Presented at the International Offshore and Polar Engineering Conference, Busan, Korea, June 15-20, 2014.
Paper no. ISOPE 2014-TPC-0110.

Eliminating the Precommissioning Hydrotest for Deepwater Gas Pipelines

Peter Carr
Peritus International Inc.
Houston, Texas, USA

Ian F. J. Nash
Peritus International Limited.
Woking, Surrey, UK

ABSTRACT

A preliminary case has been developed for using alternative integrity validation (AIV) for deepwater gas pipelines in place of the precommissioning hydrotest. The features of an AIV program that would compensate for lack of a hydrotest have been identified. Measures to prevent or detect linepipe damage or deterioration during transportation and installation are discussed. Quantitative reliability analysis methods have been explored to compare the effectiveness of AIV and the traditional hydrotest in exposing the types of flaws typically found in girth welds. Recommendations are made for the collection of data on flaw frequencies and size distributions, as lack of these data is currently an impediment to the full application of probabilistic models.

KEY WORDS: Pipeline; Middle East; India; deepwater; gas; precommissioning; hydrotest; system pressure test; alternative integrity validation; AIV; reliability; MEIDP.

INTRODUCTION

The Middle East to India Deepwater Gas Pipeline (MEIDP) will be about 1300 km in length and will be installed in water depths to approximately 3500 m. The pipeline developer, South Asia Gas Enterprise Pvt Ltd, is considering replacing the system pressure test (the precommissioning hydrotest) with an alternative integrity validation (AIV) program. A study was therefore performed with the following objectives:

- Identify the features of an AIV program that may compensate for lack of a hydrotest.
- Attempt to make a preliminary demonstration with qualitative and quantitative arguments that an appropriate AIV program is capable of providing the same level of assurance of safety as a pre-commissioning hydrotest.

HISTORICAL FAILURE RATE IN THE PRE-SERVICE HYDROTEST

The factual information in this section is based on Cosham, Eiber, Owen, & Spiekhout (2006). The pre-service hydrotesting of pipelines, as a means of revealing weak sections of pipeline by causing them to fail under test pressure, first became widespread in the 1950s in the USA. In the beginning, test failures were common, and the test had considerable value in exposing defects before the pipelines were put into service. As the decades passed, the quality of linepipe, welding, construction practices, NDE, and quality control improved, and failures during test became less frequent, causing the value of the test to be increasingly questioned.

It is highly pertinent to enquire as to the frequency of test failures for modern pipelines but there are few reliable sources. According to Cosham et al. (2006), two reliable sources are (a) the UKOPA Fault Database relating to British Gas/Transco onshore gas transmission pipelines and (b) the PARLOC database relating to North Sea pipelines. The UKOPA Fault Database records zero test failures from 1985 to 2005, during which period 2500 km of pipelines were tested, while PARLOC indicates zero test failures in rigid pipelines from 1995 to 2000, during which time 5067 km of pipelines were tested. In the absence of test failures, it is only possible to derive upper bound failure rates, which on the basis of these data are 0.0004 km^{-1} for UK onshore gas transmission pipelines and 0.0002 km^{-1} for North Sea rigid pipelines. Applying the North Sea upper bound figure of 0.0002 km^{-1} to the proposed 1300-km long MEIDP would give an upper bound test failure probability of 26%. If this were a true risk, it would certainly not be acceptable to waive the hydrotest. This very high upper bound value illustrates that there is not enough reliably reported hydrotesting experience in modern pipelines for a test failure rate to be accurately quantified. Therefore, whether test failures are rare or not in modern pipelines remains an open question, at least so far as an objective, as opposed to anecdotal, answer is concerned.

POSSIBLE MOTIVATIONS FOR ELIMINATING THE HYDROTEST

The chief function of the precommissioning hydrotest is to test the pipeline strength. For modern pipelines, the probability of failure during the test is perceived to be low (though based on anecdotal rather than objective evidence as noted above), therefore, the value of the test is sometimes questioned, particularly when the test is difficult or expensive to carry out.

A second function of the hydrotest is to check for leaks. However, alternative and possibly more sensitive leak testing methods may be available or could be developed.

Where fatigue crack growth is a concern, the hydrotest is of little use because it can allow large defects to survive in regions of high material toughness. The high material toughness does not prevent such defects from growing in service. NDE is a more reliable method for detecting defects when critical flaw sizes are small. The hydrotest will not expose small flaws unless there is an associated problem such as low material toughness.

The hydrotest pressure may cause lateral buckling of the pipeline, buckling that would not occur under the lower pressures of operations. Lateral buckling can be mitigated by means of rock berms to limit feed of the pipeline into the buckles but this can be expensive.

For a gas pipeline, the weight of the hydrotest water may cause unacceptable stresses at long free spans, spans that would be acceptable under service conditions when the pipeline is gas filled. Free spans can be rectified by placement of rock or, in suitable soil conditions, by lowering the pipeline into the seabed. Again, this can be expensive.

