Factual and Foreseeable Reliability of Tailings Dams and Nuclear Reactors: a Societal Acceptability Perspective.

Franco Oboni
Oboni Riskope Associates Inc, Vancouver, British Columbia, Canada

Cesar oboni
Oboni Riskope Associates Inc, Vancouver, British Columbia, Canada

ABSTRACT: This paper compares the “historic” rate of failure (major accidents only) of tailings dams and nuclear reactors world-wide to well known, previously published acceptability criteria and codes. It then attempts a comparison of these two very different industries (nuclear vs. mining) from a quantitative risk evaluation point of view showing unexpected results. The paper then shows that the mining industry as a whole has considerably progressed in reducing the rate of dam failures from the 1970 to date. Three questions are asked: 1) Is it sufficient? 2) Is it sufficient from the perspective of present and future societal acceptability? 3) Do common practice risk assessments offer enough depth of analysis with respect to modern requirements? To answer these questions, the authors suggest a simplified model for long-term risk evolution of tailings dams with particular emphasis on post-production – closure – long-term dam life. The paper shows how a generic modern “excellent quality” dam probability of failure can be estimated, as well as how the initial probability of failure will evolve during the dam life as care and monitoring are released (in view of closure), and during the closure phase until, for example, the first Maximum Design Earthquake. The paper is written in non-scientific language to broaden the audience beyond the engineering community.

1 INTRODUCTION

This paper compares the “historic” rate of major accidents (failure) of tailings dams and nuclear reactors world-wide to previously published acceptability criteria and defines the long-term evolution of the risks (major accidents) generated through the life of a tailings dam (TD), with particular focus on the post-production phases/closure. In contrast to hydro dams that would typically be breached at the end of their production life, the closure phase is the longest phase for TD. A TD's life can be summarized as follows (M.B. Szymanski, M.P. Davies, 2004): Production (tailings disposal); Transition (preparation for the closure phase which may include flushing out contamination); Long-term treatment (dam operation continues in the sense of regulated water levels); Closure (dam is no longer operated in the sense of regulated water levels), likely more than 1,000 years. For the purpose of this paper, we will summarize the phases as follows: Production as the phase with the highest monitoring and care, Transition and Long-term treatment as phases during which monitoring and care are gradually reduced, and Closure as the phase during which the dam is “abandoned”. Major hazard hits may occur during any of these phases.

A TD’s probabilities of failure and those generated by major nuclear reactors' accidents to date are empirically estimated, and both are compared to societal and technical acceptability thresholds to understand if present and foreseeable performances are aligned with expectations. Risk is often modeled as the expected value of an undesirable outcome by combining the probabilities of various possible events and an evaluation of the corresponding harm into a single consequence value. Accordingly, in this paper, the risk is defined as the product of the probability of failure by the related consequences expressed in casualties, leaving aside all other environmental and physical direct or indirect consequences, for the sake of simplification.

- As a first step the paper examines real life failure data (Fig. 1) for the world-wide and US TDs portfolio in an attempt to define a "base value" of the rate of failure p_f(annual probability of major accident), then an "excellent" TD rate of failure p_f.
- The second step encompasses the definition of a simplified model for the evolution of the probability of failure p_f of a TD over time, including the long-term post-production "era". In today's world, civilian/industrial facilities are generally designed to last for 50-100 years save, for example, spent nuclear fuel disposal sites. Some former mining sites (uranium tailings, arsenic disposals, etc.) are calling for "perpetual care", very much like nuclear fuel disposal. Perpetual care design means ensuring somehow survival for repeated major hazards hits, etc., certainly for more than the "classic 1,000 years". For example, the need to account larger exposure periods for TDs has been recognized in the New Zealand Dam Safety Guidelines (NZSOLD, 2000).
- Third, we compare risks generated under various conditions to different published acceptability thresholds and discuss the implications.

