitioned into independent tasks on a subsystem level and then provide coordinating control on a systems level. This means for instance that simple logic controls can be applied for the protection of components on the lowest level and that more refined controls carry out state control or start-up and shut-down sequences on the subsystem and system level.

Hierarchical control is also a typical solution for systems that have special operational modes. A power plant for example could be in normal operation or shut down, where the two modes are controlled by two different sets of hierarchical controls. Similarly the loss of a critical component or function by technical failures or human errors may necessitate a transfer to an emergency mode, where safety systems are activated and the preference relation regarding outputs is changed.

## 4. MANAGEMENT SYSTEMS

The management system is supposed to exercise organisational control of a facility that consists of three interconnected subsystems. It should provide necessary controls to make it possible for the three systems to function in a coordinated manner to ensure that organisational goals can be reached. Present views on management systems build on two earlier constructs, quality systems and organisational handbooks. These functions are today typically combined into one integrated management system [8], which in addition to performance and safety goals also includes requirements in areas such as environmental protection and labour safety. The management system is assumed to break down goals and requirements for organisational functions and to make them concrete in instructions and documents.

### 4.1. Organisational design

The design of an organisation involves definition of structure and the specification of interrelated functional units. A typical division of organisations that operate large industrial facilities is to separate between operations, maintenance and technical support. Operations is responsible for the 24/7/365 hands on operation of the facility, maintenance for corrective and preventive actions to ensure operation over extended periods and technical support for various activities that operations and maintenance may need.

Organisational structure involves two constructs, a set of processes and a line of responsibility and authority that are used for recurrent activities, task and actions. The line of responsibility and authority defines delegation and reporting from the CEO through organisational units and down to single individuals. Processes are subdivided into sub-processes, tasks and actions, which are given to organisational units and

individuals through instructions at various levels in the organisation. Many of the instructions that are used at large industrial facilities are in fact a kind of control algorithms.

### 4.2. Control loops and control agents

Organisational controls rely on control agents that can be single individuals or organisational units. Several control agents may participate in one control loop with tasks of information collection, decision making and communication. The functionality of a specific control loop will in addition to the four necessary conditions for successful control depend on an efficient interaction between participating control agents. This interaction may break down if participating control agents use different system models and/or preference relations in their control actions.

A common lesson from incidents and accidents is that prescriptions of the management systems are not always followed to the point. In some cases this can be blamed on ambiguities in instructions, but it is in most cases it can be seen as a consequence of the organisational culture [9], i.e. norms, preferences, practices and habits. This observation would propose the installation of a loop for control of organisational culture.

### 4.3. Management system structures

The organisational structure and the tasks of the organisational units should be reflected in the management system. In addition management systems will typically contain mission and value statements together with policies and strategies that give the preference relations on the highest hierarchical level in the organisation. Lower level preference relations are defined in procedures, instructions and other documentation. It has been argued that these preferences can be structured in means-ends hierarchy, where means on a higher level define ends on a lower [10].

Viewing the management system as the control system of the organisation, a number of requirements can be proposed. In addition to the four necessary conditions for successful control one may propose that the management system should be understood, accepted and used. Furthermore it should be documented, updated and fitted to its purpose. A large industrial facility can typically contain tens of systems, hundreds of subsystems and thousands of components and can be operated from hundreds to thousands of people. It is therefore easy to understand that the management system easily may contain thousands of instructions and documents, which together may account to more than hundred thousand written pages. This would suggest documentation control to be one important part in the management system.

### 4.4. Operations, maintenance and support functions

Operations and maintenance differ from the support functions in that respect that they are in direct contact with the technical process, which has been said to execute controls at the sharp end. The support functions include technical support, human resources, procurement, finances, etc that produce necessary support functions for the sharp end and they could correspondingly be said to execute controls at the blunt end.

The sharp and the blunt end differ in many aspects. Firstly failures in the controls at the sharp end are often seen immediately, where failures in the blunt end often are hidden and may be so for long time intervals. Secondly control actions in the sharp end develop in real time, where similar control actions in the blunt end may take their time. Finally instructions for the sharp end should typically be followed to the point, where instructions at the blunt end have a more guiding nature.

