

Plan Design Enable

ATKINS

Plan Design Enable



Welcome to the fourth edition of the Atkins Technical Journal.

This edition of the Journal is dedicated to Cressida Spachis who sadly passed away earlier this year. Her paper, "Delivery of bus priority projects - a partnership approach", was submitted to the Technical Journal to showcase the success of working in partnership on the Route 38 Project. Cressida was Project Manager for the Route 38 Corridor Management Pilot Study which has developed principles for intensified bus priority now being used across London. The project received many accolades and Cressida was due to be part of the client/consultant team receiving the Improvements to Bus Services award at the London Transport Awards earlier this month. A further tribute to Cressida is made by Andy Southern, Managing Director for Transport Planning and Management, on page 69 immediately prior to the text of her paper.

Once again there has been a great response to the call for papers across all businesses and it is particularly encouraging to have been able to publish papers from a wide range of staff grades covering Graduate Engineers to Technical Directors. The Journal continues to highlight the extensive range of technical disciplines Atkins can offer its clients and our Defence and Offshore Energy capabilities are prominent in this edition. With the recent appointment of three new Network Chairs we can look forward to even more effective showcasing of the diverse and complex problems that Atkins' staff solve so well in their everyday activities.

It is also particularly pleasing to see the excellent working relationships we have developed with our clients and partners being manifested in co-authored papers such as those written with the University of Bath, University of Southampton and English Heritage. This demonstrates that Atkins creates solutions that both we and our clients are proud of.

I hope you enjoy the selection of Technical papers included in this edition.

Chris Hendy

Chair of H&T Technology Board

Highways & Transportation

FOREWORD

Technical Journal 4

Papers 051 - 066

051 - Structures	
The launching of the River Esk Bridge	5
052 - Structures	
Recommendations for assessment Eurocodes for bridges	17
053 - Structures	
Limit equilibrium assessment of drystone retaining structures	29
054 - Intelligent Transport Solutions	
ITSO smart cards in Welsh public transport	39
055 - Intelligent Transport Solutions	
On balance - the see-saws of congestion charge business cases	45
056 - Intelligent Transport Solutions	
Improvements to ramp metering system in England: VISSIM modelling of improvements	51
057 - Highways	
'Scrambled' pedestrian crossings at signal controlled junctions - A case study	57
058 - Highways	
Delivery of bus priority projects - A partnership approach	71
059 - Water & Environment	
Water quality modelling as a tool for assessing new water resource management options: The case of the River Stour, Kent	81
060 - Water & Environment	
High accuracy recording for heritage applications - Dover Castle graffiti	87
061 - Aerospace	
Conjugate heat transfer study of a spin pit rig: Application to the lifing of HP turbine disc firtrees	93
062 - Aerospace	
Non-deterministic thermo-fluid analysis of a compressor rotor-stator cavity	105
063 - Energy	
The optimum position for a tidal power barrage in the Severn estuary	115
064 - Energy	
Greater Plutonio project - Subsea flowline design and performance	123
065 - Defence	
Development of an armour mass estimation tool for land vehicles	147
066 - Defence	
Explosives safety - An oxymoron?	155



Rob Liddle

Senior Group Engineer
Highways & Transportation

Abstract

The River Esk Bridge carries the southbound carriageway of the M6 motorway over the River Esk estuary and forms part of the recently constructed M6 Carlisle to Guards Mill Motorway Extension which provides the missing link in the motorway network between England and Scotland. The bridge is a 180m long four span steel composite viaduct. Due to site constraints the most economical and practical method of construction was to launch the steelwork into position rather than to lift the steelwork in by crane. The bridge steelwork, weighing over 800 tonnes, was assembled on the northern approach to the bridge and launched southwards over the River Esk estuary

This paper demonstrates that the design and construction of a launched bridge can be significantly more complicated than the design of a comparable bridge which is constructed using conventional methods, and that there are many more load cases and construction details to consider.

This paper covers the approach for carrying out the analysis, design, detailing and construction of the structure and focuses specifically on the key issues relating to launching steelwork. It serves as a useful guide and quick point of reference for the design of launched steel bridges.

Introduction

This paper covers the approach for carrying out the analysis, design and detailing of the River Esk Bridge on the M6 Carlisle to Guards Mill Motorway Extension. The Esk bridge is a four span 180m long steel composite viaduct over the River Esk and carries the southbound carriageway of the new section of M6 Motorway. The span arrangement is 31.4m/51.9m/51.9m/44.9m. The south abutment is fully integral, forming the point of fixity, and the superstructure is free to move longitudinally over the piers and north abutment.

The bridge steelwork was assembled on the northern approach to the bridge and launched southwards over the River Esk estuary. The total weight of the steelwork was over 800 tonnes.

The reason for the launch was that access to the river for steelwork assembly and lifting was not possible due to the tidal nature of the river and the high risk of flooding. Also, lifting the beams in from the adjacent A74 'Metal Bridge' was not possible due to traffic management issues. The A74 Metal Bridge was refurbished to carry the M6 northbound carriageway.



Figure 1 - River Esk Launch



Figure 2 - River Esk Launch



Figure 3 - Views of the completed bridge adjacent to the existing A74 metal bridge



Figure 4 - Views of the completed bridge adjacent to the existing A74 metal bridge

The M6 Carlisle to Guards Mill scheme provides the missing link in the motorway network between England and Scotland.

The main difference between a static bridge design and a launched bridge design is that each section of the steelwork elements is designed to work in different ways, depending on its position during the launch and also in the completed in-service condition. The main girders and splices act in both hogging and sagging and the leading front span behaves as a very long cantilever during the launch.

A temporary nose and tail assembly is also required to be spliced to the main steelwork to accommodate the large deflections that occur. Many bridge substructure and steelwork details were modified from a 'standard design' to accommodate the effects of the bridge launch and use of temporary launching equipment.

This paper focuses specifically on the launch-related aspects of the design and construction of the bridge, not on standard bridge design and construction topics.

Analysis, design and buildability

Line beam analysis of launch phases

Launch Phases

The bridge steelwork was launched southwards in two phases:

Phase 1 - The southern half of the steelwork was assembled on temporary land bases together with a launch nose and tail, and then launched southwards. The launch platform comprised the landbases, spaced at 28m/28m/28m, constructed in the partially completed approach embankment with the landbase levels set so that the steelwork could be set out and assembled to the vertical alignment required for the launch. The piers in the river were constructed in temporary cofferdams which were then converted into temporary islands to allow safe access and working at the piers without the risk of flooding. The cofferdams and islands were constructed higher than the maximum predicted flood level.

Phase 2 - The tail was removed from the Phase 1 steelwork. The northern most steelwork spans were assembled and attached to the steelwork already launched. The tail was reattached to the rear of the Phase 2 steelwork. The whole bridge steelwork was then launched. On completion of the launch the temporary launch nose and tail steelwork were removed.

The launch equipment comprised strand jacking tendons which were attached to the north abutment using a steel hauling frame which was embedded into the concrete stem of the abutment to resist the hauling load. The tendons were attached to the launch tail via a large steel cross-beam which was attached to the end of all five beams to transfer the hauling load into the beams.

Choice of nose and tail length

The longer the nose, the less permanent steelwork would be cantilevering at the leading span prior to 'touch down' at a support. A longer nose tends to reduce the cantilever moment as the nose weight per unit length is normally lighter than the permanent works.

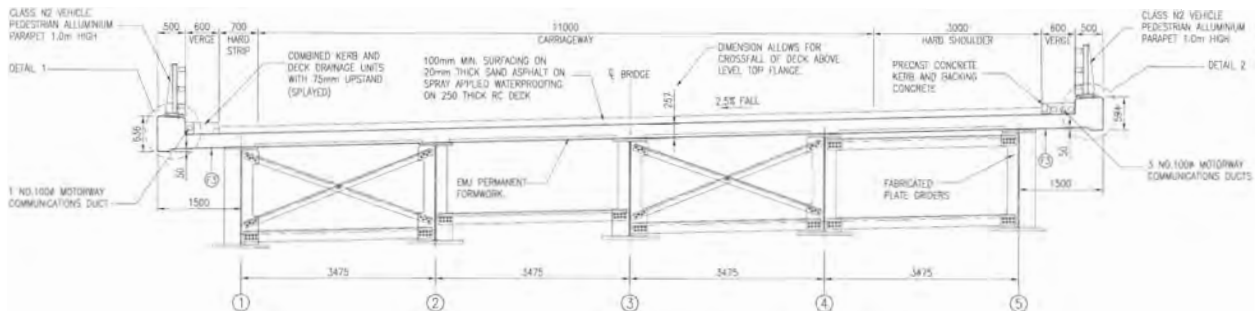


Figure 5 - Cross section of the deck



Figure 6 - View of the launch showing significant leading span cantilever deflection and temporary cofferdam/island around an intermediate pier

The nose bottom flanges need to be curved to compensate for the cantilever deflection of the leading span and to prevent the steelwork clashing with the piers and abutments as it passes over them.

The nose profile should be detailed to allow sufficient tolerance to ensure the nose travels past the support bearings a sufficient amount before 'touch-down'. 1m was specified and the vertical clearance at the tip was 165mm before 'touch-down' after considering deflections. The maximum cantilever deflection was 1112mm.

The use of a curved soffit is a different approach from that commonly used for post-tensioned incrementally launched concrete bridges in which the nosing usually has a hydraulic height adjustable tip to take up the deflection prior to touch down at a support.

The tail needed to be profiled to take account of the back span 'dropping' off the land bases and bridge supports.

Land bases are temporary supports in the steel launch platform area where the steel is assembled. A curved profile is required in the tail to gradually take up the cantilever deflection and to prevent the steelwork making a sudden drop as it moves off a support.

At each stage of the launch the 'touch-down' positions in the nose were calculated as they were required by the steelwork contractor – these being the location in the nose which would touch down first and also the angle of the nose flange at touch down.

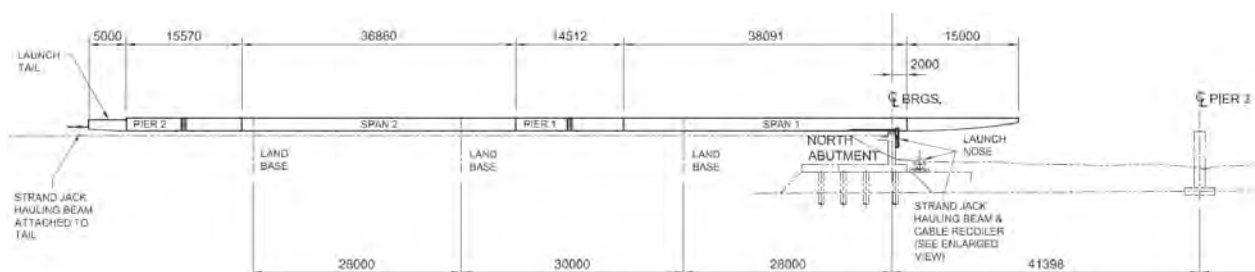


Figure 7 - Phase 1 steelwork assembled ready for launch

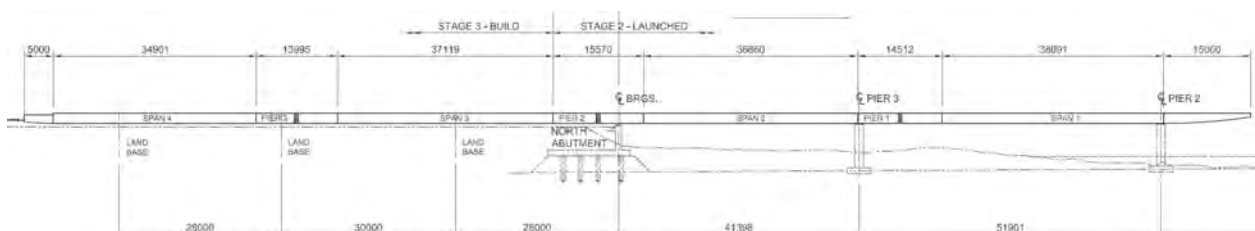


Figure 8 - Phase 1 steelwork launched and phase 2 steelwork attached ready for launch

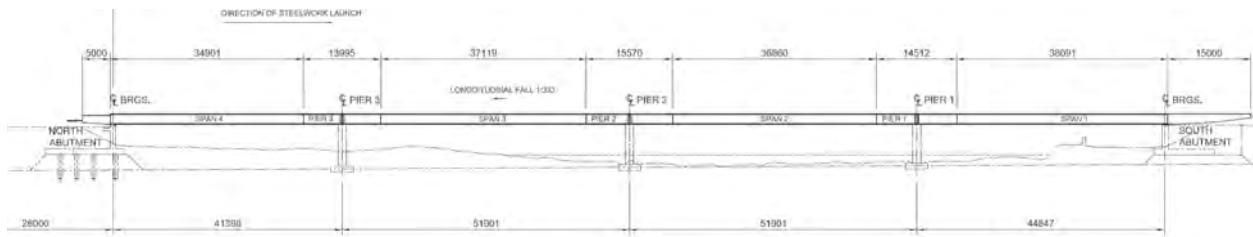


Figure 9 - Phase 2 launch complete

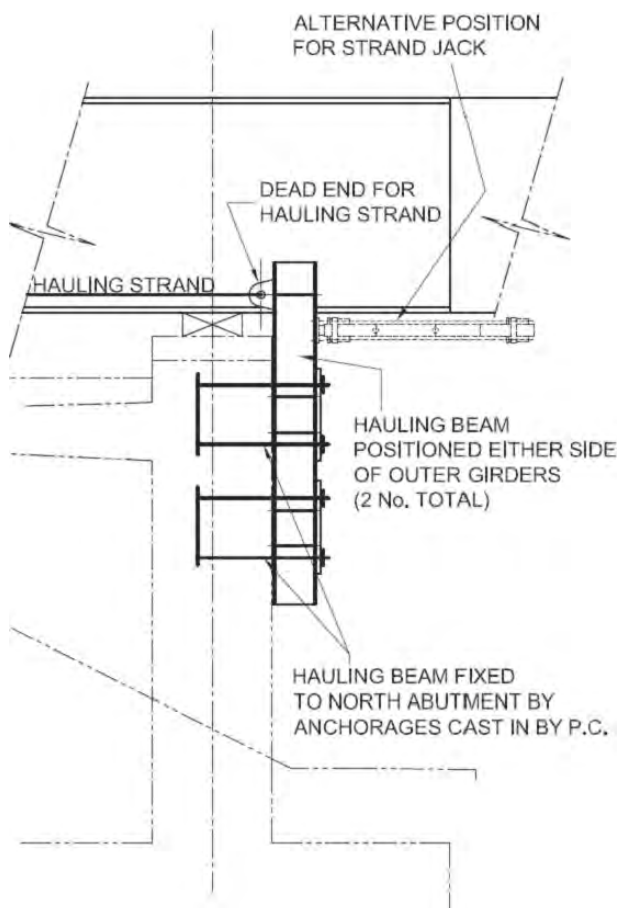


Figure 10 - Hauling frame attached to north abutment

Launch bearing type

The steelwork was launched on temporary launch bearings, set 351mm higher than the final position and then 'jacked' down onto the permanent bearings. Launch bearings are usually a steel roller/skate or PTFE pads with lubricant applied to the underside of the bottom flange. Skates can sometimes damage the paint to the underside of the bottom flange and have a higher resistance/coefficient of friction against the launch forces. PTFE bearings were chosen for the Esk launch.

Deflection analysis during launch

The software package STRAP was used to analyse the structure. Line beams were used to analyse the nose and tail deflections. At some stages during the launch uplift occurred at a support, indicated by a negative support reaction. The support was removed from the line beam model and the model solved again to obtain the correct results. It was important to accurately model the steelwork weight distribution, especially the plan bracing in the leading span, nose and tail, to obtain the correct results. Calculated deflections were required to enable accurate profiling on the nose and tail to accommodate the predicted deflections.

Analysis of steelwork

The steelwork was analysed for the cases where it was supported on the land bases as well as on the permanent supports. The former is especially important for modelling tail deflections and to check that the steelwork does not bottom-out on the ground when leaving a land base. To check for bottoming-out, the steelwork precamber also needed to be considered.

Determination of launch phases

The most cost effective solution was to launch from the north end only. On some very long bridges the steelwork is launched from both ends but this requires heavy plant, access, assembly area and welfare facilities in two locations rather than one.

Due to the curve of the road alignment and compulsory purchase order (CPO) boundary, an assembly area was prepared which was large enough for only half the steelwork and therefore the steelwork had to be launched in two phases. Launching from one end only also minimises the number of temporary land bases required.

Design details to consider

Changes between hog and sag affecting effective length calculations

At every location along the bridge each section of the steelwork acts in either sagging or hogging, depending on its position during the various stages of the launch as the steelwork is pushed over and between supports. Effective length calculations and section checks were required for every scenario as follows:

- The number of bracings in a span could vary
- The steelwork spans varying span lengths as it is moved over and between land bases, abutments and piers
- Sections changed from acting in hogging to sagging
- Both the launch phase and in-service case were considered.

For each steelwork section the most critical hog and sag case was chosen to calculate the beam effective length. Splices were also subject to both hog and sag moments and this was considered in the splice design.

Intermediate bracings were treated as discrete torsional restraints to BS 5400 Part 31 clause 9.6.4.1.2 and BD13/06² designed to clause 9.12.2.

Effects of differential precamber, setting out tolerances and fabrication tolerances

At any point along the bridge the levels of the beam bottom flanges will vary between adjacent beams. The difference in level will be mainly due to differences in precamber but also a component from setting out and fabrication tolerances. This would not normally present a problem for a bridge constructed conventionally, however, for a launched bridge as the steelwork is pushed over the launch bearings at a support, the steelwork will be forced up or down as the level of the bearings is fixed. Put another way, the difference in level between adjacent launch bearings at a support is fixed but the difference in level between adjacent flanges will vary throughout the span. This level difference or 'lack of fit' results in distortion of the bracing frames as the steelwork moves over the supports imparting large forces into the bracing members.



Figure 11 - View of leading span and nose during the launch



Figure 12 - Touch down at an intermediate pier position



Figure 13 - View of tail, temporary land base, launch bearings and strandjack launch system

If the deflections are too large to be accommodated, then beam lift-off at a support will occur causing additional large forces in the bracing members.

Plane frame analysis models of the bracing were used to model these effects as summarised below:

- (i) Precamber values were determined from the main grillage of the completed structure (for all in-service dead and superimposed dead loads)
- (ii) The differences in precamber between beams across individual braced bays were calculated
- (iii) A plane frame model including all 4 bracing bays was set up to account for any transverse distribution effects (redistribution of vertical load)
- (iv) The calculated differential deflections were applied to the plane frame model to determine the forces induced in the bracing and also calculate the altered support reactions based on transverse distribution effects
- (v) The system was rechecked for lift-off effects as follows:
 - (a) The vertical force required to produce the prescribed deflection in the plane frame was calculated by removing the relevant support and applying a force to the model iteratively
 - (b) If the force required to produce the deflection was larger than the actual support reaction that occurred then it was assumed that lift-off would occur
 - (c) The support reaction at the lifted-off beam was transferred to the two adjacent beams. The additional dead weight in the bracing due to the lifted-off beam was taken as being equal to the support reaction divided by the number of bracing bays in half the span either side of the support. This is because for a lifted-off beam the weight of the beam and bracings is transferred to the two adjacent beams via several bracing bays
 - (d) The bracing plane frame was resolved by removing the support where the beam had lifted off

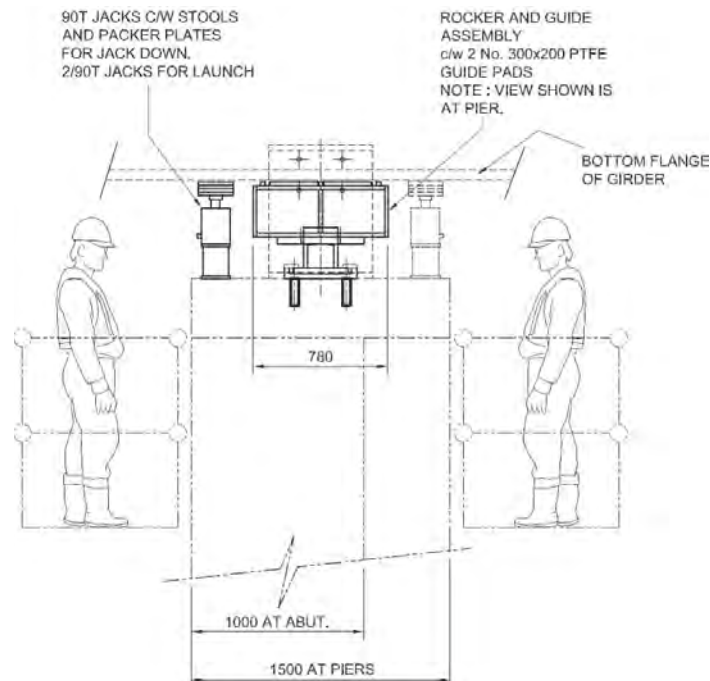


Figure 14 - Placing of jacks for removal of launch bearings prior to jack down



Figure 15 - Nose approaching launch bearings at a pier prior to touch-down

- (e) The resulting forces from the effects of differential beam levels were added to other coexistent loads effects for design of the bracing members and connections.

Of course, a more sophisticated space frame model of the whole steelwork could have been used to simplify the number of operations, but it was decided at the beginning of the scheme to only use line beam models of the girders to model the launching effects.

The total fabrication and setting out tolerance alone was +/- 10mm without taking account of differential precamber. It is important to allow for the possibility of beam uplift since applying the full deflections to the plane frames models can generate massive forces in the bracing members.

Initially the steelwork contractor suggested we provide both a top and bottom 'push-pull' member for each bracing set between beams 2 and 3 as shown in Figure 16.

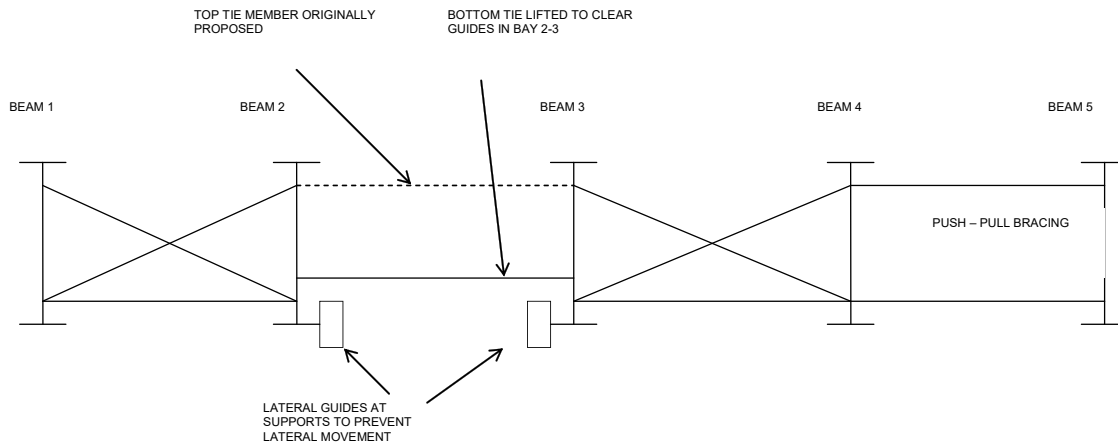


Figure 16 - Arrangement of bracing

The contractor had concerns that the bridge was not adequately braced between the group of three beams and the pair and that the two sections could deviate apart. Launching 5 beams is not a usual scenario as beams would normally be launched in pairs.

They proposed continuous bracing at all bracing locations should be used. This made the whole bracing system effectively fully rigid and so prescribing large deflections generated large forces resulting in the need for excessively large bracing members and bolted connections. By removing the top member in bay 2-3 the bracing system became more flexible as the group of 2 beams (beams 1 and 2) could rotate or flex relative to the group of three beams (beams 3, 4 and 5).

The bracing would be difficult to remove, being over the river, and with the EMJ formwork already in place. Therefore, from a health and safety and CDM perspective it was decided to leave the bracing in place. The bracing was also checked for the in-service case.

Web patch load checks on unstiffened sections of web

In the in-service condition the steelwork at supports is stiffened with web bearing stiffeners to allow the web to carry the vertical support reactions. However, for most of the time during the launch, the position of web bearing stiffeners or intermediate transverse web stiffeners would not coincide with the position of the supports. Therefore the unstiffened parts of the web were checked for carrying the vertical support reactions.

The vertical support reactions during the launch comprised the steel self weight, formwork on all but the leading span and a proportion of the deck reinforcement and so the reactions were a lot less than the in-service support reactions. However, the web thicknesses required were slightly thicker than those required for the in-service condition to accommodate the unstiffened bearing support scenarios. BS5400 Part 31 clause 9.9.6 and Annex A, 'web patch loading', were used to perform this check. It is noted that some further refinement of the design would have been possible using the more recently developed patch loading rules in EN 1993-1-53.

Splices designed to resist hauling load

Splices were checked for carrying the horizontal hauling load together with in span bending and shear effects.

Plan bracing requirements

During launching the leading span acts as a long cantilever prior to touch-down on a support. In BS5400 Part 3 the determination of the effectiveness of lateral restraints is based on how well the compression flange of a beam is restrained against lateral and rotational movement. Without plan bracing, for a maximum cantilever of 51m, the tip of the steelwork would be unrestrained and would sway significantly and so the effective length would be too large to enable a realistic plate design. Triangulated plan bracing was provided throughout the leading span to prevent this sway and fully restrain the steelwork.



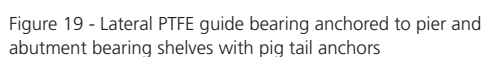
Figure 17 - View showing launch track marks on bottom flanges and plan bracing left in place



Plan bracing was also required between all 5 beams in the end bay of the leading end of the permanent steelwork i.e. just behind the nose splice. This plan bracing was required for erection purposes and during launching to prevent adjacent beams moving longitudinally relative to each other, known as 'leading'.

Launch load effects due to plan curvature of the superstructure

At first glance it might appear that the lateral launch forces on the PTFE lateral guide bearings would be equal to the force required to push the steelwork sideways over the support bearings i.e. the total vertical load at the support multiplied by the coefficient of friction.



For a simple symmetrical structure on a constant radius horizontal curve, the lateral guide loads can be easily determined by hand calculation using simple rules of equilibrium, as shown in Figure 18. The simple example assumes the steelwork is pushed and pulled at both ends whereas the more usual case, as with the River Esk, is to push the steelwork from one end only. Equal supports reactions are also assumed for simplicity.

Taking all moments about point A summed to zero:

$$F = 3P$$

Summing all loads in plan along the y axis to zero:

$$R_1 = \frac{P + 2PC \cos(\theta_4) - 3PC \cos(\theta_2)}{2 \sin(\theta_4)}$$

The above assumptions are correct for a deck on a constant radius horizontal curve. For the case where the deck is on a varying radius horizontal curve, or a combination of constant curves and transition curves, then additional lateral guide forces need to be considered. These additional forces result from changing the plan curvature of the deck from the initial constructed shape that is forced at any instant by the guides. It is worth noting that if these forces were too large then the assumption of ignoring friction on the guides would not be correct.

The additional lateral guide forces can be determined by modelling the effects of the forced deflection of the deck in a simple grillage model or by using simple beam formulae.

An impact factor of 1.5 was applied to the calculated lateral guide forces and these forces were considered in the bracing design.

Launch load effects on abutments and piers

The substructures were checked for the load effects from the launch hauling load, lateral guide forces and friction effects of the steelwork while being launched. The steelwork was also being launched uphill by a gradient 0.3% and this gradient was considered.

The lateral and longitudinal forces on the piers and abutments due to moving the steelwork were considered in the design of the abutments and piers. The coefficient of friction in the longitudinal and transverse directions was taken as 11% (0.11) and 5% (0.05) respectively.

The north abutment was designed to resist the full hauling load and the hauling frame was anchored into the north abutment stem. The hauling load effectively pulled the abutment into the fill and so passive resistance could be utilised.

Participating bracing

The optimum number of main beams was investigated at the preliminary design and value engineering stages of the design. Cost comparisons were made between 4, 5 or 6 beams. With an even number of beams the bracing is usually arranged in discrete unconnected pairs i.e. for 6 beams - three braced pairs. However, for an odd number of beams the bracing is arranged so that three beams are connected together. For the five beam case on the Esk there is a braced pair and a group of three.

When three or more beams are connected together, either by cross bracing or top and bottom 'push-pull' bracing, then load sharing can occur. This is termed 'participating bracing', where, as the deck deflects under load, some vertical load is transferred through 'transverse distribution'.

The participating bracing was modelled to determine the additional load effects in the bracing members.

Lubricant coefficient of friction values

A lubricant was brush applied to the bottom flanges of the beams during the launch to minimise friction. The lubricant was actually a soap normally used for removing grease and oil, namely 'Tufenega Green Gel' (like Swarfega).

During the approvals process of the AIP the Highways Agency (HA) SSR requested a justification for the assumed coefficient of friction values. As mentioned in the AIP, the main girders and substructure were checked for additional temporary loads due to launching of the main girders, these loads are summarised as follows:

- Wind loading to BD37/01
- Loading due to differential vertical pre-camber as specified in BS5400
- Lateral sliding friction loads exerted by temporary launching guides due to the curved horizontal alignment of the deck, assuming a friction coefficient of 0.05 based on data used for similar past schemes as detailed below
- Longitudinal sliding friction loads, assuming a friction coefficient of 0.11 based on data used for similar past schemes as detailed below
- Longitudinal launching forces applied on the tail end of the steel deck.

The same launch method had been completed successfully by the steelwork contractor shortly before the launch of the River Esk on the new Swale Crossing forming part of the 30 year DBFO contract for the A249 Iwade Bypass between the Isle of Sheppey and mainland Kent. To ensure that the structure would not be subjected to excessive loads during launching, the steelwork contractor had carried out extensive in-house



Figure 20 - Strand-jack launching system with stiff cross beam transferring loads to tail beam ends - monitoring station under cover in background

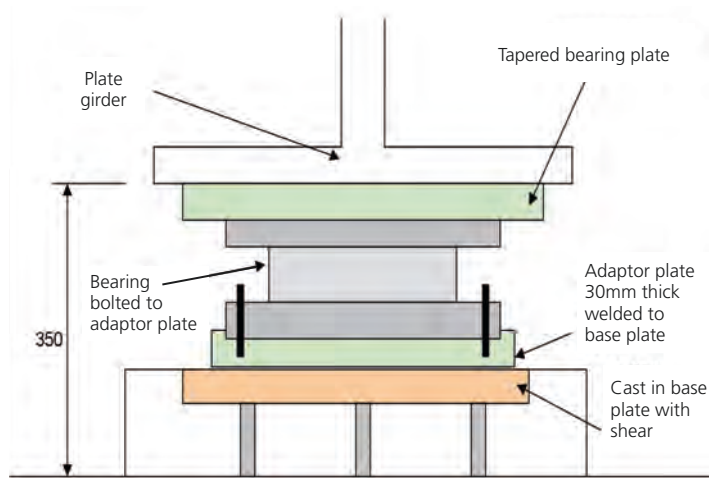


Figure 21 - Adaptor plate arrangement

testing in order to determine the most suitable friction coefficients specifically used for the scheme. The friction coefficients stated above were based on data taken from continuously monitoring the load effects on the structure during the Swale Crossing launch. The same monitoring system was adopted for the River Esk scheme described as follows:

Strand jack loads were monitored during the launch operation - calibrated pressure transmitters were installed in the hydraulic lines to the strand jacks and connected to a laptop computer to monitor strand jack launch loads. This provided a digital and graphical display of the load being applied by the two jacks every five seconds that were continuously monitored by an engineer who was in radio contact with the Launch Controller.

If unexpectedly high loads were indicated then the launch was to be halted and the cause investigated and resolved to continue the operation. The monitoring system provided a permanent digital record of the hauling loads.

Prisms were attached to each pier, at the top, middle and bottom levels in order to check for displacement and rotation of the substructure. Measurements were taken continuously during the launching operation by surveyors using EDM equipment.

Furthermore, to provide an additional margin of safety a higher load factor than that specified in BD37/01 was conservatively adopted to check the permanent works during launching to allow for dynamic and impact effects. A load factor of 1.5 was used for both dead load and superimposed dead loads, instead of 1.05 and 1.2, together with the friction coefficients stated above. Initial structural checks showed that the load effects on the substructure due to the in-service condition would be far in excess of those generated during launching, resulting in a very high safety factor during launching ranging from 4.3 for overturning up to 29 for shear; the main girders safety factor was 1.75.

The friction coefficients to be adopted during the launch were deemed safe to use and accepted by the HA SSR.

The main girders were checked in biaxial bending for the effects of major axis bending due to the launch and minor axis bending due to wind loads.

Geometry arrangements, details and buildability

The steelwork contractor would have preferred the bearing taper plates to be 'through bolted' to the bottom flange to avoid having to site weld and paint up over the piers. However, due to space limitations for bolt locations the taper plates were welded on site post launch. Obviously the taper plates could only be added once the steelwork was launched otherwise they would snag with the launch bearings.

A steel adaptor plate was used to attach both the temporary launch bearings and the permanent bearings. At each bearing location a base plate with shear studs was cast into the bearing plinth. An adaptor plate was welded to the base plate. The temporary launch bearings were bolted to the adaptor plate and then removed after the launch. The permanent bearings were then bolted to the same adaptor plate prior to the steelwork being lowered down on jacks from the elevated launch position onto the permanent bearings. This was a practical detail to allow the use of two types of bearing in the same location. All site welding was painted to full Highways Agency specification.



Figure 22 - Split bottom flange splice plates

The temporary launch bearings comprised two 300 long x 200 wide PTFE bearings in a line (600 long). The whole rocker bearing assemblies were typically 780mm x 370mm.

During the launch at least two points of lateral restraint were required at all times to prevent the steelwork shifting out of line.

During launching the steelwork was 351mm higher than its final in-service position. At each support a pair of lateral guides was provided between beams 2 and 3. The bracing in this bay was lifted higher than in adjacent bays in order to clear the guides. A distance of 350mm from underside of bracing to underside of bottom flange was required to clear the guides. Bay 2-3, being a push-pull bracing bay, was easier to modify than an adjacent cross-bracing bay.

Once the launch was complete the structural steelwork was jacked down into its final position. Initially the preferred sequence was for the steelwork to be jacked down fully, one support at a time. However, deflection calculations were carried out to determine what bending moments would be induced in the steelwork by this imposed deflection (351mm). It was found that lift-off from the jacks would occur at some support locations before reaching the required level. The jacking sequence was revised to allow only jacking in stages at some supports to prevent lift-off.

Bottom flange splice cover plates comprised two separate plates 350mm apart rather than a single plate. This was to ensure the splice plates ran either side of the launch bearings. Bottom flange butt welds were ground flush where there was to be contact with the launch bearings and the guides in order to prevent snagging and unnecessary high hauling loads. This was the central section (300mm wide) of the underside on all girders, and the edges of the two inner girders, which were in contact with the lateral guides. All other butt welds were to be dressed only.

Bottom flanges needed to have the same outstand dimension in bay 2-3 so that the lateral guides would always be in contact with the flange edge.

This also applied to the launch nose and tail. In the final design of the permanent steelwork all bottom flanges were 900mm wide and all top flanges 600mm wide. The flange sizes varied in the nose and tail for economy except in bay 2-3 where the same outstand as for the main steelwork was required.

Web to flange welds were prepared in a T & I machine and were not 'fitted'. Therefore for web patch load checks from the support reactions the weld was checked for resisting the direct patch load since the patch load would not be transferred from the flange to the web in direct bearing but via the web to flange welds.

A safe maximum launch wind speed was specified on the drawings.

The wind speed corresponded to a 10 year return period with a 3 second gust speed of 18m/s.

The permanent formwork and a proportion of deck reinforcement was added to the steelwork prior to the launch on all but the leading span to minimise construction operations working at height and also mitigate access problems. A cost comparison was made between EMJ and Omnia formwork. EMJ was chosen as it is half the weight of Omnia.

At the integral south abutment the wall reinforcement was required to run vertically up into the diaphragm and lap horizontally into the deck slab. In normal construction starter bars are usually left protruding where the diaphragm is to be cast both in the front and back face of the abutment stem/diaphragm. Because of the launch operation these bars would clash with the steelwork and nosing. A 300 wide x 1800 deep box-out was provided on the back face so that the kicker bars could be left protruding below the level of the steelwork. For the front face bars couplers were used instead of starter bars. The reduced thickness of abutment stem was checked for the temporary load cases.