The simplified models used through this paper for TDs focus on major events and use approximations: they are geared toward defining an order of magnitude of the long-term risk evolution of a generic dam undergoing variation of the standard of care, during and after the mine closure, and when hit by a major hazard. This discussion could be expanded to cover various Maximum Design Hazards but it should not be relied upon for specific assessments; only a third party, i.e. not performed by the design team, quantitative risk assessment should guide specific sites' decisions to avoid biases and prejudices (F. Oboni, C. Oboni, S. Zabolotniuk, 2013).

A generic well-designed, built and monitored "excellent" dam is used as Base Case Study (BCS), where "excellence" is defined by compliance with modern TDs codes for critical, high consequences dams (ANCOLD, 2011). The BCS p_f is estimated using the "SLM" (F. Silva, T.W. Lambe, W.A. Marr, 2008) semi-empirical relationship with the classic Factor of Safety (FoS) (Fig. 2). The first step is to define the "Category" (I to IV) of the structure under examination by examining sequentially Design, Construction and Operations aspects of the slopes.

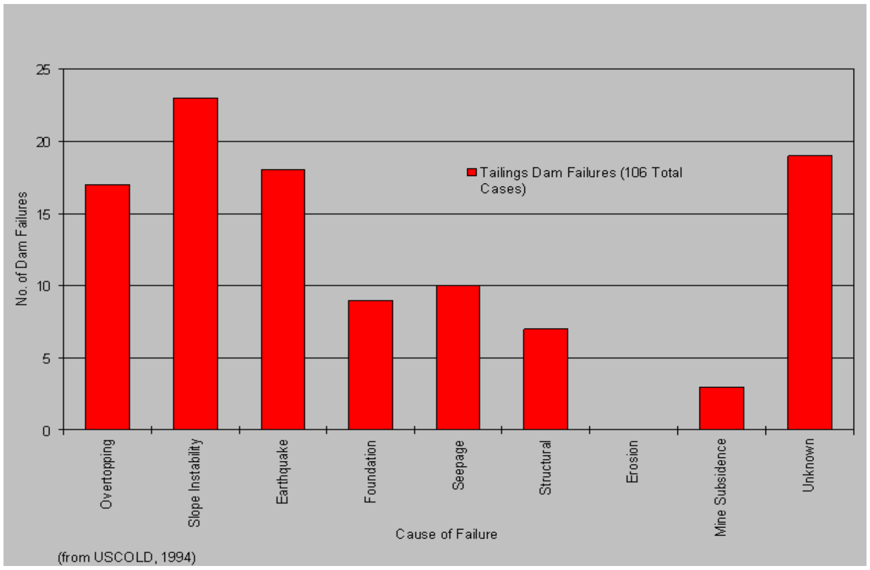
Within the frame of a specific and detailed analysis for a site specific quantitative risk assessment, probabilistic slope stability methods should be used to define the probability of failure of a particular dam.

2 HISTORIC FAILURES RATES

2.1 *Tailings Dams and Hydro Dams*

USCOLD (1994), UNEP (1996, 1998) studies have shown that Slope Instability is the highest cause for TDs failures (Fig. 1), followed by earthquake, overtopping and "unknown" (ICOLD, 2001). Following Davies & Martin (M. P. Davies, T.E. Martin, 2000; N. Lemphers, 2010) there were more than 3,500 TDs (of various kinds and construction type, with at least 50% of the upstream design) around the world. (For the US (See *1 in the Literature) we consider the total number of TDs at 984). In absence of more accurate data, and more particularly, on the evolution of this number, we will consider in this paper a portfolio of 3,500 dams globally and 1,000 dams for the US alone. Portfolios are assumed constant over the years but, as we will see later, even if that assumption was wrong by one order of magnitude, which is very difficult to believe, there would be no significant alterations in the conclusions of this paper.

From UNEP (1998) world data we read that there were 44 TDs failures in the decade around 1979 and 7 in the decade around 1999. From US sources, 984 TDs 984 (See *2 in Literature), the chronology of major failures (from 1960) gives a total of 28, with 7 failures in the decade around 1979 and 8 failures in the decade around 1999).



(from USCOLD, 1994)

Fig. 1 Causes of Failure of Tailings Dams, Number of Occurrences over a Total of 106 Recorded by USCOLD, UNEP up to 1994.