### 4.5. Performance indicators

Performance feedback can be obtained using measurable system outputs. Such components however, often characterise past performance (lagging indicators) and do not necessarily predict future performance (leading indicators). To obtain leading indicators, they should be related to system state at a given time. This implies that a state estimator should be built that can provide reliable and valid feedback of system performance. In this way a set of leading indicators may be obtained to supplement lagging indicators that are collected by direct measurements. The yearly planning of activities can be seen as a control loop with the quality circle components of plan-do-check-act. In the planning earlier performance is used to create plans for a new cycle and to define performance indicators that can be used for detecting deviations from the plan.

A state estimator can at least in principle be built by using models of how certain variables in the technical, peoples and organisational systems influence organisational performance. A common requirement is that performance indicators should be specific, measurable, accepted, realistic and timely. Measurability implies that there is a scale (nominal, ordinal, interval, ratio) on which performance can be mapped. Nominal or ordinal scales give only qualitative measurements, where interval and ratio scales in addition can provide quantifications of performance. For the technical system performance can usually be measured with quantitative indicators, but for the personnel and the organisational systems it usually necessary to be content with qualitative indicators.

### 4.6. Changes in operational conditions

The need for adaptations in the control structures occur when there are changes in the environment. Such changes may be due to innovations in technology and/or in organisational designs. Technical innovations may for example due to better methods for calculations allow for decreased margins or for scaling up of selected components. They may also make certain tasks easier to perform, due to better tools and better accessibility of information. Organisational innovations may increase personnel commitment and efficiency through changes in the division of labour and/or in the reward systems.

Adaptations to changed conditions would in principle imply an assessment of the need for changes in control structures. In many cases necessary changes would be possible to initiate simply by adaptations in used models or preference relations. In other cases however more radical changes may be necessary. If this is the case guidance from research in organisational learning may be helpful [11].

## 5. APPLICATIONS TO SAFETY MANAGEMENT

Safety management should address possible threats during the lifetime of an industrial facility. Risk analysis and safety engineering are important activities within safety management. Risk analysis is assumed to identify and evaluate different threats and safety engineering to bring unacceptable threats to a level where they can be accepted. The tools of safety engineering to eliminate, control and mitigate various threats rely on building control loops by which specific conditions can be detected and corrected. Two other important activities of safety management are experience feedback and change management, which ensure that experience and knowledge is collected and acted upon.

### 5.1. Safety among other performance attributes

Safety is the main performance attribute in high reliability organisations [12], i.e. safety takes a preference over other performance attributes. In spite of this preference it is still necessary to understand that also other attributes enter the equation. It is for example argued that a specific hazardous technology can be used if it provides societal benefits that are higher than the risks involved. The assumption is thus that risks of a facility are brought to a level, where they are considered to be acceptable.

Cost and benefit are two attributes that are important for all industrial activities and also so for hazardous facilities. Assuming that the society sets a risk level to divide between acceptable and unacceptable risks, the task for safety management is to ensure that the risk of a facility is brought below that level. This can be seen as the ultimate control loop to be enacted by safety management. As safety precautions come with some costs, the benefits of operating the facility have to be larger than these costs, because otherwise the facility would be shut down. This implies that a facility may reach its economic end of life before it is worn out, if new experience reveals design flaws that are too expensive to correct.

### 5.2. Designing for safety

When a new hazardous facility is planned it is important that the intended site and design is evaluated carefully with regard to safety. This can be done on the drawing board, where alternatives are compared and improved. At the same time needs for personnel in various categories should be evaluated. Hiring and training should be started early to have the organisation in place when the facility is commissioned. At that time organisational design, with functions, authorities and responsibilities described in a management system, should also be available. During the first years of operation focus should be set on experience feedback and change management to identify and correct remaining flaws in the design and construction.

During operation safety management should focus on the control loops that have been designed into the facility as parts in the technical, peoples and organisational systems. Correct operation of the controls should be monitored and possible malfunctions should be diagnosed and corrected. When changes are initiated their risks should be analysed and safety engineering should be applied to minimise the risks. The management system and other documentation should be updated continuously to ensure that they describe the facility as operated.