Eliminating the hydrotest eliminates the environmental impact of discharging a large volume of chemically treated seawater. It also shortens the construction schedule, the duration of temporary land use requirements, and reduces emissions to the atmosphere.

POSSIBLE MOTIVATIONS FOR KEEPING THE HYDROTEST

If the hydrotest is eliminated, the pipeline might fail while it is being brought up to operating pressure for the first time. The costs of a failure at that project stage could be considerable in terms of increased repair costs and lost/deferred production. Unless very sure that a hydrotest would reveal no defects, it may be prudent to keep the hydrotest.

As well as providing a strength test and a leak test, the hydrotest is reported to have beneficial side effects such as blunting defects, favorably modifying residual stresses, and (in certain cases only) the "warm prestressing" effect. These side effects probably do not highly influence the decision on whether to test or not for most pipelines.

In soft seabeds, the weight of the hydrotest water may reduce free spans by bedding the pipeline down.

If the hydrotest is eliminated, alternative leak testing methods will need to be developed, possibly involving new technology.

The hydrotest is useful for contractual reasons before the contractor hands over the pipeline to the operator.

DNV REQUIREMENTS FOR WAIVING THE HYDROTEST

The design standard for the MEIDP is DNV-OS-F101, which allows the system pressure test (precommissioning hydrotest) to be waived subject to certain requirements. The primary requirement is defined in Section 5, B203: *"Alternative means to prove the same level of safety as with the system pressure test is allowed by agreement given that the mill pressure test requirement of Sec. 7 E100 has been met and not waived in accordance with Sec. 7 E107."* Table 5-1 of the standard imposes additional requirements, including the following: *"An inspection and test regime for the entire submarine pipeline system shall be established and demonstrated to provide the same level of safety as the system pressure test with respect to detectable defect sizes etc.; Records shall show that the specified requirements have consistently been obtained during manufacture, fabrication and installation."* Table 5-1 also includes the requirement *"Less than 75% of the pressure containment design resistance shall be utilized"*, which implies a 33% wall thickness penalty where the wall thickness is governed by the burst limit state. The other requirements of Table 5-1 reflect measures that would probably be taken anyway for most long-distance deep-water pipelines.

The requirement to demonstrate *"the same level of safety as the system pressure test with respect to detectable defect sizes"* is perplexing. The hydrotest demonstrates with almost perfect certainty that there are no defects present that would cause failure on subsequent application of the smaller in-service pressure loading. (The certainty is not quite perfect because of the pressure reversal phenomenon and time dependent factors discussed later.) Since it directly tests strength, the hydrotest takes into account not only defect size but also the coexisting material toughness, wall thickness, residual stresses, geometrical imperfections, etc. NDE techniques, on the other hand, only take into account defect size and then only within the limitations of their detection capabilities and measurement errors. Therefore, the confidence that there are no defects present that would cause failure on application of the service loading is expected to be slightly lower when the hydrotest is waived.

Where fatigue is an issue, initial defect sizes need to be kept relatively small to allow for their subsequent future growth. Unless located in a region of unexpectedly low material toughness, these small defects would not fail in a hydrotest. NDE techniques, on the other hand, can be very effective for detecting small defects. Therefore, an NDE-based approach would be expected to achieve a higher level of safety with respect to the fatigue limit state. The same may be true if defects need to be limited in size in case of large strains such as might occur, for example, in a submarine landslide, turbidity current, or fault movement situation.

Methodologies for probabilistic fatigue analysis are well established (e.g. Petruska, Ku, Masson, Cook, McDonald & Spong, 2006) and a comparative reliability analysis for the screening power of a hydrotest in relation to a small crack could easily be done to demonstrate the superiority of NDE over hydrotesting where fatigue life is the issue.

PRECEDENTS FOR AIV

TransCanada eliminated the pre-service hydrotest for various onshore pipelines (Zhou, Murray, & Abes, 2008). The strength test portion of the hydrotest was replaced with a more comprehensive QA/QC process. The pipelines were leak-tested by filling with natural gas to 50% of the design pressure and conducting leak detection by a helicopter-mounted laser system and/or walking the route of the buried pipeline with portable leak detection equipment.

The GulfTerra Phoenix Gas Pipeline is a 126-km, 18-inch diameter pipeline originating in 1600m of water at a spar-type production facility in the Gulf of Mexico and designed to DNV-OS-F101. It conveys gas to another subsea pipeline system. GulfTerra obtained a waiver from the Code of Federal Regulations hydrotesting requirements (Federal Register, 2004) based on alternative risk control activities.

The Gulf of Aqaba pipeline was completed in 2003. It is a 36-inch pipeline, 34.9 mm wall thickness, pressure 100 barg, in water depths to 860 m. The offshore section is only 15 km in length. Reported motivations for eliminating the hydrotest (Williams, 2007) were that the hydrotest condition would require extensive intervention due to long freespans on the rough seabed and the environmental impact of discharging treated seawater in the restricted waters of the Gulf of

Aqaba, which is designated as an area of high environmental importance. It is reported that additional NDE was performed in the mill and offshore to justify the elimination of the hydrotest and that the pipeline was confirmed as being free of local buckles by passing a gauging pig though the pipe propelled by dry compressed air.