Where	When (decade)	p_f	Approx p_f	Verbiage for p_f
World-wide	Around '79	44/(3,500*10)	10^{-3}	One in one thousand
World-wide	Around '99	7/35,000	$2*10^{-4}$	Two in ten thousands
US	Around '79 & Around '99	7 or 8/(1,000*10)	7 or $8*10^{-4}$	Seven or eight in ten thousands

2.1.1 Annual Probability of Failure vs. Factor of Safety

As slope instability is the greatest source of TDs failures (Fig. 1), for the sake of simplification, this paper looks only at this particular failure mode. The proposed methodology could, however, easily be expanded to cover other failure modes.

The SLM methodology examines sequentially the following *aspects* of the Design (D1 Investigation, D2 Testing, D3 Analyses and documentation) and Construction “CO” as well as Operations and Monitoring “OM” of embankments and slopes to determine the Category for a structure. Each aspect is described by various detailed specifications. The less stringent the specifications, the lesser the quality of the considered structure; thus, SLM defines four Categories ranging from I (Best) to IV (Poor). Experience shows that structures with high failure consequences are generally designed, built, and operated in such a way that they fall in Category I. Of course, if a structure has received little or no engineering it will fall in Category IV. Accordingly SLM's Category I has OM described as “complete performance program including comparison between predicted and measured; no malfunctions; continuous maintenance” whereas a Category IV will have “occasional inspection, no field measures”. We will assume in this paper that Category I OM applies to Production Phase (undergoing full monitoring), and Category IV OM to Closure Phase with “abandonment” (no monitoring).

Figure 2 shows the FoS-p_f relationship for the four categories. If we go back to the rate of failure p_f empirical estimates (Section 2.1) for TDs = 10^{-3} to $2*10^{-4}$ (we will not mention the US value as it lies within that range) and assume that the average original FoS was in the area of

1.3 (ANCOLD, 2011), we would come to the conclusion that TDs are generally Category I or slightly inferior structures.

The statement above is in good agreement with the common understanding and empirical knowledge that TDs are generally of “lesser quality” than hydro dams. Indeed, if we consider the hydro dam failures in the decades around 1989 and 1999 and evaluate p_f as we did for TDs, based on an “average number of dams” of 30,000, we get $p_f = 3*10^{-6}$ to 10^{-5}, values which are compatible, considering usual FoS, with a Category I (three failures in a million to one failure in one hundred thousand dams per year). By using the SLM methodology, we can estimate the p_f of an “excellent” TD at 10^{-5} to 10^{-6} as “excellence” would mean “top of the class” Class I structure, well engineered, undergoing serious QA/QC, with a minimal FoS of 1.4-1.5. Should inspections become occasional, measurements/monitoring not be performed, the probability will raise to $6*10^{-5}$. Interestingly, many different industries around the world consider values below 10^{-6} to 10^{-5} (below one in one hundred thousand to one in a million) as the boundary of what is humanly credible (meaning that below that range of probability any lay person or expert would imagine that an accident is “incredible”).