At the north abutment the inspection gallery starter bars were required to lap into the back of the gallery wall. These bars were kept as low as possible by slightly lowering the gallery floor.

At the piers sufficient room was available to jack down the steelwork from under the bottom flanges of the main beams as the piers are 1500mm wide. However, at the north and south abutment the bearing shelves were much narrower and so no room was available to place the jacks under the main flanges. At the south abutment no permanent bearing plinths were to be provided and so jacking was made more difficult. To solve the problem, the bracing adopted at these locations was mid-height horizontal I sections, with stiffened jacking points provided either side of each main beam to enable jacking.

Conclusions

This paper demonstrates that the design and construction of a launched bridge can be significantly more complicated than the design of a comparable bridge which is constructed using conventional methods. There are many more load cases, load effects and details to consider. However, in situations where access to the bridge site or traffic management issues make conventional bridge construction difficult, bridge launching can offer an economical solution. This paper summarises some of the key issues relating to launching steelwork.

The first phase of the launch was successfully completed on 16 May 2007 and the second phase on 11 June 2007. The scheme was opened to traffic in December 2008

Acknowledgements

The author would like to acknowledge Fairfield Mabey Ltd, the Steelwork Contractor, for sharing their knowledge and expertise on launching steelwork. The author would also like to acknowledge the Main Contractor - Carillion, the Lead Consultant - Capita and also the Sub-Consultants (Bridge Designers) Grontmij.

References

1. BS5400 : Part 3 (2000): Design of steel bridges. British Standards Institution, London.
2. BD13/06 (2006). Design of steel bridges. Use of 5400-3:2000. Highways Agency, UK
3. BS EN 1993-1-5 (2006) Design of Steel Structures. Part 1.5: Plated structural elements. British Standards Institution, London.

Recommendations for assessment Eurocodes for bridges



Chris R Hendy

**Head of Bridge Design
and Technology**
Highways & Transportation



Jessica Sandberg

Group Engineer
Highways & Transportation



Navil K Shetty

Director
Rail

Abstract

The new Structural Eurocodes offer increased economy in design over most existing codes of practice. This is achieved through codification of more advanced calculation methods, such as non-linear analysis, and also through codified rules which give increased economy in many situations compared to the old British Standard rules because they are based on recent testing and numerous non-linear finite element parametric studies.

This presents great opportunities for improving sustainability in design. However, the assessment of existing structures is specifically outside the scope of the Eurocodes, so many of the benefits these new codes bring cannot be directly applied when reviewing old structures. Increased assessments of strength would be highly beneficial in demonstrating that old structures with increased loading or small amounts of deterioration were still adequate to continue in operation without costly refurbishment, modification or reconstruction. The Eurocodes lend themselves to adaptation for assessment of structures because the rules are generally based around the application of first principles; this gives greater scope for determining the true ultimate strength of a structure and for the incorporation of structure specific information on loads and material properties in a rigorous manner. The assessment criteria can also be differentiated to take account of differences in consequences of failure of a bridge.

This paper investigates the areas where the steel and concrete Eurocodes give increased resistances compared to existing codes and makes recommendations for sections which could be directly applied to the assessment of existing structures. Areas where the Eurocodes cannot be applied directly to existing structures are also identified together with the reasons why, such as reliance on modern material and execution specifications. Recommendations are then made for how the Eurocode design rules could be modified for assessment situations, including the use of measured strengths and imperfections in calculation. Actions and the format for combining actions, are also investigated and recommendations made for modifications to these aspects for use in assessing bridges.

Finally, an overall assessment of the scope of work required to produce an assessment suite of Eurocodes is made.

Introduction

The new Structural Eurocodes offer increased economy in design over most existing codes of practice. This is achieved through codification of more advanced calculation methods, such as non-linear analysis, and also through codified rules which give increased economy in many situations compared to the old British Standard rules because they are based on much more recent testing and non-linear finite element parametric studies. This presents great opportunities for improving sustainability in design. However, the assessment of existing structures is specifically outside the scope of the Eurocodes, so many of the benefits these new codes bring cannot be directly applied when reviewing old structures. Increased

assessments of strength would be highly beneficial in justifying that old structures, with increased loading or small amounts of deterioration, were still adequate to continue in operation without costly refurbishment, modification or reconstruction. The Eurocodes lend themselves to adaptation for assessment of structures because the rules are generally based on the application of first principles; this gives greater scope for determining the true ultimate strength of a structure than do more prescriptive or empirical rules.

Traffic data (collected using weigh-in-motion systems for long span bridges or from known traffic counts for short span bridges), measured dimensions and material properties derived from

in situ tests on a structure can be rigorously incorporated into the limit state verification method given in the Eurocodes. In addition, assessment criteria can be differentiated by taking account of the differences in consequences of failure of bridges. The approaches to be used for developing the relevant rules for assessment are discussed in this paper. The main area of difficulty in applying the Eurocodes to the check of existing structures is that the assumptions of particular material properties and workmanship inherent in the codes may not be borne out by the actual construction. Many of the enhanced resistances in Eurocode 3 for example stem from the ductility and strain hardening characteristics



Figure 1 - BSALL is used for the operation of the Forth Road Bridge



Figure 2 - Imperfections in longitudinal stiffeners following an impact

of modern steels that are produced to the specification EN 100251. Older steels, or indeed wrought and cast irons, may not comply with the requirements of this specification. Similarly, all steelwork designed to Eurocode 3 should be fabricated and erected in accordance with EN 1090-22. This document sets out, among many other things, the tolerances for steelwork such as bow imperfections in columns. The strut curves in Eurocode 3 assume that these bow imperfections will not be exceeded but this may not be the case for old bridges which were not constructed to the same tolerances; it is necessary to use the actual tolerances in calculation. The Eurocodes may not even cover design with certain materials. Plain round bars, for example, are not covered by Eurocode 2 because of their poor and unpredictable bond

characteristics. An additional factor is that the Eurocodes do not directly deal with reduction in strength due to deterioration arising from unexpected behaviour or lack of maintenance.

In the absence of assessment versions of the Eurocodes, a number of problems and inconveniences will arise for UK engineers. First, there is the obvious problem that engineers will need to remain conversant with two design standards: the Eurocodes and the old Highways Agency (HA) assessment standards which were heavily based on BS 5400. Whilst this is not unmanageable, it is unfortunate because a prime driver for the Eurocodes was to avoid the need to use more than one code of practice. Second, there is the more significant problem that structures designed to the Eurocodes will mostly fail their assessment when the old

HA standards are applied to them because resistances to Eurocodes tend to be higher on average. The avoidance of this problem leads to a potential third. Inevitably, engineers will seek to use large parts of the Eurocodes in their assessment as a departure from assessment standards, thus creating a pseudo assessment version of the Eurocodes, but without the control, expert input, consistency and consensus that would be present if an appropriate committee had assembled the assessment code.

The remainder of this paper investigates areas where the steel and concrete Eurocodes give increased resistances compared to existing codes and makes recommendations for sections which could be directly applied to the assessment of existing structures and the modifications required to the sections which could not be directly applied. Recommendations are also made for the modifications needed to action calculation and the combinations of actions when assessing bridges.

Key features of an assessment version of a Eurocode

The process of assessment is of crucial importance for maintaining bridges in a safe and serviceable condition. The objective of assessment is to evaluate the safety of an existing bridge quickly and with a minimum of effort. However, the assessment rules and criteria need to be established rigorously and judiciously. If assessments are unduly conservative, structures will be unnecessarily strengthened, or needless load restrictions will be imposed. Conversely, if the rules are too lax, the safety of bridges could be compromised.

Eurocodes employ characteristic actions and material properties combined with load and material partial factors to ensure an appropriate level of safety and reliability. These factors guard against extreme variations in design parameters (e.g. material properties and applied loads) which could occur during service. In order to ensure that the design rules are simple for routine use, the values of the partial factors have been chosen such that they cater for a wide range of

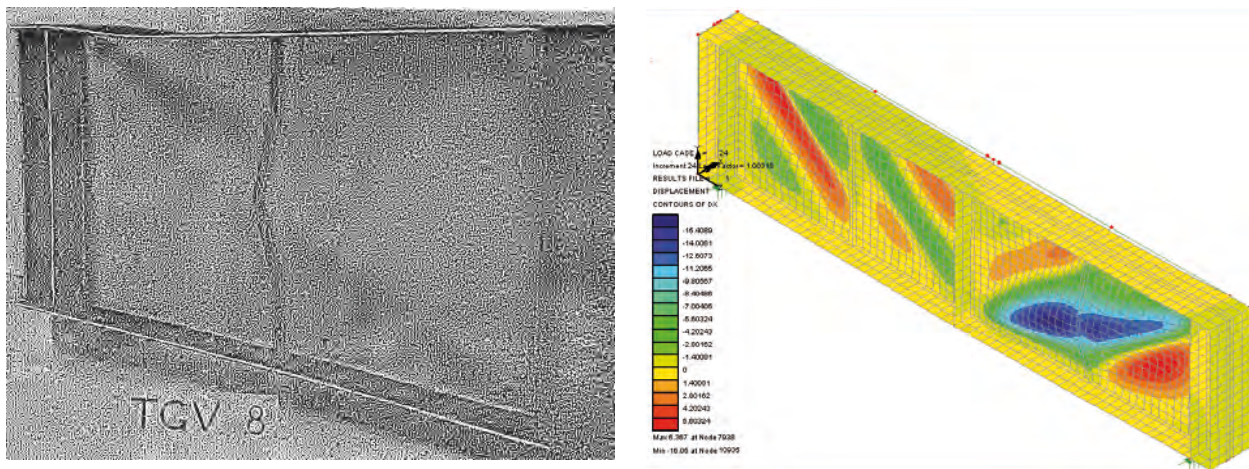


Figure 3 - Physical test and non-linear model of a plate girder

structure/component types and failure modes. It follows, therefore, that the rules tend to be conservative for the majority of bridges and the level of conservatism varies considerably from structure to structure.

The objective of the assessment code would be to produce a more realistic assessment of a structure than is possible at the design stage. This in part is achieved through using current state of the art analysis techniques and design rules based on physical testing where appropriate, and in part by taking advantage of the information available to an assessing engineer in respect of the loading, material strength, geometrical properties and imperfections which necessarily have to be conservatively estimated at the design stage. The assessment code must therefore enable account to be taken of:

- Traffic loading appropriate to vehicle flows over the bridge
- Bridge specific partial factors accounting for consequences of failure
- Measured imperfections and sizes different from those assumed in the code
- Measured material strengths
- Construction not fabricated or erected in accordance with the relevant current execution specifications
- Material not complying with the relevant current product standards
- General configurations and shape limitations not complying with the limits in the design code
- Minimum stiffnesses and strengths not complying

with the design code

- Outmoded forms of construction not covered by the design code
- Deterioration of structural members through condition factors or modification to material factors.

A feature of previous UK assessment codes has been the introduction of modified, and usually significantly more complicated, calculation rules whereby some of the simplifying and conservative assumptions in the equivalent design formulae are removed. The need to do this in a Eurocode assessment standard would often be negated because the rules are usually already more general and computer modelling can often be used directly to obtain the required results.

Overview of modifications to account for bridge specific loading and resistance information

General

In assessing an existing bridge, account may be taken of a number of factors that are specific to a bridge and the assessment criteria modified to reflect these, for example:

- Bridge specific loading
- Bridge specific resistance properties (material and geometrical)
- Consequences of failure.

The work needed to develop assessment provisions for the above are discussed in this section. The remainder of the paper discusses the evaluation of the ultimate strength of bridges based on modifications to Eurocodes 2, 3 and 4.

Bridge specific assessment loading

The procedures for the derivation of Bridge Specific Assessment Live Load (BSALL) for long span bridges with spans greater than 50m are already well established in BD 503, and can be readily extended to Eurocode traffic loading. BSALLs are often vital in justifying continued use of bridges without the need for strengthening to meet full design standard loading see Figure 1. The derivation of BSALL involves collecting traffic data at the bridge site through weigh-in-motion systems and using this data to simulate extreme load distributions. The live load partial factors given in the UK National Annex to Eurocode BS EN1990: Annex A2 can be used unchanged with the BSALL.

For bridges of spans less than 50m, live load reduction factors to the load models LM1 and LM2 given in the UK National Annex to BS EN 1991-24 can be derived for assessment purposes in a similar manner to those given in BD 215 depending on traffic density and road surface roughness for different levels of assessment loading. The surface roughness factors in BD21 may be readily applicable but the live load reduction factors will require some calibration.

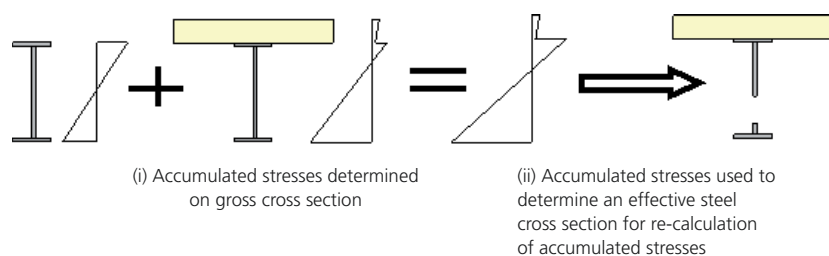


Figure 4 - Illustrative procedure for determining effective cross section in Class 4 composite beams

The load models LM3 (abnormal loads) and LM4 (crowd loading) given in the UK National Annex to BS EN 1991-2 can be used for assessment of existing bridges without further modification, although the derivation of bridge specific crowd loading models is also possible.

Where the structural dimensions and the thickness of road surfacing are measured for an existing bridge, the measured dimensions, together with measured values of material density, can be used to calculate more realistic values of self weight and surfacing load for the bridge. For surfacing, this will allow the tolerances on range of surfacing depth given in the Eurocodes to be reduced, provided controls are in place to limit changes to surfacing depths in the future. Depending on the level of variability in measured values and the controls applied, the partial factors for self weight, super-imposed dead load and surfacing load given in the UK National Annex to BS EN 1990 Annex A2 could be reduced further.

Bridge specific resistance parameters

Where material properties, such as yield strength of steel or cylinder strength of concrete, have been obtained through intrusive in situ testing on an existing bridge, the in situ characteristic strength can be calculated using the Bayesian updating methodology given in

Section D7.2 of BS EN 19906. Further details of this methodology are given in reference 7. The calculation of characteristic strength values takes into account the number of test results, variability in test results and prior knowledge about the statistical distribution of material properties. The in situ characteristic strength values can then be used together with the material partial factors given in the UK National Annexes of material Eurocodes EN 1992, EN 1993 and EN 1994 for bridges. Where the ultimate resistance of a component or product is established through testing, the design resistance can be calculated using the procedures given in Annex D of BS EN 1990.

In some cases, construction tolerances (imperfections) or impact distortions may also affect the structure's strength and will therefore need to be considered explicitly if they are greater than those assumed in the Eurocode design rules Figure 2 shows an extreme case of post-impact deformation. Greater discussion on this is included below.

Consequences of failure

Consequence of failure reflects both the economic impact of the structural failure (in terms of rebuild costs and disruption costs) and also the potential for loss of life. The UK National Annex to BS EN 1990 Annex A2 recommends that all bridges up to 200m span should normally be regarded as

'medium consequence' structures (Consequence Class 2). Designing for a higher Consequence Class may be considered for bridges with span greater than 200m. However, small culverts or bridges with span less than 10m or medium span bridges on a minor road with light traffic may be regarded as 'low consequence' structures (Consequence Class 1). The annual target reliability (as defined in BS EN 1990 Annex B) for a low consequence structure can be taken as 4.2 as opposed to 4.7 used for a medium consequence structure. As a result the partial factors on live load can be reduced by a factor of 0.9 for Consequence Class 1 as given in Table B3 of BS EN 1990, Annex B. In assessment, such a reassessment of the consequences of failure may be appropriate.

Reliability based assessment (Level 5 Methods)

The Eurocodes allow the direct use of probability based structural reliability analysis methods for the design of structures. It is important that their use is permitted for the assessment of existing bridges as they provide significant benefits for assessment. These methods are referred to as Level 5 methods in the UK⁸. In addition to accounting for bridge specific loading, resistance parameters and consequences of failure, the reliability based assessment methods enable account to be taken of a bridge's previous exposure to known abnormally heavy loads (e.g. Special Order loads in excess of 180 tonne gross weight), warning of failure related to structural form and material and quality and frequency of inspections. It is, however, important that the reliability-based assessment procedures are fully standardised, including the probability distributions for the load and resistance random variables and the target values of reliability index. The necessary background research in this regard has already been undertaken during the calibration of load partial factors and load combination factors for use in the UK National Annex to BS EN 1990, Annex A2.

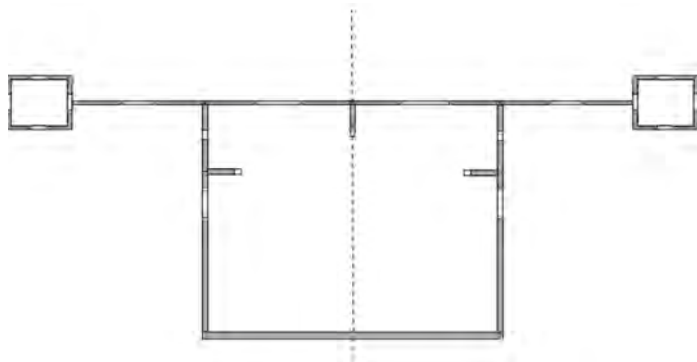


Figure 5 - Typical effective cross section for design to EN 1993-1-5

Overview of the economic benefits of Eurocodes 3 and 4 and the modifications required to produce a steel assessment code

The following discussions are based around Eurocodes EN 1993-29 and EN 1994-210 and the other Eurocodes they cross reference and identify, in outline, the modifications required to produce an assessment standard. The various sections of the codes are reviewed one by one.

General

The scope of EN 1993-2 in its section 1 would need fairly extensive modification because of the number of cross references to modern product and execution standards, but this would mostly involve deletions. The terms, definitions and symbols defined in the section however would mostly be retained. Guidance would have to be provided also on inspection for assessment and the information on material properties, imperfections and condition that should be obtained and recorded. This could largely be taken from existing documentation such as BD 5611.

Basis of design

The material on the principles of limit state design and design assisted by testing would require very little amendment for use in assessment; indeed design assisted by testing would be a useful assessment tool.

Materials

The section on materials would require significant modification. This section of the Eurocode sets requirements for minimum ductility and strength. The relevant parameters would need to come from as-built records and specifications or from intrusive investigation and testing of the structure.

Non-compliance with the prescribed limits on these parameters, however, would mean that some of the remaining Eurocode resistance rules could not be used so the implications of non-compliance on the remaining design rules would need to be addressed.

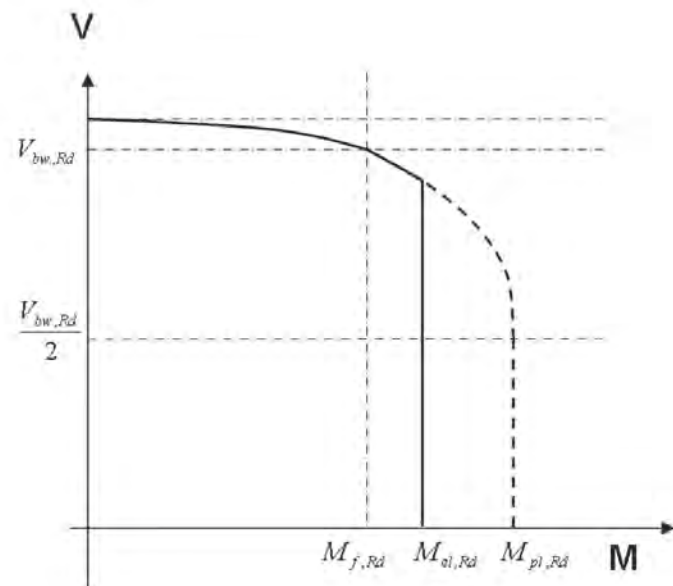


Figure 6 - Typical interaction diagram for shear and moment to EN 1993-1-5

This will require either modifications to the strength equations and procedures or the use of non-linear analysis using known material properties for the structure. Other requirements for minimum toughness to guard against brittle fracture would need less modification, but the Charpy energies for use in calculation would again need to be determined via the means above. The sections on brittle fracture referenced in EN 1993-1-1012 permit the use of fracture mechanics methods to check adequacy, so actually lend themselves to assessment.

The corresponding section in an assessment code could usefully give key material characteristics such as yield strength, ultimate tensile strength, ultimate strain, toughness and modulus of elasticity for historic steel grades. This could be readily provided by extracting the required properties from old superseded steel specifications. An assessment National Annex would be a good place for this information to be recorded.

Durability

Most of the design requirements involving durability would simply be deleted in an assessment version of the standard as inspections would determine to what extent the design had succeeded in that regard. Deterioration would need to be accounted for via a condition factor in the resistance equations in the same way as is currently required in UK assessment standard BD 21.

Structural analysis

Much of what is written on structural analysis would require very little amendment as the principles given are to model the structure in a way that reflects its behaviour; this applies equally to design and assessment. Elastic analysis itself, with subsequent member cross-section and codified buckling checks, needs little adaptation except that facility would be needed to enable the buckling curves to be modified to allow use of measured imperfections.

Non-linear analysis, particularly of slender structures, will often give the most accurate and realistic picture of true structural strength and is thus a very effective assessment tool. The model in Figure 3, for example, predicted the physical test failure load to within 4%. Non-linear analysis methods are, however, influenced by imperfections (residual stresses and tolerances on straightness and flatness) and those given in Eurocode 3 may not be appropriate for an existing structure. Requirements for non-linear analysis would therefore require either magnitudes of equivalent imperfections to be provided for old structures or for imperfections to be measured in the existing structure and residual stresses determined by calculation based on, for example, weld details. Often, doing the latter will in any case be beneficial compared with the alternative of using the design code values. The rules would need to provide guidance on the application of imperfection patterns to the structure. Guidance would also be needed on the material properties to use in analysis where the material behaviour does not conform to the assumptions and limitations in the design code, but non-linear analysis is, in principle, the best form of analysis provided that the material characteristics are known and can be modelled.

Plastic analysis would also be more readily justifiable for use in assessment since in-service performance can be visually assessed. The Eurocodes set out the requirements for plastic analysis so only a relaxation in the conditions of use is likely to be needed in the assessment version, provided that the requirements for ductility are met.

Ultimate limit states

Material factors and cross section resistances

An important feature of the design code is the application of the partial safety factor for material strength, γ_M . Eurocode partial factors could generally be used with the characteristic material properties determined from as built records or from tests. Additionally, EN 1990 sets out means of determining material resistances and factors in its section on design assisted by testing.

This aims to fix material factors such that a base level of reliability is obtained when using the assessment rules, based on the true resistances obtained in tests. The rules in EN 1993 utilise relatively few values of γ_M for simplicity. Background papers 13 are, however, available in a number of cases to support the use of reduced values where adequate reliability would still be obtained and advantage of this could be taken in assessment.

The Eurocode cross section resistances make some allowance for strain hardening in a number of areas (for example in determining the reduction of strength due to holes and in the shear-moment interaction curves) so it would be necessary to set out what to do if the ductility limits were not met. The simplest situation in this case would be to revert back to the more conservative provisions of BD 56, adapted to make them fit the general Eurocode terminology in terms of notation. A similar exercise has already been done in the drafting of reference 14 for salvaging non-contradictory material from BS 5400 Part 3:15 to be used in design with Eurocode 3.

Bending cross section resistance

For Class 1 to 3 cross sections, the calculation procedure and results for bending resistance are very similar to that for previous practice. One small area of economy for Class 3 cross section design is that where the bending resistance is based on first yield at an extreme fibre, EN 1993 defines an extreme fibre as the mid-thickness of the flange, rather than the outer surface as was previous UK practice. This gives a small increase in economy over BD 56, particularly for shallow beams with thick flanges. For steel-concrete composite design, EN 1994 employs a similar rectangular stress block for concrete for plastic design as is used in BD 6116, although the resisting compressive stress is slightly higher. This leads to slightly increased resistances.

The treatment of Class 4 cross sections and beams with longitudinal stiffeners (which are treated as Class 4 cross sections in EN 1993) differs significantly, however, from that in BD 56 and brings significant benefit. Class 4 beams without stiffeners are treated by making reductions to the compression areas and then checking

stresses against yield when calculated on the resulting reduced cross-section. The procedure for composite beams is to first calculate accumulated stresses on the gross cross section, following the construction sequence, and then to determine the effective areas of the compression elements based on this stress distribution. Finally, the accumulated stresses are recalculated using the reduced effective steel cross section at all stages of construction. This is illustrated in Figure 4.

Class 4 beams with longitudinal stiffeners are treated in the same way as beams without longitudinal stiffeners in EN 1993, unlike in BD 56 where a completely different approach to calculation was employed. In BD 56, individual panels and stiffeners are checked for buckling once stresses were determined in them, generally using gross cross sections other than for flange plates where some allowance is made for loss of strength due to local plate buckling. There is therefore limited load shedding between components and a single overstressed component can govern the design of the whole cross section. In EN 1993-1-5, effective widths are again used to allow for buckling of web and flange elements, as for unstiffened Class 4 cross sections, but the same approach is also used for stiffeners see Figure 5. This effectively allows load shedding between all the various elements such that their combined strengths are optimally used. This represents a significant change from previous UK practice and can give rise to a significantly increased prediction of load carrying resistance. These rules could be used with very little modification other than to allow adjustment to the reduction factors for plate buckling to allow for measured imperfections.

Shear buckling resistance and web transverse stiffeners

The rules for shear buckling in EN 1993-1-5:17 and BD 56 are based on quite different theories but produce similar results for shear resistance. Little change would be required for an assessment version of the Eurocodes when the required material properties were met in terms of ductility; the shear resistance of webs is not very sensitive to imperfections, so no adjustment would be necessary.

The implications of the different theories are, however, more significant in the design of transverse stiffeners as Höglund's theory (used in the Eurocode) places less demand on their strength. This is reflected in EN 1993-1-5, which allows lighter transverse shear stiffeners to be designed than would be permitted to BS 5400 Part 3. The rules in EN 1993-1-5 have still been shown to be conservative and could be refined further in the assessment Eurocode, as suggested in Reference 18.

Shear - moment interaction

EN 1993 produces a more economic check of shear and moment interaction than does BD 56 and consequently has been used in assessment to justify not strengthening existing bridges. It is more economic for three reasons:

- Shear does not interact with lateral torsional buckling resistance. It only interacts with cross section resistance
- The interaction diagram is a continuous curve, rather than a series of straight lines as was the case in BD 56, as shown in Figure 6

- Even if the cross section is in Class 3 or 4 (so that the bending resistance is limited to first yield), the interaction is performed using the plastic bending resistance. The interaction is truncated by the requirement to limit the moment to the elastic moment. This has the effect of permitting almost full web shear resistance with full bending resistance, as shown in Figure 6, which reflects the findings of recent non-linear parametric finite element studies.

The rules do rely on ductile behaviour and an amount of strain hardening being available. If relevant limits for these were not met, the simplest resolution in this case would be to revert back to the more conservative provisions of BD 56, adapted to make them fit the general Eurocode terminology.

Longitudinally stiffened cross sections are treated in essentially the same way in EN 1993-1-5 as for unstiffened cross sections, so the same economic benefits can be obtained. To BD 56, the check of the cross section would have to be performed on a panel by panel basis in such a way that any shear stress at all has the effect of reducing bending strength.

One important caveat is that longitudinal stiffeners must not be susceptible to torsional buckling if this method is to be used, otherwise adequate ductility does not exist. If stiffeners are susceptible to this mode of buckling, a method analogous to that in BD 56 would again be required. EN 1993-1-5 Section 10 provides such a method. The rules in EN 1993-1-5 for checking susceptibility to torsional buckling are themselves a little conservative as they ignore restraint from the parent plate to which they are attached. Rules could easily be provided to include this restraint and also measured imperfections and make a less conservative estimate of torsional buckling strength. Modified versions of the rules in BA5619 could be used.

Lateral torsional buckling and distortional buckling

Whilst BD56 gave extensive empirical guidance on lateral torsional buckling, EN 1993 takes a more theoretical approach. EN 1993, as a general approach, gives only an expression for slenderness,

$$\bar{\lambda}_{LT} = \sqrt{\frac{W_{pl,y} f_y}{M_{cr}}}$$

where M_{cr} is the elastic critical buckling moment. No guidance is given on the calculation of M_{cr} , which tends to lead the designer towards performing a computer elastic critical buckling analysis for its determination. The advantage of this in an assessment context, is that it usually produces a less conservative prediction of slenderness and strength reduction than does a simplified hand calculation see Figure 7. The reduction factor slenderness curves themselves would need to be adapted to allow use of measured imperfections in the assessment code, but this would be straightforward to do and could be adapted from provisions in BD 56. For quicker estimates of strength, EN 1994 gives simpler methods of calculation of slenderness without the need for computer analysis.

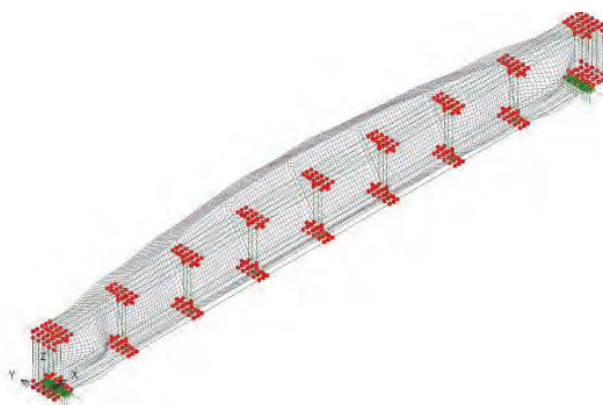


Figure 7 - Elastic critical buckling analysis of historic bridge edge girder

Flexural buckling of struts

Flexural buckling slenderness is similarly written in terms of a critical buckling force N_{cr} so that,

$$\bar{\lambda} = \sqrt{\frac{N_{pl}}{N_{cr}}}$$

Once again, this allows the assessor to use a computer elastic critical buckling analysis which usually produces a less conservative prediction of slenderness and strength reduction than does a simplified hand calculation of slenderness, particularly in cases where otherwise simplified effective lengths would need to be used. A tapered arch with hangers and transverse bracing, see Figure 8, provides a good example of a case where an effective length would be difficult to determine without a computer analysis. The reduction factor-slenderness curves themselves would again need to be adapted to allow use of measured imperfections in the assessment code, but this would be straightforward to do and could be adapted from provisions in BD 56.

Serviceability

Most of the design requirements involving serviceability would be retained in an assessment version of the standard, but inspections would determine to what extent the design had succeeded in that regard.

Fasteners and welds

Most of the provisions for bolts and welds would be able to remain in an assessment standard but greater coverage of the resistance of other connectors, such as rivets, would need to be added from older codes such as BD 56.

Fatigue

The approach to calculation of fatigue life is unlikely to be different between design and assessment, but it may be appropriate to use a different vehicle spectrum for assessment than for design based on measured traffic conditions. EN 1993-1-9 already allows reduced factors of safety to be used where damage tolerant conditions exist (i.e. the structure's details are regularly inspected for fatigue cracks and there is some redundancy available in the event of a component failure) and this approach is particularly suitable for assessment where monitoring of the structure is possible. It would also be possible to incorporate modifications to the S-N curves where fracture toughness properties were measured directly in the structure.

Overview of the economic benefits of Eurocode 2 and the modifications required to produce a concrete assessment code

The following discussions are based around Eurocodes EN 1992-220 (and its cross-references to EN 1992-1-121) and identify, in outline, the modifications required to produce an assessment standard. The various sections of the code are reviewed one by one.

General and basis of design

The modifications required for the general and the basis of design section would be similar to those discussed for steel above.

Materials

The section on materials would require modification for concrete to remove the need for compliance with current execution standards and reference would need to be

added to the determination of characteristic material properties, as discussed previously, where no as-built records are available. The section on creep and shrinkage would also require some modification as usually the detailed concrete composition information required for these calculations will not be available for older structures. In such cases it would be possible to give recommended values based on common concrete mixes that were used at the time of construction.

For reinforcement, requirements for compliance with current execution standards will need to be removed and the properties assumed for validity of the design rules, such as ductility, clearly stated. Where reinforcing steels do not comply with EN 1992, alternative rules such as those found in BD 442² may need to be used, although generally it should be possible to modify the EN 1992 rules to suit. This might include removing consideration of strain hardening in the flexural resistance check where the existing steel is brittle. Guidance will also need to be included for the treatment of plain round bars which are not covered by EN 1992 because of their very poor bond characteristics. The properties of commonly used historic reinforcing steels could usefully be given in this section. An assessment National Annex would be a good place for this information to be recorded.

Durability and cover

As for structural steel, most of the design requirements involving durability would simply be deleted in an assessment version of the standard as inspections would determine to what extent the design had succeeded in that regard. The implications of low concrete cover would, however, need to be considered at the ultimate limit state where this adversely affected anchorage and lap lengths; the EN 1992 rules assume a minimum cover equal to the bar diameter in this regard. Deterioration would need to be accounted for via a condition factor in the resistance equations in the same way as is currently required in UK assessment standard BD 21.



Figure 8 - Typical tapered arch



Figure 9 - Non-linear analysis of slender piers (deflected shapes shown for various load cases)

Structural analysis

As for structural steel, the structural analysis section would require little modification as many of the assumptions would remain valid for assessment. For non-linear analysis however, measured imperfections would need to be used if the actual construction tolerances used were not known. Non-linear analysis is a particularly appropriate assessment tool because it produces the most realistic prediction of structural behaviour and resistance. However, the practical application of the technique is more limited for concrete structures than for steel ones because of the added complexity. For members where the predicted failure is a flexural one, such as for beams and slender piers, see Figure 9, the analysis is relatively straightforward. For more complicated cases where concrete failure under a multi-axial stress field is concerned, the analysis can be much more time consuming, iterative and difficult to agree with checkers because of the number of concrete models that exist. Guidance on the latter could usefully be taken from *fib* bulletin 45²³ among other sources.

Plastic analysis would also be more readily justifiable for use in assessment since in-service performance can be visually assessed. The Eurocodes set out the requirements for plastic analysis so only a relaxation in the conditions of use is likely to be needed in the assessment version, provided that requirements for rotation capacity are met. To that extent, it would be necessary to give guidance on rotation capacity for reinforcing steels not covered by EN 1992.

Ultimate limit states

Bending

The rules for flexural resistance would require little modification but reinforcing steels would need to be classified for ductility; there is a small amount of benefit that could be gained from these rules because of the consideration of reinforcement strength beyond yield, though this would depend on the ductility of the reinforcement and the strain hardening characteristics. Moment redistribution would also provide benefit, as allowed by EN 1992, though there was previously allowance in BD 44 for this also.

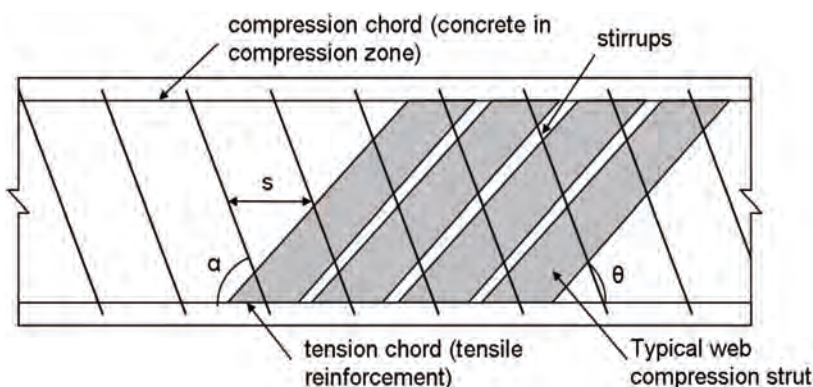


Figure 10 - Variable angle shear truss model from Eurocode 2

Shear

The variable angle shear truss model provided in EN 1992, see Figure 10, provides significant benefit for reinforced concrete sections, and using this model for assessment will provide an important tool. The resistance model is essentially a plastic solution that allows the resistance to be tailored to the reinforcement provided. Where concrete shear stresses are not high, greater link resistances can be developed by using a flat truss angle. Greater shear stresses can also be allowed compared to BD 44, which will often give improved assessed load rating for beams with slender webs.