The South Stream Offshore Pipelines will comprise four 32" gas pipelines running from Russia to Bulgaria across the Black Sea, a distance of 931 km in water depths to 2200 m (Meijer & Ethembabaoglu, 2014). It is reported that the pipelines to 30 m water depth will be hydrotested, while the remaining pipeline sections received a hydrotest waiver from DNV in February 2013. The conditions of the waiver include that the pipeline is to be designed to safety class High between 30 m and 345 m water depth. The allowable defect sizes for girth welds are to be more restrictive than permitted by engineering criticality analysis (ECA), making “a balanced judgment between installation practicability and the pipeline integrity”. The DNV-OS-F101 Table 5-1 requirement for 33% extra wall thicknesses in sections governed by the bursting limit state was relaxed to 23% to avoid certain manufacturing and installation issues. Leak inspection is to be performed by ROV after start of operations for at least the pipeline sections between 30 m and 345 m water depth.

All three of the above offshore pipelines were designed to DNV-OS-F101 but none of the projects appears to have published details on how it satisfied the fundamental DNV requirement to demonstrate *“the same level of safety as the system pressure test”*.

THE MIDDLE EAST TO INDIA DEEPWATER GAS PIPELINE (MEIDP)

The MEIDP will run from Oman or other origin in the Middle East to the Indian coast at Gujarat, a distance of approximately 1300 km. The design code is DNV-OS-F101. The linepipe steel will be DNV SAWL 485 DUF, equivalent to API 5L grade X70. The linepipe will be formed by either the UOE or JCOE process and double submerged arc welding.

The project is considering several design options. The AIV study considered one of these options in which the nominal pipe diameter would be 27", with a constant internal diameter of 610 mm, and with nominal wall thicknesses of 32.9 mm, 36.6 mm and 40.5 mm depending on location and water depth. Buckle arrestors (joints of heavy walled pipe at intervals) are required in water depths greater than 659 m.

The design pressure is 400 barg with a minimum delivery pressure at Gujarat of 50 barg. The hydrotest pressure required by the design standard is 441 barg at reference elevation (50 m above sea level), corresponding to 446 barg at sea level or below.

The hoop stress in the subsea pipeline during hydrotest as a fraction of the specified minimum yield stress (SMYS) would be 90%, 81% and 51%, according to whether the wall thickness is 32.9 mm, 36.6 mm, or 40.5 mm, respectively.

The axial stress in the hydrotest is tensile and depends on the water depth. Assuming full axial restraint by soil friction, the ratio α of the axial stress σ_z to the hoop stress σ_θ may be shown to be:

$$\alpha = \frac{\sigma_z}{\sigma_\theta} = 0.3 - 0.2\,\frac{p_o}{\Delta P} \qquad (1)$$

where p_o is the external hydrostatic pressure of the sea and Δp is the pressure differential across the pipe wall. For the MEIDP, with Δp = 446 bar, α would range from 14% at 3500 m water depth to 30% at sea level.

REQUIRED FEATURES OF AN AIV PROGRAM

The AIV study for the MEIDP adopted as a premise that the required features of an AIV program should be determined by reference to a hypothetical precommissioning hydrotest. On this premise, the AIV program should contain features that would ensure that the pipeline would be highly unlikely to fail in a precommissioning hydrotest, were one actually to be performed.

Required Features for Linepipe

Mill Pressure Test

DNV-OS-F101 requires that linepipe be pressure tested in the mill to a von Mises equivalent stress of 96% SMYS. The pre-service hydrotest would normally involve a hoop stress in the most highly stressed section of 90% SMYS. The conditions of end restraint are different in the mill test and the field test. The mill test can be arranged such that the axial stress is zero. In the field test, there will be a tensile stress present, the magnitude of which will depend on the degree of end restraint, on the water depth, on residual lay tension, and on temperature difference between installation and testing. The tensile stress enhances the burst capacity of the pipeline, e.g. axial tensilestress of 30% of hoop stress enhances the burst strength by about 9%. Therefore, a mill test to 96% SMYS is equivalent to a hoop stress of significantly more than 96% SMYS under field test conditions.

For the moment, let it be assumed that no defects are introduced into the linepipe between the completion of the mill test and a pre-service hydrotest. It might be expected that there would be no possibility of linepipe failure in the field stress at hoop stress of 90% SMYS given that a hoop strength of upwards of 96% SMYS was demonstrated in the mill test. However, this expectation is not entirely accurate. Two reasons have been identified why mill pressure tested linepipe could fail in the hydrotest (even in the absence of subsequently introduced defects): (a) the phenomenon known as pressure reversal and (b) time-dependent factors.