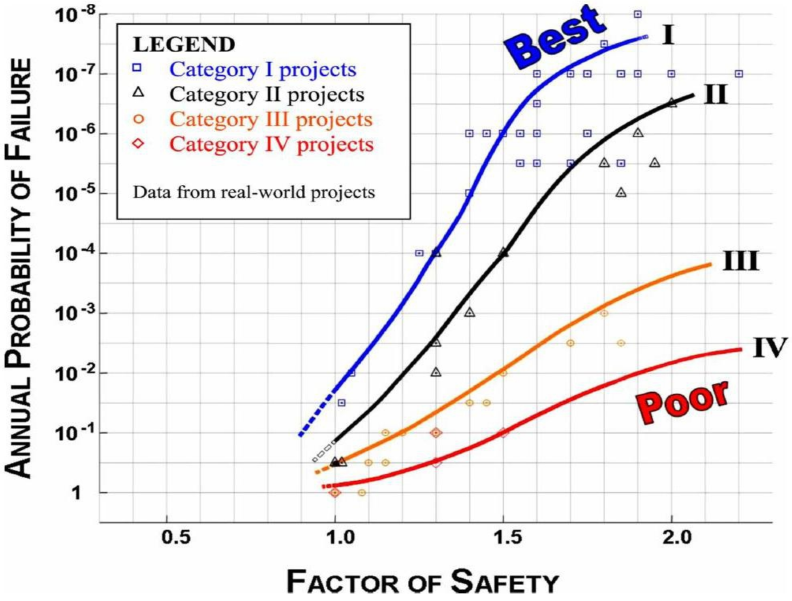


Fig. 2 Annual Probability of Failure vs. Factor of Safety following Silva, Lambe, Marr (SLM) (F. Silva, T.W. Lambe, W.A. Marr, 2008) methodology.

In case of partial fulfillment of a Category's qualifications, the original SLM paper suggests to interpolate values (on Fig. 2) after defining weights for the specifications/Categories. Thus a stepped increase of the probability of failure is suggested if one of the *aspects* is gradually worsened for a dam belonging to a specific Category.

In the case of TDs, it is possible to “simulate” long-term complete abandonment. Thus the p_f could actually reach the value of the lower Category IV (and even higher). For example, if we look at the case of OM standards phased release from Category I (operational life) down to IV (long-term closure), the probability of failure will increase each time the OM standard attains a lower category. Interestingly, for initial FoS in the 1.3 to 1.5 range, the difference between Category I (we are assuming that the dam under examination is initially an “excellent” structure) and II varies respectively between 1.5 and 2 orders of magnitude: for FoS=1.5, $p_{fCatI} = 10^{-6}$, with possible increase to 10^{-4} and higher if the same structure falls in total neglect.

2.1.2 P_f with Seismic Event

Following ANCOLD (2011) an “excellent” TD should be designed to remain serviceable under Operating Basis Earthquake (OBE), i.e. to undergo limited damage/deformation repairable without significant service disruptions (this moves the structure from Category I to Category II-III following SLM). Therefore p_f will increase from 10^{-6} to 10^{-4} or higher under OBE if the initial FoS is still attained (no expected liquefaction, well-drained structure, no change in shear strength).

In case the post-seismic shear strength would be lower than pre-seismic values, then ANCOLD 2011 declares an acceptable minimum Fos=1.1 and therefore the p_f could increase to $1.25*10^{-2}$ or higher. Of course this is again quite simplistic, insofar we are considering "a single earthquake" instead that the whole set of aftershocks on a weakened or damaged structure.

Under Maximum Design Earthquake (MDE) damage will be more extensive and disrupt operations, but the structural integrity of the dam needs to be maintained and uncontrolled release of tailings/water should not occur. MDE for Significant Consequence Dams should be those with an Annual Exceedance Probability (APE)=1/100=10^{-2}; APE=1/1000=10^{-3} for High & Extreme Consequences dams (ANCOLD, 2011). Thus we can estimate that after a MDE the dam could fall down to Category III during operations, and most likely Category IV or higher during the very long closure phase, during which the occurrence of at least one MDE is almost certain to occur.

Such low "acceptable" FoS and respective high p_f will lead to deformations and other defects of the dam which will be weaker when the next seismic event occurs, but more importantly an increase of risks. If we use the SLM suggested interpolation technique, if after an OBE the dam does not undergo routine maintenance and has uncorrected malfunctions, the probability of failure will rise significantly. After multiple hits, or if the first hit provokes severe damages equivalent to the cumulated damage of several quakes, we could see a possible post-seismic increase to $p_{fCatIPS}=1.2*10^{-1}$ and higher if the dam has fallen in total neglect.

2.2 *Major Nuclear Accidents*

As of February 2, 2012, 435 nuclear power plant units with an installed electric net capacity of about 368 GW were in operation in 31 countries and 63 plants with an installed capacity of 61 GW under construction in 15 countries. The cumulative nuclear reactor operating experience amounted to 14,745 years by February 2012 (See *3 in the literature).