Torsion

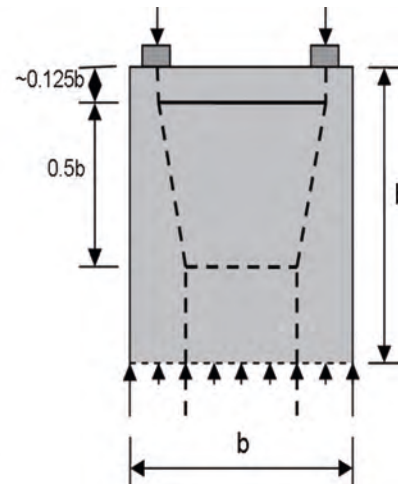
As for shear, the rules for torsion provided in Eurocode 2 would require very little modifications and could be directly adopted for assessment. Once again, the variable angle truss model gives the designer an important tool to tailor the resistance model to the reinforcement provided both longitudinally and transversely.

Strut and Tie

The strut and tie rules could also be directly adopted for assessment, see Figure 11. Since strut and tie analysis is a special case of plastic analysis, once again the rules allow models to be developed which suit the actual arrangement of reinforcement provided. Some additions to the Eurocode for assessment would, however, be useful. The stress limit in struts where there is transverse tension is currently rather conservative in cases where the tensile strains are small and the stress limit could be refined to include the effect of tension angle in relation to the strut axis and the tensile strain. Further, more guidance should be included on by how much the idealised model can depart from the elastic flow of forces because concrete has limited ductility and excessive departures from such conditions will reduce the allowable strut and node compression limits. Guidance on quantifying this reduction could also be usefully added.



Figure 11 - Simple application of strut and tie rules for bridge pier



Membrane rules and sandwich models

The membrane rules in EN 1992-2 Clause 6.109 would be useful for assessment where shell elements are required in analysis due, perhaps, to the complexity of structure geometry and the lack of applicability of other codified assessment rules. These rules allow stresses to be used directly from an FE analysis model for the assessment calculations. However, the significant limitation of the approach is that design is done element by element and hence no redistribution between elements is possible. This makes the use of the membrane rules conservative compared to other member design rules (such as those for shear and torsion) and these member rules should therefore always be used where they are applicable. Within an element, however, there is some scope to tailor the analysis to suit the reinforcement directions.

Serviceability

Most of the design requirements involving serviceability could be retained in an assessment version of the standard, but inspections are generally more useful in determining serviceability performance. This again would be consistent with previous UK practice.

Fatigue

Checks of fatigue in concrete and steel were not covered to any significant degree by previous UK assessment standards, but fatigue can be important in some structures where the details are unusual. Details that may need assessment include bars connected with threaded couplers and bent bars used at locations where the bars are highly stressed. Both these situations can lead to fatigue problems^{24,25}. EN 1992 provides a realistic assessment of fatigue in both these situations and would require minimal modification. The coupler failure in Figure 12 was predicted by the rules in Eurocode 2 for example.



Figure 12 - Broken reinforcement coupler due to fatigue

Detailing Rules

Section 8 of EN 1992, dealing with detailing, would have to be modified substantially to deal with non-compliant details. Detailing rules tend to either be provided to ensure adequate workmanship or to guarantee compliance with assumptions in the design rules.

Detailing rules relating to workmanship could simply be deleted as poor workmanship could be accounted for via a condition factor. This would include, for example, minimum bar spacing requirements; non-compliance will potentially lead to poorly compacted concrete but this should be possible to identify from examining the surface or taking a core sample from the concrete.

Detailing rules which affect the applicability of design rules will need to be maintained and guidance provided on how to handle non-compliant details. For example, minimum cover rules are required to ensure that the rules for bond strength can be used. Where the minimum cover requirements are not met, additional rules will need to be provided to determine the effects on the bond strength.

Rules on minimum bend radii for bars are also given in the design code to prevent crushing of the concrete inside the bends and to limit bending cracks in the reinforcement to acceptable levels. Where these minimum radii are not complied with, explicit checks on concrete stress can still be made inside the bends but additional information may need to be provided for fatigue strength calculation of such bars.

Conclusion

The new Structural Eurocodes lend themselves to adaptation for assessment of structures because the rules are generally based around the application of first principles. This not only means that the rules can be adapted relatively easily but they also provide scope for determining an accurate prediction of the true ultimate strength of a structure based on structure specific information on loads and material properties in a rigorous manner. This paper has identified areas where the Eurocodes could be directly applied to the assessment of existing structures and those which would need modification. It has also identified areas where the Eurocodes give increased resistances compared to existing codes.

The amount of work to produce an assessment version of the Eurocodes would not seem to be prohibitive, even given a modest budget, if the scope was limited to adapting the existing rules to assessment without refining them further for economy. This should be the priority for the work because the Eurocodes themselves provide a better estimate of strength than many of the UK assessment standards and the lack of a Eurocode assessment standard is likely to lead to inappropriate application of the design Eurocodes to assessment. The secondary task of updating the rules and the partial factors would be considerably more time consuming and would mostly be equally applicable to design. It is therefore unlikely that this refinement would be undertaken solely for an assessment code.

References

1. BS EN 10025 (2004): Hot rolled products of structural steels. British Standards Institution, London.
2. BS EN 1090-2 (2008): Execution of steel structures and aluminium structures - Part 2: Technical requirements for steel structures. British Standards Institution, London.
3. BD 50/92: Technical Requirements for the Assessment and Strengthening Programme for Highway Structures Stage 3 - Long Span Bridges, Design Manual for Roads and Bridges, Highways Agency (1992)
4. BS EN 1991-2 (2003): Eurocode 1: Actions on structures - Part 2: Traffic loads on bridges. British Standards Institution, London.
5. BD 21/01: The Assessment of Highway Bridges and Structures, Highways Agency (2001)
6. BS EN 1990 (2002): Eurocode - Basis of structural design. British Standards Institution, London.
7. Shetty, N. and Chubb, M. (2001), Probabilistic Methods for Improved Bridge Assessment, International Symposium on Bridge Management, Singapore.
8. Shetty, N. Chubb, M.S. and Manzocchi, G.M.E. (1998), Advanced methods for the assessment of bridges, Int. Symposium on Management of Bridges, Institution of Civil Engineers, London.
9. BS EN 1993-2 (2006): Design of Steel Structures. Part 2: Steel bridges. British Standards Institution, London.
10. BS EN 1994-2 (2005): Design of Composite Steel and Concrete Structures. Part 2: General rules and rules for bridges. British Standards Institution, London.
11. BD 56/96: The assessment of steel highway bridges and structures, Highways Agency (1996)
12. BS EN 1993-1-10 (2005): Design of Steel Structures. Part 1.10: Material toughness and through-thickness properties. British Standards Institution, London.
13. B. Johansson, R. Maquoi, G. Sedlacek, C. Müller, D. Beg (2007), Commentary and worked examples to EN 1993-1-5 "Plated structural elements", JRC Scientific and Technical Reports, Italy
14. PD 6695-2:2008 Recommendations for the design of structures to BS EN 1993-2:2006, BSI (2008)
15. BS 5400:Part 3 (2000): Design of steel bridges. British Standards Institution, London.
16. BD 61/96: The Assessment of Composite Highway Bridges and Structures, Highways Agency (1996)
17. BS EN 1993-1-5 (2006): Design of Steel Structures. Part 1.5: Plated structural elements. British Standards Institution, London.
18. Presta, Hendy et al Numerical validation of simplified theories for design rules of transversely stiffened plate girders, The Structural Engineer, Volume 86, Number 21 pp 37 - 46 (4/11/2008)
19. BA 56/96: The assessment of steel highway bridges and structures, Highways Agency (1996)
20. BS EN 1992-2 (2005): Design of Concrete Structures. Part 2: Concrete Bridges. British Standards Institution, London.
21. BS EN 1992-1-1 (2004): Design of Concrete Structures. Part 1.1: General rules and rules for buildings. British Standards Institution, London.
22. BD 44/95: The assessment of concrete highway bridges and structures, Highways Agency (1995)
23. fib bulletin 45 (2009), Practitioners' guide to finite element modelling of reinforced concrete structures, fib, Lausanne, Switzerland
24. Hendy, C.R and Smith, D.A, Design of the Dubai Metro Light Rail Viaducts - Substructure, ICE Bridge Engineering, Volume 162, June 2009
25. Sandberg, J.M and Hendy, C.R, Replacement of the stays of a major cable stayed bridge, to be published ICE Bridge Engineering

Limit equilibrium assessment of drystone retaining structures



Chris Mundell

Graduate Engineer
Highways & Transportation



Paul McCombie

Senior Lecturer
University of Bath



Claire Bailey

Project Engineer
Clarke Bond



Andrew Heath

Senior Lecturer
University of Bath



Peter Walker

Senior Lecturer
University of Bath

Abstract

A limit equilibrium analysis program has been developed as part of an investigation into the stability of drystone retaining structures. Initial verification of the programme's function was in relation to field trials conducted in 1834 by Lieut-General Burgoyne, which have been the main reference to date for checking numerical modelling of drystone retaining walls. Parametric studies and investigations of bulging mechanisms are reported and analysed. Program predictions have been compared with the initial results from new small scale and full scale drystone retaining wall tests.

Introduction

Drystone technology is an ancient form of construction, suitable for applications ranging from simple field walls to large earth retaining structures several metres high. Typically utilising undressed stone and constructed without mortar, structural integrity is maintained through self weight, inter-block friction and overlapping of stones. The technique relies upon the skill of the mason in selecting a suitable stone for each location in turn, placing each appropriately.

There are estimated to be some 9000km of drystone retaining structures lining the road and rail networks¹ of the UK, whilst globally the total length is many times this figure, with walls found throughout Europe, parts of Asia, Africa and the Americas.

Most construction in the UK dates to the 19th and 20th centuries. Though poorly constructed walls presumably collapsed shortly after their construction, the majority of walls have remained perfectly stable over decades of usually steadily increasing loading and weathering of the constituent stone.

However, many otherwise stable walls have deformed or bulged and with little guidance currently available to assist engineers in assessment of these structures², the authorities responsible for any potentially unstable walls are often forced to replace or rebuild them - in many cases unnecessarily and at great cost. The total replacement cost for the walls lining the UK's highways is estimated to be in excess of £10 billion³.

Such figures highlight the need for the means to assess these structures adequately, as current design standards often deem them insufficiently safe⁴. There are several difficulties when attempting to assess drystone structures; often, very little is known about any particular wall, as their construction usually pre-dates the strict design guides which are adhered to today, leaving uncertainties regarding wall thickness, age, construction quality, foundation capacity and the mechanical properties of the material retained. There are also regional differences as many styles of wall construction exist, either necessitated by the material properties of the stone or for purely aesthetic purposes.

There is in any case an important philosophical difference between assessing an existing structure and designing a new one, in that many of the uncertainties that are to be covered by factors of safety in design have been resolved by the fact that the wall has been standing and serviceable. While this fact gives no assurance that the structure has ever experienced a full design applied loading, there remains the important fact that the assessor is most concerned by possible changes from the status quo, in which the factor of safety must at least exceed one under permanent loading conditions. Inappropriate interventions such as pointing become of greatest concern, because while they may increase the compressive strength and stability at the face of the structure, they can lead to catastrophic changes in the pore water pressure regime. It is therefore very desirable to be able to assess the possible impact of changes in geometry and loading on the structural stability of an existing wall, especially given that old structures often appear to have departed from their originally constructed geometry.

It is also important to understand the extent to which structural stability is dependent upon precision in geometry and quality of construction. Standards for modern drystone retaining wall construction are very high, with good practice resulting in very strong structures with a high degree of integrity. However, such construction is time consuming and expensive. A proper understanding of drystone retaining wall stability could lead to narrower structures requiring less volume of carefully placed stone and significantly less construction time. Similarly, an understanding of their tolerance of deformation and of the actual sensitivity to variability within the construction could lead to faster construction. These factors would make it easier to repair drystone structures rather than replace them, and easier to replace with a new drystone structure which will be sustainable, re-using materials where possible, or using locally sourced materials, and resulting in structures that are in keeping with their surroundings. It would also make this highly sustainable form of construction a more attractive proposition for new constructions.

Objectives and scope

As part of an ongoing investigation, several unmortared retaining structures of both large and small scale have been built and tested to failure. These experiments are carefully monitored and both the stress changes and deformations at critical locations are recorded, and then used to determine the underlying mechanisms behind the failures. In parallel with these studies, the authors are developing alternative means of analysing drystone structures which are then verified with the gathered physical test data.

Current analysis techniques for drystone walls are either simplistic by considering the static equilibrium of the wall as a monolithic structure, or too complicated, using time consuming numerical packages to model each element within the wall and backfill. Numerical packages such as UDEC (Universal Distinct

Element Code) may provide precise details regarding wall stability and the potential failure mechanisms given sufficient data and careful modelling, but the analysis can take several hours, making parametric studies of any particular structure a lengthy and expensive process.

Neither option is acceptable for routine use. A computer program has therefore been developed to explore efficient approaches to analysis and design which might be carried out by hand calculation or by a range of simple computing approaches. The program is based on a rapid 2-dimensional limit equilibrium (LE) appraisal for structures of any size, with the ability to account for any deformations or bulges that might occur. In addition this program is being utilised to further understand the mechanisms behind the deformations within drystone walls, as well as the critical factors that affect this particular construction.

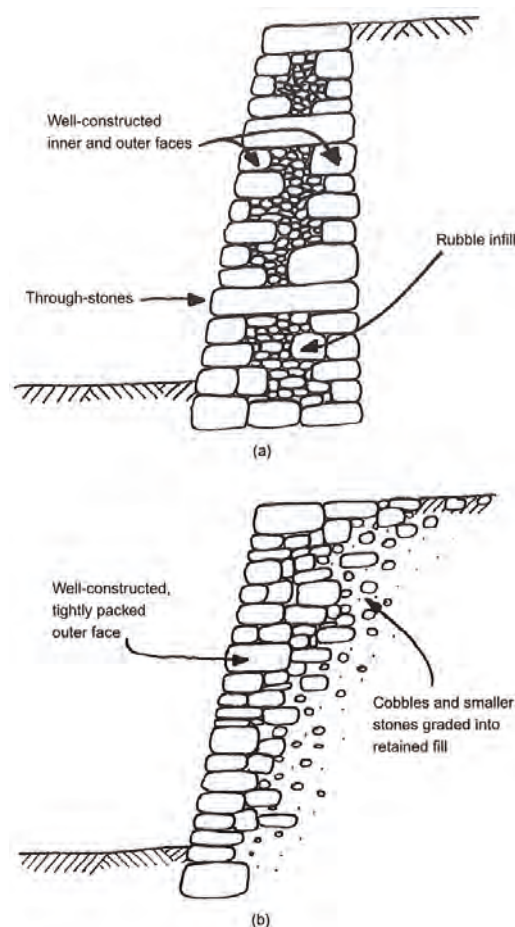


Figure 1 - Comparison between single and double faced walls.

Drystone construction

Although many differences exist between the various drystone construction styles, several common features are usually exhibited. Typical drystone walls are built in horizontal layers or 'courses', with each course ideally consisting of stones of a 'uniform' thickness, retaining a straight and level appearance. The cross section of the wall usually consists of a well made, tightly-packed outer face, with a core of smaller random material packed behind. Some drystone retaining walls follow this core material directly with the retained backfill material, whilst others have a second inner face, usually less well finished than the outer face. 'Tie-stones' span from the outer to the inner face, binding the wall together. Where there is no inner face, tie-stones are often used to anchor the outer face further back into the packing fill. Coping stones can act in a similar manner, spanning the entire width of the wall at the crest, see Figure 1.

Each block within the wall should ideally be in contact with several other stones, and pressure upon any part of a freshly placed stone should not cause any rocking or lifting at the opposite corner. In practice it is usually necessary to wedge in small pieces of rock known as pins to prevent rocking. The unavoidable presence of these pins presents a weakness for all drystone structures, especially as weathering of these smaller elements has a substantial effect much more quickly than for the larger stones. Pins are often used to allow a more even appearance to the face and assist in drainage by tilting stones so that their outer surface is in the plane of the face. Thus the face of a structure can often give a misleading impression of a very tight well-ordered construction, whilst behind the face there are substantial voids held open by a large number of small pins.

Depending on the quality of workmanship and the material used, the density of the walls can differ vastly. Void percentages within drystone walls have been generally considered to range between 10% and 20%⁵, however, measurements carried out within this project on a range of test sections showed these values to be too low. Sections of wall

were built to various standards by the project masons within timber boxes. As the density of the wall material is known, together with the overall volume of the sample, a stone/void ratio can be easily determined. A very carefully constructed tightly packed double faced wall with almost ideal Cotswold limestone has around 20% voidage, while over 40% is possible within poorly built walls. The consequences of a high void ratio are more extensive than just reduced weight of the structure: a loosely packed wall gives the blocks within it a much greater opportunity to rotate and slide, facilitating bulging and other deformation or even collapse.

Previous work

To date, despite its widespread use, only a limited number of investigations into drystone behaviour have been conducted. Until recently, the only physical test data for full scale drystone walls dated back over 170 years to work conducted by Lieut-General Burgoyne in 1834⁶. Burgoyne built four full scale granite walls, up to 6.1m tall, 6.1m long and of varying thickness, in an attempt to quantify the effect that the wall profile has upon stability. These walls were then gradually backfilled until either full retention was achieved or collapse occurred. Movements and general observations were recorded upon the placement of each layer of fill but only reported posthumously in 1853 from Burgoyne's notes.

Based on these field tests, several numerical studies have recently been conducted. UDEC has been used by various authors^{1,5,7-10} to test both the validity of various modes of analysis and to study further the various parameters at work within drystone structures. Although highly informative, these investigations are both complex and time consuming, often requiring several hours to run a single cycle of analysis. Work is currently being carried out in conjunction with this project to develop three dimensional models of the full scale tests described in this paper.

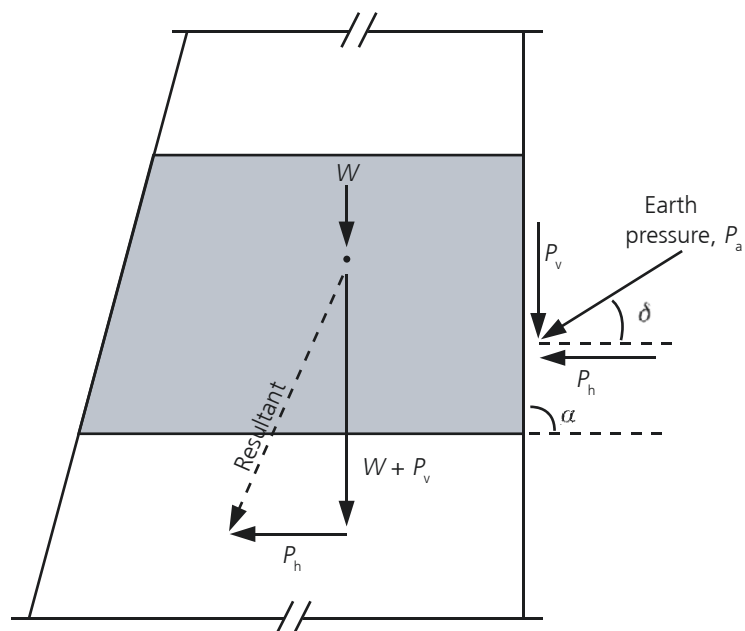


Figure 2 - Determination of resultant forces

Program operation

By analysing the stabilising forces within the wall and using Coulomb's earth pressure coefficients to determine the horizontal and vertical stresses acting at each level up the back of the wall¹¹, the magnitude and direction of the overall thrusts are determined, see Figure 2. The initial wall geometry is entered along with the material properties of both the wall and the backfill (mass, friction angles, etc) and the eccentricity (ϵ) is calculated at a number of levels to generate a thrust line as shown in Figure 3.

In addition to forces arising from the self-weight of the backfill, patch surcharging may also be applied. Additional pressure is then applied to the backfill, spreading out by a ratio of 1h:2v. This is clearly a simplification, as used for example in BS8006¹², compared with the more rigorous approach of Bolton¹³ as suggested in BS8002⁴, but for the present purposes this approximation allows the combination of rapid calculation and reasonable accuracy required here. Further justification of this approach was given by Corte¹⁴. It is currently assumed that the surcharge will have no effect upon the calculated thrust line until the expanding area over which it is distributed crosses the boundary of the wall. Although the analysis is two-dimensional, three-dimensional load dissipation can be assessed as spreading in both horizontal dimensions. The most problematic loading is wheel loading from a heavy vehicle, so it is important to model the three-dimensional distribution, even if crudely. A more sophisticated stress distribution calculation is simple to implement, but given the uncertainties in wall and backfill stiffness and anisotropy, this may overcomplicate the analysis without adding value.

Standard masonry construction recommends that for stable construction the eccentricity of the thrust line must remain within the middle third of the structure ($\epsilon = 1/6$ of the base width away from the neutral axis). If the masonry were to behave in a linear elastic manner and deformations were very small, this would result in no tension being taken at the back of the structure. However, even if the tensile strength

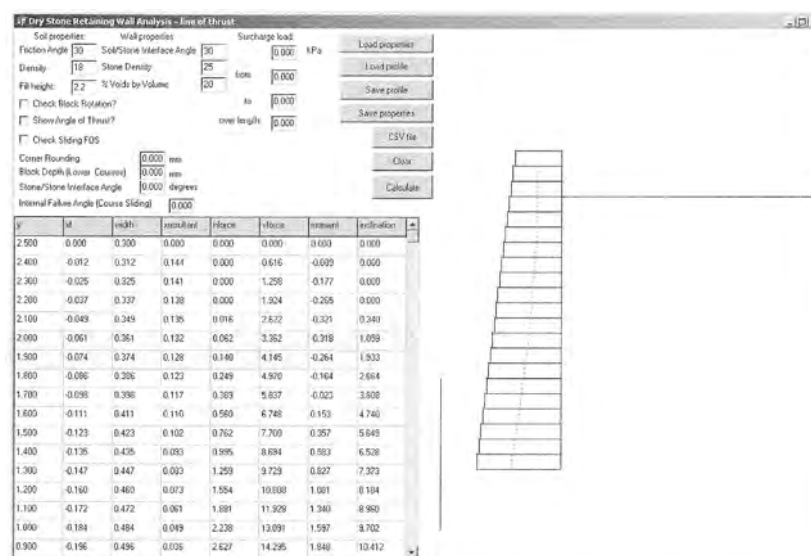


Figure 3 - Program operation screen

of the masonry were zero, a thrust line in front of the middle third would simply result in the stone at the back of the structure being progressively unloaded, which need not have any immediate serious consequences. As the thrust line moves further forwards, the area carrying the vertical load reduces, so increasing the stress. The compressive strength of most masonry, including drystone, is usually relatively high compared with the stresses acting. Therefore, compression failure of the main stones is very unlikely, but a concentrated thrust may cause localised crushing of weakened pins or a flexural fracture of some stones leading to further deformation. In addition, foundation settlement might give rise to significant deformations. Given sufficiently strong masonry and foundation, failure would only occur once this thrust line breaches the wall face ($\epsilon > 1/2$). However, individual block rotation will occur before this,⁵ as the block at the face must carry all of the lateral thrust as well as a vertical load which is moving closer and closer to its leading edge. This could result in an immediate rotational failure of the entire structure, and indeed such a mechanism accounts for the heights reached by the Burgoyne walls which failed⁵. It may also be noted that a crushing failure at a point of contact or of a pin may lead to collapse before the line of thrust reaches the front face.

A significant advantage of the program is the ability to allow the user to deform the wall model and induce

bulges. Once the parameters are set and the wall profile is generated, the user may move the individual blocks to any position, either by typing in a new co-ordinate, or by clicking the cursor at a new position for the front face. Recalculation and redisplay of the new thrust line is virtually instantaneous, see Figure 4. In such a manner, idealised wall structures can be deformed to represent commonly observed drystone bulging patterns, to assess their effect upon overall stability. Conversely, existing walls which have bulged may be quickly recreated using the program to ascertain their stability.

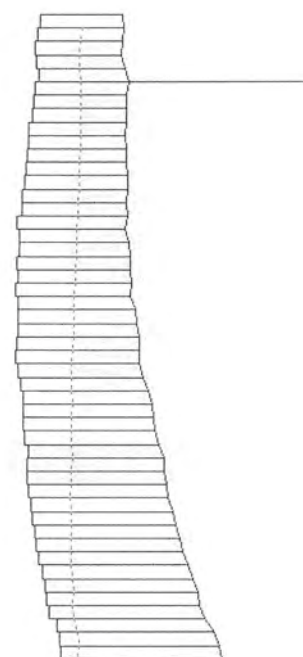


Figure 4 - Example of user defined bulging

Program validity check

Initially, the program was validated against Burgoyne's four test walls. The geometries of each wall were recreated, and the material properties entered from Burgoyne's tests⁶. Backfill heights were then systematically increased by 300mm (simulating Burgoyne's test procedure) until the thrust line reached the wall face, indicating failure via toppling. The final simulated heights were very similar to those recorded by Burgoyne, and indeed also similar to previous attempts using other more sophisticated and complex numerical packages⁸, see Table 1.

Both the first and second of Burgoyne's test walls were backfilled to their full height without excessive movement, and by using the LE program it can be demonstrated that the thrust line lies within the boundaries of the wall. It should be noted that for both these walls the eccentricity is outside of the middle third at the base, indicating uplift at the heel. The third and fourth walls both fell before full height of retention was achieved. For both these wall geometries, the LE program predicted failure at a height similar to that found by Burgoyne, although it has been demonstrated that consideration of individual block rotation gave a tighter correlation with actual failure heights⁵. To allow this to be seen in the program, the direction of the resultant force at each level is also shown at the point at which it acts. A resultant which points in front of the toe of an individual block may result in rotation of that block. With regard to the results shown in table 1, this would indicate a failure at 5.2m for Wall C, bringing it in line with the observed results.

Parametric analysis

Due to the nature of the program, a parametric analysis of any structure is a rapid process. This has a two-fold application; firstly it allows users to quickly grasp which parameters have the greatest impact upon wall stability and secondly it allows engineers in the field a greater flexibility when assessing existing walls. Once the cross section of a wall has been recreated within the program, each variable may be altered to examine the safety factors at the worst possible conditions.

From the program it is apparent that for any given geometry of wall, several parameters are dominant for stability. For example, the assumption of a 1h:2v load spread from surcharges means that the loads must be close to a wall to have an effect, but it is also found that loads must be relatively large, corresponding to wheel loads from the heaviest trucks. This corresponds with anecdotal evidence, confirmed by numerical modelling studies⁹, that there is a relationship between increasingly heavy traffic and failures of walls which had been safe for many years.

The friction angle of the backfill material is critical to wall stability. This angle determines the coefficient of active pressure (k_a), which in turn determines the magnitude of the horizontal forces upon the retaining wall. A stiff, tightly-packed backfill material might have a high density, but its consequently high angle of friction is likely to result in a lower horizontal pressure than a much looser, yet less dense material.

There is also the friction between the wall and backfill to consider. Due to the nature of drystone construction, the wall faces are generally rough, which allows the inner wall face adjacent to the retained material to attract some of the vertical load from

the backfill. As this vertical force acts against the overturning forces and stabilises the wall, this is a value which would ideally be as high as the interface allows, although in reality it is not always guaranteed that the full friction angle will be achieved.

One of the most variable and difficult to ascertain parameters is the density of the walls themselves. Non-destructive testing methods, such as ground penetrating radar, or horizontal coring can be used to give some indication of wall depth, profile and even voidage. Whilst the density of the rock will not vary greatly, its age and the construction style, and the skill of the mason will all affect a wall's overall density and hence the total volume of voids within. Whilst this voidage has little impact on wall stability when changed by a few percent, the value may vary by much more than this, as mentioned above. Low density reduces the wall's stability in terms of both sliding and overturning. Perhaps most critically, a reduction in density allows easier movement and rotation of the individual blocks, determining the flexibility of the wall and the amount of bulging that may occur.

Bulging investigation

Bulging is common in drystone walls, usually occurring at roughly a third to half the height of the wall, creating a distinctive 'belly bulge' shape, see Figure 5. Upon investigation of the effects of bulging and wall deformation, it was discovered that far from causing instability, a moderate bulge may indeed increase the safety of a wall against certain failure modes when subject to surcharge loading conditions.

Wall geometry	In situ observations	UDEC analysis	Limit equilibrium analysis	
	Maximum fill height: m	Maximum fill height: m	Maximum fill height: m	Eccentricity at base of wall: mm
Wall A	Full height	Full height	Full height	102 from toe
Wall B	Full height	Full height	Full height	156 from toe
Wall C	5.2	5.2	5.5	N/A
Wall D	5.2	5.2	5.2	N/A

Table 1 - Comparison of Burgoyne tests

Bulging begins when the loads behind the wall cause blocks or entire sections of wall to move, and the resulting movement causes both the forces acting on the wall and the equilibrium of its own mass to change, such that a new equilibrium position is found. Were this not the case, the wall would continue to move resulting in collapse. This rearrangement usually occurs lower down the wall, and can be due to slips in the retained earth, increased pore water pressure or an increase in loading conditions, or the equilibration of negative pore pressures within the backfill. Bulging probably occurs much more commonly than is appreciated, but is usually on a scale too small to be noticed.

Bulging and movement can also occur much higher up a wall – usually caused by localised surcharging, or disturbances to the wall itself, such as growth of vegetation, although this is generally detrimental to wall stability and can easily lead to partial or full wall collapse.

Once a bulge is formed, the pressures acting upon the wall must change in response to the new geometry. A section of a typically bulged wall is shown in Figure 6, highlighting the common features. Above the bulge, the wall is leaning back somewhat, having a two-fold effect. Firstly, it stabilises the wall by moving its centre of gravity away from the toe of the wall, which is usually the overturning point. Secondly, it reduces the magnitude of the forces applied to the wall by the backfill.

Below the bulge, the wall is leaning forwards, causing the active pressures within the backfill to have a much greater effect upon this portion of the wall. The magnitude of the force will be greater, but the downwards component will be most increased, so increasing the stability of this portion of the wall, provided that the face has not moved so far forwards that individual blocks are no longer supported. Overall, these changes tend to be in favour of increasing wall stability, and walls have commonly remained perfectly safe for years whilst displaying this type of bulge without any detrimental effects. However, new works, such as excavations for services at the toe of deformed walls, or changes in loading, are common factors attributed to triggering collapse.

Due to the flexible nature of these walls, significant movements may take place before a failure occurs, giving visible warning signs. Final collapse can occur either by toppling or bursting, but is usually a combination of both.



Figure 5 - Deformed wall

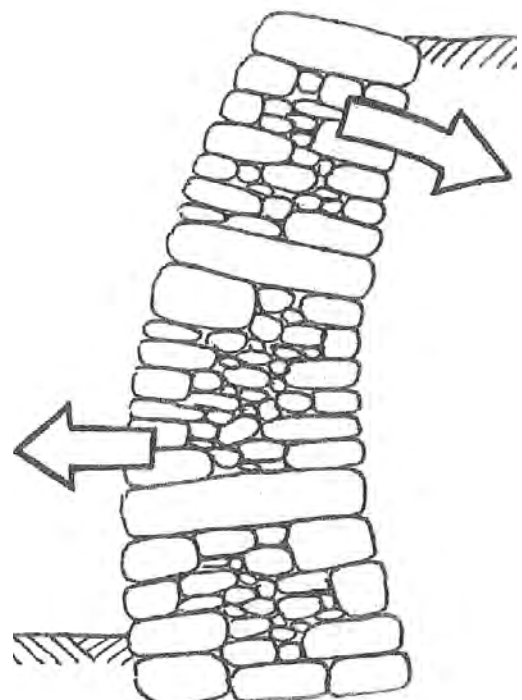


Figure 6 - Section of typically bulged wall

Small scale testing

A series of small scale tests has been conducted to determine if observed drystone behaviour can be recreated in smaller, simpler experiments. To house the tests, a steel box was constructed, with the capacity to hold scale walls 500mm in height and 500mm in width, see Figure 7. The box was lined with a double layer of plastic sheeting, to help reduce friction at the edges and hence minimise end effects¹⁶.

As the goal of these experiments was to reproduce full scale drystone behaviour, small pellets (2-3mm diameter) of lead shot were used as backfill to induce sufficient pressures to cause deformations and failures. The lead shot used has an uncompacted unit weight of 50kN/m³ and an internal friction angle of 31°, allowing the generation of sufficient lateral pressures to overcome the stabilising forces within the test walls.

To overcome three-dimensional effects long blocks were used, each spanning almost the whole width of the steel box but with gaps at either end to allow small rotations of the wall elements. Both timber and concrete block walls were tested independently; the timber blocks were quickly discarded as their densities proved too low for realistic modelling of drystone behaviour (5.5kN/m³ as opposed to 24kN/m³ for the concrete blocks) although the data proves useful for comparison with the LE program results.

For each test, the scale walls were fully constructed without any retained backfill, and then slowly backfilled. Results from the small scale tests are shown in Table 2 together with the backfill heights predicted by the LE program.

From table 2 it is clear that the program is accurately predicting

the collapse heights of these small scale tests. It was assumed that the interface friction between the wall blocks and the backfill was $\frac{2}{3}$ the full value of the backfill's internal friction angle. Evidence gathered by the small scale tests supports this assumption, although in practice it is difficult to ascertain precisely how much of the backfill's full friction angle has been mobilised against the wall. This obviously has a large impact upon wall stability, although it is expected that ground settlement over time and the rough nature of drystone structures results in the full friction angle being mobilised for in-situ walls.

Large scale testing

In addition to small scale testing, full scale walls have been built and tested to failure to validate modern theories and analysis tools. To test the drystone walls, a bespoke test facility has been constructed, allowing recreation of localised or general foundation settlement, backfill settlement and localised surcharging, see Figure 8.

Each wall was constructed of Cotswold limestone by skilled masons. At 2.5m high and over 12m long, the test walls are large enough to be representative of many of the walls found throughout the country, and are built using traditional methods, including regular placement of through-stones and a line of coping stones at the peak of each wall. The first wall varied in thickness from 600mm at its base to 400mm at its peak, and was constructed to a high standard. The second wall was of an intentionally poorer quality, and 100mm more slender throughout.

A large range of instrumentation is used to monitor each wall, including extensive surveying, multiple transducers, load and pressure cells, high resolution imagery and video footage. With this vast range of data, the critical events that lead to the failure of each wall can be better understood and incorporated into both our general understanding and the theories and programs which are used to evaluate a wall's stability.

For each test the LE program was used to check initial stability. Ideally, each wall should be able to retain the full height of fill intended



Figure 7 - Test Setup

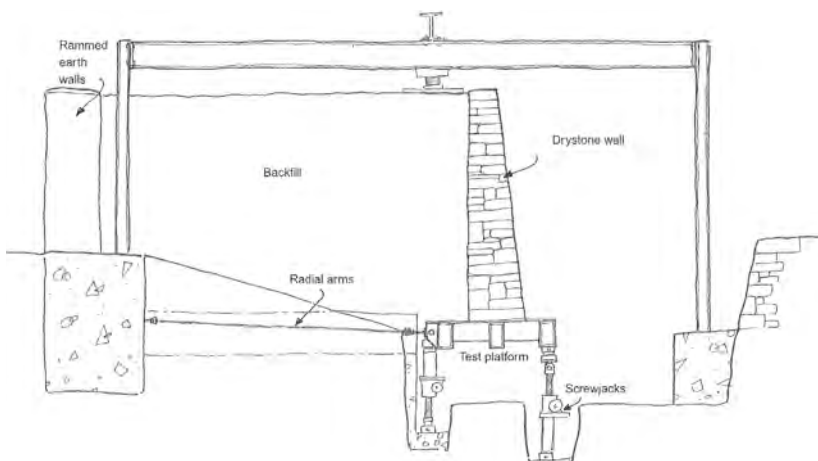


Figure 8 - Bespoke test facility

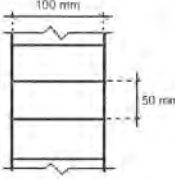
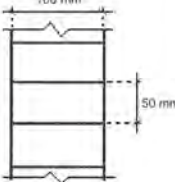
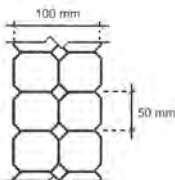
Cross-section profile	Wall material	Failure details
	Softwood timber Wall height: 500 mm Density: 5.5 kN/m ³ Friction angle: 24°	Recorded backfill height at failure via toppling: 245 mm Predicted backfill height at failure: 240 mm
	Concrete blocks Wall height: 500 mm Density: 24 kN/m ³ Friction angle: 29°	Recorded backfill height at failure via toppling: 350 mm Predicted backfill height at failure: 350 mm
	Concrete blocks (10 mm chamfer) Wall height: 500 mm Density: 24 kN/m ³ Friction angle: 29°	Recorded backfill height at failure via toppling: 300 mm Predicted backfill height at failure: 315 mm

Table 2 - Small scale testing results

with a sufficient margin of safety during installation of remaining instrumentation and loading plates, whilst being close enough to its ultimate conditions that the proposed movements and loadings can take the structure to failure. Both walls complied with these criteria; however the second wall was significantly less stable, with the eccentricity on completion lying in front of the wall's middle third (fig 11c). The initial phase of each test to date has involved the raising of the platform up to ensure that the maximum possible friction is generated at the wall/backfill interface. In real walls, full friction is likely to be achieved due to settlement of the backfill following construction of the wall. In order to take full control over this important parameter the inverse movement is applied, and the wall is moved upwards in relation to the backfill via the jacked platform. Load cells supporting the wall show a steady increase as the platform is lifted, until the full friction is mobilised when the loads level out, see Figure 9. Following the initial raise of the platform, Wall 1 was subjected to a combination of forwards rotation and surcharging, whilst Wall 2 was simply surcharged until failure occurred. Both walls failed by toppling, though each displayed a great deal of movement prior to collapse, including block rotation and sliding, see Figure 10a, b and c.