Pressure reversal is a phenomenon whereby a pressure test causes the growth of an already large defect, which then fails in a second pressure test at a lower pressure than it withstood in the first test (Brooks, 1968; Kiefner, Maxey, & Eiber,, 1980). The percentage strength reduction due to the pressure reversal phenomenon is usually small, less than 5% or 10%. The pressure reversal phenomenon can also mitigated by ensuring that there are no large flaws present at the time of the mill test since a flaw must be close to being critical in order to grow at all. If the hydrotest is eliminated, the large margin between the high hoop stress in the mill test and the relatively low hoop stress under operating conditions should be sufficient to prevent the possibility of a pressure reversal type failure.

The duration of the mill pressure test is usually about 10 seconds compared with 24 hours in the pre-service hydrotest (DNV-OS-F101 Section 10, O.515). This difference is generally assumed compensated by the lower stress in the field test.

The DNV Table 5-1 requirement that less than 75% of the pressure containment design resistance shall be utilized when the hydrotest is waived would provide an additional margin of safety with respect to the pressure reversal phenomenon and time-dependent factors.

In summary, linepipe that has been mill pressure tested in accordance with the DNV standard is confidently expected to have close to zero failure probability when the lower precommissioning hydrotest pressure or in-service pressure is applied, provided that the linepipe has not been damaged or suffered deterioration since the mill test.

Prevention of Damage/Deterioration of the Linepipe after the Mill Test

As a precaution against delayed hydrogen cracking of the seam weld, seam welding qualification testing typically includes a 72-hour wait period to allow any hydrogen cracking to develop. An additional precaution during linepipe production would be to delay the final NDE of the seam weld until a similar period of time after welding.

Damage of the linepipe after the mill test could be caused by handling, producing dents and/or gouges. If the linepipe is visually inspected with great thoroughness before coating, after each handling operation, and at each arrival point during its transportation, any handling damage of this nature should be easily detected.

Improper stacking of linepipe could cause ovalization or bowing but this is easily preventable through normal good practice.

Transit fatigue cracking can be prevented by proper stacking and support arrangements during transportation. The API has provided recommended practices to avoid transit fatigue cracking for various modes of transportation. Transit fatigue cracks cannot be detected visually but abrasion or peening, where fatigue cracks often start, can be detected visually.

Corrosion of linepipe is preventable by means of internal and external coatings. Internal corrosion is also preventable by use of high quality end caps that prevent condensation or moisture ingress inside the pipe. Corrosion is also easily detected by visual inspection if sufficiently advanced to be a problem.

Thus, qualitative considerations suggest that it should be straightforward to prevent or detect damage or deterioration of linepipe subsequent to the mill test and prior to installation.

Required Features for Girth welds

The axial stress in a pipeline during hydrotesting is tensile and typically equal to 30% of the hoop stress in shallow water. In deeper water, the ratio of the tensile axial stress to the hoop stress decreases due to the effect of the external hydrostatic pressure, e.g. to 14% of hoop stress at 3500 m water depth for the MEIDP. Due to this axial stress, the hydrotest is capable of revealing circumferentially oriented as well as axially oriented girth weld defects.

The pipe/pipe girth welds of the MEIDP will be subject to automated ultrasonic testing (AUT). Pipe/buckle arrestor welds will be inspected by radiography. The frequency with which a substandard girth weld finds its way into a pipeline is expected to increase in proportion to the number of girth welds in the pipeline, all other factors being equal. The MEIDP will contain around 110,000 girth welds, therefore the reliability per weld must be extremely high to ensure that the overall pipeline reliability is acceptable.

The rejection thresholds for different types of girth weld defects must be set sufficiently low that each accepted weld will have an extremely low probability of failure. Defect acceptance criteria are often set by engineering criticality analysis (ECA) using codes such as BS 7910. With application of the maximum recommended safety factors, an ECA to BS 7910 aims to produce a failure probability of less than 10^{-5} per flaw. This may not be sufficiently low for a very long pipeline such as the MEIDP. The rejection thresholds should probably be less than indicated by a conventional ECA when the hydrotest is waived. The possibility of rationally determining the rejection thresholds using probabilistic models is discussed in a later section.

Appropriate welding procedures, NDE procedures, and flaw acceptance criteria should be established for each type of girth weld – pipe/pipe, pipe/buckle arrestor, pipe/riser, pipe/subsea assembly etc.

All detected flaws in girth welds should be measured and reported, not just unacceptable flaws, to verify that welding process parameters are under control and that the acceptable flaw size criteria have been consistently met. These measurements will also provide data on AUT detection and measurement capabilities under field conditions. Very importantly, these measurements will also provide statistical data on numbers, types and size distributions of girth weld flaws to support the development of rational flaw acceptance criteria using probabilistic models.

Essential girth welding variables (current, voltage, travel speed etc.) should be verified and signed off regularly during production welding.

Welding qualification tests for girth welds should include a 72-hour wait period to allow any hydrogen cracking to develop. For stalking operations (e.g. doubles, triples, quads), a similar wait period should be included if reasonably practicable between welding and final girth weld NDE to allow any hydrogen cracking to develop. With this strategy, only a minority of welds may not be subject to inspection for delayed hydrogen cracking.