To date, the world has seen the occurrence of a number of major nuclear reactors accidents (rated 5 and above on the International Nuclear Event Scale by the International Atomic Energy Agency). For Fukushima we consider one accident (although more than one reactor was involved) to ensure the list is made of "independent" accidents. The reason Level 5 and higher is selected lies in Level 4 definition as "Accident with local consequences" which might not be comparable to a TD major failure discussed in the prior sections. Finally, to counter any possible comments depicting the nature of Fukushima quake magnitude and resulting tsunami as "exceptional", we note that:

a) in many areas of the world, seismic potential was/is poorly understood and therefore many structures were/are under-designed,
b) it is difficult, if not impossible, to structurally upgrade a working reactor,
c) power plants' risk assessments apparently focused on the safety of the nuclear reactor, but often underestimated the impacts and consequences on its environment and ancillary structures/ equipments.

Assuming seven accidents, the "historic" world average rate of Scale 5+ accidents is: 7/14,745 total operating time= $4.75*10^{-4}$ Scale 5+ accident/annum.

Level 5	Level 6	Level 7
Accident with wider consequences	**Serious accident**	**Major accident**
Chalk River (1952) Windscale (1957), Lucens (1969), Three Mile Island (1979)	Kyshtym (1957)	Chernobyl (1986) Fukushima (2011)

3 FAILURE RATES VS. HISTORIC AND ANCOLD ACCEPTABILITY THRESHOLDS

Figure 3 displays acceptability criteria developed independently by various authors over a peirod of more than forty years, together with some examples of industrial, transportation and dams accidents dating back to the 1960s (G. Morgan, L. Lave, 1990, R.V. Whitman, 1984, ANCOLD, 2003). Among these accidents, depicted as "bubbles" to include scatter of data, it is of particular interest to observe the "(hydro) dams bubble". If we compare these to the values defined in Section 2, we notice immediately that world-wide (but, surprisingly, not so much in the US) the situation has improved considerably with the p_f of TDs creeping down to values similar to those of hydro dams in the 1960s, and hydro dams lowering their p_f by one order of magnitude.