### 5.3. Operating safely

Safe operation relies on maintaining the state of the technical, peoples and organisational systems in safe regions of their state spaces. For the technical system it implies that the requirements of the safety technical specifications are maintained and that the engineered safety systems are available. For the peoples system it implies that the personnel is knowledgeable and experienced and that there are enough people to handle both normal and abnormal situations. For the organisational system it implies that tasks are defined, instructions are available and accountabilities are understood.

Safe operation implies furthermore that there are alarms to signal deviations and abnormal occurrences. For the technical system these may be engineered alarms in the control room or information collected from periodic tests that indicate failures or slow degradations. For the peoples system such alarms may be connected to regular surveys and performance appraisals. For the organisational system audits, assessments and reviews may be used for that purpose. If deviations and abnormal occurrences are detected they should be analysed to determine if they could be considered to fall within limits of normal system variability or if they should be considered to signal a move of system state outside defined safe boundaries.

### 5.4. Audits, assessments and reviews

Audits, assessments and reviews are organisational activities that aim at a continued verification and validation to ensure that agreed safety precautions are in place and are efficient. Audits have a function of a control loop by which agreements between actual and prescribed practices can be assessed. Findings from audits are typically documented under headings of observations and deviations. If deviations are found, they should be acted upon either by changing actual or prescribed practices. An important high level feedback loop is exercised by the senior management at regular intervals in which observed deviations are evaluated to initiate corrective actions.

General organisational reviews should also be a part of the controls performed by senior management to evaluate the overall performance of the organisation and its functional units. If problems are detected they should result in adaptations in the line of organisational authorities and responsibilities. In addition to their own reviews the senior management may call in peer reviews that are carried out by outsiders, who due to their own experience have a working knowledge of managing similar organisations. Such a practice has the benefit of looking at organisational performance with fresh pairs of eyes.

### 5.5. Regulatory control

Regulatory control can be seen as an outer control loop by which the society ensures that facilities are operated within accepted frames [13]. Regulators have an important task as the representatives for the society to define safety requirements and to ensure that they are fulfilled. From a safety point of view however, it is not enough to fulfil regulatory requirements at a specific point of time, which can be seen from the fact that all major nuclear accidents have brought new regulatory requirements into place.

The most important principle in regulatory control is that the operators of hazardous facilities have an undivided responsibility for safety. This means that there in an invisible line between actions the regulator may require and questions the licensee would be allowed to ask. The regulator should define the frames of acceptability, but should never prescribe solutions. Similarly the licensee should argue for why a specific solution fits into agreed frames, but should never ask for solutions that can be accepted. If controversies arise they should be resolved through in depth discussions regarding the believability of the arguments used.

## 6. DILEMMAS IN CONTROL OF SAFETY

So far we have discussed the control problem and safety management in general terms. Now we would like to address some remaining dilemmas of safety management. One is connected to the fact that risks associated with hazardous facilities have their gravity centre in low probability high cost events. There are also dilemmas connected to the selection of a suitable model of safety and the system to model. The search for safety indicators and the concept of safety culture have some dilemmas involved and agreeing on proper preference relations for the controls requires the resolution of certain balances.

### 6.1. Low probability high cost events

A specific dilemma in the control of safety is connected to low probability high cost events. For such events there may be large uncertainties in both probability and consequence estimates. This makes it difficult to give reliable risk estimates to set priorities for improvements. Establishing proper margins in the control of safety is therefore important in the organisational control loops.

In responding to this dilemma it is necessary to ensure that the risk analysis is reasonable complete, correct and consistent. In comparing risks in different domains, it is important to have quantitative risks estimates that are connected to specific events chains. If risks within a specific domain are assessed and different event chains lead to the same ultimate consequence, a risk estimate given on a quotient scale may be used to set priorities between alternative changes in the technical, the peoples and the organisational systems. The quotient scale gives the possibility to require that an improvement that is twice as costly as another alternative should give at least a two times higher risk reduction.

### 6.2. Models of safety

In the discussion above we have assumed that the MTO-model is appropriate in explaining how safety is constructed. The so called Swiss cheese model [14] is another model, which puts emphasis on safety barriers and their availability. This model can be seen as complementary to the MTO-model and it suggests the function of barrier controls, to be implemented in the technical, peoples and organisational systems. We argue that also other models are possible, which may take different, but still complementary views on the control of safety. Important however, is that the models are used in their region of validity and that they are mutually consistent.