Required Features for the Pipeline During and After Installation

During installation, it is possible that the pipeline could be damaged by local buckling, by denting resulting from laying over large boulders, or by girth weld fracture resulting from overstress or overstrain. Consideration should be given to monitoring the touchdown point by ROV during pipelay. The QC personnel should verify and sign off in the daily log that installation criteria have not been violated. A gauging pig should be used to check for local buckles. Consideration should be given to conducting inline inspection using smart pigs after installation.

Post-lay trenching/burial operations can cause pipe damage. Contactless equipment should be selected for trenching/burial operations if possible.

Pipelines are at higher risk of anchor damage immediately after installation because shipping is not yet familiar with the presence of the pipeline. Visual inspection and/or smart pig inspection should be made to check for third party damage prior to commissioning.

Consideration should be given to hydrotesting the onshore and nearshore pipelines in recognition of potentially higher risks of mechanical damage in these sections during or shortly after installation. Hydrotesting the sections where wall thickness is controlled by bursting should be considered to avoid the 33% wall thickness penalty imposed by the DNV standard.

Procedures should be developed for detecting and locating any leaks in non-hydrotested sections. Such procedures could be based on detecting seawater ingress. It is also possible that leak inspection could be done by ROV following start of operations. Where the operating pressure exceeds the external hydrostatic pressure, the ROV-based leak inspection might be based on visual inspection. Where the external hydrostatic pressure exceeds the internal operating pressure, ROV-based acoustic techniques may be possible. Further research may need to be done on leak detection methodologies.

DEMONSTRATING "SAME LEVEL OF SAFETY"

One method of meeting the DNV requirement to demonstrate the same level of safety as the system pressure test would be to apply probabilistic models to determine the probability of failure due to mechanical defects were a hypothetical hydrotest to be performed. This is an exercise needed mainly for girth welds since, as noted, there is essentially zero probability of mill-tested linepipe failing in the absence of damage or deterioration subsequent to the mill test.

Different types of girth weld defect would need to be the subject of different probabilistic models. As an example, a probabilistic model has been developed for circumferentially-oriented, semi-elliptical, planar flaws located at the toe of the weld cap, which can be a critical location due to high tensile welding residual stresses and localized bending stresses caused by misalignment and ovality. The fracture resistance of the weld is assessed using the BS 7910:2005 Level 2 Failure Assessment Diagram (FAD), shown schematically in Figure 1.

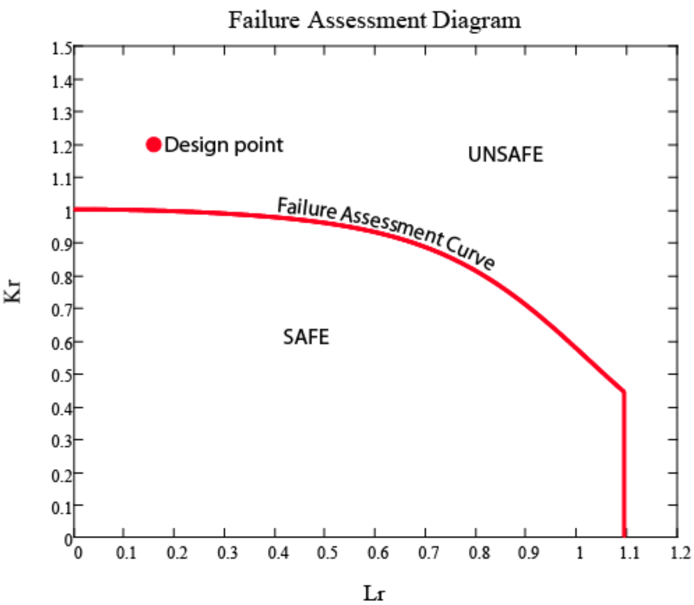

Figure 1: BS 7910 Level 2 Failure Assessment Diagram

The fracture load normalized by the fracture resistance (K_r) and the primary stress normalized by the material yield strength (L_r) provide a design point. For a deterministic analysis, the design point may be plotted on the FAD (see Figure 1) to determine whether the flaw is safe or unsafe. The fracture load is measured by the stress intensity factor, K_I, which is a measure of the magnitude of the stress field at the crack tip. The fracture ratio (K_r) is calculated as

$$K_r = \frac{K_I}{K_{mat}} + \rho \qquad (2)$$

where K_{mat} is the material toughness and ρ is a plasticity correction term. In this example, K_{mat} is assumed to be inferred from Charpy test results. Plastic failure rather than fracture may be the dominant failure mode. The reference stress (σ_{ref}) is a function of the membrane stress and bending stress in the pipe wall. The propensity to plastic failure is measured by the strength ratio L_r calculated as

$$L_r = \frac{\sigma_{ref}}{\sigma_Y} \qquad (3)$$

where σ_Y is the yield strength. σ_{ref} is calculated using the model given in BS 7910:2005, section P.4.3.2, for surface flaws in cylinders. K_I in equation (1) was calculated using the model given in BS 7910:2005, section M.3.2, for surface flaws in flat plates. In applying the flat plate model to determine K_I, the width of the plate W was taken as $\pi\, r_m$, where r_m is the mean pipe radius.