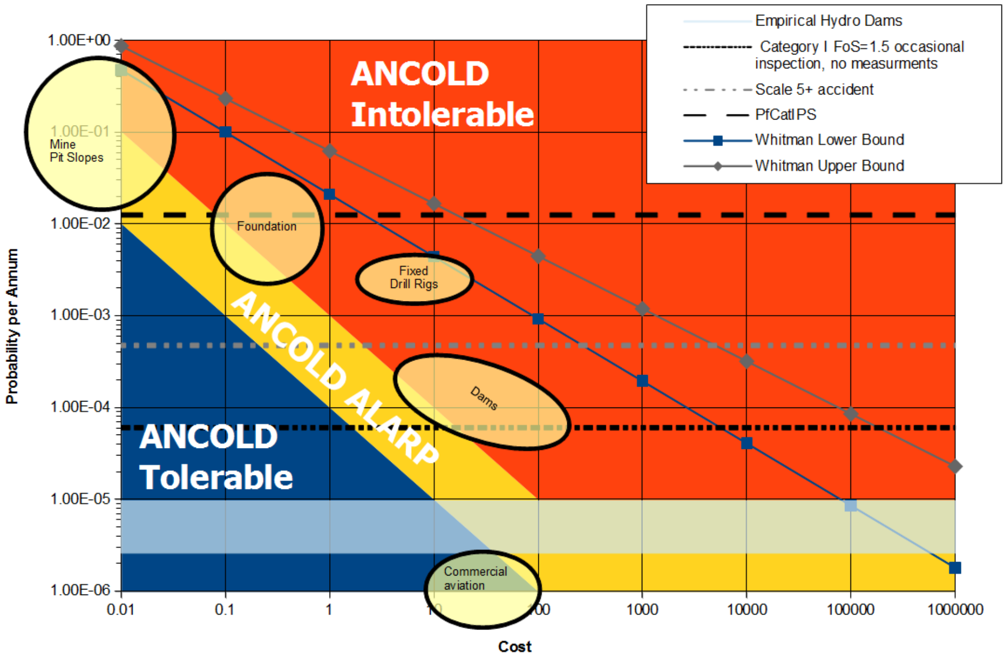


Fig. 3 A probability (vertical)-Consequence (Horizontal) graph showing p-C "bubbles" from various industrial, transportation and dams accidents in the 1960s (R.V. Whitman, 1984 & G. Morgan, L. Lave, 1990), Whitman tolerability curves (upper bound and lower bound) as well as the ANCOLD (2003) "acceptability zones" (Tolerable, ALARP, Intolerable). The horizontal lines correspond from bottom to top: to the p_f *of a Category I dam with occasional inspections and no monitoring/measurements, to Scale 5+nuclear accidents (See Section 5), and a Category I dam under total neglect and after seismic events.*

The recent ANCOLD 2003 acceptability criteria (Fig. 3) are compatible with Comar, Wilson (Comar, 1987, Wilson & Crouch, 1982) and later criteria published in the field of chemical industry, such as those from Renshaw (Renshaw, 1990), who defined simple societal risk acceptability criteria expressed as probability of fatality of one individual per year of risk exposure. Many publications from reputable (governmental, research) sources point at a probability (of a casualty per annum) of 10^{-4} (similar to ANCOLD lower bound) as being the limit of "safe", however with a lower limit of 10^{-6} for unwillingly exposed public.

The p_f of:

- a Category I dam with occasional inspections and no monitoring/measurements,
- a Scale 5+nuclear accidents, and a
- Category I dam under total neglect and after a number of seismic hits

can be compared to ANCOLD and other published historic tolerability thresholds (Fig. 3). It becomes immediately apparent that any accident with more than 1 to 15 casualties is considered intolerable by ANCOLD or other modern “technical” tolerability thresholds when paired with the selected cases' p_f ,whereas historic societal thresholds would have been complied with.

4 COMPARING RISKS OF TWO VERY DIFFERENT INDUSTRIES

4.1 Empirical probabilities of mishaps

The Table below (Data from Section 2.1, 2.2) compares TD and nuclear 5+ accidents including occurrence decade and empirical rate of return.

Where	**What**	**When (decade)**	**p_f**	**Approx. p_f**
World-wide	TD	Around '79	44/(3,500*10)	10^{-3}
World-wide	TD	Around '99	7/35,000	$2*10^{-4}$
US	TD	Around '79 & Around '99	7 or 8/ (1,000*10)	7 or $8*10^{-4}$
World-wide	Nuclear 5+ accidents	Since inception	7 Scale/14,745	$4.75*10^{-4}$

Despite the fact that nuclear industry is highly regulated, its safety record to date is far from optimal. One would like to see severe mishaps to be, as for hydro dams, border-line credible, but historic data unfortunately prove the contrary.

4.2 Cost of Consequences

In order to evaluate risks we need now to evaluate consequences using the selected metric, i.e. casualties. Even such a “simplified” metric is difficult to apply as there might be for nuclear accidents a long delay between exposure and health effects, as explained below.

4.2.1 Nuclear

In some class 5+ accidents the estimate of casualties has been equated to nil. It is not within the scope of this paper to discuss that evaluation.

According to a June 2012 Stanford University study by John Ten Hoeve and Mark Jacobson, the radiation released at Fukushima could cause 130 deaths from cancer (the lower bound for the estimates being 15 and the upper bound 1,100) and 180 cancer cases (the lower bound being 24 and the upper bound 1,800), mostly in Japan. Radiation exposure to workers at the plant was projected to result in 2 to 12 deaths. An additional approximately 600 deaths have been reported due to non-radiological causes such as mandatory evacuations. Evacuation procedures after the accident may have potentially reduced deaths from radiation by 3 to 245 cases, the best estimate being 28; even the upper bound projection of the lives saved from the evacuation is lower than the number of deaths already caused by the evacuation itself. We can assume based on these numbers that a class 5+ accident to date has caused between nil and a maximum of 3,500 casualties, with a “best estimate” at 890 casualties.