In this paper we have proposed that the state space of the system model gives the opportunity to consider characteristics of safe and unsafe regions in the state space. We have also argued for application, characterisation and modelling of interconnected controls on several hierarchical levels, with appropriate feedback and feed forward loops to suggest requirements on safety management. In addition one may introduce the concept of specialised control agents on different hierarchical levels, which have a task of detecting selected safety threats and responding to them with protective actions. Taken together these would suggest an integrated control structure to be applied, where low level automatic functions signal concerns, which are responded to with higher level control loops.

### 6.3. Selecting the system to model

An important task is to select the system one want to model. It can be a single component at a plant that is maintained by some group of people, it can be a complete plant with its people and organisation, it can be an international company with controls from its corporate level or it can be a regulatory agency on the national level in a country. In selecting the system to model it will have an inner structure of interacting subsystems and a state space, where necessary requirements for safety can be pondered to identify safe and unsafe regions. By selecting different systems to model it is possible to take different views that in combination can provide an approach towards establishing sufficient conditions for safety.

In this connection we argue that the control structure we have suggested above can provide a frame for modelling safety at several level of detail from the society down to single components at the facilities. A person responsible for safety on some level in that hierarchy should have a good understanding of the models used at her/his level and the levels immediately above and below. Only then the possible safety implications for decisions s/he makes are possible to assess.

### 6.4. Safety indicators and safety culture

Since the TMI accident there has been on-going discussions of safety indicators [15], which unfortunately have not yet converged. Most of the efforts have been based on ad hoc models of safety and not on a discussion of necessary conditions for safety. A more sound approach to safety indicators would go through a discussion of appropriate models of safety and safe regions of system states.

The Chernobyl accident in turn brought in safety culture as a new component of safety [16]. In retrospect the interest in safety culture can in a dialectic perspective be seen as the antithesis of the older thesis that safety is a construct of the technical system. In a consideration of recent discussions of the concept, it however seems difficult to ensure that the four necessary conditions for a successful control of safety culture can be fulfilled [17].

Indicators to be used in assessments of safety culture are difficult to defend if they are not based on a believable model that is simple enough to be practical [18]. An assessment of recent regulatory interventions suggests that a deterioration of safety culture is one focus of regulatory concern. To what extent this focus is based thorough assessments is hard to say. Anyhow with an attention on recent discussion of resilience engineering [19], one observation is that variability in the technical, peoples and organisational systems are expected to generate deviations and that they may not be of too large concerns, provided that main safety control loops are functional and efficient.

### 6.5. Balances in preferences

The preference relation has to do with the quality of control and trade-offs in performance. The preference relation is usually set up as a function in which a suitable balance including state, resources and time is sought. There are many different balances to be considered in the definition of a preference function. One such balance is addressed in the question what is safe enough [20], because the society has to set a level for risks that can be considered as acceptable.

The ETTO principle [21], illustrates another balance that has to be resolved by providing actors with proper instructions and decision support for various situations. The recommendation to apply conservative decision making is another case, where a proper balance should be found. The needs for conservative decisions arise in situations, where large costs differences are found between erroneously accepting one or the other of two hypotheses. An important balance can also be seen in the strictness of the controls, because there are organisational trade-offs between very strict and more flexible control. A strict control uses more resources and can be demotivating, but may be necessary for safety critical tasks. Finally to become a learning organisation a proper balance between traditions and renewal should be found [22].

## 7. CONCLUSIONS

We find the control metaphor helpful is setting up requirements on safety management. We argue also that a full benefit of the metaphor requires a closer look on applied control structures. Our discussion above puts a strong emphasis on systems modelling on all hierarchical levels from controls implemented with international agreements down to control of single safety critical components. To keep the modelling effort within reasonable limits, the graded approach to safety should be applied.

In modelling it is important to remember that a model is a simplification that disregards phenomena which in other cases may be important. A model should be complex enough not to be trivial, but still simple enough to be practical. A model does not need to be quantitative to be useful, because also simple qualitative models with highly aggregated state spaces may provide useful insights. We see the model of safety management, which we have developed in this paper as a frame into which more detailed models can be placed.