The interaction between the fracture and plastic collapse mechanisms is addressed by the failure assessment curve (FAC) shown in the FAD. The FAC in its curved portion is given by

$$K_r = (1 - 0.14\, L_r{}^2)\{0.3 + 0.7\, exp(-0.65 L_r{}^6)\} \qquad (4)$$

The straight part of the FAC is given by the ratio of flow stress to yield stress. BS 7910:2005 defines the flow stress as the average of the yield strength and the ultimate tensile strength. To avoid having to consider correlations between yield strength and tensile strength, the flow stress was here approximated as 1.094 times the yield stress.

The primary stresses considered are the membrane stress (the axial tensile stress that occurs in hydrotest) and the bending stresses across the pipe wall resulting from ovality and pipe wall thickness differences at the field joint. The primary stresses affect the calculation of both K_I and σ_{ref}. Welding residual stresses are also considered and affect the calculation of K_I only. The residual stresses are conservatively taken equal to the material yield strength.

The probability of failure of a given girth weld flaw when the pre-commissioning hydrotest pressure is applied is determined by the limit state function

$$g = d_2 - d_1 + M_U \qquad (5)$$

where

- d_1 = radial distance from the origin of the failure assessment diagram (FAD) to the design point
- d_2 = radial distance, measured at the same angle as line d_1, from the origin of the FAD to the failure assessment curve (FAC)
- M_U = modelling uncertainty
- $g < 0$ corresponds to failure

The modelling uncertainty term addresses the fact that design points falling outside the failure assessment curve (FAC) do not necessarily represent failure. Conversely, points within the FAC do not necessarily represent safety. An investigation into the modelling uncertainty associated with the BS 7910 Level 2 FAD approach has been described in UK HSE report, OTO 2000/21 (Muhammed, Pisarski, & Sanderson, 2002). The modelling uncertainty was found to vary around the different regions of the FAD. The best-fit probability distributions developed in OTO 2000/21 to represent the modelling uncertainty are given in Table 1. Separate distributions are given for the elastic (60-90º), elastic-plastic (30-60º), and plastic (0-30º) regions of the FAD, where the angle is measured from the horizontal. Note that including the modelling uncertainty reduces the computed risks since the BS 7910 model is inherently conservative.

As an example, the probabilistic model has been applied to a section of the MEIDP, with internal diameter 610 mm and nominal wall thickness 32.9 mm, which wall thickness will be used between 9 m and 2690 m water depth. The linepipe will have an SMYS of 485 MPa. The operating pressure is taken as 400 barg at a reference elevation of 50 m. With an incidental pressure ratio of 1.05 and a test pressure factor of 1.05, the required hydrotest pressure is 441 barg at the reference elevation. Allowing for the internal and external heads of water, the net pressure across the pipe wall would be 446 bar at or below sea level, and the hoop stress would be equal to 90% SMYS in this section. For this example, uncertain variables were described by probability distributions as shown in Tables 1 and 2.

Table 1: Best-fit probability distributions describing modeling uncertainty per OTO 2000/21

FAD region	Distribution	Parameters		
		Location	Scale	Shape
Elastic (60º-90º)	Weibull	-0.06	1.90	2.13
Elastic-plastic (30º-60º)	Weibull	-0.06	0.55	1.08
Collapse (0º-30º)	Weibull	-0.06	0.89	1.18

Table 2: Probability distributions adopted for uncertain variables

Variable	Distribution	Mean	Standard Deviation
Wall thickness	Normal	32.9 mm	0.25 mm
Young's modulus	Normal	210,000 MPa	4,200 MPa
Yield stress	Normal	523 MPa	21 MPa
Ovality	Lognormal	0.002	0.001
Charpy value	Lognormal	150 J	60 J

To compute the probability of a failure in a hypothetical hydrotest, it is also necessary to know the probability distribution of defect sizes after AUT and any repairs. This would be a joint probability distribution relating defect height and defect length.

Published data on girth weld flaw size distributions are very scarce. Kimura et al. (2007) report that 5467 flaws were detected in 4073 girth welds using AUT but present the height and length data as a three dimensional histogram from which little statistical information can be extracted. The detected flaws ranged from 1 mm to more than 6 mm in height. The most common flaw height (rounding to the nearest millimeter) was 3 mm. Based on an ECA, some flaws were repair-welded, reducing the number of flaws to 5291 after repair welding.

For lack of a joint probability distribution relating flaw height and length prior to repair welding, our example will be continued using fictitious data in order to draw out a few points. Let it be assumed that the flaw height distribution before repair welding is lognormal with mean of 3 mm and standard deviation of 3 mm and that the flaw length to height ratio is uniformly distributed in the range 4 to 20.