Throughout each experiment the geometry of the wall face is constantly monitored, allowing the wall profile to be recreated within the LE program, so that stability can be assessed as the loads are changed and the walls deform. The images of the thrust lines generated immediately prior to failure for the first two tests are shown in Figure 11. From Figure 11 it is evident that the structures are both on the verge of collapse, although three-dimensional effects may have given added stability, especially in the case of the first test wall. Both walls developed bulges only through the central region of the wall adjacent to where the surcharge loading was applied. The high friction generated between the courses allows a tensile strength to develop along the length of the wall, so the

relatively lightly loaded end sections help support the central section. This load shedding subsequently allowed the first wall to deform to a greater extent than would have been possible had the wall been acting purely in plain strain, with the wall profile showing the coping overhanging the toe by some 500mm prior to failure. Following this experience, the second wall was intentionally of a poorer quality than the first, showing a number of 'running joints'; a term used to describe when the joint between two adjacent blocks is similar in position to a similar joint above or below. These prevented the transfer of load along the length of the wall, and opened up as loading proceeded to allow the central section to move more easily relative to the adjacent sections. The result of this was that the wall behaved in a manner more akin to that represented in the LE program, and probably more akin to sections of wall which actually fail in practice. Most wall failures are localised at weak points, rather than being general failures along the full length of a wall. Thus three-dimensional effects, in particular the transfer of load along the line of a wall, can certainly help support a weak section, but failure is much more likely to occur where loss of such support results in a behaviour which is nearer two-dimensional or plain strain and two-dimensional analysis is therefore recommended for most situations. The final observations were recorded just minutes before failure, when the wall was remaining stable in the unloaded state but giving indications of imminent collapse (increased movement for only minimal surcharge loading), and this was successfully reproduced

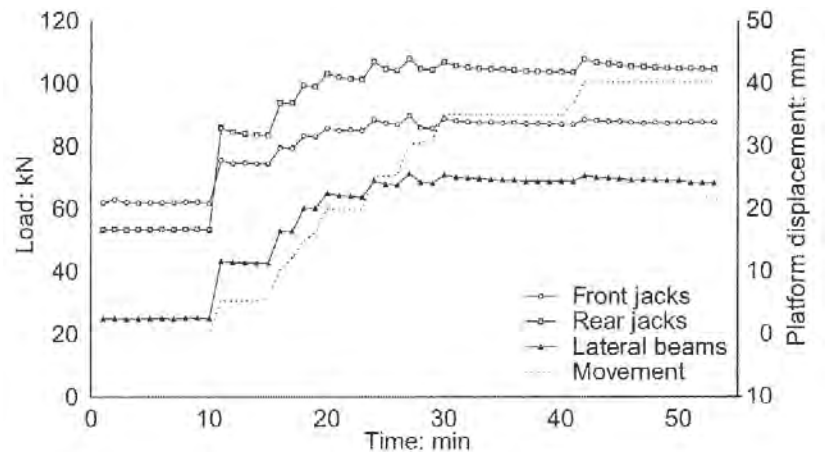


Figure 9 - Load cell readings during platform raise



(a)



(b)

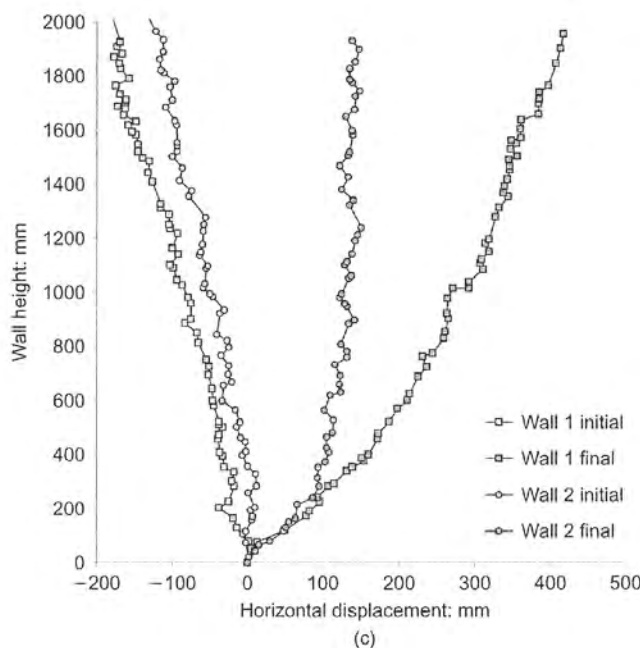


Figure 10 - Physical test results: (a) wall 1 prior to failure; (b) wall 2 prior to failure; (c) graph of wall displacements

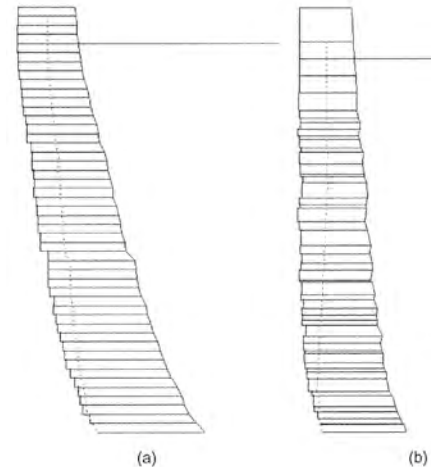


Figure 11 - Limit equilibrium analysis of physical tests: (a) wall 1; (b) wall 2

by the LE program. The surcharge loading was displacement rather than load controlled, allowing a progressive deformation past peak load with full control, allowing safe collection of data, until eventual collapse associated with excessive distortion of the geometry of the structure. Both structures remained absolutely stable, with no ongoing deformation, when the surcharge loads were removed, even though they were very severely distorted.

Conclusions

The limit equilibrium analysis program described in this paper has enormous potential compared to numerical analysis packages; its simplicity allows any engineer with a basic knowledge of a wall's geometry and material properties to obtain a reliable understanding of the factors influencing its stability, without the need for the detailed knowledge, advanced design parameters, time and expertise that are needed for reliable numerical analyses. The program's flexibility in use allows walls of any geometry with variable backfills to be analysed, and the application of surcharging can be applied to represent circumstances such as new constructions in the proximity of the wall or increased vehicle loading.

Acknowledgements

The work carried out in this paper was conducted at the University of Bath in conjunction with Southampton University, funded by the Engineering and Physical Sciences Research Council (EPSRC).

References

1. POWRIE, W., HARKNESS, R. M., ZHANG, X., and BUSH, D. I. Deformation and failure modes of drystone retaining walls. *Geotechnique*, 2002, 52, No. 6, 435-446.
2. O'REILLY, M. P. and PERRY, J. Drystone retaining walls and their modifications - condition appraisal and remedial treatment. 2009, RP723,
3. O'REILLY, M. P., BRADY, K. C., and BUSH, D. I. Research on masonry-faced retaining walls. 2nd European Road Research Conference, 1999,
4. BRITISH STANDARDS INSTITUTION(BSI) Code of practice for earth retaining structures. 1994, BS 8002,
5. CLAXTON, M., HART, R. A., MCCOMBIE, P., and WALKER, P. J. Rigid block distinct-element modelling of drystone retaining walls in plane strain. *ASCE*, 2005, 131, No. 3, 381-389.
6. BURGOYNE, J. Revetments or retaining walls. *Corps of royal engineers*, 1853, 3, 154-159.
7. DICKENS, J. G. and WALKER, P. J. Use of distinct element model to simulate behaviour of drystone walls. *Structural engineering review*, 1996, 8, No. 2/3, 187-199.
8. HARKNESS, R. M., POWRIE, W., ZHANG, X., BRADY, K. C., and O'REILLY, M. P. Numerical modelling of full-scale tests on drystone masonry retaining walls. *Geotechnique*, 2000, 50, No. 2, 165-179.
9. BRADY, K. C. and KAVANAGH, J. Analysis of the stability of masonry-faced earth retaining walls. 2002, TRL Report TRL550,
10. WALKER, P. J., MCCOMBIE, P., and CLAXTON, M. Plane strain numerical model for drystone retaining walls. *Geotechnical Engineering*, 2006, 159, No. GEI, 1-7.
11. COOPER, M. R. Deflections and failure modes in drystone retaining walls. *Ground Engineering*, 1986, 19, No. 8, 28-33.
12. BRITISH STANDARDS INSTITUTION (BSI) Code of practice for strengthened/reinforced soils and other fills. 1995, BS 8006,
13. BOLTON, M. D. Geotechnical stress analysis for bridge abutment design. 1991, Transport and road research laboratory, report 270,
14. CORTE, J. F. Reinforced earth retaining walls under strip load: Discussion. *Canadian Geotechnical Journal*, 1981, 18, 324-326.
15. WONG, H. N. and HO, K. K. S. The 23 July 1994 landslide at Kwun Lung Lau, Hong Kong. *Canadian Geotechnical Journal*, 1997, 34, 825-840.
16. BAILEY, C. Model tests of dry stone retaining walls. MEng Dissertation, University of Bath, 2008,

Vivien Collins

Integrated Transport
Division, Welsh
Assembly Government

Joanna Scott

Senior Consultant
Intelligent Transport
Systems



Sarah Warlow

Project Support
Intelligent Transport
Systems

Abstract

This paper will discuss the progress of the Welsh Assembly Government project to introduce smart card technology to public transport in Wales and addresses the achievements of the project to date, how the lessons learnt during its initial phases can benefit other smart card schemes looking to implement within the Integrated Transport Smartcard Organisation (ITSO) environment and how the further exploration of smart card technology will move forward in Wales.

Wales' ITSO environment for smart cards

Wales is one of the four nations that make up the United Kingdom and has a population of just over 2.9 million. The Welsh Assembly Government is the devolved government for Wales and is responsible for health, education, economic development, culture, the environment and transport. Residents aged over 60 and disabled passengers of all ages have been entitled to free travel on local bus services across Wales since 2002, and the concessionary fares scheme remains part of the Top 10 Assembly Government commitments to the electorate, where over £50m of public money is provided per annum in revenue support to bus operators each year. In 2008, the Welsh Assembly Government began to issue standardised contactless smart cards to over half a million concessionary pass holders, to further enhance the concessionary fares scheme, provide quicker boarding times, accurate data capture and more efficient and auditable reimbursement.

The benefits of ITSO standards in Wales

The scheme makes use of integrated systems that follow a common open specification developed by ITSO², an organisation formed by leaders in the UK transport industry to establish standards in smart card technology with the intention to make national multimodal interoperability a reality. In contrast to other "closed" or "black box" smart card schemes that act within self-prescribed boundaries, the interoperable ITSO specification is built on the vision that multiple schemes (both commercial and public sector) will be able to share and interlink acceptance of their ticketing products and services so that passengers may experience a seamless journey on any mode of public transport they choose using electronic ticketing. In addition to this, the chip embedded within the smart card allows opportunities for extension beyond ITSO and the transport environment, with enough memory to hold additional applications that may be used for services such as prepayment, citizen entitlement privileges, loyalty bonuses, and beyond. Services of this kind are already being used effectively by ITSO schemes in Scotland and England.

The ITSO operating licence

After becoming a member of ITSO, signature of the ITSO Operating Licence is the first step in setting up a live transport smart card environment, and forms the contractual basis of a scheme's relationship with other scheme owners and ITSO themselves. It represents the agreement between all parties to protect the integrity of the environment, and includes clauses that cover use of the ITSO systems themselves, service level agreements and fees. Although called a Licence, it is a binding contractual document that must be treated as such, and adequate time should be given to allow a thorough understanding of the implications of signature by senior members of the operator's organisation and, more than likely, lawyers. Because the document is owned by the ITSO membership, changes to clauses are not easily negotiable and would have to go through a full consultation period with all ITSO members if it were to change - meaning potential delays to the start of any scheme. In Wales the Operating Licence took months to agree and signature was ultimately raised to Cabinet level.

Our recommendation would be to have early sight of this document at the project inception stage, identify any potential issues at the start and to discuss and understand them with ITSO at the earliest opportunity. This may well help to expedite a potentially lengthy process, and is very important as the systems cannot be built without signature of this document.

Setting up the HOPS

The central hub of the ITSO system is known as the AMS/HOPS (Asset Management System / Host Operator Processing System - usually known simply as "HOPS"), and acts as a passenger journey database through which all secure ITSO transactions are processed. Setting up a HOPS is one of the most technically challenging aspects of entering the ITSO environment and requires experienced suppliers and advisors as it is the mission-critical system that must be addressed before any other. Schemes must first understand how they wish their Operator Identification (OID) structure to work - this is essentially a set of identifiers that allow definition of the structure of your scheme, to allocate ticketing product ownership, provide the ability to uniquely encode the ITSO artefacts and configurations, and future proof for changes in that structure. Wales has an "umbrella" OID structure, where the Welsh Assembly Government retains the top level (or parent) OID identification, with a series of other OIDs below. Every scheme has its own requirements, different types of flexibility and different goals. However, the consideration given to an OID structure, and the relationships held therein, is critical to ensuring that an ITSO scheme remains flexible enough to change in the future with minimum impact on time, resource and cost. It is from this that the HOPS can then be built, the system can be commissioned and linked to a Card Management System (CMS), enabling personalised card-encoding, ISAM commissioning, and POST messaging to begin.

POSTs, ISAMs, PersoPOSTs, and the CMS

The HOPS communicates with the front office bus depot systems and Point of Service Terminals (POSTs); these are the Electronic Ticket Machines (ETMs) on-board buses which allow the flow of data to and from the ticket machine to the bus depot system to the HOPS on a daily basis. Inside every POST is a specially commissioned ITSO Secure Application Module (ISAM), a literal "key" to interoperability, which determines which ticketing products and services may be accepted and includes the ability to securely add additional ITSO ticketing products and services to the smart card on bus. The ISAM also allows the ITSO transaction to be identified for operator reimbursement. When the chip is encoded, as well as containing the unique Wales identification, it is assigned to a public transport operator so that it may be individually managed and tracked throughout the course of its life, and so that each operator's transactions can be differentiated from the next - especially important for reimbursement.

The personalised smart card issuance equipment (Perso-POSTs) in Welsh Local Authorities enable new live smart cards to be produced and encoded on demand, and for products and services on the card to be added, confiscated, hot-listed or deleted by the issuer through a direct and secure web-based interface with the centralised CMS. The CMS can be accessed by all 22 Local Authorities in Wales with hierarchy based security partitions to protect both the integrity of the data and to ensure that certain functions are only accessed by those with appropriate system privileges. The system also has comprehensive reporting capabilities which allow the Welsh Assembly Government and Local Authorities to analyse and model the passenger data held in the system. During the bulk card re-issue and rollout in Wales, a temporary remote central bureau was also formed using multiple concurrently running perso-POSTs to allow thousands of cards to be encoded and switched on each day, whilst maintaining real-time communication with the HOPS and CMS.

Reimbursement system

The centralised reimbursement system, also web-based and securely partitioned, receives detailed extracted data from the HOPS on journey transactions so that Local Authorities can accurately calculate the amounts owed to bus operators for concessionary travel. It is at this point that the ISAMs allocated to operators allow transactions to be identified, grouped and analysed using time parameters or geographical boundaries, and then translated into a monetary value using a reimbursement factor set by the Welsh Assembly Government. Currently in Wales, operators are reimbursed for each concessionary journey based on the passenger's point of boarding (the boarding point signifies the Local Authority responsible for the administration and payment for the travel). Therefore, if a passenger uses multiple operators across multiple boundaries or regions, at the point of reimbursement the transaction is grouped according to location and service provider and reimbursement occurs for the journey (or part journey) that has taken place. The system also retains enough flexibility to adjust the parameters to protect against change in reimbursement principles, for example adjustments in geographical boundaries or routes.

ISMS, external HOPS, and ITSO compliance

Finally, to complete the ITSO architecture and provide interoperable transactions, the Welsh architecture is linked to external ITSO systems. The ITSO Security Management System (ISMS) facilitates multi-HOPS communication with other schemes; and a "HOPS to HOPS" capability allows Wales to communicate directly with HOPS owned by schemes that share our back-office systems supplier. So, if two schemes agree to accept and/or recognise each others' ticketing products but have different systems suppliers, the ITSO transaction is forwarded to its HOPS via the ISMS as and when that smart transaction is recorded. The ISMS also provides the ability to send certain types of messaging to ISAMs in POSTs across the ITSO environment, which is especially important when

dealing with “hot-lists” that protect all schemes against fraudulent card or product use. Hotlisting is described in further detail later in this paper.

To maintain system security and integrity, each aspect that sits within the Wales architecture that requires reciprocal messaging from the HOPS must be ITSO compliant, and tested and certified to this standard. This is extremely important, not only for the immediate needs of Wales’ scheme, but also for the integrity of any operator that sits within the ITSO environment. True interoperability means that all the secure systems are interacting constantly, so the ITSO community as a whole must remain confident that any new scheme owner or “ITSO Licensed Operator” protects the security and integrity of the entire estate and not merely its own portion of it. On signature of the ITSO Operating Licence, this also becomes a contractual obligation, and although this certification comes at a premium in terms of contractual and financial commitment, the ultimate benefits of cross-scheme capability for passengers merits the investment.

Project Implementation Phase 1: Commissioning systems and migration

The first major phase of the project has been system commissioning and the bulk reissue of bus passes to eligible concessionary travellers. Prior to this project, concessionary travel pass data were held on local databases by each of Wales’ 22 Local Authorities. For a cost-effective implementation of an all-Wales scheme the migration to the centralised CMS database was necessary to benefit from economies of scale; the systems were procured in 2006, commissioned in 2008, with card rollout finishing in 2009. The project worked through many challenges, including technical problems mainly experienced during the integration and testing phases in the first two months after system commissioning. Building in contingency weeks for testing and for each migration phase also proved invaluable and although it has taken 18 months in total this has been an acceptable timescale for the phased authority-by-authority approach.

The avoidance of a “big-bang” rollout has been beneficial to Wales, especially in the confidence that now exists with regard to the quality and security of our national passenger database. The ability to show flexibility in addressing 22 different sets of needs in each location has also been essential to retain enthusiasm, momentum and buy-in from all the Welsh authorities.

Prior to migration, the quality of the data held by Local Authorities varied widely and, as part of their responsibility as data owners, extensive data cleansing was carried out by each Authority prior to migration to the new database. For the majority, this phase usually worked through two cycles over two years to ensure maximum benefit and minimum risk. Usually data were sent to a contracted cleansing service and the remaining records were manually, sometimes painstakingly, searched and updated in-house based on the up-to-date information available. The relationships between local registrars during this process also both helped and hindered progress depending on the strength of information that could be obtained at each Local Authority. Recommendations have therefore been taken forward to further improve internal relationships and an understanding of the real and sometimes assumed boundaries around sharing data.

In addition to these steps, during the migration process itself authorities have been able to use a “holding area” within the CMS for any record which may have had a question mark on its validity. This has allowed Authorities to continue checking their data immediately after their migration, ‘flushing out’ erroneous data after the initial reissue, although this has been minimal. This has ensured that a proportion of cards that could have been reissued in error have been managed over a designated period of time without any undue negative impact on the card holder. This has saved time, money, and any unnecessary heartache for families with recently deceased relatives whose records may not have been identified.

Mandatory field requirements and data handling business rules within the system also ensure that data recorded accurately and are consistently across the country, ensuring ongoing quality.

These extra steps in addressing data quality have given the Welsh Assembly Government a great deal of confidence in the class of the data held in the system - very important given that each Local Authority is grant-aided based on the number of live concessionary travel card users they have on their system each year. Again, the accuracy of this process and ability to audit this information centrally if required enabled the government to ensure that public money is spent appropriately (and saved where possible), and apportioned fairly. It has also meant that the requirement for large investment in the additional integration of large-scale data cleansing tools has been avoided as the scale of any issues remains relatively small and data quality can continue to be managed in-house. For any scheme looking to migrate to live smart card use, ITSO or not, this aspect can be endorsed as a major benefit.

As well as renewed confidence in the quality of the national database, records are also partitioned within the CMS allowing Local Authorities access to the data of residents within their area, with access to others’ partitions being granted on a per-request basis, which allows updating of records to take place locally. Allowing this responsibility to remain local, with residents contacting their own Authority directly in all related matters, retains a sense that a local service is being provided, even within the larger national scheme. This is arguably particularly significant given the vulnerable demographic of the concessionary fare scheme. The economies of scale brought by the migration to one centralised national system have therefore been achieved without sacrificing the benefits of local services.

Project implementation phase 2: Electronic Ticket Machine (ETM) rollout

The Welsh Assembly Government's aim of encouraging local business and inward economic investment is also promoted through the project by the award of sliding-scale grants to bus operators for purchase agreements for ITSO compliant ticket machines and provision of regionalised depot systems to support smaller public transport businesses. This was particularly significant to Wales' ninety locally based small to medium sized operators, where the cost of smart equipment might otherwise have precluded involvement. While a small number of larger operators have installed their own back office systems that allow data to be transferred directly to the HOPS, central systems have been installed and are managed within each transport region by a lead authority on behalf of the four Welsh transport consortia for use by its smaller operators.

Rollout of ticket machines in Wales began in November 2008 and has been subject to significant delay beyond original expectations. Rollout is planned as a phased activity due to the number of stakeholders involved, both to minimise the risks of wholesale migration and to ensure that all operators are adequately prepared for the change. Delays have occurred mainly due to the complex set of dependencies that need to be met before progress can be made. Software needs to be tested to a very detailed level to ensure that all former capability for ticketing is retained alongside ITSO capability. Integration of the many other software products that communicate with the ticket machine (e.g. GPS systems) also requires sufficient time to be tested; multiple depot systems need either to be created or updated, and drivers need to be trained and educated to a sufficient level to provide confidence for the operator and the general public. The relative immaturity of schemes in this area has meant that delays have often impacted the project plan as new previously undiscovered problems are resolved. This has been common and impacted all ITSO implementations in the UK.

Managing the logistics involved in delivering hundreds of machines to buses across many operators in short installation timescales with limited resources has also been difficult. Scheduling around weekend services has proven the best option, although this has meant complex planning to ensure that all buses (often across a very large geographical area) can maintain day to day business with minimum or no impact during installation. Good relationships and partnership working between Local Authorities, Regional Transport Consortia, Government and each bus operator has been critical to success, especially where timescales have been affected and plans changed.

The first operators in Wales went live with ITSO compliant electronic ticket machines in November 2008. At the end of April 2009, approximately 25% of the Welsh ISAM estate was live and fitted to ETM POSTs and Perso-POSTs. Within this initial six month period, each ETM supplier achieved live smart transactions on their equipment, giving greater confidence for a smoother and quicker rollout over the course of the summer. In south east Wales we have been able to immediately see the transactions that have passed through the HOPS and bus operators have submitted their claims for reimbursement for concessionary journeys based on the auditable data that the system has gathered.

The project has also had to take into account policy decisions that manage scenarios where the technology on-bus does not work as expected, and how drivers should deal with scenarios where a card malfunctions or is fraudulently used. Unfortunately, the root cause in each of these scenarios cannot be investigated and identified clearly and quickly enough to facilitate resolution for the user at the point where they wish to use their entitlement on board a bus. Parallel use of the new smart-card as a flash pass will mean that the difficulty in dealing with scenarios where the card does not work will be abated in the short term, but this will only be true up to a cut-off point when all transactions are smart. This will happen when all ticket machines are in the field and smart transactions are used everywhere by default.

Using a flash pass indefinitely is not an option as it will undermine the purpose of using smart technology to collect valuable data, both for reimbursement and for planning purposes. Part of the work currently being implemented involves identifying real examples of these scenarios, quantifying non-functioning cards vs. fraudulent cards and defining errors with ticket machines. Although experience to date indicates that the problem is not large, we will not have a clear picture across Wales until all ticket machines are installed and smart transactions become the default position.

Project implementation phase 3: extending the Welsh ITSO environment

Once the ITSO infrastructure is in place Wales has a selection of opportunities to pursue beyond a simple concessionary ticketing structure, as well as ongoing challenges in keeping up to date with changes in the technology. In the short term, hotlisting requires implementation. This is a form of identifying fraudulent or invalid cards on-bus via ITSO messaging to the ticket machine via the HOPS and ISAM. Cards reported lost or stolen are flagged within the CMS and become part of a 'warm-list' in the HOPS. If a card is subsequently used on-board a bus, it will become part of the 'hotlist' when the transaction data is uploaded each day. The 'hotlist' is sent to each ticket machine and further use of the card will result in a message indicating to the bus driver that the card is not valid for travel.

Since the ITSO specification was formed, many theories on how hotlisting might work have been discussed but there are no definitive guidelines on implementation. In Wales, the initial applications of simple business rules around hotlisting are based on prioritising types of products (ticket types) and shells (the ITSO partition of a smart card). The decision to hotlist based either on a product or a shell originates from a principle that takes future commercial implementations into account.

Although at the moment an invalid concessionary card's shell would be hot listed if it was lost or stolen, in the future it may be that a product sits within an ITSO shell containing multiple products. If one product became invalid - for example, a cancelled prepaid product - no other product would be affected and there would be little point in incurring the additional cost and time involved in cancelling and reissuing a full card if only one aspect of it is affected. Products could also then be prioritised on the hotlist based on their value: a season ticket has inherently more value than a prepaid day ticket.

We plan to use one hotlist that is broadcast to our OID group on a weekly basis with incremental updates of new or prioritised shells or products, avoiding the additional messaging traffic that would incur from a point to point approach (HOPS to individual POST). There are also certain constraints within this simple implementation: the capacity of memory on ticket machines dictates the maximum size of the list; the size of warm-list within the HOPS needs to be managed and pruned as the scheme grows; software changes will require implementation in the back office as well as changes to the ITSO specification itself; and interacting with other ITSO scheme hotlists in the future will become more complex as more operators become live and implement their own hotlisting policies. In Wales we are planning to implement our initial hotlisting structure by the end of 2009 and are currently likely to be the first ITSO scheme to do this.

As the Welsh scheme progresses towards a fully-smart environment, further benefits of the technology will also have been highlighted. Before the end of the year we will have also migrated the scheme to a new card media type where all new passes will be issued using a different specification of card. The ITSO platform has a selection of certified card media available for use, and Wales will shortly be using two card types in the scheme, retiring one card type to allow for better technology to take its place through another.

The flexibility of the system will also have been demonstrated by the introduction of a commercial smart card scheme by Cardiff Bus, one of Wales' major bus operators. The expansion into a commercial scheme will go further towards highlighting the key benefits of the replacement of paper ticketing by smart technology, such as the invaluable ability to capture data on passenger travel habits to better inform marketing practices and the potentially vast reduction in fraud. Wales remains committed to ITSO, interoperability, and the implementation of further ticketing types within the smart card environment which will extend across other modes of public transport in the future.

On balance - the see-saws of congestion charge business cases



Paul Grayston

Technical Director
Intelligent Transport
Systems

Abstract

The design of urban congestion charge schemes within business case proposals to the UK Department of Transport's Transport Innovation Fund has presented a variety of competing and sometime conflicting considerations for designers. This paper looks at how balances need to be struck between geographic coverage, revenue, effect on traffic, technology choice, and the treatment of discounts and exemptions.



Andrew Ferry

Senior Consultant
Intelligent Transport
Systems

Adrian Bourne

Graduate Engineer
Intelligent Transport
Systems

Introduction

Overview

Many transport practitioners see the use of a congestion charge as a potential part of their armoury to change travel patterns and reduce congestion. The success of schemes in London and Stockholm has provided a degree of interest, even support from previously wary politicians. In the case of London the charge was part of Mayor Livingstone's mandate, and in Stockholm a referendum was only held once citizens had actually experienced road pricing. However, as was seen in the "no" vote Manchester in 2008 (for a scheme not due to go live until at least 2013) political support was not enough. This was despite indications that a positive business case for the charge could be made. Perhaps the biggest lesson from Manchester is not for politicians to be pushed into

a referendum when the general public is unable to comprehend what the whole package of measures, never mind the congestion charge scheme! A full, well thought out business case may help with this in other cities considering road pricing such as Cambridge and Bristol.

When announced in 2006, the UK Department of Transport (DfT) Transport Innovation Fund (TIF)¹ offered English Local Authorities the opportunity to bid for major funding for packages of transport investment that would enable step changes in provision (new bus, tram and rail services, better interchange facilities, improvements in information and smartcard ticketing etc). In return schemes would implement some form of congestion charging. The process attracted many bids for pump-

priming money for studies to develop proposals. The original enthusiasm has dampened somewhat. In some areas politics put paid to the idea, but in others the clear benefits and/or positive business cases through traffic modelling have not been proven.

The basis of this paper

In general sponsors of congestion charging schemes will want to keep their schemes simple to understand, and aim for the low operational costs and efficiency seen within the tolling industry. From Atkins work for some of the TIF authorities wanting to develop draft business cases, we have learned that this can be a very difficult balancing act given that there are actually few operating schemes in place and comparisons are difficult to make. This paper gives some

insight into those “see-saws” and how to account for them within the business case for charging schemes.

The graphical plots used in this paper have been generated using cost and volume models developed by Atkins for road pricing scheme design option analysis and sensitivity testing. They are used to give a typical indication of results and do not represent any specific scheme.

The see-saws

See-saw 1 - What to charge for versus the political requirement

The political instruction to investigate the options for introducing a local road pricing scheme is almost certain to contain some caveats. They could be about: the level of charge, the revenue expectations, the discounted or exempted users, geographical coverage and so forth. More often than not (and quite understandably) to the scheme designer this initial ‘back of an envelope’ concept will be fraught with complication. Furthermore, at an early stage little thought is usually given to how the charge will be presented to end users. Keeping it simple is a byword that is often adhered to, but the consultation processes for the Manchester and Cambridge schemes suggest that a large proportion of people didn’t even understand that the proposals were for peak hour charges only, not all day.

The only way to demonstrate robustly to decision makers what is likely to work and what is not is to carry out comprehensive traffic modelling. But, because of its theoretical nature, it can sometimes cloud incorporation of the obvious such as natural boundaries and gates to funnel traffic in and out of a scheme area, or that a more simple tolled crossing solution might be more appropriate.

Traffic modelling provides the key input to any congestion charge scheme design. Identifying the number (and type) of users of the scheme per day and relating it to a particular level of charge is intense and complex work.

The level of charge (for a local charge) is driven by a need to generate a certain level of revenue. As well as covering the operating costs, this could be to provide funds for paying off loans, paying for additional works, or even general contribution to authority funds. To avoid overloading the see-saw with unachievable objectives it is essential to derive an early indication of:

- (a) the average level of charge users will be required to pay
- (b) the transaction cost target.

See-saw 2 - The scheme size and boundary

Schemes will either be charging for crossing one or more cordons (they could be concentric as the Manchester proposals or adjacent zones), or for travel within an area.

In considering the size and shape of a scheme boundary, the extent of the traffic modelling will set the basic coverage. However, reference to detailed maps and on the ground surveying are needed to produce a set of “gateways” into the scheme.

They may be junctions, bridges, or other physical constructions. It is essential that traffic funnelled towards the gateways is the traffic that will be liable for the charge and can be charged. Similarly communities should not be isolated by the boundary of the scheme.

The issue of splitting conurbations by a scheme boundary is a vexed one, bound to cause local resentment. This may result in considerations for discount schemes which will complicate the operation of the scheme. The two ring Manchester concept was an interesting variant. The fact remains though that in traffic modelling terms the most attractive options, both financially and in terms of affecting traffic are often the simplest.

If the scheme is a cordon then the cordon should not be too long, otherwise it might result in too little effect or be too difficult to divert away from. But a long boundary for an area scheme may mean that a relatively large number of internal charge points will be required, increasing the capital cost for setting up the scheme. Also detection/charge points with an area scheme will need to cover both directions of traffic in most locations.

But, reduce the geographic coverage of an area scheme and there comes a point where it might as well be a cordon because there are few if any places where an internal detection point would be worthwhile.

And then, as scheme concepts go out for consultation, local objections will raise their head. Often they will have an emotional aspect to them (for instance giving exemptions to lower paid workers such as nurses) that will be persuasive to local politicians to champion. Cost implications can be severe, even to the extent that exemptions might result in a net cost to the scheme because of the human administration involved.

See-saw 3 - The technology choice

For the moment, despite perhaps justified claims by suppliers, Time Distance Place is not a viable option for an urban type of scheme to be implemented. Under the UK TIF programme the technology choice see-saw for the local authority scheme promoters is weighted down at one end with a cautious, risk averse attitude.

Realistically this means any business case must propose the use of technology already proven on a wide scale. This leaves Automatic Number Plate Recognition (ANPR) or a combination of ANPR and tag and beacon solutions as the two main options.

Tags have attractions in that they have a high accuracy, as a detection technology the monetary transaction begins immediately, and interoperability can provide users and operators with incentives for use.

But on the other end of the see-saw the permanent scheme in Stockholm has elected to use ANPR as its main solution, with technology improving the accuracy of the system to a level deemed acceptable by the authority.

ANPR must be considered for the purposes of identifying non-payers. This creates a further balancing point relating to the expected number of Penalty Charge Notices (PCN) to be issued. In the London scheme PCN income is a considerable contributor to the overall revenue. If there is a high take up of tags within a scheme there is potentially less PCN revenue because of the high level of identification of scheme users.

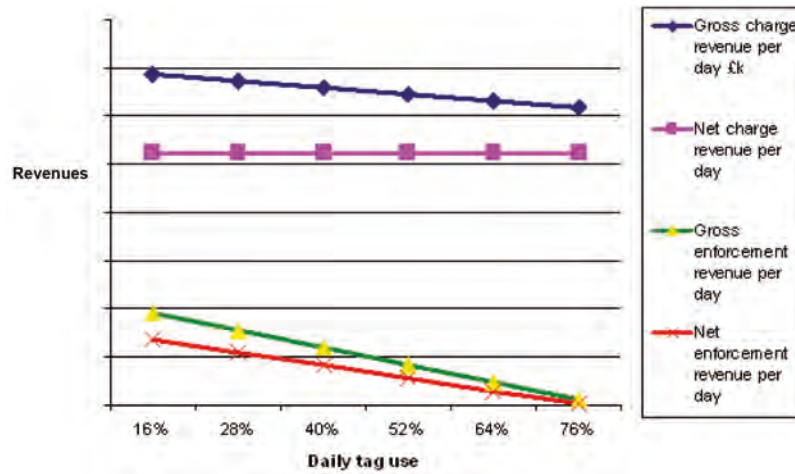


Figure 1 - Example of balancing revenue sensitivity to tag take up

In the draft proposals for the Cambridge scheme (October 2007)², it was considered that a tag solution offered the most efficient scheme option. To encourage tag take up, discounted charges would be offered. Different scales of charges and degree of take up of the tag option were analysed using the scheme cost model. Even small differences in the charges caused quite large (20%) differences in the overall annual revenue for the scheme.