4.2.2 Dams

If we look at the history of TD accidents (221 accidents: USCOLD, 1994; UNEP, 1996; 2001) in terms of victims, we record a few accidents with approximately 260 casualties (2008, Taoshi, Linfen City, Xiangfen county, Shanxi province, China; 1985, Stava, Trento, Italy). In

1966, the Mir mine disaster in Sgorigrad, Bulgaria, killed 488 people. In 1965, the El Cobre New Dam in Chile killed more than 200 people. Very few cases killed around 100, and most cases had 1-15 casualties. This leaves us with a minimum of nil, a maximum of ~500, and an expected value of ~80 casualties.

4.3 Risks and their Evolution

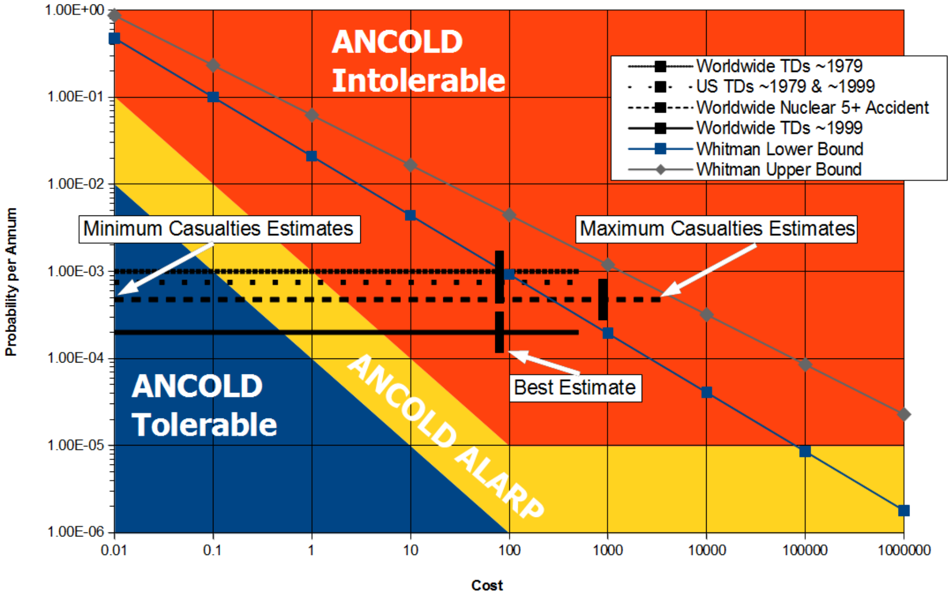


Fig. 4 Same pf-C (Casualties) diagram as in Fig. 3, but "historic bubbles" have been removed and Casualties Estimates for TDs' and Class 5+ nuclear accidents have been added.

If we look at Whitman tolerability thresholds, most of the recent (1999) TD failures fall within societal acceptability lower bound, unlike the more ancient ones or US failures. However, should a TD failure result in a number of casualties greater than 1000, we should expect a very serious re-think of the industry world-wide similar to the one that the nuclear industry has just seen. Should a TD be abandoned and undergo in the long run a number of natural hazards hits, the risks will become socially intolerable even if there are less than ten casualties.

These results lead to the following set of conclusions:

a) Unless a TD is built as a Category I geo-structure, then is highly monitored and maintained, the ANCOLD tolerability guidelines suggest no one should ever be exposed to it.
b) unless a TD is designed, built and cared for like a hydro dam, which means "at perpetuity" high-level monitoring and care (TD cannot be breached, unlike hydro dams) no residents should be allowed downstream of the structure (this paper does not consider environmental damages), within reach of possible run-out from a breach, to ensure ANCOLD compliance.
c) risk assessments have to be sufficiently sophisticated to allow p_f estimates compatible with the ANCOLD tolerability thresholds.
d) Standard practice matrix-based risk assessments (F. Oboni, C. Oboni, S. Zabolotniuk, 2013, C. Oboni, F. Oboni, 2012) cannot be used as they lack the necessary finesse and resolution and could actually severely mislead TD owners/operators to the point of exposing them to severe liabilities.