Concerns connected to complexity and increasing interconnectedness in the large scale industrial systems of today has proposed that we are forced to accept even large accidents in the future [23]. It is true that unexpected connections, tight coupling and nonlinear responses together with trivial failures and latent deficiencies may trigger sequences that lead to accidents. We still believe that safety management with its models and practices can create resilience at all levels in the hierarchy of systems to ensure that incidents are rare and their consequences small.

One important lesson however, is that it may be necessary to apply the precautionary principle [24], when new technologies are introduced, scaled up or built at an increasing rate. A slightly slower rate, may in such cases be appropriate, to allow for the accumulation of operational experience to make them safer. A combination of forward looking risk analysis and experience based improvements should make it possible to manage the unknowns even when new technologies are introduced.

**Acknowledgements**

The support of Vattenfall AB in enabling this research through interactions with their nuclear power plants is greatly appreciated. The opinions expressed in the paper are the authors only and do not necessarily reflect any Vattenfall policies or practices.

**References**

[1] Rasmussen J, Svedung I. Proactive risk management in a dynamic society, Swedish Rescue Services Agency, Karlstad, Sweden, 2000.
[2] Zadeh L, Desoer C. Linear systems theory, McGraw Hill, New York, 1963.
[3] Aven T. Some recent definitions and analysis frameworks for risk, vulnerability, and resilience, Risk Analysis, Vol.31, No.4, 2011.
[4] Möller N, Hansson SO. Principles of engineering safety: Risk and uncertainty reduction, Rel.Eng. & Syst.Safety, Vol.93, Issue 6, 2008.
[5] Argyris C, Schön D. Organizational learning: A theory of action perspective, Reading, Mass: Addison Wesley, 1978.
[6] Kontogiannis T. Modelling patterns of breakdown (or archetypes) of human and organizational processes in accident dynamics, Safety Science, doi:10.1016/j.ssci.2011.12011, 2012.
[7] Mesarović MD, Macko D, Takahara Y. Theory of hierarchical, multilevel, systems, Academic Press, 1970.
[8] International Atomic Energy Agency. The management system for facilities and activities, GS-R-3, Vienna, 2006.
[9] Schein E. Organizational culture and leadership, Jossey Bass, San Francisco, 1992.
[10] Elrod E, Hubbard CL. Applying means – ends decision trees, Business, Jan-Feb, 17–25, 1979.
[11] Easterby-Smith E, Crossan M, Nicolini D. Organizational learning: Debates past, present and future, Journal of Management Studies, 37:6, Sept, 2000.
[12] La Porte TR, Consolini PM. Working in practice but not in theory. J.Publ.Admin.Research and Theory, 1, 19–47, 1991.
[13] Wahlström B. Reflections on regulatory oversight of nuclear power plants, Int.J.Nuclear Law, 1, No. 4, 2007.
[14] Reason J. Managing the risk of organizational accidents, Ashgate Publishing Company, Brookfield, VT, 1998.
[15] International Atomic Energy Agency. Operational safety performance indicators for nuclear power plants, TECDOC-1141, Vienna, 2000.
[16] International Atomic Energy Agency. Safety Culture, INSAG-4, Vienna, 1991.
[17] Silbey SS. Taming Prometheus: Talk about safety and culture, Ann.Rev.Sociol, 35:341-369, 2009.
[18] Mohaghegh Z, Mosleh A. Measurement techniques for organizational safety causal models: Characterization and suggestions for enhancements, Safety Science 47, 1398–1409, 2009.
[19] Hollnagel E, Woods DD, Leveson N. Resilience engineering : Concepts and precepts, Ashgate, 2006.
[20] Starr C. Social benefits versus technological risks, Science, 165, 1232-1238, 1969.
[21] Hollnagel E. The ETTO Principle: Efficiency-Thoroughness Trade-Off, Ashgate, 2009.
[22] Wahlström B. Organisational learning – reflections from the nuclear industry, Safety Science 49, 65–74, 2011.
[23] Perrow, C. Normal Accidents: Living with high-risk technologies, Basic Books, New York, 1984.
[24] Sandin P, Peterson M, Hansson SO, Rudén C, Juthe A. Five charges against the precautionary principle, Journal of Risk Research, 287–299, 2002.