The possible scheme illustrated in the draft proposals for Cambridge used the output of traffic modelling that suggested an average charge of £4 in the modelled year would provide the benefits in traffic reduction required. An objective was for more efficient, more accurate operation through high take up of tags. An analysis of the optimum scale of charges to give a flat net daily revenue profile was undertaken, see Figure 1. This was to ensure the scheme income would not be dependent on a particular split between tag users, ANPR declaration and PCN revenue, essentially making sure discounts for tag use did not end up as a cost liability. It is important to note that tags increase compliance, therefore reduced enforcement revenue from PCN's drops as tag use increases.

Using the cost model also showed that if there was low take up of tags (perhaps because the discount incentive was too low) then there was a point where, instead of the discounted rate, a standard rate and an ANPR declaration scheme (i.e. like the London scheme) became the most attractive to operate.

See-saw 4 - The discount and exemption see-saw

Public opinion expressed during scheme consultation will often result in the need to provide discounts, even exemptions to certain types of user. These discounts may be quite substantial. The to be discontinued Western Extension to the London scheme gives residents a 90% discount for the whole scheme area. Has this level of discount actually contributed to a rise in traffic within the zone? We might never know, but the level of (discounted) residents' charge needs to be sufficiently high to persuade at least some proportion of the residents not to use their cars and continue to contribute to the congestion.

In Stockholm the Lidingö exemption is managed using tags. Tags are useful for this application, as they provide a certainty of registration with the scheme. However, there is a cost for their provision and operation – in some cases this may not be recovered if those having a tag do not actually interact with the scheme (for instance they may have a tag but not interact with the scheme during its hours of operation).

Once discounts or exemptions are introduced into a scheme, the dynamic of the cost modelling and the traffic modelling is severely tested. Every interaction with the scheme has a cost, even the detection of a vehicle to be exempted. The discount has to be commensurate with the overall transaction cost.

The group of users qualifying for a discount can be given stricter interaction requirements. For example, in London paying for a weeks worth of residents' discount is one method to reduce scheme operation costs.

Using expected transaction volumes the cost and volume models created can be used to carry out various analyses on the sensitivity of take up of particular solutions on the overall revenue. Figure 2 illustrates an example where a residents' discount scheme using a tag might be in place as well as a general tag scheme, an 'Express' registered user declaration channel where automation of the process is maximised (e.g. payment by text message), and a general declaration channel (e.g. payment at a shop terminal) where some form of human interaction may need to take place.

It must be noted that this graph is based on a mix of solutions being used (so when there are 80% users of a solution the other 20% is split between the other three), and that confidence levels are lower below 10% and above 85% because where a different scheme design philosophy is likely to be more appropriate. As would be expected, a greater proportion of declarers increases the transaction cost, and a greater proportion of tag detection reduces the transaction cost. But, when restrictions are put on users to create the 'Express declaration' option (typically some form of payment guarantee) the slope of the line is actually downward.

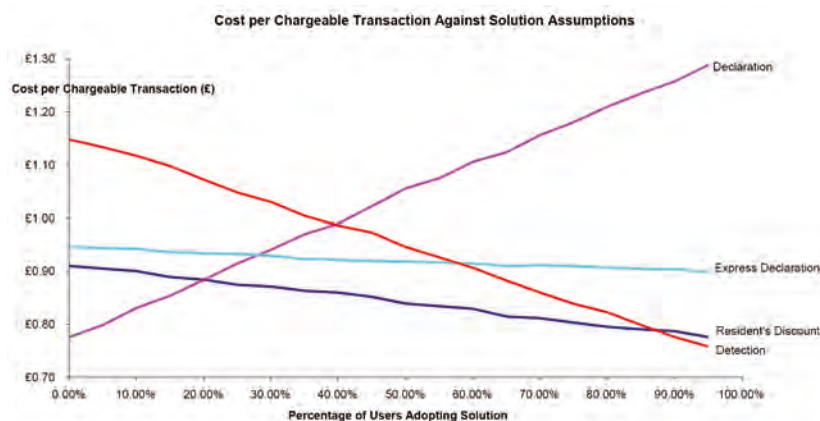


Figure 2 – Typical plot of transaction cost against take up of solutions

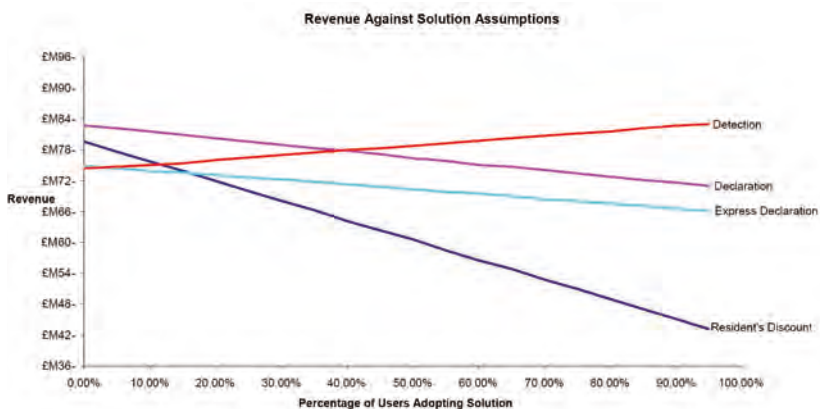


Figure 3 – Typical plot of revenue against solution assumptions

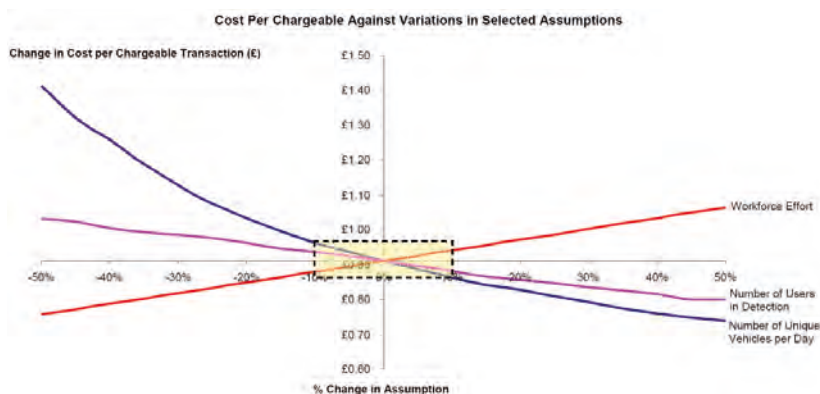


Figure 4 - Typical plot of transaction cost against scenario assumptions

The effect of this scheme option on the annual overall revenue raised is shown in Figure 3. Applying similar techniques as used in the example in Figure 1 could minimise revenue variation for general users, but, as the plot shows (as would be expected) increasing the numbers using the residents' discount solution causes a distinct reduction which is a function of the slope of the residents' discount plot in Figure 2.

Tailoring the scheme by running different tariff and solution mixes through the models can help to balance the discount and exemption see-saw. But it must be noted that such work has a theoretical basis and is designed to support options with a business case for road pricing. If in future the scheme is implemented it will be the operational model of the system supplier that drives the plot, and the contractual arrangement should apportion the risk accordingly.

See-saw 5 - Transaction cost and the number of daily users

It would be reasonable to assume that transaction costs decrease with higher numbers of users. A selection of scheme scenarios has been examined using the cost and volume models to see what the effect of variance in conditions produces on the transaction costs.

In Figure 4 the intersection of the x and y axis represents a typical scheme example default condition with a transaction cost calculated as £0.91 with these characteristics:

- it uses a set average number of daily users derived from the transport model
- 50% of users are expected to be using tag detection.

A range of schemes has been tested (including using figures relating to the London scheme) and all have produced plots with similar characteristics.

The three plots are independent of each other, but plotting them together shows which of three factors (number of users, proportion of users using tags and cost of workforce to administer the system) is most sensitive to variation. The workforce effort plot is linear, and reflects the sensitivity of staff salaries on transaction cost. The number of users in detection plot is reasonably linear, some slight wobbles occur reflecting the step increases in IT provision as the number of users increases. The number of users plot, however, is a curve that shows low numbers of users can result in significantly higher transaction costs resulting in low revenues and threatening the required profitability of the scheme.

For this particular scheme the plot also demonstrates (see the dotted line central box) that variations of $\pm 10\%$ on the design case only result in transaction cost sensitivity of less than $\pm 5\%$. At business case derivation stage this is an important indication of how risk is being limited in the scheme design.

Conclusion

Designing an urban congestion charge scheme is a matter of balance between practicalities and politics. What seems straightforward at first impression often turns out not to be. In developing a business case for a scheme a wide range of tests and scenarios needs to be analysed in order to progress to procurement of a system that will meet the economic and operation targets required of it.

References

1. UK DfT Transport Innovation Fund Guidance (2006/7)
<http://www.dft.gov.uk/pgr/regional/tif/>
2. Cambridgeshire CC Outline Proposal for Funding (October 2007)
<http://www.cambridgeshire.gov.uk/transport/strategies/tacklingcongestion/ourproposals/bid.htm>

Improvements to ramp metering system in England: VISSIM modelling of improvements



Jill Hayden

Managing Consultant
Intelligent Transport Systems



Roger Higginson

Senior Systems Engineer
Intelligent Transport Systems



Matthew Hall

Managing Consultant
Intelligent Transport Systems

Sukhvinder Ubhi

Project Sponsor
Highways Agency

Abstract

The Highways Agency has installed ramp metering at over 80 sites in England, with evaluation results showing an average 13% reduction in journey times¹. Despite the success of the existing system, some potential improvements have been identified. Due to the complexity of the algorithms, the improvements were modelled using the micro-simulation package VISSIM².

This paper describes the method and results of the modelling. The modelling exercise has shown that all the algorithms have worked as expected. It has also shown that all the algorithms designed could provide additional performance benefits over the existing system.

Background

Ramp metering was first trialled in England in 1986, with a small number of sites around Birmingham and then Southampton. From the outcome of these trials the Highways Agency (HA) commissioned Atkins to design a system for wider roll-out, and IPL & PEEK to implement it. This system was first installed in 2005 and there are now over 80 sites in England. The existing system was not modelled prior to implementation and no sites were modelled before being installed, for two main reasons. Firstly the basic concepts of the algorithms (ALINEA on the main carriageway, queue management and queue override on the slip road) were well understood and had been successfully used elsewhere, giving confidence that the system would work. Secondly, the timescales did not allow it.

Since the system was first installed, several improvements have been identified with the following potential benefits:

- A reduction in the amount of calibration required
- Improvements to the operational benefits
- Widening of the applicability of technique

These improvements are described in detail in the paper "Improvements to Ramp Metering System in England: Detailed Description of Algorithm Development"³. As these are new ideas and some of the solutions are quite complex, it was decided that modelling should be used to check their feasibility and that they would work as expected, before producing the final specifications.

Although there are a range of micro-simulation modelling packages available such as PARAMICS, AIMSUN and DRACULA, the VISSIM software from Visual Solutions was chosen for this assessment due to its strength in junction modelling, both signalled and un-signalled. VISSIM is a microscopic traffic flow simulation model based on car following and lane change logic. Furthermore, it includes the Vehicle Actuated Programming (VAP) integrated package that enables complex algorithms to be modelled.

Methodology

The project made use of an existing VISSIM model of the current ramp metering system at M1 Junction 33 southbound, which comprised:

- A geometrically correct road layout of the junction
- MIDAS loops and traffic signals accurately located compared to the real system on the road
- The ramp metering system algorithms, coded using VAP programming to represent the system

The model was improved and re-calibrated for the purpose of this project. This provided the base ramp metering model, against which the improvements could be compared. The improvements were each implemented in the VAP code and then calibrated. To compare the performance of the ramp metering system, with and without each improvement, various measures were used such as:

- Network performance
- Travel/delay times
- RM controller outputs
- Speed profile

This enabled conclusions to be drawn about the success of each improvement.

Calibration and validation of existing ramp metering

The existing model was re-calibrated, for both pre and existing ramp metering conditions. The validation showed excellent comparisons between actual on-road and modelling results. For example, Figure 1 shows the comparison of upstream occupancies for existing ramp metering conditions, for real MIDAS data over two days, and the model. These results gave confidence that the model was suitable for the purposes of the project.

The evaluation of the ramp metering system implemented at M1 J33 southbound, showed a 9.1% reduction in journey time on the main carriageway. The model results showed a 7.4% decrease which is of a similar order; this gave confidence in the model. The model also showed that an overall journey time reduction of 4.2% was achieved for both main carriageway and slip road.

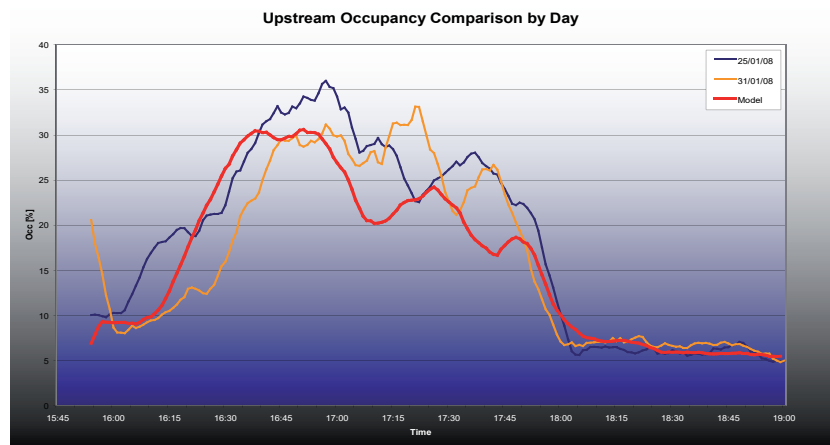


Figure 1 - Comparison of upstream occupancies for model and real data

Modelling of improvements

The following sections describe the results of modelling the four key algorithms. Although the algorithms are only briefly explained here, detailed descriptions are available in the paper "Improvements to Ramp Metering System in England: Detailed Description of Algorithm Development".

Auto-calibration of signal timings (Auto-Sig)

The main purpose of this algorithm was to reduce the significant time and effort required to calibrate the traffic signal timings for each release level. If the actual flow achieved by a release level is lower than the required flow, then the red time is reduced, and vice versa. The modelling showed that overall performance in terms of journey times through the junction is significantly improved, as shown in Table 1.

Table 1 - Comparison of total vehicle journey times for auto-signal timings

Scenario	Mainline		Slip Road		Total	
	Veh Hrs	%Diff	Veh Hrs	%Diff	Veh Hrs	%Diff
Pre-RM	1008.6	-	84.5	-	1093.1	-
Base RM	934.2	-7.4%	112.6	33.3%	1046.8	-4.2%
Auto-Sig	900.8	-10%	105.6	24.9%	1006.3	-7.9%

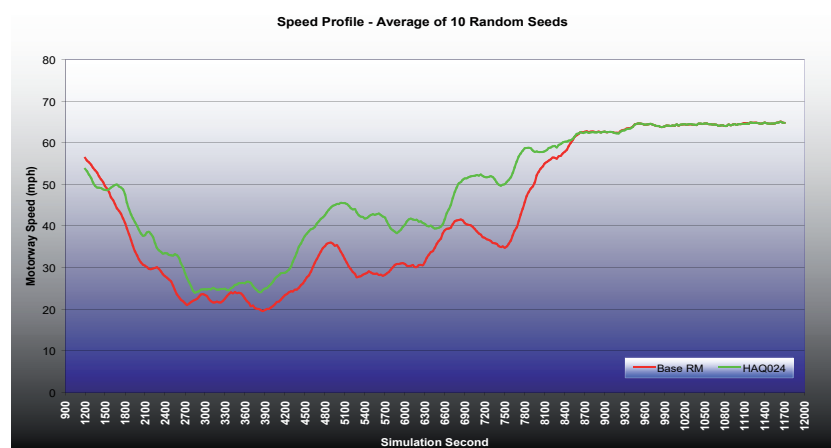


Figure 2 - Comparison of speeds with existing RM and with auto-sig

Improvements to ramp metering system in England: VISSIM modelling of improvements

It can be seen that both the main carriageway and slip road journey times are reduced, compared to the existing ramp metering scenario. This is because having accurate release flows allows more efficient operation of both the ALINEA and queue management algorithms. The modelled overall journey times are improved by 7.9% from the pre-ramp metering scenario, compared with a modelled improvement of 4.2% with existing ramp metering. This is an excellent result because as well as fulfilling the main objective of successfully reducing calibration time, the algorithm should also significantly reduce journey times.

Figure 2 shows the comparison of the speed throughout the peak, for the existing ramp metering (in red), and the auto-signal timings scenario (in green). It can be seen that speeds are higher with auto-signal timings for the whole of the peak, which explains the significant improvement in journey times. The modelling showed that even when the red times for all release levels are initially set to the same value (15s), the algorithm calibrates the signal timings fairly successfully by the end of one day of operation (iteration 1 in Figure 3). The system improves further to converge to sensible signal timings within a week of being installed (iteration 7).

Auto-calibration of Queue Management Parameters (Auto-QM)

This algorithm dynamically updates the values of two key parameters in the QM algorithm; the desired combined queue occupancy ρ_{descq} and the queue management gain factor K_{pqm} . This is designed to remove the significant effort required to calibrate the proportional occupancy queue management algorithm. Figure 4 shows the comparison of the speed throughout the peak, for the existing ramp metering (in red), and the auto-QM scenario (in green).

The figure shows that the Auto-QM scenario is slower to break down than base ramp metering, makes a quicker minor recovery but is then slower to recover in the latter periods. This algorithm has an opposite effect to the Auto-sig algorithm; the main carriageway journey time is slightly worse than with base ramp metering but the slip road is significantly better (because the queue is managed more efficiently). This results in a slight benefit overall, as shown in Table 2.

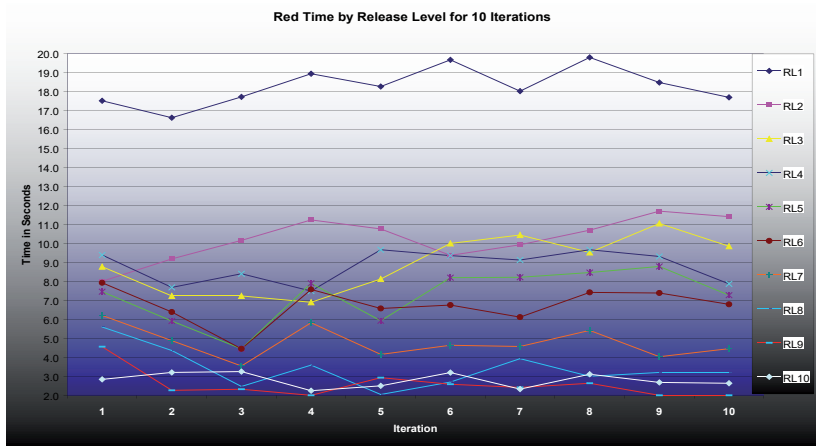


Figure 3 - Variation in red times per release level for first 10 days operation

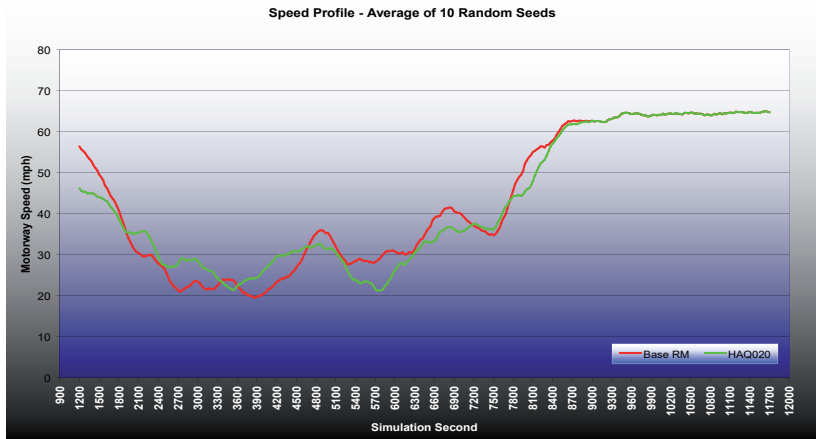


Figure 4 - Comparison of speeds with existing RM and with Auto-QM

Table 2 - Comparison of total vehicle journey times for Auto-QM

Scenario	Mainline		Slip Road		Total	
	Veh Hrs	%Diff	Veh Hrs	%Diff	Veh Hrs	%Diff
Pre-RM	1008.6	-	84.5	-	1093.1	-
Base RM	934.2	-7.4%	112.6	33.3%	1046.8	-4.2%
Auto-Sig	938.4	-7.0%	103.6	22.6%	1041.9	-4.7%

Merge Control

The Merge Control algorithm attempts to prevent saturation of the merge area which can be caused by the high flows requested by queue management and queue override algorithms at certain sites. A new set of loops is placed downstream of the stop-line within the nose area (approximately 200m downstream of the stop line for this model) to determine when the merge area is congested. When this occurs, the merge control algorithm limits the release levels used by the traffic signals, i.e. it limits the flow from the signals.

The journey times on both the main carriageway and on the slip road are lower than for existing ramp metering, with an overall very significant benefit (see Table 3).

Figure 5 shows the comparison of speeds throughout the peak, for the existing ramp metering (in red), and the Merge Control scenario (in green). The figure shows that the Merge Control model does not break down as far and is quicker to recover.

This algorithm showed the largest overall benefits in the modelling. These benefits will only be gained at a relatively small number of sites where the merge area is regularly saturated by the slip road flow. At junctions with no regular merge problem, incidents may occasionally cause a saturated merge area and on those occasions Merge Control would also be beneficial.

AINEA Cascaded with Demand-Capacity (ACDC)

The ACDC algorithm combines the best characteristics of the ALINEA and Demand Capacity (DC) algorithms, by cascading them (putting the output of the ALINEA equation into the DC equation). This was expected to provide performance benefits.

Figure 6 shows the comparison of speeds throughout the peak, for the existing ramp metering (in red), and the ACDC scenario (in green). This shows that the ACDC model is slower to break down and slightly quicker to recover, although it does break down further than the base.

The journey times on the main carriageway are slightly lower and on the slip road are marginally lower than for existing, with an overall slight benefit (see Table 4).

Table 3 - Comparison of total vehicle journey times for Merge Control

Scenario	Mainline		Slip Road		Total	
	Veh Hrs	%Diff	Veh Hrs	%Diff	Veh Hrs	%Diff
Pre-RM	1008.6	-	84.5	-	1093.1	-
Base RM	934.2	-7.4%	112.6	33.3%	1046.8	-4.2%
Auto-Sig	888.1	-11.9%	102.5	21.3%	990.6	-9.4%

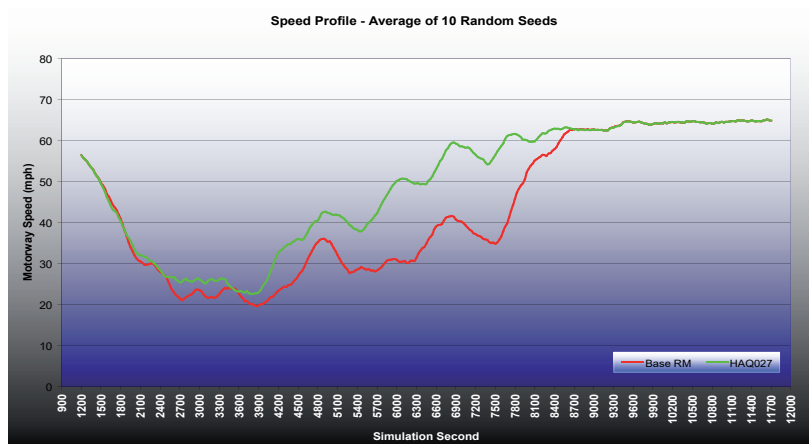


Figure 5 - Comparison of speeds with existing RM and with Merge Control

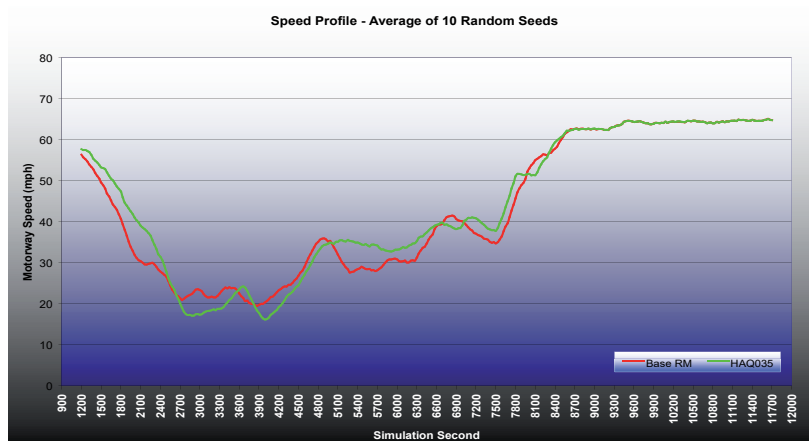


Figure 6 - Comparison of speeds with existing RM and with ACDC

Improvements to ramp metering system in England: VISSIM modelling of improvements

Table 4 - Comparison of total vehicle journey times for ACDC

Scenario	Mainline		Slip Road		Total	
	Veh Hrs	%Diff	Veh Hrs	%Diff	Veh Hrs	%Diff
Pre-RM	1008.6	-	84.5	-	1093.1	-
Base RM	934.2	-7.4%	112.6	33.3%	1046.8	-4.2%
Auto-Sig	922.0	-8.6%	112.2	32.7%	1034.2	-5.4%

Table 5 - Predicted journey time savings with four improvements

Algorithm Improvement	JT Improvement in model - above 4.2% with existing RM	Number of sites
Auto-Sig	3.7%	80
Auto-QM	0.5%	80
Merge Control	5.2%	12
ACDC	1.2%	80

Conclusions

The VISSIM modelling has shown that all algorithms developed work as expected in the model, which provides confidence that they could be implemented successfully. It has also been shown that all four algorithms provided additional performance benefits over and above the existing system. Table 5 shows the predicted journey time savings for each algorithm. While the results of the modelling do not guarantee the same results on the road, they do give increased confidence that greater journey time benefits can be achieved.

Acknowledgements

This paper draws on the work undertaken for Highways Agency Network Services Directorate and is published with the permission of the Highways Agency. The views contained in this paper are those of the authors and not necessarily those of the Highways Agency.

References

1. "Ramp Metering Operational Assessment", Highways Agency Report, April 2008, <http://www.highways.gov.uk/knowledge/17375.aspx>
2. VISSIM Visual Solutions software, www.vissim.com
3. Higginson R, Hayden Dr J, Charton T, Ubhi S, "Improvements to Ramp Metering System in England: Detailed Description Of Algorithm Development", in Proceedings of ITS World Congress 2009 (Stockholm, 21-25 September 2009).

'Scrambled' pedestrian crossings at signal controlled junctions - A case study



Chris Greenwood
Principal Consultant
Highways & Transportation

Abstract

'Scramble crossings' for pedestrians at signal junctions are widely used in Japan and have been reintroduced in Canada and the United States as a way of prioritising pedestrian movement by stopping all traffic movement and allowing pedestrians to cross in every direction at the same time. Some examples exist in the UK but their use is not widespread. One potential reason for this is there is limited guidance on when 'scramble crossings' should be considered as potential schemes for promoting pedestrian priority.

This paper focuses on the proposed scramble or diagonal crossing scheme at Oxford Circus in central London. This has been designed by Atkins on behalf of The Crown Estate, Transport for London (TfL), Westminster City Council (WCC) and the New West End Company (NVEC) and was implemented during 2009.

The purpose of the paper is to identify some of the existing examples in the UK and overseas and review the existing guidance on their application of scrambled crossings. The paper then describes the design process and key features of the Oxford Circus scheme and based on this experience concludes by suggesting potential future applications of scramble crossings in the UK.

Introduction to scramble crossings

For the purpose of this paper the term 'diagonal crossing' will be used as well as 'scramble crossing' to describe signalised crossings which have an 'all red' stage and where pedestrians are encouraged to cross in all directions.

Scramble crossings for pedestrians involve stopping all traffic movements at signalised junctions and allowing pedestrians to cross in every direction at the same time. Their origins are unclear but they are also known as a 'Barnes Dance' after Henry Barnes, a traffic engineer in the United States, who popularised the concept after overseeing their introduction in Denver and New York.

The general consensus is that the first scramble crossings were introduced in Vancouver and Kansas City in the 1940's and spread to other cities in North America including Los Angeles in the 1950's which at one point had 25 examples. Over 300 scramble crossings now exist in Japan including the 'Shibuya crossing' situated close to Tokyo's Shibuya railway station which is said to be used by over 250,000 pedestrians every day. This is probably the most widely known example of a scramble crossing in the world.

The author considers that the popularisation of the motor car in the post-war period resulted in the withdrawal of some of the scramble crossings particularly in North America. However, the acknowledgement of the need to promote more sustainable modes of travel together with increased urbanisation has led to scramble crossings being considered once again as a way of dealing with the conflicts between pedestrians and vehicles in the centre of our cities.

New scramble crossings have recently been introduced in Oakland (2002 at 8th Street / Webster Street) and Calgary (2003 in Quartier International district) and Toronto (2008 at Yonge Street / Dundas Square).

Potential advantages and disadvantages of scramble crossings

A number of papers have been published on the theoretical and actual benefits and disbenefits of introducing scramble crossings including papers by; Ahuja et al 2008¹, Ishaque and Noland 2006² and Bechtel et al 2003³.

Through review of these papers and the author's experience of developing the proposed scheme for Oxford Circus the key advantages of introducing scramble crossings include the following:

- Promotion of pedestrian priority and the relief of pedestrian congestion on more traditional orthogonal crossings and footways particularly where pedestrian volumes are very high.
- Reduction of walk distances and times particularly where pedestrians would otherwise use two orthogonal crossings to reach their intended destination and can now complete their journey through the junction by making a single diagonal crossing movement.
- Potential improvements in safety by reducing conflicts between pedestrians and vehicles. Bechtel et al 2003 demonstrated in their analysis of the Oakland scramble crossing that safety benefits were realised by the introduction of a scrambled crossing. The North American examples of scramble crossings have particular

benefits of reducing the conflict between pedestrians crossing and vehicles turning left and right ‘on red’ which is legally permitted in the United States. But this is not permitted when a scramble crossing operates.

Ahuja et al 2008 concluded that scramble crossings are beneficial only when both pedestrian and vehicular volumes are relatively high and the junction is characterised by significant delay for both modes of travel. In free flowing conditions for vehicles and pedestrians, they concluded a conventional crossing arrangement would be a better solution as scrambled crossings generally lead to an increase in delays at the junction. This may in turn lead to an increase in pedestrians not complying with the signals and therefore an increase in conflict between pedestrians and vehicles.

The potential drawbacks of introducing scramble crossings for pedestrians including the following:

- Increased delays to vehicles particularly where an ‘all red’ signal stage has to be introduced.
- The effectiveness of scramble crossings relies on pedestrian compliance with the signals and several examples are considered to suffer from pedestrians failing to ‘clear’ the junction at the end of the crossing periods leading to potential additional delays to vehicles.
- Increase in the perceived risk to pedestrians by the removal of barriers and guardrails to allow crossing in all directions. The key consideration is the potential risk of pedestrians waiting to cross on diagonal movements at the corners of junctions being struck by vehicles turning left around the corners.

Status and examples in the UK

The basic prerequisite for introducing scramble crossings exists at a large number of signal junctions in the UK through the provision of an ‘all red’ traffic stage during which all traffic movements are stopped and pedestrians use all marked crossings at the same time. However, in the UK controlled pedestrian

crossings at signal junctions are usually of the orthogonal / straight across or staggered type: these provide crossings of the traffic approaches to the junction.

In the UK the Department for Transport (DfT) sets out the design standards and guidelines for pedestrian facilities at signal controlled junctions. Design Manual for Roads and Bridges (DMRB) TD50/04⁴ states that at traffic signal junctions ‘where a pedestrian need is established then appropriate signal controlled facilities should be provided’. Traffic Advisory Leaflet 5/05⁵ describes the main options for providing pedestrian facilities at traffic signal junctions. This states that:

“the designer has to consider the pedestrian flow patterns, degree of saturation and the topographical layout to decide on which option is best suited to a particular site”.

The options identified include underpasses and overbridges; no pedestrian phase or stage; full pedestrian stage; parallel pedestrian facility; staggered pedestrian facility or displaced pedestrian facility. With respect to providing a full pedestrian stage, the advice from the DfT in terms of diagonal crossings is as follows:

“Diagonal crossings (crossing the centre of the junction, say, from north east to southwest) are largely untried but a small number do exist. There are important design aspects to be incorporated. Diagonal crossings are not considered appropriate for many disabled users, particularly those who are visually impaired. Also road safety education generally teaches children not to cross diagonally at junctions. Conventionally orthogonal crossing places should therefore always be provided with flush dropped kerbs, tactile paving and audible / tactile signals. Flush dropped kerbs, tactile paving and audible and tactile signals should NOT be provided on the diagonal crossing part. If a lowered kerb is provided there should be a minimum upstand (after possible resurfacing) of at least 25mm. Careful thought also needs to be given to the use of markings or coloured surfacing at the junctions so that partially sighted pedestrians are not misled”.

Although the advice is therefore prescriptive in terms of the design,

little guidance is given to allow the designer to determine whether a diagonal crossing layout should be used as opposed to a more traditional arrangement with orthogonal crossings. This along with the perceived risk of not following guidance may be a key factor as to why diagonal crossings are still the exception to the norm in the UK.

Current UK examples of scramble or diagonal crossing layouts at signal junctions include:

- Aberdeen – Scrambled crossings at several locations across the town centre.
- Balham High Road / Balham Station Road junction, Balham, South London – urban signal junction located directly outside Balham London Underground Station with two diagonal crossings as shown in Figure 1. This is the first example of a diagonal crossing arrangement introduced in London.
- A240 / A2022 Drift Bridge, Epsom Downs, Surrey – suburban signal junction providing access to a local centre with single diagonal crossing from north-east to south-west corner only.
- Balliol Road / Pembroke Road, Bootle, Merseyside – urban signal junction with single diagonal crossing connecting two separate sites of Hugh Baird Higher Education College (introduced 2004-2005).
- Oxford Road / Dover Street, Manchester – urban signalised staggered crossroads in University area.

The original justification for providing diagonal and scramble crossings at these locations is not known. However, justification appears to be either a result of the existing layout (such as a staggered crossroads) which lends itself to a diagonal crossing and / or because they are in local and town centres, where pedestrian connectivity is important and pedestrian flows are relatively high.

There are also a number of existing junctions where traditional orthogonal crossings are provided at signalised crossroads and where the absence of barriers to movement such as guardrail permit ‘informal’ diagonal crossing. This includes the A24

'Scrambled' pedestrian crossings – signal controlled junctions – a case study



Figure 1 – Diagonal crossing in Balham, South London



Figure 2 – Informal diagonal crossing in Epsom town centre

High Street / Waterloo Road / Ashley Road junction in Epsom town centre close to Atkins' offices and shown in Figure 2. This is situated in an outer London suburban town centre but also on the A24 London to Worthing Trunk Road. Such examples represent potential sites where formal diagonal or scramble crossings could be introduced.

The Oxford Circus case study

London's West End has been one of the UK's dominant retail and entertainment centres since the early 19th century. It continues to be a vibrant place with a distinctive retail offer and range of attractions and is renowned on an international level. However, its popularity

presents unique problems for its infrastructure and public realm including congested streets, vehicular and pedestrian conflicts and a tired looking street environment.

Existing junction layout (2008)

Oxford Circus is at the heart of the West End and one of the most renowned junctions in the world, marking the convergence of London's two most famous retail streets, Oxford Street and Regent Street.

The junction is a signalised crossroads with traditional orthogonal crossings for pedestrians on all four arms, each of which include central islands as shown in Figures 3 and 4. The crossings are 'straight across' with pedestrians crossing either Regent Street or Oxford Street in

one continuous movement – this is facilitated by the provision of an 'all red' traffic stage during which every pedestrian crossing operates. Each of the four crossings are set back between 7 and 15 metres from the junction which means they do not serve the key pedestrian desire lines along the footways on either side of the two streets. Stairways to Oxford Circus London Underground station are located at each of the four corners of the Circus and stone balustrades mark the kerbs around each of the corners preventing crossing by pedestrians anywhere other than at the formal crossings. The layout of the London Underground stairways and the balustrades means that pedestrians cannot walk between them creating areas of 'dead space'. Guardrails are used in the vicinity of the orthogonal crossings on the central islands and further down each approach to the junction.