5 CONCLUSIONS

Fukushima, with a “best estimate” casualties count at ten times the average of past TD accidents and the specter of radiations, brought the Japanese nuclear industry to an immediate halt and caused a very serious re-think of the industry world-wide. No TD failure has yet created such a socio-political shift. However, in the present climate of social awareness and social licensing to operate, should a severe accident happen anywhere in the world, we could expect “industry-wide” repercussions of some kind as failure will be deemed socially intolerable. To give an example of such repercussion, following the 2010 British Petroleum/Deepwater Horizon oil spill in the Gulf of Mexico, President Obama issued a moratorium on offshore drilling that halted work on 33 exploratory drilling rigs in the Gulf of Mexico. The ban was lifted in October 2010 but, by February 2011, no one had received a permit to drill because those applying had to prove the ability to contain a spill.

Perpetual care design means ensuring survival for repeated major hazards hits, thus design criteria should be more stringent for the closure phase than for the production phase, especially if the consequences of dam failures will increase because of population, land-use, etc.

Especially in the case of TDs located in areas where demographic pressure leads to settlements in the downstream areas, social and legal consequences of a failure will dramatically increase. This will particularly be the case if the methodologies used to perform the risk assessments prove to be in disconnect with the needs of our modern society.

REFERENCES

ANCOLD, International Commission on Large Dams, 2003, Australian National Committee. Guidelines on risk assessment

ANCOLD, Guidelines On Tailings Dams Planning, Design, 2011, Construction, Operation And Closure, Draft-

M.P. Davies, T.E. Martin, January 2000, Upstream Constructed Tailings Dams- A Review of the Basics, Tailings and Mining Waste, Balkema Publishers

Dams Safety Committee, June 2012, DSC, DSC3F, Tailings Dams, New South Wales

ICOLD, 2001, Tailings Dams: Risk of Dangerous Occurences (Lesson Learnt from Practical Experience, bulletin 121

N. Lemphers, Oct. 12, 2010, Could the Hungarian Tailings Dam Tragedy Happen in Alberta?, www.pembina.org

G. Morgan, L. Lave, 1990, Ethical considerations in risk communication practice and research, Risk Analysts 1o, 355- 358-

NZSOLD, 2000, New Zealand Society on Large Dams, New Zealand Dam Safety Guidelines.

Oboni, F., Oboni, C., Is it true that PIGs fly when evaluating risks of tailings management systems? Short Course and paper, Tailings and Mine Waste '12, Keystone Colorado

Oboni, F., Oboni, C., Zabolotniuk, S., 2013, Can We Stop Misrepresenting Reality to the Public? CIM 2013, Toronto

F. Silva, T.W. Lambe, W.A. Marr, December 2008, “ Probability and Risk of Slopes Failures”, JGGE, ASCE

M.B. Szymanski, M.P. Davies, 2004, Tailings Dams Design Criteria And Safety Evaluations At Closure, BC Mine Reclamation Symposium,

UNEP, 1996, Environmental and safety incidents concerning tailings dams at mines: results of a survey for the years 1980-1996 by the Mining Journal Research Services, Report prepared for United Nations Environment Programme, Industry and Environment, Paris

UNEP, 1998, Case Studies on Tailings Management, United Nations Environment Programme, International Council on Metals and the Environment, November, ISBN 1-895720-29-X

UNEP, 2001, Tailings Dams - Risk of Dangerous Occurrences, Lessons learnt from practical experiences, Bulletin 121, Published by United Nations Environmental Programme Division of Technology, Industry and Economics (DTIE) and International Commission on Large Dams (ICOLD), Paris, 144 p.

USCOLD,1994, Tailings Dam Incidents, U.S. Committee on Large Dams (USCOLD), Denver, Colorado,

R.V. Whitman, 1984, Evaluating calculated risk in geotechnical engineering. J. Geot. Engineering, 110, 2

*1 http://www.wise-uranium.org/mdaf.html

*2 http://www.wise-uranium.org/mdaf.html

*3 http://www.euronuclear.org/info/encyclopedia/n/nuclear-power-plant-world-wide.htm