The probability of detecting and rejecting a defect is taken to be as follows:

$$POD(a, a_o) = \Phi\left[\frac{\beta_0 + \beta_1 \ln(a) - a_o}{\tau}\right] \tag{6}$$

The variable a is the flaw height and a_o is the maximum allowable flaw height. Φ denotes the inverse cumulative normal distribution. From limited AUT qualification data from another project, the model parameters are β_0 = 1.411, β_1= 1.354, τ = 0.819. For the purpose of the example, it is assumed that a_o = 3 mm and that flaws will be repaired either when the measured flaw height exceeds a_o or when the flaw length $2c$ exceeds 25 mm. In reality, there could be different allowable flaw heights depending on flaw length but a single criterion is assumed here to keep the example simple.

The model is solved by Monte Carlo simulation. By post-processing the simulation results, several flaw height distributions can be developed as shown in Figure 2. The initial lognormal flaw height distribution, describing the flaw population, is shown in black. When flaws are detected and found to exceed the assumed acceptance criterion of 3 mm x 25 mm, the weld is assumed repaired, and the simulation yields a post-repair flaw height distribution as shown by the blue curve. Note that the post-repair height distribution is not truncated at 3 mm, because there is a possibility that defects less than 3 mm might be measured as more than 3 mm and wrongly rejected, or conversely that flaws exceeding 3 mm might be measured as less than 3 mm and wrongly accepted.

The critical flaw height distribution corresponding to failure in the hydrotest is shown in red. Failure can occur in the hydrotest for a wide range of flaw heights, depending on flaw length, local material toughness, local wall thicknesses, misalignment, ovality, etc. This model shows that it is not possible for the NDE regime to provide literally the same level of safety as the hydrotest with respect to detectable defects, because the blue curve extends to infinity while the red curve extends to zero, therefore there will always be some area of overlap between the curves, representing failure risk.

For this particular analysis case, the probability of a defect remaining after AUT/repair welding and capable of causing failure were a hydrotest to be performed is found to be about 2 x 10^{-6} per flaw at 2690 m water depth and about 8 x 10^{-6} per flaw at sea level. The difference in these values reflects the influence of external hydrostatic pressure on the axial tension produced in the pipeline.

If the analysis is repeated assuming a more severe distribution of the population flaw heights, the failure probability in a hypothetical hydrotest is found to increase even though the AUT detection capabilities remain the same. This shows that workmanship has a profound effect on risks and that reliance cannot be placed entirely on AUT in order to achieve reliability.

Increasing the wall thickness by 33% (as per DNV-OS-F101 when the wall thickness is governed by the burst limit state) lowers the computed failure probability per flaw only slightly in this example, e.g. to 6 x 10^{-6} per flaw at sea level.

The above probabilistic model was for circumferentially-oriented, semi-elliptical, planar flaws located at the toe of the weld cap and it was based upon BS 7910. Probabilistic models for other types of girth weld flaw could also be developed on the basis of BS 7910.

The greatest challenge in conducting this type of analysis is the lack of published data on the joint probability distribution of flaw height and length (prior to repairs) for the different types of flaws that could be encountered in a girth weld.

A traditional inspection records readings only when there is an above-threshold measurement response. To provide statistical information on the girth weld flaw population, all measurement responses should be recorded during production welding. When there is an above-threshold response, and a repair is performed, the flaw size should be determined and recorded during the repair operation. If the repair involves progressively removing material until the defect is no longer present, the estimate of the defect size will be interval censored. Interval censored observations are easily taken into account in standard statistical methods. By collecting data in this manner during production welding, it would be possible to simultaneously collect data on flaw population statistics and on NDE tool detection and sizing accuracy capabilities. This would allow comprehensive probabilistic models to be developed for use in future projects. When data of this type are available from several projects, information on flaw population statistics within and between projects would be available.

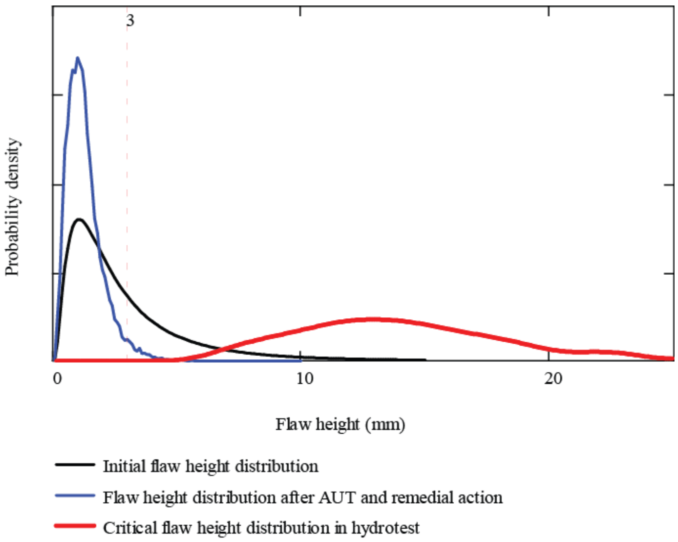

Figure 2 Flaw height distributions

CONCLUSIONS

The DNV-OS-F101 requirement to demonstrate “the same level of safety” as the traditional hydrotest is probably unachievable, at least for the limit state of bursting. Screening out larger mechanical defects by means of NDE cannot expose defects that are critical for the burst limit state with complete certainty. This is for a number of reasons: defect detection probability is finite; NDE involves measurement errors; and defect criticality depends upon factors other than defect size (notably material toughness). There is also opportunity for defects to be introduced after each NDE stage has been completed, e.g. defects can be introduced during linepipe handling and during/after installation. The precommissioning hydrotest on the other hand exposes low burst strength pipeline with near perfect certainty. It is recommended that the industry should engage in a discussion with DNV over the “same level of safety” rule because it is unclear how it is supposed to be demonstrated and probably unachievable with respect to the limit state of bursting. The difficulty of making the demonstration is evidenced by the fact that no project as far as is known has yet published details of its approach to doing so.