Both Regent Street approaches have two lanes of traffic, as does the approach from Oxford Street west which has one lane dedicated to right turning buses. There is a single lane on the Oxford Street east approach. A number of turning movements are banned including the right turns from both Regent Street approaches and Oxford Street east as well as the left turn from Regent Street north. The Oxford Street approaches are generally perceived to be for the use of buses and taxis only during 0700-1900 hours but in actual fact service vehicles are permitted to make certain turns to and from Oxford Street. Separate traffic stages are provided for Regent Street and Oxford Street with an extended phase for the right turning buses from Oxford Street west.

Current use and operating conditions

Oxford Circus is dominated by pedestrians and is extremely busy at most times of day, particularly at weekends. Typically, pedestrians account for 64 percent of all people passing through the junction. Bus passengers account for the next highest proportion, at 32 percent, with around 1 percent of people passing through in private cars, taxis, on motorcycles and by cycle.

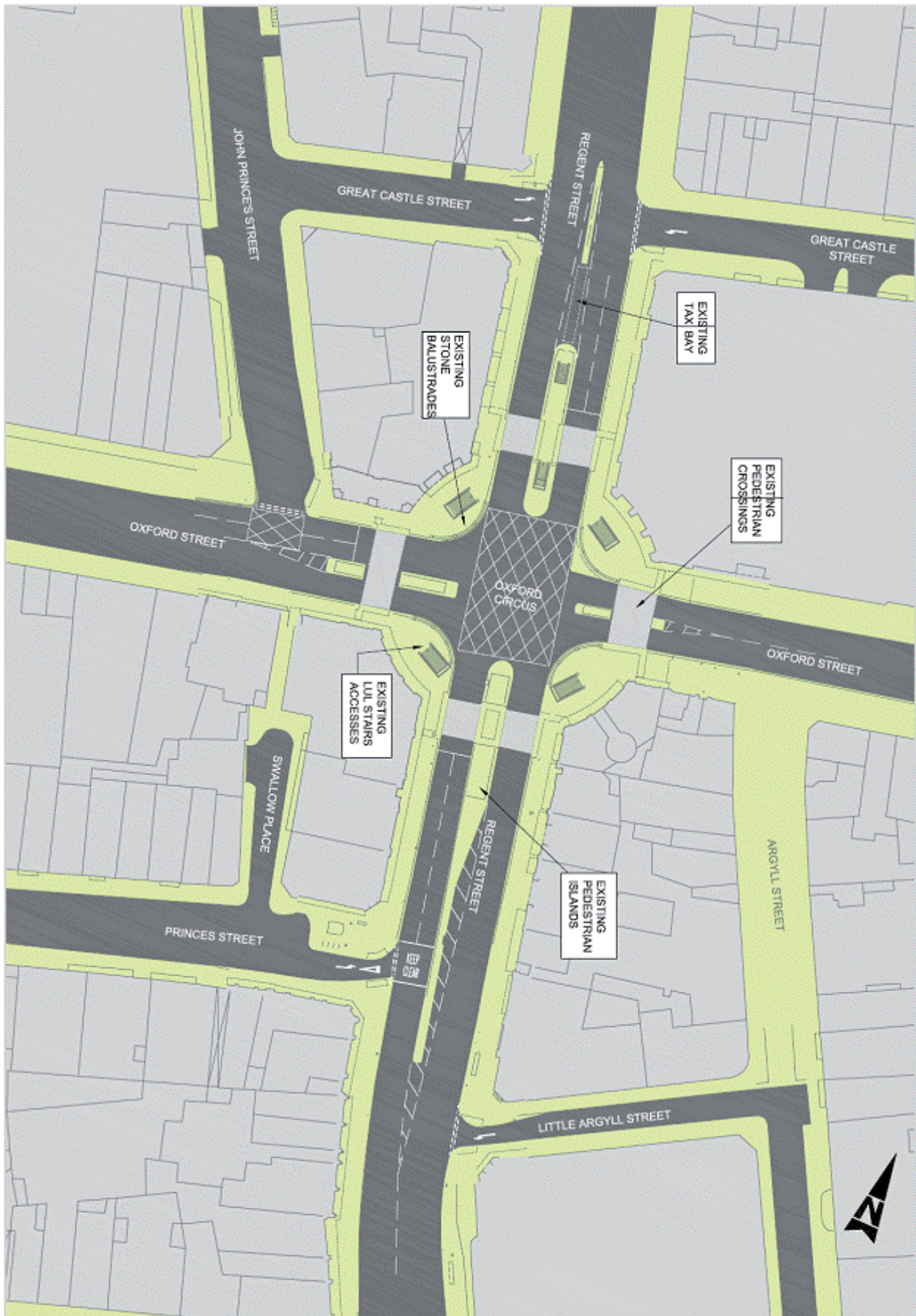


Figure 3 – Plan of existing Oxford Circus layout (before scheme)

'Scrambled' pedestrian crossings at signal controlled junctions – a case study



Figure 4 – Oxford Circus layout (before scheme) – view towards north east corner (Nike Town)

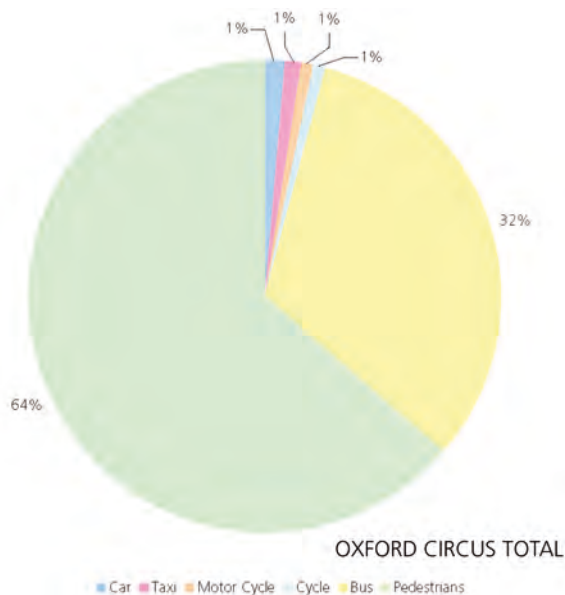


Figure 5 – Surface mode share at Oxford Circus: Weekday PM Peak Hour (Source – 2006 traffic and pedestrian counts)

Oxford Circus has the highest pedestrian volumes recorded anywhere in London. At the busiest times, over 40,000 pedestrians per hour pass through the junction which includes those accessing the London Underground station. The existing crossings are each used by between 5,500 and 11,000 pedestrians per hour with the crossing of Regent Street south generally the busiest. A combination of the high volumes of pedestrians, together with the constrained nature of the space and the prevalence of street clutter and

stationary pedestrians means high levels of pedestrian congestion are a normal part of the Oxford Circus experience for large parts of the day.

Fruin⁶ Levels of Service (LoS) D to F have been calculated on all of the footways in the vicinity of Oxford Circus including those between the London Underground station accesses and building lines as shown in Figure 6. These LoS represent extremely congested conditions for pedestrians which dissuade walking at and through the Circus and ultimately dissuade visitors to the West End.

The four pedestrian crossings are also extremely busy and do not always clear of pedestrians at the end of the crossing stage. Many of the existing pedestrians use two crossings as part of their journey through the junctions and would therefore be well served by a diagonal crossing.

A number of vehicle turning movements are banned which results in 64 percent of pedestrians crossing during the 'red man' phase when either Regent Street or Oxford Street traffic movements are on 'green'. Conflicts between pedestrians are also an issue, for example between those accessing the London Underground and others passing through the junction using the footways and crossings. Faced with these conditions a small number of pedestrians opt to walk in the road on the 'traffic side' of the balustrades and guardrail as shown in Figure 7.

Both streets are major traffic thoroughfares - typically, hourly traffic flows peak at approximately 2,000 vehicles per hour during the week with around 500 to 700 vehicles per hour on Regent Street in each direction and 200 to 400 vehicles per hour on Oxford Street in each direction. The junction is generally considered to be operating at or close to capacity for a large part of the day with queues evident on all approaches but particularly Oxford Street west. The junction is a key 'node' in the network of traffic signals in this part of the West End.

Buses account for a significant proportion of the existing traffic at Oxford Circus: a total of 24 bus routes pass through the Circus with over 400 buses per hour on a typical weekday. Several of the routes terminate in the area and pass through the junction more than once in accessing and egressing from their stands. Oxford Circus is highly accessible by bus and is also a key interchange between the various bus services and between buses and London Underground. Bus stops are located within 200 metres of the junction on all four approaches.

Scheme development

Westminster City Council, The Crown Estate, Transport for London and the New West End Company, together with other key stakeholders, developed and adopted the 'Oxford,

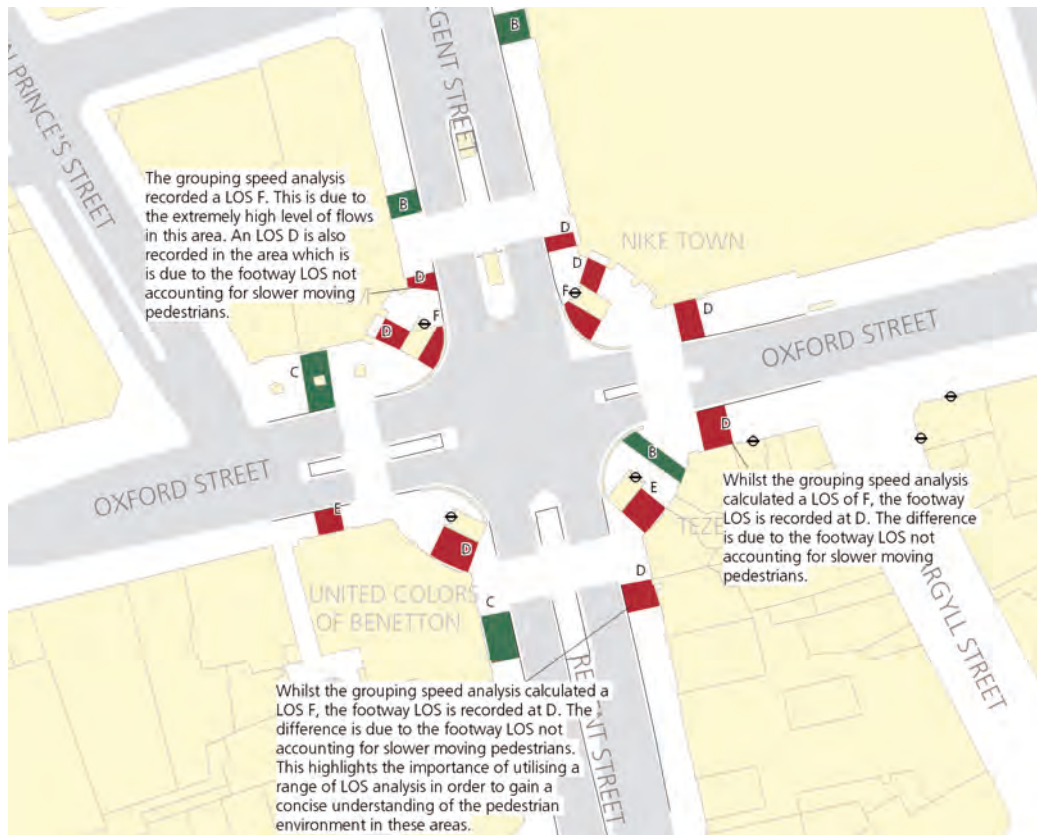


Figure 6 – Fruin level of service analysis of existing layout



Figure 7 – Pedestrians walking in the carriageway

Regent and Bond Street (ORB) Action Plan, a vision for the future development of the West End to safeguard its position as one of Europe's finest shopping districts. The action plan puts forward a number of proposals to improve the quality of the public realm and to enhance the visitor experience.

As part of the ORB Action Plan, Atkins (working for The Crown Estate) proposed a comprehensive redesign of Oxford Circus to provide more space for pedestrians through reduced street clutter, selective footway widening and diagonal (scramble) crossings. The overarching objective was to transform Oxford Circus into a rejuvenated space befitting its status as gateway to the West End and address its problems in terms of discouraging visitors, shoppers and tourists.

A plan and image of the proposed scheme are presented in Figures 8 and 9 - the key elements are as follows:

- Two new diagonal crossings for pedestrians.
- Straight across crossings realigned to better serve the pedestrian desire lines with narrowed central islands.

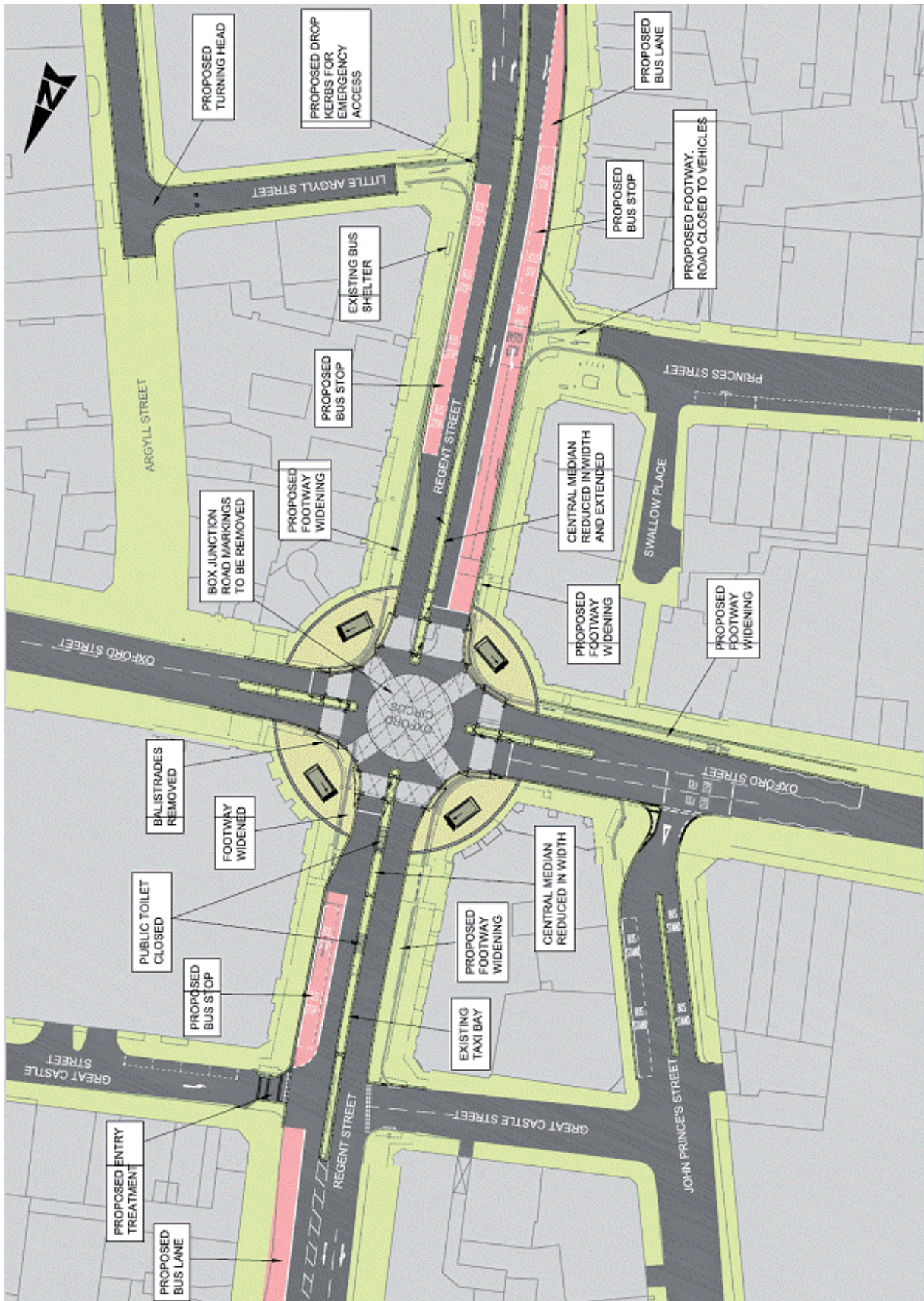


Figure 8 – Plan of Oxford Circus scheme



Figure 9 – image of Oxford Circus scheme – view towards north east corner (Nike Town)

- Footway widening on Regent Street both north and south of the junction which together with the removal of street clutter to create 63 percent more usable space for pedestrians.
- Removal of balustrades and guardrail to facilitate diagonal and scramble crossing.
- Introduction of or improvement to bus lanes on the Regent Street approaches.
- Closure of Princes Street and Little Argyll Street, two side streets south of the Oxford Circus junction.

The new diagonal crossings are the centrepiece of the scheme and were identified as potential solutions to the existing pedestrian congestion on footways and crossings, particularly in the light of the fact that pedestrians outnumber those travelling in vehicles twenty to one and that a large number of the existing pedestrians use more than one crossing to complete their journey through the junction. The perception also was that as an 'all red' traffic stage already existed, that diagonal crossings could be implemented without a significant impact on traffic conditions since the broad order distribution of signal 'green' time between vehicles and pedestrians and the overall cycle time could be maintained. Minimising the traffic impact and the effect on buses in particular was also a key objective of the scheme.

The new diagonal crossings are approximately 4 metres wide and between 25 and 27 metres in length. This compares with the realigned straight across crossings which are 4 to 6 metres wide and between 11 and 15 metres in length. In common with the realigned straight across crossings, the diagonal crossings have dedicated signal heads and studs marking the width of the crossing. It should be noted, however, that as with other 'scramble' crossings already in existence there is an expectation that pedestrians will also cross outside of the stud markings making full use of the space in the centre of the junction between opposing stop lines and their parallel stud marks.

Extensive public and stakeholder consultation was carried out on the design philosophy to be applied to the diagonal crossings. The public consultation gave feedback which was very positive with 95 percent of respondents expressing their support. Consultation with accessibility groups such as Guide Dogs for the Blind helped to influence the design and as a result of this consultation it was decided that different coloured skid resistant surfacing would be used to mark out the diagonal and straight across crossings. It was also decided that whilst tactile paving and dropped kerbs would be used for the straight across crossings, none would be used for the diagonal crossings which would feature a 50mm upstand at the kerb. A compromise was

therefore sought between providing access to all crossings for all users and the desire to not encourage use of the diagonals by the partially sighted who could become confused by the arrangement which is not yet standard practise in the UK.

A pedestrian 'countdown' system was also considered for implementation in parallel with the scheme that would visibly and audibly count down the time in seconds until the end of the pedestrian crossing stage. As this technology is also untried in the UK and requires approval from the DfT it was decided to omit countdown from the scheme for the time being although it will be considered for implementation at a later date.

Guidance and approval was sought from Transport for London's Directorate of Traffic Operations (DTO) who manage the signal junctions within the Greater London Authority area on the appropriate intergreen and pedestrian crossing timings. These have been derived with reference and in accordance with the appropriate guidance (TA 16/81⁷ and TTS6⁸).

It was decided to reduce the pedestrian invitation from the standard 9 seconds to 8 seconds. This decision was taken in order to minimise the impact of the scheme on traffic conditions and therefore bus journey times. A comparison of the existing and proposed signal settings including traffic stage lengths, inter greens and pedestrian stage is presented in Table 1 (for the two signal cycle times currently used in the weekday Inter Peak Hour (1300-1400 hours) and the cycle time used in the weekday PM Peak Hour (1700-1800 hours). The table indicates that the overall length of the pedestrian crossing stage (inclusive of subsequent intergreen) has been increased from 26 to 32 seconds. This increase is associated with the increased length of the longest single crossing to approximately 26 metres (the diagonal from south east to north west) with the scheme in place compared to the existing longest crossing (Regent Street South) which is approximately 17 metres in length. The new diagonal crossing therefore requires a pedestrian crossing stage which is 6 seconds longer than the existing stage inclusive of the subsequent intergreen period.

'Scrambled' pedestrian crossings at signal controlled junctions – a case study

Table 1 - Proposed signal settings

	Cycle Time	Stage 1 Green (Oxford Street Traffic)	Inter stage	Stage 2 Green (Oxford Street West - Right Turn)	Inter stage	Stage 3 Green (Regent Street Traffic)	Inter stage	Stage 4 Green (Pedestrians)	Inter stage
Existing Inter Peak Hour	104 second cycle	21	4	5	9	24	15	9	17
Scheme Inter Peak Hour		16	5	4	6	33	8	8	24
Difference		-5	+1	-1	-3	+9	-7	-1	+7
Existing Inter Peak Hour	112 second cycle	24	4	5	9	29	15	9	17
Scheme Inter Peak Hour		19	5	4	6	38	8	8	24
Difference		-5	+1	-1	-3	+9	-7	-1	+7
Existing PM Peak Hour	112 second cycle	24	4	5	9	29	15	9	17
Scheme PM Peak Hour		22	5	7	6	32	8	8	24
Difference		-2	+1	+2	-3	+3	-7	-1	+7

Scheme appraisal

This increase is offset by reduced inter greens between the traffic-traffic phases and the traffic-pedestrian phase achieved by moving the orthogonal crossings and their associated vehicular stop lines closer into the centre of the junction. Overall, a saving of 8 seconds has been achieved through these changes which (taking into account the change in pedestrian 'invitation' and crossing time) provides a net gain of 2 seconds per cycle for vehicular green periods.

It should also be noted that the proposed closure of the Princes Street side road is predicted to remove approximately 100 vehicles per hour from passing through the Oxford Circus junction as this street is eastbound (towards Regent Street only) and traffic can only currently turn northbound towards Oxford Circus. This is a reduction in total junction throughput of approximately 5 percent.

Whilst the saving in inter green times and the removal of the Princes Street traffic from Oxford Circus has undoubtedly helped to achieve a scheme that minimises the effect on operating conditions, it should be noted that these measures are actually offsetting the effect of the footway widening (and corresponding narrowing of Regent Street northbound and southbound from two lanes to one lane) as well as the effect of the longer diagonal crossings.

A significant investment was made in undertaking detailed appraisals of the proposed scheme for a number of reasons:

- The innovative nature of the scheme particularly the diagonal crossing and scramble zone.
- The high profile nature of Oxford Circus and its role as the key junction in the West End and the surrounding network of traffic signals.
- The large number of stakeholders involved in any such scheme in central London.

Although the development of the scheme was predominantly funded by The Crown Estate, the costs of its implementation will be shared between The Crown Estate, Transport for London and Westminster City Council. For this reason a standard TfL Business Case was required to evaluate the financial benefits of the scheme. This, in itself, required detailed appraisals of the effects of the scheme on the various users that currently pass through the Circus, particularly, pedestrians, buses and other vehicles.

Atkins Intelligent Space undertook a detailed manual analysis of pedestrian movements including Fruin Level of Service (LoS) as well as developing Legion pedestrian micro-simulation models of the Circus. Atkins Transport Planning developed a parallel VISSIM vehicle micro-simulation model to evaluate the impacts on overall

junction operating conditions and the impact on particular vehicle classes. This encompassed the prediction of before and after journey times for buses and other vehicles, predicted queues and predicted Degrees of Saturation for traffic. The VISSIM model included the full length of Oxford Street from Marble Arch in the west to St Giles Circus in the east and Regent Street from Langham Place in the north to Great Marlborough Street / Maddox Street in the south. Both models were audited and approved by the relevant departments within TfL.

Detailed analysis of the pedestrian movement issues predicts that the scheme will have the following effects:

- Average walk times through the junction will be reduced by 52 seconds by virtue of reduced pedestrian congestion and the provision of the diagonal crossings which will allow many pedestrians to undertake a single crossing movement rather than use two of the existing straight across crossings in succession.
- Fruin LoS on the footways in the immediate vicinity of Oxford Circus are predicted to improve from Level of Service between D and F to Levels of Service between A and C as shown in Figure 10. This represents a significant improvement from the exiting congested environment to busy but generally free flowing conditions for pedestrians.
- Typically the new diagonal crossings will attract use by up to 6,500 pedestrians per hour who will be abstracted from the four straight across crossings improving conditions there.
- The configuration of the space on the footways and around the London Underground stairways will provide separate 'wait' zones for pedestrians congregate to use all of the crossings and 'through' zones for walking pedestrians. Movement either side of the London Underground stairways will also be possible. This is illustrated by the images from the Legion pedestrian model presented in Figure 11.

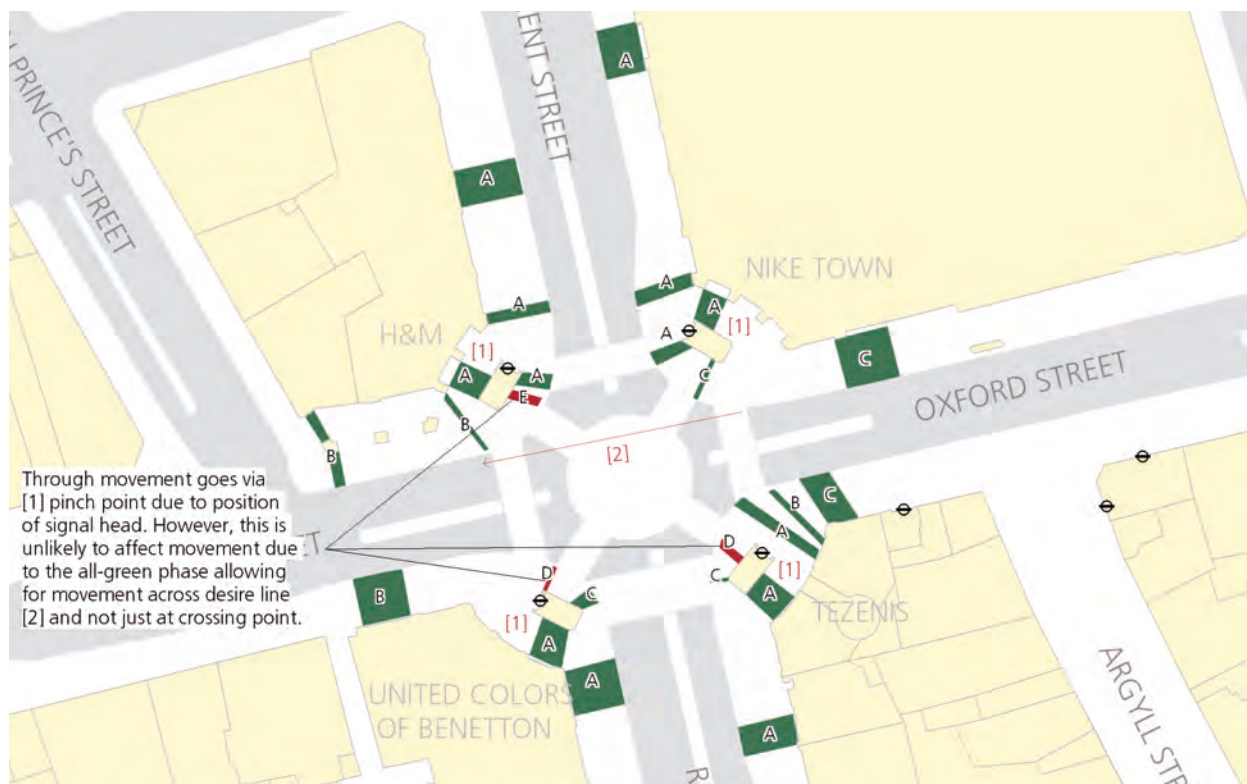


Figure 10 – Fruin level of service analysis for Oxford Circus scheme

Whilst the diagonal crossings are only one element of the scheme, they are key to its success: without them the realigned crossings are predicted to be overcapacity and function in much the same way as the existing straight across crossings. It is notable that the Legion model can only simulate the current conditions if pedestrians are permitted to cross during the red man phase (as actually happens) whilst the scheme can operate effectively assuming 100 percent compliance by pedestrians.

A VISSIM micro-simulation model, an image from which is shown in Figure 12, has been used to demonstrate that the scheme could deliver the significant benefits for pedestrians and the public realm without a significant impact on junction capacity and knock-on effects on journey times for traffic and buses.

Case studies of other schemes have indicated that the introduction of diagonal crossings can have a significant impact on traffic. This is not predicted to be the case at Oxford Circus. Although the signal setting standards have been relaxed by reducing the ‘invitation to cross’ to 8 seconds this has in effect been one of the measures used to minimise the impact of the overall scheme. The proposed footway

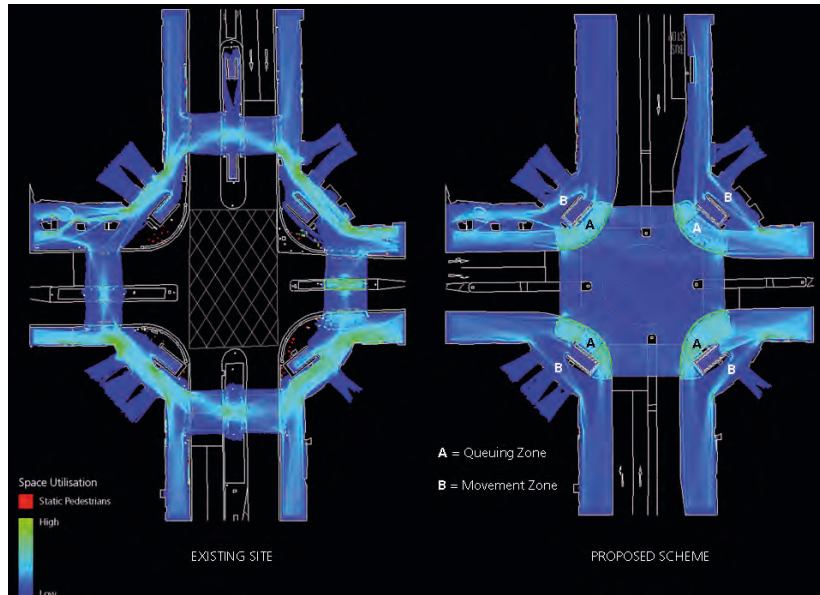


Figure 11 – Legion space utilisation plans

widening is considered to have the greatest impact as it reduces the Regent Street southbound movement from two lanes of traffic to one. The reduction in traffic-traffic and traffic-pedestrian inter green times and the removal of the Princes Street side road traffic has also helped to mitigate this potential effect. New or improved bus lanes have been provided on both sides of Regent Street to provide buses with priority on these approaches to the Circus.

In this context and based on the signal settings presented in Table 1, the VISSIM model has predicted the following impacts on traffic and buses of the scheme:

- With the exception of the Princes Street side road traffic which will be removed from Oxford Circus, all other traffic currently passing through the junction will still be able to do so on a typical day.

'Scrambled' pedestrian crossings – signal controlled junctions – a case study

- Very little change in bus journey times is likely to occur – average journey time improvements of 7 to 15 seconds are predicted for buses during the weekday Inter Peak Hour and weekday PM Peak Hour. Journey time improvements are predicted on at least five of the eight possible bus turning movements during these periods.
- Very little change in the journey times for other vehicles is likely to occur – journey times are predicted to change by between +3 and -8 seconds during the two peak hours. Journey time improvements are predicted for around half the turning movements during these periods.
- The length of traffic queues is predicted to be similar to existing. Although small increases in queue lengths are predicted for some approaches this will be offset by the relocation of the stop lines at Oxford Circus which will provide greater 'stacking space' between the Circus and adjacent junctions. The probability of 'blocking back' of queues occurring is therefore similar to the existing.

Notwithstanding that the impact on traffic and buses is essentially 'neutral', increasing the cycle time from the current 112 seconds to 116 or 120 seconds is now being considered by TfL to provide the highway network with greater resilience to un-planned events or accidents at Oxford Circus or nearby.

The overall financial benefits and disbenefits generated by the scheme as estimated by the Legion and VISSIM models and subsequently included in the TfL Business Case results in £6m of benefits per annum which can be disaggregated as follows:

- £5.1m - benefits for pedestrians per annum
- £800,000 - benefits for bus users per annum
- £10,000 - disbenefits for car users per annum
- £10,000 - benefits for taxi users per annum

The overall result is a Benefit – Cost Ratio (BCR) for the scheme of 12.9:1. This is a very healthy BCR as typically TfL is aiming to achieve a BCR of at least 1.5:1 for schemes that it is funding. The pedestrian

Figure 12 – VISSIM Microsimulation model of the scheme

benefits significantly outweigh the benefits and disbenefits for other modes given the huge number of pedestrians passing through the junction. Despite this the business case results do suggest that diagonal crossing solutions could represent viable schemes elsewhere even recognising that most locations in the UK will have significantly fewer numbers of pedestrians.

Overall conclusions & future applications of scramble crossings

The scheme at Oxford Circus includes an innovative scramble crossing arrangement which is untried as a solution in such a high profile location in the UK where a constrained environment is coupled with very high pedestrian and vehicular demand.

Limited guidance exists in the UK on scramble crossing solutions and this tends to focus on the design aspects of scramble crossings instead of identifying the circumstances in which they should be considered as a way of promoting pedestrian priority. Based on the experience of developing the scheme for Oxford Circus and reviewing previous case studies, suitable locations in the UK for scramble crossings could potentially include:

- Town or local centres where there are crowded footways or crossings; where pedestrian flows are significant and pedestrian connectivity needs to be improved but where an existing junction for vehicles also needs to be maintained.

- Locations where a diagonal pedestrian desire line exists but is not served by the existing crossings, for example between two trip generators or attractors located on diagonally opposite corners of a signalised cross roads. In these circumstances, the scramble crossing as the potential to reduce walk times and distances as well as reduce the pressure on existing footways and crossings.
- Other locations where a signalised crossroads exists with an 'all red' traffic stage where further improvements for pedestrians would be welcome perhaps in relation to a new development. In this situation, benefits can be provided for pedestrians whilst potentially avoiding significant impacts on vehicle flows and delays. It is accepted, however, that this is likely to be highly site specific.

The Oxford Circus scheme was implemented at the end of 2009. It is anticipated a monitoring exercise will be undertaken in accordance with the approvals for the scheme obtained from Transport for London. This exercise will examine and record the actual impact of the scheme on junction operating conditions, particularly impacts on buses and other vehicles as well as the changes in pedestrian behaviour. The results of the monitoring exercise will help to further clarify the potential of introducing scramble crossings elsewhere.



References

1. Ahuja. S, S Bose, T van Vuren and D Ragland, Towards the Provision of Ultimate Pedestrian Priority: Guidelines for Installation of Scramble (all red phase) Pedestrian Signals at Intersections Paper submitted for presentation at the European Transport Conference, the Netherlands 2008
2. Ishaque, M.M and R.Bm Noland, Micro-simulation Comparisons of Alternative Signalized Pedestrian Crossings, Paper submitted for presentation at the Transportation Research Board Annual Meeting, Washington D.C 2006
3. Bechtel, A.K, K.E. MacLeod and D.R. Ragland, Oakwood Chinatown Pedestrian Scramble: An Evaluation, UC Berkeley Traffic Safety Centre, December 2003
4. TD 50/04, The Geometric Layout of Signal-Controlled Junctions and Signalised Roundabouts, Design Manual for Roads and Bridges (DMRB), Vol. 6, Section 2. The Stationery Office (TSO). November 2004
5. Traffic Advisory Leaflet (TAL) 5/05, Pedestrian Facilities at Signal Controlled Crossings, Department for Transport (DfT) (March 2005)
6. Fruin, J.J, Pedestrian Planning and Design. Alabama: Elevator World Inc, 1987
7. TA 16/81, General Principles of Control by Traffic Signals, Design Manual for Roads and Bridges (DMRB), Vol. 8, Section 1. The Stationery Office (TSO) September 1981
8. TTS6, Design Standards for Signal Schemes in London, Transport for London, August 2002

Delivery of bus priority projects - A partnership approach

"Tragically the author of this paper, Cressida Spachis, passed away on 6th March at the age of 33 following a pulmonary embolism. The paper is published not just in memory of Cressida, but as a tribute to her contribution to Atkins and the industry.

Cressida was a Group Leader and Deputy Business Manager for the London Networks & Traffic (N&T) business unit within Transport Planning & Management, and had been with Atkins since 1999. She joined us as a first class honours graduate in Civil Engineering from Imperial College, and during her career worked on a variety of projects for public and private clients across the UK as well as a number of secondments. Over the years, her clients included Transport for London, BAA, LOCOG and a range of local authorities.

Cressida was Project Manager for the Route 38 Corridor Management Pilot Study which has developed principles for intensified bus priority now being used across London. The project received many accolades and Cressida was due to be part of the client/consultant team receiving the Improvements to Bus Services award at the London Transport Awards earlier this month.

Cressida was enormously talented, bright and well-respected, not only as a group leader but also as a team player, displaying a passion for her work and a vitality that endeared her to all who worked with her. She will be greatly missed. Her contribution to Atkins will be remembered for many years to come."



Dr Andy Southern

Managing Director
Transport Planning & Management
Highways and Transportation

Delivery of bus priority projects - A partnership approach



**Cressida
Spachis**

Managing Consultant
Highways & Transportation

Abstract

Population and employment growth in London is projected to continue as part of a national and Europe wide trend seeing an increased proportion of the population living in cities. In London, the residential, commercial and non-commercial densities are already high and the historic road layouts can not accommodate increased trips by private vehicle and this situation presented transport planners with a challenge.

Introduction

Population and employment growth in London is projected to continue as part of a national and Europe wide trend seeing an increased proportion of the population living in cities. In London, the residential, commercial and non-commercial densities are already high and the historic road layouts can not accommodate increased trips by private vehicle and this situation presented transport planners with a challenge.

In 2003, Transport for London (TfL) London Buses carried out a Strategic Review showing that the bus service in London provides an indispensable and increasingly valuable service to the whole of London, particularly in social inclusion terms. Key conclusions included, but were not limited to:

- The (at the time) current subsidy of London Buses delivers real 'value for money';
- The bus is the only means of increasing capacity rapidly to meet the challenges of delivering the Mayor's London Plan and forecasted population and employment growth;
- The service planning guidelines used to determine the London bus network are

based on a robust benefit/cost framework which reflect passenger requirements; and

- Other than the further extension of Congestion Charging (or road pricing) and other real traffic restraints, the clearest potential cost savings are in Intensified Bus Priority (provided that enforcement is effective and sustained) and delivery of the Cashless Bus strategy.

(Source "London Buses, MAYOR OF LONDON, The case for investing in London's buses, Presenting the Results of the London Buses Strategic Review, September 2003")

In line with this Strategic Review, TfL Surface Transport Bus Priority Team (BPT) promoted two Intensified Bus Priority projects on the bus Routes 149 and 38, aiming to introduce a 'step change' in bus priority. Bus Route 149 was mostly operated on the Transport for London Road Network (TLRN), whilst bus Route 38 was mostly operated on Borough roads. The Route 149 project was not proceeded with.

In 2003, the Route 38 "Intensified Bus Priority" Pilot Project was set up and jointly promoted by TfL BPT and the London Bus Priority Network (LBPN) Central Sector. The focus of the

project changed in 2005 to that of a "Whole Route Corridor Management" Pilot Project (now referred to as the Route 38 Pilot Project).

This paper highlights the importance of a strong partnership approach in developing and delivering the Route 38 Pilot Project, a project that has undergone changes in management, approvals and network conditions throughout its life.

Route 38 operational conditions

Route 38 operates in London, between Victoria (Westminster) and Clapton Pond (Hackney) through the Boroughs of Camden and Islington. This area of the capital is very densely populated and the services are in heavy demand for commuting and social use including tourism. Route 38 is a 12.5km long, high-frequency corridor with 25-30 Route 38 buses per hour, as well as other routes, and serves up to 7000 passengers per hour.

Route 38 was converted from Routemasters to articulated buses in October 2005, and it was announced that it will be converted back to double deck buses at the end of 2009.

The reliability and journey time of Route 38 and other bus services on the corridor was poor and buses were subject to excessive delay and disruption at many locations. Much of this was due to the lack of highway capacity, conflicting demands of other road users, inappropriate bus stop layouts, limited waiting and loading restrictions, junction and side road disruptions, inadequate enforcement, and the fact that bus priority measures were present in some sections of the corridor only, and often limited to peak hours.

The Route 38 Pilot Project

Key objectives

The Route 38 Pilot Project was set up to provide a 'step change' in bus priority through Intensified Bus Priority schemes, building on measures previously implemented under the London Bus Initiative (LBI) and the LBP. Following the change of focus to that of Whole Route Corridor Management, the project aimed to improve bus reliability and journey times, whilst adopting a holistic approach as far as possible through balancing the competing user demands and providing benefits to other users as well as buses. Some examples include to:

- Enhance the environment along the route corridor;
- Improve safety and facilities for pedestrians and cyclists;
- Create a more pleasant and safe bus stop environment to make it easier for all passengers to get on and off the bus; and
- Provide additional benefits, where possible, for all road users, communities and businesses.

Key characteristics

The Route 38 Pilot Project started in May 2003 and is due for completion in March 2010. During its lifetime, the project has encountered a number of changes. Partnership working with the Client, the various Partners and Stakeholders proved to be key to the change management and successful project delivery.

The project involved approximately a £20m investment programme incorporating measures for bus priority, streetscape, pedestrians, cyclists, bus stop accessibility, enforcement, safety, signals (SCOOT, AVL), Selective Vehicle Detection (SVD) and iBus, converting the Routemaster to articulated bus operation, and innovative ideas for traffic management.

The importance of the project has been demonstrated by the fact that TfL has continually used it as a test bed for the Third Generation of Bus Priority (3GBP). The link has been achieved through TfL's 3GBP Project Manager attending the fortnightly formal progress meetings of the Route 38 project, and when relevant also the fortnightly technical progress meetings, and feeding back to the 3GBP the experience gained and lessons learnt.

The Project Partners

The Route 38 Pilot Project is being promoted jointly by TfL BPT and the City of London, the Lead Borough of LBP. Central Sector, and in partnership with the four London Boroughs the route goes through, namely Westminster, Camden, Islington and Hackney, who are the highway authorities for much of the network concerned.

The project structure is shown in the Figure 1 below.

When initially set up, the SRO for the project was the Head of TfL BPT. However, in order to foster a partnership approach, the SRO role was transferred to the LBP. Central Sector, and was undertaken by the Head of Transportation for City of London.

In addition to the formal Partners of the project, relationships were developed with other partners, referred to as Stakeholders. These included TfL departments such as the Public Carriage Office, Cycling Centre of Excellence and London Underground (LU), as well as Borough departments other than the delivery Partners, and Developers.

The existing road layout has emerged through a process of refinement and competition between the modes and the various interests of frontages and the Local Authority. In order to make changes it is essential that all interests are properly reflected in any planned changes.

Key delivery Partners included TfL's Network Assurance Team (NAT) and Department of Traffic Operations (DTO) as well London Buses Area Traffic Controllers and the Borough's Street Managers. In addition, close working was important at a local level with town centre managers, other consultants and, vitally, Borough Councillors and the public.

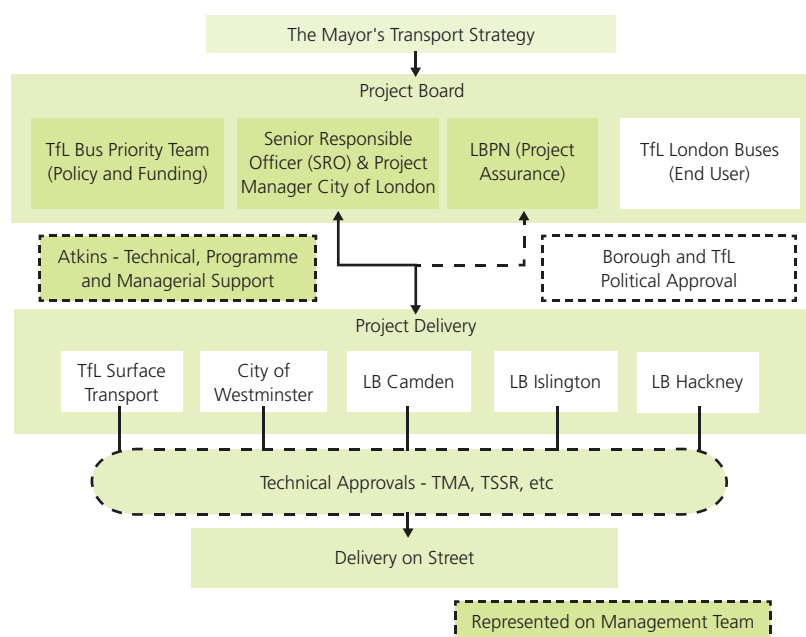


Figure 1 - Project structure



Figure 2 - Theobald's road footway and bus stop improvement

The regular day-to-day management team of the SRO, the LBP Senior Route Co-ordinator, the 3GBP Project Manager and Atkins was small. However, it was vital that the management team was fully committed and involved with all project Partners and Stakeholders, whether end users or delivery.

By following a partnership approach working closely with Stakeholders and Partners, the project was able to tailor schemes, approaches and methodologies to the preferences, requirements and needs of the ultimate end users, the general public, whether travelling or just using the street.

Where a Borough had, for example, a safety priority to introduce an All Red Pedestrian Phase at a junction, the scheme was adjusted to include it whilst at the same time maintaining bus benefits. Or, where a Borough required provision of loading bays for businesses, the scheme was developed to remove the bays from the carriageway in order to improve bus and traffic flow, but bays were provided as inset or shared surface with widened footways, hence not only accommodating businesses, but also providing additional pedestrian facilities. Another instance would be where a Borough had inspirations to create an attractive town centre area, a scheme was developed diverting general traffic away from the town centre, but retaining buses through it, hence reducing bus journey times whilst improving the public realm and pedestrian environment.

Change control and management

During its lifetime, the project encountered a number of changes in management, approvals and network conditions, which influenced the project progress, programme, and delivery. Some of the changes encountered and managed include, but are not limited to, the following:

- The introduction of new approvals processes, including the introduction of the Traffic Management Act 2004 (TMA), and subsequently the introduction of the TfL Network Assurance Team (NAT); and the new TfL Department of Traffic Operations (DTO) processes and signals guidelines;
- The project focus changed in 2005 from an Intensified Bus Priority to a Whole Route Corridor Management. This meant that the project was no longer purely a bus priority project, but needed to adopt a holistic approach as far as possible, hence balancing the competing user demands including those of pedestrians, cyclists, and general traffic;
- The conversion of Route 38 buses from Routemaster to articulated in 2005, requiring the redesign of all schemes and bus stops along the route to cater for them;
- The project Governance was transferred in June 2006 from TfL to the City of London, highlighting the importance of the Borough partnership;

- Local Council Elections in 2006 leading to a change of political lead in certain Borough Partners, hence a change of Council priorities;
- The introduction of the Western Extension Zone to the London Congestion Charging in 2007, leading to changes in the traffic patterns of Central London; and
- The development of a new TfL Business Case Assistant in 2007 requiring new business cases to be developed for all sections along the corridor.

The sections below describe the project stages and how they have been impacted upon by these changes, and highlight how partnership working was key to the change management and to the project delivery.

Project stages

Although a higher profile project than some, the Route 38 project still covers the lifecycle of a standard traffic engineering scheme, divided into the following stages:

- (1) Scoping;
- (2) Feasibility (including public consultation);
- (3) Preliminary Design (also including consultation);
- (4) Detailed Design;
- (5) Implementation; and
- (6) Monitoring.

The route was initially divided into 13 sections for ease of scheme development, consultation, approvals and implementation.

Scoping

The scope of the project was set by TfL at the initiation, in 2003, and was to investigate the use of intensive measures to give priority to buses. The scope was changed to a demonstration of corridor management and a pilot for 3GBP in 2005.

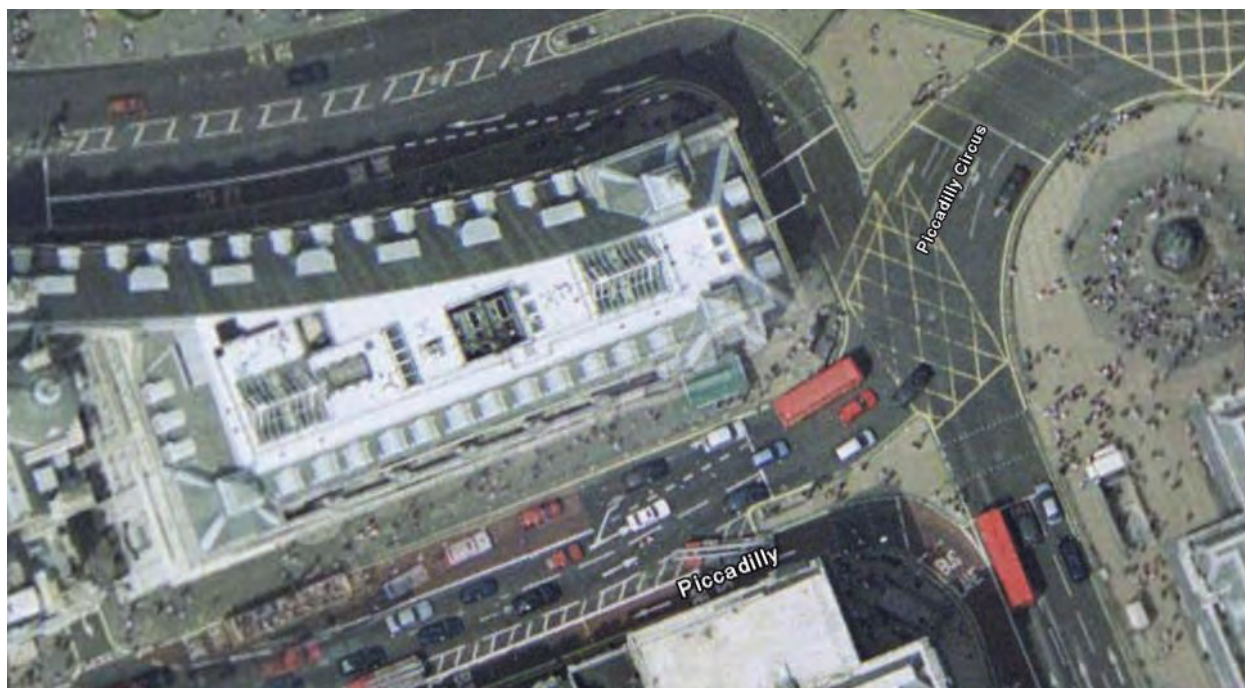


Figure 3 - Piccadilly Circus existing streetscape

Feasibility

Based on the scope, options were developed by Atkins for the route and set out into scheme-based logical entities. Subsequently, the chosen schemes were grouped to form sections for the purposes of consultation, design and implementation.

The options included 'blue sky' measures, developed as a demonstration of all ideas with perceived benefit and application in London. Some of those considered were:

- Infrastructure measures; such as guided bus ways, bus ways, high occupancy vehicle (HOV) lanes, heavy goods vehicle (HGV) lanes, median bus lanes, bus segregated strips, inflatable humps, converting disused railway into bus ways or general traffic links, guided bus flyovers, raised surfacing, raised carriageway/ junctions, grade separated carriageway, tidal bus/ traffic lanes, underpasses, bus only roads, bus lane colour changing road studs (lights), central bus lane – 24hrs – route corridor, bus dedicated junction lanes, diversion of traffic / buses to Royal Parks roads, rail bridge modifications, and tube station access links;
- Technology measures; such as GPS tracking of buses, dual SVD detection, bus door activated signal detection, and signal/ stop integration;
- Mechanisms; such as corridor road user charging, multiple zone based road user charging, statutory authorities/builders renting carriageway, roads policy for bus corridors, signal policy and design guidance review, identification and classification of highest priority bus routes, and review of Mayor's Transport Policy; and
- Additional ideas and concepts; such as route based loading strategy, permitted bus lane variants, and removal of bus / pedestrian conflict.

The choice of which options to develop was made in 2004/5. At the time, the political will was not sufficiently strong to take many of the 'blue sky' schemes forward although some have re-emerged in later programmes. Elements of these measures, however, were incorporated in the preferred options to be taken forward to preliminary designs. It is possible that opportunities for 'blue sky' measures will arise again in the future, as traffic conditions worsen and the political imperatives change with the need to address the major environmental concerns facing society.

Consultation on the preferred options

Once Borough Partner officers were satisfied with the schemes jointly developed, approval was sought to carry out public consultation from the Lead Member, Cabinet Members at Committee, or using delegated authority. The approval to consult was sought by the Borough officers, and following the processes of the individual Council. Public consultation was undertaken on a section basis to allow all consultees to appreciate the extent and effects of the options.

A common project approach and format for consultation were followed for all sections, whilst also tailoring them to the specific requirements of the Partner.

Common consultation approach

All public consultations consisted of a consultation pack including a leaflet questionnaire, bus stop drops, a telephone contact number, postal address, on-street posters, advertisement in the local paper, and exhibitions on weekdays, weekends, day and evening. In addition, a dedicated webpage was created for each consultation showing all information on the consultation pack and giving consultees the opportunity to comment online. The dedicated Route 38 webpage was also linked to the Borough's web page.



Figure 4 - Piccadilly Circus proposed streetscape

During the early stages of the consultation stage, discussions were held with the Borough Partners and the Client regarding the branding of the consultation material. Should each delivery Partner's branding be used for the schemes in their Borough? Or should the Route 38 Pilot Project develop dedicated branding? It was soon agreed by all that a dedicated Route 38 project branding clearly showing the partnership would reinforce the identity of this Whole Route project, whilst also speeding up the process by not requiring a full review and approval by each Partner's consultation team.

This common approach and methodology was very well received by all the Partners. In a true partnership way, good practices of the Partner's consultations were adopted into the Route 38 one; and vice-versa some Partners adopted elements of the successful Route 38 consultations in their own consultations or notification communications. An example would be Camden requesting Atkins to produce their works notification for Bloomsbury Way, using the easy-to-understand graphical maps developed.

Tailored consultation approach

The public consultation, however, was also tailored to the requirements and processes of the individual Partner. This included agreed scheme proposals and leaflets, the extent of the consultation pack distribution areas, stakeholder lists, exhibition locations, appropriate local papers, local meetings with associations such as traders and residents, etc. In addition, where the Partner had a local scheme in the vicinity of the Route 38 section being consulted on, and if appropriate, it was included in the Route 38 consultation material in order to reduce timescales and costs. An example would be the reversal of Air Street or the pedestrianisation of Glasshouse Street in Westminster being included in the Route 38 Piccadilly Circus consultation leaflet.

Atkins took the lead on all public consultations carried out, however, involving the Borough officers as much as they wished to be involved.

Following the public consultation, a consultation report was compiled and issued to the Borough for each section. This report analysed all responses received, in the form of leaflet questionnaire returns, comments received by telephone, post or posted online, by proposal type, area, etc to ensure the local

knowledge available within the local community was captured. In addition, the report presented post-consultation scheme amendments and additional supporting technical information.

Preliminary design

Following consultation and approval to proceed, preliminary designs were developed by Atkins in liaison with the five delivery Partners to demonstrate how the schemes operate. The designs developed for the route were presented to stakeholders within the relevant Partner for feedback and input, through meetings and correspondence. Using their knowledge of the local network and their Council's strategy and priorities, the delivery Partners commented on the schemes presented, and advised on the elements that needed to be included, adjusted, or all together removed, in order for the scheme to be approved by their Council.

In some cases, the requested changes were outside the scope of the project. In these cases, a process of communication, negotiation, and compromise with Partners was followed until a scheme acceptable to all was developed.

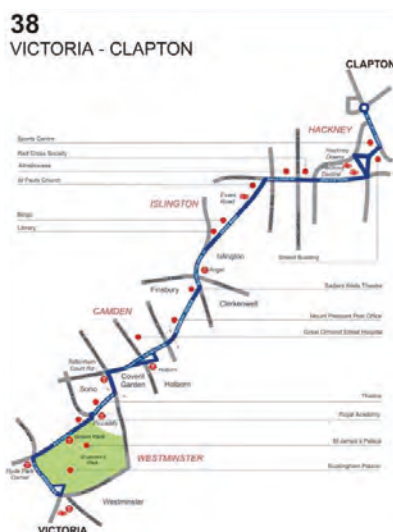


Figure 5 - Route 38 map

For the Upper Street (TLRN) section, for instance, the original proposals included a bus lane leading to a bus pre-signal. The proposals were subsequently revised as a result of partnership working with TfL Road Network Development (RND) to consist of relocated and new pedestrian crossings, signal coordination adjustments, and bus stop layout changes. Finally, they were further revised to widen the footways, introduce shared surface loading bays and widen the bus lanes to accommodate cyclists.

This iterative process of revising and fine tuning the schemes until an acceptable solution was found set the Partners' relationships. The relationships developed varied from Partner to Partner, depending on the officers, their Council's priorities and buy-in of the project, and the negotiations held to reach an agreed scheme.

Scope changes:

From an Intensified Bus Priority Project to a Corridor Management Project

The designs were initially developed whilst Route 38 was an Intensified Bus Priority project building on the LBI and LBPN schemes; however subsequently they had to be adjusted to reflect the new focus of a Whole Route Corridor Management and therefore the bus priority schemes needed to be adjusted to provide more holistic solutions.

This was an opportunity for Partners to provide improvements on their network, other than bus related. Urban designers were also involved at this stage to advise on the feasibility of urban realm opportunities and improvements, such as the creation of open spaces for pedestrians, etc. Schemes were therefore adjusted or redeveloped together with the Partners to reflect their aspirations as far as possible.

An example would be the recently implemented Shaftesbury Avenue scheme, which included in addition to bus measures, the widening of footways, shared surface loading bays to address local businesses' needs, improvements to safety and an element of public realm improvements linked to the Theatreland project, in line with Westminster's aspirations. Another example would be the introduction of an All Red Pedestrian Phase at the junction of Gray's Inn Road/ Rosebery Avenue in Camden, or the dedicated inset parking and loading provision on Essex Road in Islington to accommodate businesses' and visitors needs.

London Congestion Charging and its Western Extension Zone

Whilst the Route 38 project commenced in 2003, the year the London Congestion Charging Zone (CCZ) was introduced, the baseline data available for use to carry out feasibility and develop preliminary designs was collected prior to the CCZ. This and the subsequent introduction of the Western Extension Zone (WEZ) in 2007, altered the baseline conditions in terms of traffic patterns, and required the schemes to be reassessed and adjusted.

Conversion from Routemaster to Articulated Bus

A further change was the conversion of the Route 38 Routemaster buses into articulated buses in 2005.

This required bus stop and scheme designs to be adjusted once again to reflect the new bus type.

Working closely with the Borough Partners helped considerably to manage these changes, and hence minimise as far as possible delays to the project programme, and risk in terms of the feasibility and buildability of the schemes.

Borough Partner approvals

At the end of public consultation, Borough Partners needed to take the proposals through their Council's approval process prior to commencing detailed design. Assistance was provided to officers including Committee or Member reports and material, attendance of and presentations at Committee meetings, Committee requests and follow up actions, additional meetings with Members and any other requirements until scheme approval was granted. The level of assistance provided varied from Borough to Borough based on the officers' requirements, the Council approval process and proposals in question.

For instance, in the case of Islington, we met with Council Members, drafted Committee reports, presented schemes at Committee meetings and we dealt with the follow up actions and Committee requests; effectively relieving the officers from the additional workload as far as possible. For Camden, information was provided to officers to draft the Committee reports but we attended the Committee meetings in order to provide any background technical assistance to officers; the officers took the proposals through the Council approvals process and our role was to assist them technically. For Hackney, information was provided to officers who met the Lead Member. For Westminster, we provided information to officers for the Member reports and we presented the proposals and their benefits to the Lead Member during meetings. Where consultation was required on sections of the TLRN, similar processes were followed but in liaison and collaboration with TfL RND and in consultation with the relevant Borough.



Figure 6 - Shaftesbury Avenue footway widening

The partnership approach followed allowed each Borough Partner to request, and receive, the level of assistance they required at that moment in time. This tailor-made approach led to a good exchange of knowledge and assisted the further development of our existing relationship with the Partners. In most cases this approach also minimised delays to the project programme, since the officers were able to request our assistance on any task they were too busy to carry out. There was only one instance where Borough officers did not take advantage of our offer to assist in order to alleviate their workload; and this led to a delay to the programme of a few months. However, it is understood that part of that delay was due to the sensitive nature of the scheme.

Local Council Elections

Local Council Elections in May 2006 imposed a *purdah* restriction in terms of carrying out public consultation three months prior to and three months after the elections, hence causing delays to the project programme. In addition to this delay, the elections led to the change of political leads in certain Borough Partners, and hence the change of priorities. In those cases, we worked very closely with the Borough Partner officers to understand the new priorities as quickly as possible, and try to adjust the schemes in a way that addresses the new priorities, whilst still delivering on the Route 38 project objectives.

In a couple of cases, the partnership was more difficult to develop and maintain. Part of the reason was internal Borough politics, where perhaps bus priority was not their priority. This caused considerable delays to the schemes.

On the contrary, where the partnership working was successful the schemes were quickly adjusted and successfully taken through the new Borough Lead Members.

Transfer of project governance

During the consultation stage, project governance was transferred from TfL BPT to the City of London, highlighting the importance of the Borough partnership, although TfL remained the funding authority.

The transfer of governance created some confusion as to who was the Client. As relationships became more informal with time, it was very easy to take instructions from a partner who did not have the appropriate authority. This required a conscious effort from all Partners to operate in line with the responsibilities of their new roles, instead of taking forward those already established in the first few years of the project.

Furthermore, for all parties involved the transfer of governance translated, amongst other things, into different project and financial reporting requirements and processes to follow. More specifically, we now needed to directly comply with the requirements and processes of the City of London as the new Client, but indirectly also with those of TfL who remained the budget holder. Project instructions and new briefs, for example, were now to be issued by the City of London, instead of TfL. However, elements of TfL's "Spearmint" project management reports were required, but adapted to the requirements of the City of London and Route 38 project.

The transfer of governance to the City of London seemed to assist positively the relationship of the Borough Partners with the Client and the project. The Partners seemed more willing to be involved in the development of the project, and happier that the project was being moulded to their wishes and requirements.

The decision making process, however, seemed more complicated in some instances, with the City of London being the Client but TfL the budget holder.

The already in place good relationships between the Client and the Partners allowed open communication to take place, and assisted greatly in understanding and dealing with the new requirements in an effective way.

Detailed engineering design

Following the granting of Council approval, schemes moved to detailed (constructional) design. This stage possibly involved the greatest variance in the assistance provided to the Partners, again responding to their requirements, preferences and available resources.

For Islington and Hackney for example, Atkins carried out full detailed design in line with the Borough's design guidelines, and handed over the designs to the Borough for their contractor to implement the schemes. This included preparing the Scheme Brief and Traffic Signals Supplementary Reports (TSSR) required for London schemes and taking them through to signals approvals by TfL DTO.

It also included preparing scheme TMA submissions for the Borough to submit on the London Works computer database and liaising with TfL NAT until approval is granted. Regarding the works TMA, we assisted the contractor as required. Road Safety Audits were carried out by Atkins auditors independent of the scheme engineers.

Under this arrangement work was carried out in house with regular progress meetings with the delivery Partner. Designs were issued to the Partner for final check and approval at the end, prior to being handed over to the Partner's contractor. Beyond that stage Atkins' involvement mostly consisted of covering the designer's responsibilities. In the case of Islington, Atkins were also granted the role of CDM Coordinator; whilst Hackney appointed one of their own engineers.

Camden, on the other hand, developed the detailed designs in house with technical assistance from Atkins as and when required. This included for example signals design, traffic modelling, and generally signals approvals, as well as road safety audits and scheme TMA submissions. Regular progress meetings were in place to manage progress, deal with queries and decide on the best way forward. The relationship developed with Camden was a good partnership, where all parties contributed their stronger skills in order to achieve the best possible solution, as quickly as possible. An example would be, in addition to the assistance mentioned above, Camden's request for Atkins to prepare their notification leaflets, using the easy-to-understand graphical format developed for the Route 38 consultation, but adapted to Camden's branding.

Lastly, for Westminster, Atkins developed the designs to an agreed level and handed them over to the Borough's term consultant to complete and take forward to implementation. The partnership here was extended to include the Borough's term consultant. Regular Client/ Partner project meetings were in place to ensure appropriate coordination of works and appropriate awareness of progress; in addition, communication and collaboration took place between the two consultants to achieve best outputs.

Change in technical approvals processes

The TMA was introduced in 2004 giving highway authorities the powers and responsibility to preserve and promote uninterrupted traffic flow on their roads, i.e. broadly ensuring that the best balance for all users is achieved.

In London, this led to the creation of the TfL NAT reviewing TMA submissions for both schemes and works. These submissions have to be made by the highway authority through the London Works computer system.

In addition, all schemes involving traffic signals in London require to be assessed, and the design and modelling approved, by TfL. During the Route 38 project, the TfL department dealing with signals changed from the Traffic Technology Systems (TTS) to the DTO and with that changed the signals approvals process which became significantly more thorough and time consuming. Any scheme requiring alteration to traffic signals needed to be included in TfL DTO's Capacity And Resource Planning (CARP) programme, and subsequently submitted to DTO as part of a Scheme Brief and TSSR for the design and modelling to be reviewed and approved before submission for scheme approval under the TMA. The submission needed to include, amongst other things, traffic modelling developed based on DTO's new Modelling Guidelines.

The DTO scheme approval process and modelling guidelines were altered at a time where the modelling assessments of Route 38 schemes were close to completion or completed. All models had to be rebuilt in line with the new DTO Modelling Guidelines, and approvals submissions had to be in line with the new processes.

In order to help manage change, and minimise project delays as much as possible, TfL DTO offered to provide a DTO Route 38 coordinator acting as the main point of contact for all Route 38 schemes. This coordinator attended Route 38 progress meetings weekly, and was able to advise the Route 38 team regarding signals, but also gave feedback to the TfL DTO engineers working on Route 38 schemes.

This arrangement assisted greatly to develop an informal partnership with TfL DTO and reflected the importance of the signals approvals to the Route 38 project programme. Through this partnership DTO engineers were involved in the development of signals schemes as early as possible and assisting in the scheme adjustment to address any concerns they may have.

While the Route 38 Project was considered a priority project by TfL, it also had to "compete" against other major projects or resources; in terms of both design and implementation but also in terms availability of highway space. One example would be LU's Tottenham Court Road Station Upgrade, including integration of Crossrail. While partnership working functioned well with the delivery Partners, TfL BPT and DTO, LU and their advisors were rail focused, rather than streets and highways focused. This added a layer of complexity to the partnership working leading to significantly greater input from the Route 38 team than with other sections and partnerships.

Implementation

Implementation was always intended to be managed by the delivery Partners and carried out by the Partners' contractor on a section by section basis. Atkins retained the designer responsibilities for those sections we developed detailed design for. In addition, Atkins retained further involvement where we were appointed CDM Coordinator. In addition to the above, assistance was provided as required by the individual Boroughs. To date, implementation of sections has been progressed as follows:

- Implementation has been completed for the following sections:
 - Gray's Inn Road/ Rosebery Avenue, in Camden;
 - Theobald's Road, in Camden; and
 - Shaftesbury Avenue, in Westminster.
- Implementation is in progress for the following section:
 - Bloomsbury Way, in Camden.

- Implementation is planned to commence imminently for the following sections:
 - Upper Street, TLRN in Islington;
 - Essex Road, in Islington; and
 - Graham Road, in Hackney.
- Implementation will commence in the near future subject to final approvals being granted for the following section:
 - Piccadilly Circus, in Westminster.

The remaining sections of the corridor were dropped from the Route 38 programme due to being in an area of major redevelopment works, overlapping with other bus priority schemes more progressed than Route 38, following consultation outcome, or for political reasons. The remaining schemes will not be deferred indefinitely and will be “packaged” to enable the Borough’s to progress and deliver them through a different funding process.

Monitoring

The main monitoring parameter set at the beginning of the Route 38 project was bus journey time, in line with the Intensified Bus Priority nature of the project at the time. ‘Before’ data was provided by TfL at the beginning of the project in 2003. However, network changes such as the introduction of the Congestion Charging and later its Western Extension Zone, the conversion of Route 38 from Routemasters to articulated buses, major development works such as the LU Tottenham Court Road Station Upgrade, as well as smaller schemes such as pedestrian crossings, meant that traffic patterns had changed prior to even commencing implementation of the Route 38 schemes. In order to set a more accurate, updated baseline, additional ‘before’ surveys were carried out.

Various methods of collecting journey time data were considered for the additional ‘before’ data, including TfL’s Automatic Vehicle Location (AVL) system and iBus. Unfortunately the latter was not available at the time. An alternative Global Positioning System (GPS) and Graphical Information System (GIS) was therefore used for undertaking travel time, speed and delay studies.

Bus travel data was collected through a versatile GPS device which collected speed and position data on a second-by-second basis. This frequent ‘logging’ of data ensured a large database of route information was recorded which, in turn, allowed for versatility in analysis which was not achievable through other methods of data collection.

Post implementation or ‘after’ journey time data will, hopefully, be collected at the end of the Route 38 project in 2010/11 through TfL’s new iBus system. iBus appears to have similar functionality to the system used for the Route 38 Project. It is expected that it will be key to the promotion and delivery of 3GBP, and other bus schemes, through providing a good understanding of the existing on street conditions, and assisting for accurate and representative modelling.

Conversion from articulated to double deck Bus

It was previously mentioned that Route 38 was converted from Routemasters to articulated buses in 2005. In 2008 it was announced that bus Route 38 would once again be converted, from articulated to double deck buses at the end of 2009.

This conversion is expected to impact on traffic patterns, hence altering the baseline once again. This, and other changes yet to take place, will need to be taken into account when analysing the ‘before’ and ‘after’ journey time data and concluding on the impact and benefits of the schemes.

Business case

As the Route 38 project was a TfL BPT funded project that started as an intensified bus priority project, the business case was based on the funding granted by TfL and focussed on bus passenger journey time savings. Monitoring was agreed to be mainly on bus journey times, however, other monitoring parameters were also considered where appropriate such as accidents, impact on pedestrians, cyclists etc.

A business case was developed by TfL BPT, and revised at various stages of the project as more information became available and accurate. During the consultation and detailed design stages of the Route 38 project, TfL developed and issued a new TfL Business Case Assistant, catering for benefits to all modes and safety, rather than just journey times. TfL funded schemes were now assessed using this new Assistant. This is a fixed model, requiring though the development and interpretation of assumptions.

The 13 Route 38 schemes were now required to be re-assessed using this Assistant. The urgency to complete the assessments would have put a considerable strain on the resources of the TfL BPT and it was therefore agreed, that the schemes would be shared with Atkins.

In order to achieve consistency in developing the business cases of all schemes and the development of assumptions Atkins worked closely with TfL BPT to understand their new model and develop and interpret assumptions in a similar way.

Conclusions

The delivery of the Route 38 Pilot Project was a challenge due to many reasons including the five different delivery Partners and many Stakeholders involved, the length of the route, the difficulty of balancing conflicting user demands with bus benefits in the very constrained road network of Central London, political will, etc. The biggest challenge faced by the project, however, was probably the many changes experienced throughout its lifetime, including management, approvals and network conditions changes.

All these changes posed real time and cost control challenges and impacted on the project. The project programme was extended by approximately three years. Whilst the project budget was not increased, scheme costs increased due to price inflation, scope creep and general increase in expectations. Project costs also increased due to the extended design, assessment and approvals processes.

So was the Route 38 Pilot Project successful? Considering the complexity and challenges faced, delivering schemes on the ground is a success in itself. The schemes implemented to date seem to have already delivered considerable benefits to bus journey times, pedestrians, cyclists and other corridor users, as well as safety. The exact extent of these benefits, however, will be quantified and assessed when the outcome monitoring is carried out at the end of the project.

Furthermore, the project demonstrated that a strong partnership approach assists greatly in managing and controlling change, and delivering a good scheme acceptable to all Partners, minimising time delays and costs.

Water quality modelling as a tool for assessing new water resource management options: The case of the River Stour, Kent



Vera Jones

Environmental Scientist
Water & Environment



Brian A. Cox

Senior Environmental Scientist
Water & Environment



Neil Upton

Environmental Scientist
Water & Environment



Eliot Simons

Principal Consultant
Water & Environment

Abstract

Atkins was commissioned by South East Water to investigate a new water resource scheme in the South East of England. As part of the water quality aspect of this work, a Mike 11 Ecolab model of the tidal Lower River Stour was developed. The model was calibrated against observed data, leading to developments on the standard software template and improving our understanding of water quality processes in the river. The model was subsequently used to test a number of abstraction scenarios. The model results are anticipated to feed into the wider environmental impact assessment and ultimately help shape strategic decisions regarding water resources management for the area.

Introduction

New water resources options are being investigated in several parts of the south east of England in order to secure supplies and meet the increasing demand for water. The South East Water 'Strategic Water Resources Investigation' project aimed to deliver a successful and cost-effective planning application for a new scheme, aiming to increase water availability for supply and the security of water supply in the area. The work considered a wide range of aspects, for example heritage, recreation, aquatic ecology, hydrology, and water quality.