If the goal of demonstrating “the same level of safety” as the hydrotest is unachievable for the burst limit state, then the best that can be done is to demonstrate that the hydrotest would not be cost-effective for a particular pipeline. This means that the hydrotest would be assessed like any other risk reducing measure on the basis of whether it is a reasonably practicable measure, taking into account its perceived benefits and its costs. A qualitative assessment may suffice for linepipe to demonstrate that the QA/QC system will prevent or detect damage and deterioration taking place after the mill test and before commissioning. A quantitative assessment may be desirable for girth welds to rationally determine the acceptable defect sizes. Acceptable defect sizes calculated deterministically using BS 7910 may not necessarily provide high enough reliability when the pipeline is very long and accordingly may contain a vast number of girth welds and hence a vast number of flaws.

Currently, the application of probabilistic fracture mechanics models is hindered by lack of data on the joint probability of flaw height and length for the different types of flaws that may be encountered in girth welds.

A further impediment is that the fracture mechanics models needed in the probabilistic analyses are subject to ill-defined modeling errors. The outcomes of probabilistic analysis will depend upon the tails of the distributions describing the modeling errors and experimental data are invariably lacking to accurately define the tails. This is probably always going to be an issue.

It is recommended that reliable data on numbers and size distributions of defects encountered in pipelines should be systematically collected during linepipe manufacture, transportation, handling, and girth welding operations and ideally published for the industry to use.

For the fatigue limit state, NDE is expected to be more effective than hydrotesting for screening defects. This is because hydrotesting can allow large defects to survive in modern high toughness pipelines and such defects could be critical for fatigue life. NDE is also expected to be superior to hydrotest if it is intended to limit flaw sizes in girth welds such that the pipeline could resist high axial strains such as could occur in a subsea landslide, debris flow, turbidity current, or fault movement.

REFERENCES

Brooks, L. E., 1968, *High-Pressure Testing – Pipeline Defect Behavior and Pressure Reversals*, ASME, 68-PET-24.

Cosham, A., Eiber, R. J., Owen, R., & Spiekhout, J., 2006, *A Historical Review of Pre-commissioning Hydrotest Failures*, IPC2006-10333, Proceedings of IPC 2006, International Pipeline Conference, American Society of Mechanical Engineers, Calgary, Canada.

Det Norske Veritas, 2012, DNV-OS-F101, *Submarine Pipeline Systems*.

Federal Register, 2004, Volume 69, Number 124, Department of Transportation, Research and Special Programs Administration, *Grant of Waiver; GulfTerra Field Services LLC*.

Kiefner, J. F., Maxey, W. A., & Eiber, R. J., 1980, *A Study of the Causes of Failure of Defects that have Survived a Prior Hydrostatic Test*, Pipeline Research Committee, American Gas Association, NG-18 Report No. 111 (November 3, 1980).

Kimura, F., Hobbs, R.E., & Wadee, M.A., 2007, *Reliability Assessment for Failure from the Weld Defects in Pipelines Considering the Accuracy of Flaw Detection by AUT*, Proceedings of the Seventeenth (2007) International Offshore and Polar Engineering Conference, Lisbon, Portugal, July 1-6, 2007.

Meijer, A. & Ethembabaoglu, S., 2014, *South Stream Hydrotest Waiver – Another First*, Offshore Pipeline Technology Conference 2014, February 26-27, 2014.

Muhammed, A., Pisarski, H.G., & Sanderson, R.M., 2002, *Calibration of Probability of Failure Estimates made from Probabilistic Fracture Mechanics Analysis*, HSE Books 2000/021.

Petruska, D, Ku, A., Masson, C., Cook, H., McDonald, W., & Spong R., 2006, *Calculation of Reliability-Based Safety Factors for Establishing Defect Acceptance Criteria for Deepwater Riser Welds*, Deep Offshore Technology (D.O.T.) Conference 2006.

Williams, K. 2007, *Gulf of Aqaba Pipeline Crossing: A Story of Success*, Presentation to Society of Petroleum Engineers, Cairo.

Zhou, J., Murray, A., & Abes, J., 2008, *Implementation of Alternative Integrity Validation on a Large Diameter Pipeline Construction Project*, IPC2008-64479, Proceedings of IPC 2008, International Pipeline Conference, ASME, Calgary, Canada.