The way in which water quality might be impacted, as a result of the operation of new water resources schemes, is a major concern for planners. Water quality models are a key tool in investigating these impacts and, therefore, can form an important

part of related environmental impact assessments and planning applications. The first stage of the water quality work for this project involved an appraisal of baseline conditions in the Lower Stour catchment. Automatic water quality monitors were deployed at several locations throughout the catchment and spot sampling was also carried out by Atkins. This monitoring programme was designed to complement the Environment Agency's data collection in the area and feed into the construction of the water quality model.

This paper will provide an overview of the development and calibration of the Lower River Stour water quality model, illustrating how it was employed as a tool in the wider environmental impact assessment, relating to South East Water's proposed water resource scheme.

Water quality model development

Modelling software

In consultation with the Environment Agency (as regulator), the software Mike 11 (DHI; 2007 Service Package 1), with its Ecolab water quality module, was selected as the most appropriate modelling software for simulating flows and water quality in the Lower Stour. In contrast to simpler models, MIKE 11 includes detailed heat balance calculations, and hence reliable calculation of water temperature, and a comprehensive set of interactions between water quality determinands.

Model extent

The basis for the water quality work was the Mike 11 hydraulic model of the river, which was also developed as part of the same Strategic Water

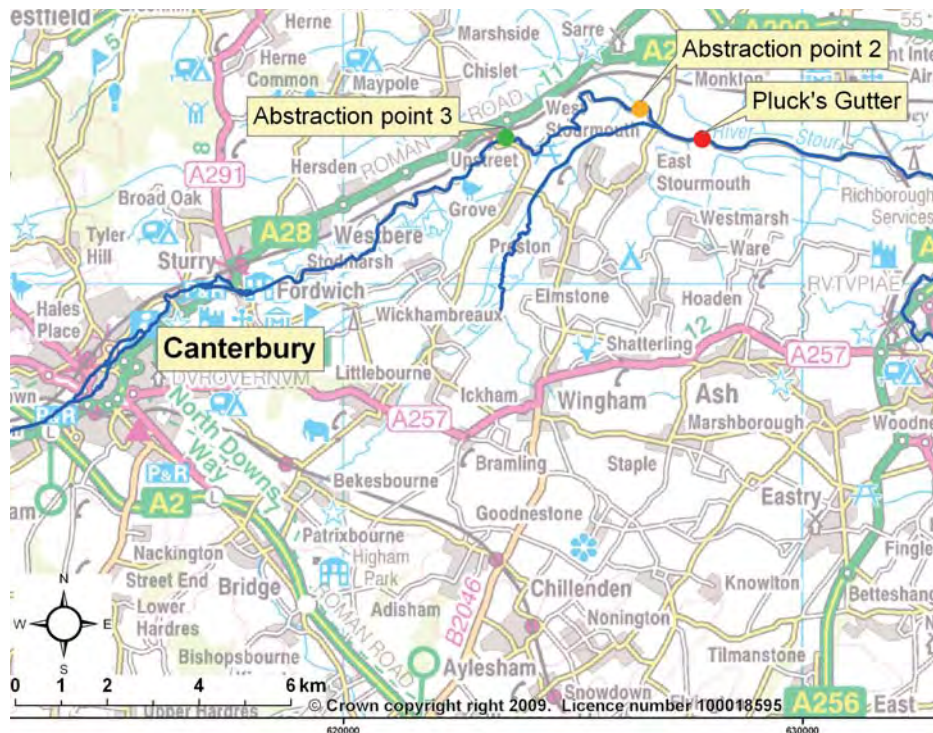


Figure 1 - Map of the Lower Stour area. The Section of the Lower Stour modelled is highlighted in blue. The locations of the upstream boundary (downstream of Canterbury) and downstream boundary (Pegwell Bay) are shown. Calibration points Sandwich and Plucks Gutter and abstraction points 2 and 3 for the scenario runs are also noted on the map.

Resources Investigation, to simulate potential hydrology and flood risks¹. The upstream boundary of the model was located at Canterbury and the downstream boundary at Pegwell Bay, see Figure 1. Additional boundaries, such as natural tributaries, a diffuse groundwater inflow, artificial discharges, abstractions and marsh feed channels were also included.

From hydraulic model to water quality model

The first step in moving from the hydraulic model to a water quality model was to simulate water temperatures, by enabling the 'heat balance' and 'wind' options within the hydrodynamic (HD) module. This allows for detailed heat calculations, which take into account humidity, air temperature, sunshine hours and wind speed data to simulate water temperatures along the river.

The advection/dispersion (AD) module was subsequently added. This determines the movement of water quality determinands in the system through advection and dispersion processes. Advection is the simple movement of mass from one point in space to another; dispersion is the movement of mass due to random water motion or mixing⁵, see Figure 2.

The AD module controls the transport of all substances and is used directly for simulating 'conservative' substances such as salinity that are not subject to biogeochemical interactions.

Once the heat balance and AD simulations were enabled and tested for water temperature and salinity respectively, the Ecolab water quality module was introduced. This module simulates the biogeochemical cycling of water quality determinands. The Mike 11 Ecolab module offers a number of standard templates, containing different combinations of water quality determinands and processes. The standard template, entitled 'WQlevel4ColiPhos.Ecolab', was selected and modified to suit the specific requirements of this project. The following seven water quality determinands were simulated: water temperature, salinity, dissolved oxygen (DO), biochemical oxygen demand (BOD), ammonia, nitrate and ortho-phosphate. Together, these determinands represent key indicators for the health of an aquatic ecosystem:

- Water temperature, and in particular changes in water temperature, have a critical impact on aquatic life, as biochemical reactions commonly experience a doubling in reaction rate with a rise of 10°C. Additionally, key

constituents of water either change their form (as in the ionisation of ammonia) or alter their solubility (as with DO) when temperature changes.

- Similarly, changes in salinity can have a big impact on the aquatic ecosystem, with a number of aquatic species having very narrow salinity tolerance. Additionally, salinity also affects the solubility of oxygen in water. Salinity is assumed to remain unaffected by biogeochemical reactions and is hence a very useful calibration tool for the AD element of the water quality model.
- DO concentrations in the water provide an excellent indicator of the overall water quality of the system. Sufficient DO is crucial for a healthy ecosystem, as fish kills are often due to asphyxia, when concentrations fall to very low levels as a result of organic pollution. If levels persist around zero, anaerobic or septic conditions yield products, such as methane or toxic un-ionised ammonia.
- BOD is a measure of the potential oxygen demand exerted within waters, generally arising from decaying organic matter. Pristine water exhibits low values of

Water quality modelling as a tool for assessing new water resource management options: The case of the River Stour, Kent

BOD, whereas sewage effluent typically exhibits very high BOD levels. BOD is measured by collecting a water sample and measuring the decrease in DO concentrations over the course of a 5-day incubation.

- Ammonia contains nitrogen, an essential nutrient for algal growth. It is present in all natural waters in very small amounts, however, levels in excess of 0.1mg N/l can be indicative of some sewage or industrial contamination. It is generally measured as total ammonia, which accounts for two aqueous forms: ammonium ions (NH_4^+) and un-ionised ammonia (NH_3). Their relative abundance changes with pH and temperature. High levels of un-ionised ammonia are toxic to invertebrates and fish, causing respiratory stress, conditions such as gill hyperplasia and reduced resistance to parasites and disease. Finally, nitrification, the process of via which ammonia is converted to nitrite, consumes oxygen and can therefore be an important control on oxygen levels in the water.
- Nitrate also contains nitrogen and is hence essential for algal growth in water. High concentrations of nitrate can lead to excessive growth rates or eutrophication and, in particular, the preferential growth of potentially toxic blue-green algae over less problematic algal species. Due to the solubility of nitrates, high concentrations can enter fluvial sources during high rainfall events from agricultural sources.
- Ortho-phosphate is also an important nutrient for algal growth in aquatic environments and likely to be the limiting nutrient in fluvial environments. The primary sources of ortho-phosphate are agriculture and human effluent. High concentrations of ortho-phosphate can lead to eutrophication.

Boundary and calibration data

Flow, level and water quality boundary and calibration data were derived from Environment Agency and Atkins spot sampling records and automatic monitoring. Where data were not available for a specific

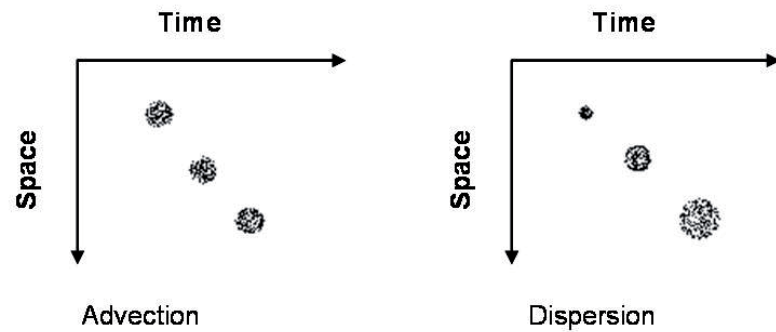


Figure 2 – Transport of solutes through advection and dispersion processes. After Chapra (1997).

Table 1 - Summary of calibration results and key issues (associated) with the calibration of each water quality determinant.

Determinand	Acceptable calibration achieved?	Key points
Temperature	✓	Heat balance and wind module applied, excellent fit with observed data.
Salinity	✓	Some gaps in the calibration data, but overall very good fit with observed data.
BOD	✓	Satisfactory fit with observed data, but only monthly spot samples available for calibration. Note the low accuracy of observed data.
Ammonia	✓	Good fit with observed data, but only monthly spot samples available for calibration. Cyclic pattern linked to the tidal flows and subsequent movement of seaward STW inputs up the river.
Nitrate	✓	Model results slightly lower than observed data. Overall satisfactory fit with observed data, but only monthly spot samples available for calibration. Cyclic pattern linked to the tidal flows and subsequent movement of seaward STW inputs up the river.
Ortho-phosphate	✓	Good fit with observed data, but only monthly spot samples available for calibration. Cyclic pattern linked to the tidal flows and subsequent movement of seaward STW inputs up the river.
DO	✓	Simulation affected by boundary data limitations during certain periods. Overall satisfactory fit with observed data. The photosynthesis and respiration expressions within the DO equation were modified to allow for changes in the size of the algal population.

*R² results from regression analysis **based on the results of analysis of variance, at the 95% confidence level *** no statistical analyses were carried out for these determinands due to the limited number of observed data points; instead an assessment was made based on the visual agreement between simulated and model data.

determinand or location, a suitable proxy has been identified. In most cases this meant data from a nearby location or representative values from a record of previous years. Where the data set was so limited that this was not possible, a set of 'generic values' was used, based on an overview of all the available data and experience from other similar studies.

Meteorological boundary data were obtained from the Met Office for the nearest suitable monitoring location (Manston; 63°24'E 16°61'N; approximately 11.2km from Plucks Gutter) and fed into the heat balance and wind elements of the model. Air temperature (hourly), humidity (three times a day), sunshine hours (daily) and wind

speed (daily) data were set as global boundary data and hence applied to the whole Lower Stour model.

Model calibration

Calibration methodology

After consideration of data availability, it was decided to calibrate the water quality model for the period 01/02/2007 to 30/06/2007. This calibration period includes a range of flow levels and is also sufficiently long to allow any initial model conditions to have no significant effect on model simulations.

Temperature, BOD, ammonia, nitrate, ortho-phosphate and DO were calibrated against data collected at

Water quality modelling as a tool for assessing new water resource management options: The case of the River Stour, Kent



Figure 3 - Observed and simulated temperature at Plucks Gutter (15-min data; full calibration period).

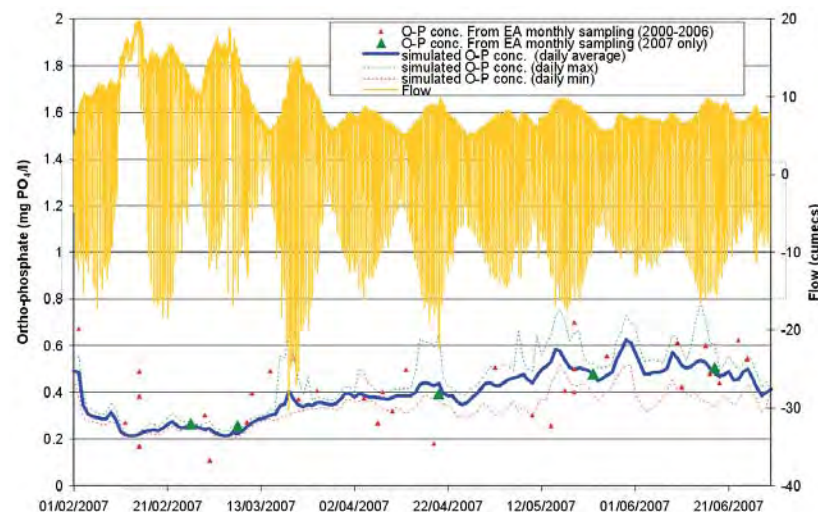


Figure 4 - Observed (Environment Agency spot samples; 2007 only) and simulated ortho-phosphate (daily averages, min and max) concentrations at Plucks Gutter. Also shown: observed data for 2000-2006 (Environment Agency spot samples) to provide an indication of interannual variability in spot measurements; flow measurements to illustrate pattern linked to tidal cycle.

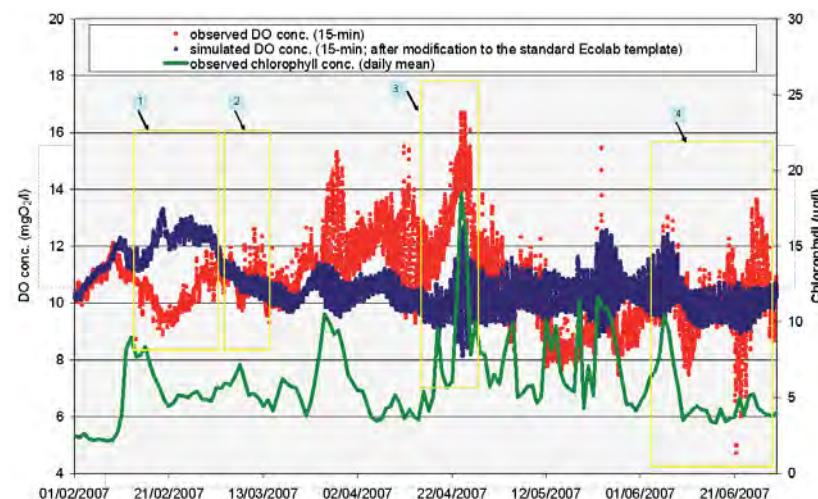


Figure 5 - Observed and simulated dissolved oxygen (DO; 15-min data; full calibration period) concentration. Chlorophyll concentrations (daily averages) at Plucks Gutter are also shown. Periods of potentially inaccurate boundary data are noted (boxes 1-4)

Plucks Gutter, see Figure 1, where the most extensive water quality data set was available. Additionally, Plucks Gutter would serve as an abstraction point under some of the water resource scenarios investigated. As salinity at Plucks Gutter is too low to resolve any calibration issues, salinity was calibrated against data collected at Sandwich, which lies just below the saline limit, see Figure 1. This also ensured that AD processes associated with the tidal movement of water in the river were correctly simulated.

The calibration process followed an order of increasing process complexity. Firstly, temperature was calibrated, then the conservative determinand (salinity), followed by nutrients and BOD and finally DO. Calibration was carried out aiming to obtain the best visual agreement with observed data, while staying within the constraints of acceptable determinand values. Past experience with similar models, relevant literature and advice in the DHI manual (DHI, 2007a; DHI, 2007b; DHI, 2007c) was taken into account.

Overview of calibration results

Calibration for DO, temperature and salinity was carried out against 15-minute data (01/02/2007-30/06/2007); calibration for BOD, nitrate, ammonia and orthophosphate was carried out against monthly spot samples (01/02/2007-30/06/2007). Interpretation of model performance acknowledges that a spot sample is only representative of that point in time and space rather than being representative of the whole reach of conditions in that month. Thus, when using spot data for calibration, it was not generally possible to judge whether extreme events or small-scale variations (e.g. diurnal cycles) were accurately simulated, other than by applying common sense and comparing with the results of past water quality models.

Some localised and short-lived instabilities were observed in the model during calibration, leading to negative salinities or concentrations. DHI informed us that this can be caused by steep local gradients and could be avoided by locally increasing the AD coefficient. However, the instabilities observed here were very much restricted in time and/or space and so considered to be trivial and are

Water quality modelling as a tool for assessing new water resource management options: The case of the River Stour, Kent

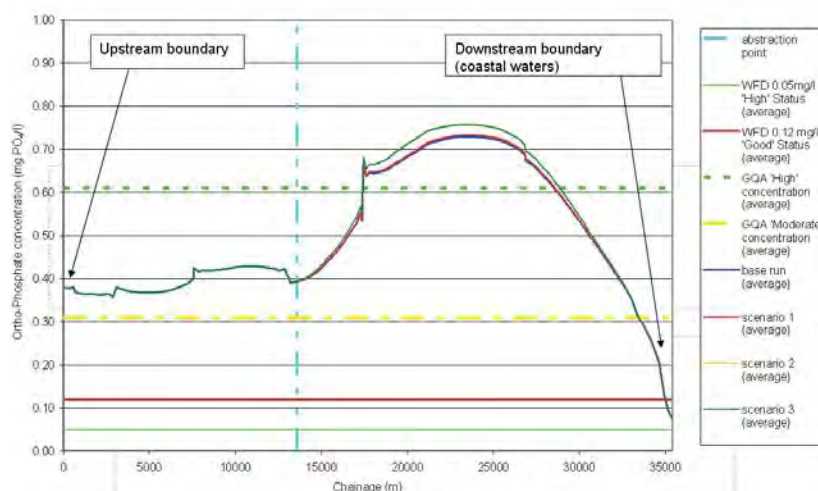


Figure 6 - Longitudinal profile of average ortho-phosphate concentration under base run conditions and scenarios 1-3 in the main river channel.

not expected to affect the prediction of impacts in the main river channel.

Despite these limitations, model simulations showed a very good fit with observed data for all seven water quality determinands considered. The calibration was, therefore, deemed to have been successful in producing a robust model of the Lower Stour. Table 1 summarises the calibration results, noting key points relevant to the calibration of each determinand. Two example calibration plots are included as Figure 3 (temperature) and Figure 4 (ortho-phosphate).

Other issues raised during the calibration process

Our understanding of a number of water quality processes occurring in the river was significantly improved through the calibration process. For example, the simulations suggested a cyclic diurnal pattern in nutrient concentrations. Closer examination of modelling results showed that this was due to the combined inputs from several STWs being swept up and down the river with the diurnal tidal movement of water. On a larger time scale, concentrations of nutrients at the calibration location (Plucks Gutter) appeared to peak during spring tide events, suggesting again movement of nutrients loads controlled by the tide, see Figure 4.

Additionally, the model calibration led to a significant improvement in the DO simulation, compared to previous applications of the Mike 11 Ecolab template used here. The modified template produced included a chlorophyll

parameter in the dissolved oxygen equation^{6,10}. This has allowed for changes in the size of phytoplankton population and subsequent impacts on dissolved oxygen levels to be modelled within the same model run, see Figure 5. The success of the DO calibration was hindered in part by boundary data limitations, see Figure 5, and further work is needed in the future to improve the fit with observed data further.

Scenario runs

Description of scenarios

After calibration of the model, preliminary scenario runs were carried out to make an assessment of water quality impacts under different scenarios of abstraction from the River Stour. Each scenario is based on a separate water resource modelling investigation². The following nine scenarios were selected:

- Scenario 1 represents abstraction at Plucks Gutter under a 180 -200 MI/day Minimum Residual Flow (MRF) with restrictions based on permitted levels of abstraction under the Water Framework Directive.
- Scenario 2 represents abstraction at Plucks Gutter under a seasonal MRF regime (160 - 220 MI/day).
- Scenario 3 represents a smaller volume of abstraction at Plucks Gutter under a 160 220 MI/day MRF.
- Scenarios 4, 5 and 6 are identical

to 1, 2 and 3, respectively, with abstraction, however, taking place at the confluence of the River Wantsum and the Stour, ~1km upstream of Plucks Gutter.

- Scenarios 7, 8 and 9 are identical to 1, 2 and 3, respectively, with abstraction, however, taking place just downstream of Grove Ferry, where effluent enters from Herne Bay STWs.

Scenarios were run for the same period as the calibration, so as to include a range of flow levels. Results from the first six days of model runs were excluded from the analysis to ensure that the model had stabilised and the impact of initial conditions was minimal.

Overview of scenario results

The results of the scenario runs were assessed against relevant environmental quality standards, in particular focusing on the the Water Framework Directive (WFD), Freshwater Fish Directive (FFD), the Surface Water Abstraction Directive (SWAD) and the Environment Agency General Quality Assessment (GQA). It should be noted that environmental quality standards generally apply to a specific length of data record (e.g. annual record for the FFD or a three-year sampling cycle for GQA), while the model results discussed here represent a 5-month-long time-series. Environmental quality standards were hence used to put the results into context, rather than assess compliance.

In summary, the model results suggested that there would be no significant impact on temperature, salinity, BOD, ammonia, nitrate or DO under any of the preliminary abstraction scenarios considered. A slight, but insignificant, increase in average concentrations (up to 2.2% when considering the whole river and up to 2.8 % when considering the main river channel downstream of the abstraction point) was simulated for ortho-phosphate and this was greatest under the scenarios with the highest maximum abstraction rate.

These small simulated changes in average ortho-phosphate concentrations become noticeable approximately 4km downstream of Plucks Gutter, where there was reduced dilution of discharges from two STWs. Nevertheless, nutrient concentrations in the Stour are generally elevated with respect to environmental quality standards, even under baseline conditions. As a result, the increase in concentrations under the abstraction scenarios was not large enough to affect current performance against those standards. A plot of some of the scenario results for ortho-phosphate is shown here as an example, see Figure 6. This illustrates the longitudinal variation in average ortho-phosphate concentrations along the main river channel, for the baseline and scenarios 1, 2 and 3. Note that only maximum and minimum results are readily accessible within the Mike 11 programme. A bespoke tool was developed for this study in order to efficiently extract other statistics from the model results.

Conclusion

The River Stour water quality model has been used as part of a Strategic Water Resource Investigation to appraise a suite of options for increasing the amount of water available for supply by South East Water. It is anticipated that the model could be used in the future as a means of testing a wider range of water resources options and scenarios. This includes the potential construction of a reservoir at Broad Oak and creation of new abstraction points on the River Stour. Separate models of a proposed reservoir⁴ and the river proposed to be impounded to form the reservoir³ have already been constructed by Atkins. These will be used in combination with the River Stour model described in this paper to provide a comprehensive assessment of water quality impacts for the catchment. The modelling approach was designed to provide information to feed into an Environmental Impact Assessment and subsequent Environmental Statement.

References

1. Atkins, 2008a. South East Water Ltd., Strategic Water Resources Investigation. Lower Stour and Sarre Penn Hydraulic Modelling Report, Draft for Client Review. (5044447 / 7502 / DG / 395).
2. Atkins, 2008b. Results of MISER modelling & Sensitivity runs on Demand Profiles for Optimisation of Broad Oak Reservoir size, Memo to Jon Holland (5044447 / 72 / DG / 429).
3. Atkins, 2008c. South East Water Ltd, Sarre Penn River Water Quality Model Calibration Report Draft for Client Review. (5044447 / 77 / DG / 436)
4. Atkins, 2008d. South East Water Ltd, Broad Oak Reservoir Model Report_ Draft v6, Draft for Client Review. (5044447 / 77 / DG / 548)
5. Chapra, S.C., 1997. Surface Water-Quality Modelling. McGraw-Hill International Editions.
6. Cox, B.A, 2002. Dynamic modelling of dissolved oxygen. A case study for the River Thames. PhD Thesis, University of Reading.
7. DHI, 2007a. Mike 11. A modelling system for Rivers and Channels. User Guide.
8. DHI, 2007b. Ecolab, WQ templates, Scientific Description.
9. DHI, 2007c. Ecolab, User Guide.
10. Whitehead, P.G., Williams, R.J., Lewis, D.R., 1997. Quality simulation along river systems (QUASAR): model theory and development. The Science of the Total Environment: 194/195: 447-456.

High accuracy recording for heritage applications - Dover Castle graffiti



Cory D Hope

**Principal
Photogrammetrist**
Water & Environment

Abstract

The 21st century has seen the adoption of digital technologies as the standard equipment and methodologies when undertaking close range Photogrammetry for the recording of Cultural Heritage. The use of calibrated digital SLR cameras, REDM total stations and softcopy Photogrammetry has reduced capture time of source imagery and photo control, improved consistency of stereo images, enabled alternative mensuration methods to the traditional individual stereo pair to be adopted and enabled very high accuracies to be achieved for reasonable cost. This paper will describe the method and application of close range digital photogrammetric techniques in order to enable millimetre accuracy measurement and 3D visualisation of historic graffiti at Dover castle. The project was undertaken by Atkins Geospatial on behalf of the English Heritage Metric Survey Team.

Introduction

In recent decades there have been many fine examples of photogrammetric recording of ancient rock art, such as carvings, petroglyphs and pictographs and the like utilising various methods to varying degrees of accuracy. One of the earliest examples was work undertaken to record stone carvings at Stonehenge (Atkinson 1968). There have been many examples following this undertaken all over the world on a variety of subjects using varying methodologies. A description and discussion of the variety of examples is a rather large task in itself and hence falls outside of the scope of this paper.

The results achieved to date have helped to drive the use and development of close range photogrammetry in Heritage and Archaeological recording. It follows on that in today's well established digital age of Photogrammetry, acceptable accuracies are achievable for a variety of subjects, large or small for reasonable or low cost.

Due to the typically limited funding available for this type of recording, there has been a focus on finding the most cost effective photogrammetric solutions (Chandler and Fryer 2005, Bryan and Chandler 2008).

These methods have proved very successful and the methodology utilised in this project follows on from this valuable work. The key difference in this project is that measurement accuracy not low cost is the key factor in determining methodology and equipment selection. This paper will describe the high accuracies achieved using a fully digital workflow for very close range Photogrammetry.

Note: The word accuracy is used in this document in general terms and relates to the combined precision and accuracy of any measurements made and statistical results achieved.

The project

Atkins Geospatial was commissioned by the English Heritage Metric Survey team to undertake a 'close up' photogrammetric survey of 7 areas of historic graffiti at Dover castle. The specification included for the provision of 'close up' stereo photography at a maximum taking distance of 1m from the subject and photo control to be measured to an accuracy of $\pm 0.001\text{m}$. The photogrammetric survey was undertaken by Cory Hope in November 2008 who has over 10 years experience of undertaking close range

photogrammetric surveys for English Heritage, Historic Royal Palaces, CADW and many other institutions and conservation architects.

The survey formed part of an ongoing English Heritage research project at Dover Castle focusing principally on the Great Tower (Keep) and Inner Bailey that is contributing to the current representation works within this area of the site. Although the work commissioned by English Heritage was limited to just the provision of the stereo photography and photo control, Atkins Geospatial has undertaken further processing of the raw survey data themselves to confirm and describe the accuracies achievable using this methodology.

The subject

The subject of the survey is a variety of inscribed elements (graffiti) on the internal walls of the Keep at Dover Castle. The graffiti is of varying age and style with much of it dating back to the 1700s when the Keep was used as a prison. Some of the graffiti is carved relatively deeply into the fabric of the wall (up to a couple of centimetres), which typically consists of names and dates but there are also some areas



Figure 1 - Image sample of historic graffiti at Dover Castle

where the graffiti consists of large areas of lightly inscribed patterns and shapes. The total coverage of this initial photogrammetric survey includes 7 separate areas totalling in approximately 70 square metres of historic graffiti.

Project planning

The decision making process as to the most appropriate method and equipment for undertaking the survey was influenced by the following: (Luhmann et al., 2006)

- The standards required for stereo photography as described in the English Heritage Metric Survey Specification for Cultural Heritage (Bryan et al., 2008)
- The required accuracy and photo scale of survey products as required by the project specification
- The nature and location of the subject
- The environmental conditions around the subject
- The available equipment held by Atkins Geospatial.

Methodology

The nature of the subject is roughly planar i.e. the subject consists of inscribed elements on the internal walls of the building. The traditional image configuration for close range stereo photography is individual stereo models which is the general requirement of the English Heritage Metric Survey specification. As most of the areas to be covered are much larger than a single stereo

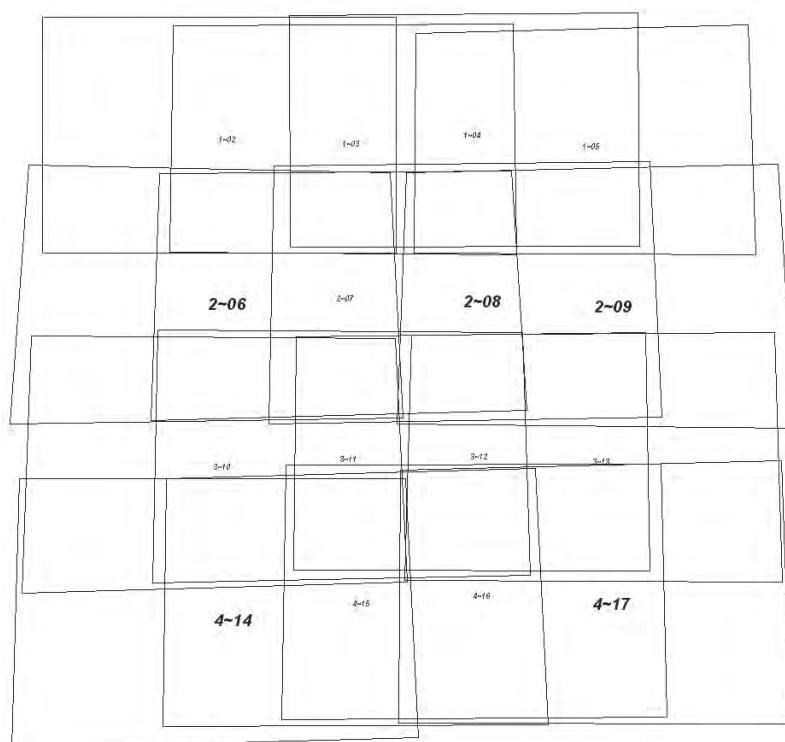


Figure 2 - Typical layout of the stereo photography

model, utilising the traditional image configuration would result in a very large amount of individual stereo pairs. In order to minimise the amount of photography required to cover the subject in stereo, it was decided that the best option would be to undertake the stereo photography using an image configuration similar to that used for an aerial photogrammetric project.

The result is a series of horizontal strips of photography with an approximate side lap of 60% and an end lap of 20-30%.

This image configuration enables a block triangulation bundle adjustment to be undertaken in order to assess and prove the potential measurement accuracies achievable from the stereo photography.

Equipment

- Kodak DCS Pro SLR/n with 28mm lens calibrated by Dr S Robson UCL
- Tripod with Manfrotto head with level bubbles
- Leica TCR703 with Leica traverse kit
- 2x Bowens Esprit 1500W flash units with soft boxes attached
- Gretag Macbeth colour chart

- Small adhesive stickers to be used as photo control.

Placing photo control grid

The design of the photo control point layout consists of a roughly regular grid of points covering the subject area. Sufficient points were placed such that for each stereo model there were 4 control points placed within each stereo model hence satisfying the English Heritage Metric Survey Specification.

Due to the requirement that the stereo photography was taken at 1 metre from the subject, the estimated Ground Sampling Distance (GSD), or the estimated pixel size in reality is 0.0003m. This small pixel size in reality offers the potential for very fine measurement, so the photo control points needed to be very small to enable accurate measurement of the photo control points from the digital images. The method of marking the photo control was to use easily available circular sticky labels of 8mm diameter. Red and orange colours were chosen as they would clash with the subject surface making each point easily identifiable. A small pen mark was made in the centre of each label to act as the point of measurement.



Figure 3 - Example of photo control point 'A20'.

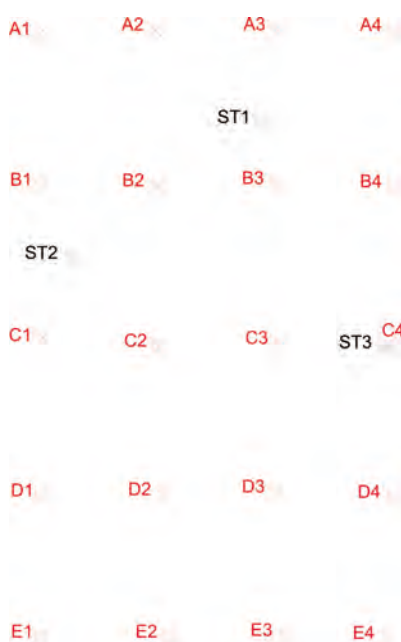


Figure 4 - Typical layout of the grid of photo control. The points in red are the photo control points and the points in black are the steel monitoring studs for Area 6_1 of the project

To aid in the identification of each photo control point, both for the photo control survey and for subsequent office processing, each horizontal line of control was given a representative letter and each vertical line of photo control was given a representative number. Thus each photo control point had a grid reference for eg. A3, B2 etc.

Easily available stickers with numbers and letters on them were then placed adjacent to the control points but outside of the area of interest to aid in the identification of these points.

There are always conservation issues when placing photo control targets and permanent monitoring points to any historic fabric. In this case the Ancient Monument Inspector Judith Roebuck was consulted about the proposals and once agreement had been sought the stainless steel markers were located by the Properties Curator Roy Porter. The stickers used as temporary photo control points were chosen as no glue would be required for fixing and they are easily removable without damaging the historic fabric.

Acquisition of stereo photography

Once the grid of photo control points was placed, the stereo photography was undertaken. The two main challenges for this part of the project were achieving consistent camera position and even lighting of the subject.

As each project area is reasonably small, the lighting was setup such that the entire subject area could be photographed under the same lighting conditions. Wherever possible the two flash units were placed such that all of the subject area was covered with as near as possible an even distribution of light and minimal shadows. Setting the lighting this way minimised the need to move the flash units when undertaking the stereo photography and in many cases the stereo photography was completed without the need to move the flash units at all. This resulted in reasonably consistent colour and image quality across the majority of the project areas which helped to ensure the high quality of the products of the survey. However, due to the nature of the subject and site limitations there were some instances where surface reflectance created image variations within some stereo models.

Before undertaking the stereo photography the colour chart was placed near the centre of the project area and photographed under the same lighting conditions that were to be used for the stereo photography. The colour chart was used to check the colour balance of the images during processing and act as a colour reference for any future processing of the photography. Due to the fact that great care was taken in the lighting and camera setup, the colour balance

of the resultant images was good and only minor colour adjustments were required and in some cases no colour adjustment was necessary.

Traditional record photography of carved stone work such as this uses the lighting method of raking light. Photography using this method typically contains high contrast shadows. This helps to highlight the carved features for visual interpretation but is not suitable for photogrammetric processing as the shadow areas lack detail and pixel correlation in the shadows is uncertain or in the worst cases, not possible.

The exposure setting for all photography was f22 at 1/125 second. This enabled a depth of field range of 0.8m to infinity to be achieved with the camera lens focus locked at approximately 2m. Sufficient tests were undertaken prior to the survey to ensure that the lens configuration described above would produce the desired results without the need to adjust the focus on the lens and hence perform a new camera/lens calibration.

Where circumstances allowed, the distance of the camera from the subject is maintained by placing a range pole on the ground at the taking distance of 1m. Prior to each exposure, the camera is mounted on the tripod and levelled using the spirit bubbles on the tripod head. Using the grid lines visible in the camera viewfinder, the camera is rotated to near orthogonal to the subject and the photo is taken.

Each strip of photography is taken from left to right, starting with the uppermost strip. The camera height is adjusted using the tripod for each strip as required.

For some of the subject areas the position for taking the stereo photography was limited by walls opposite the subject, proximity to the floor or other unavoidable obstructions. In these instances there was no alternative but to undertake the stereo photography hand held.



Figure 5 - Lighting position for recording the graffiti above fireplace



Figure 6 - Setting camera position

Photo control survey

The photo control survey was undertaken using the technique of intersection including REDM measurements to each control point. Most of the project areas allowed for all points to be measured from two survey stations. The survey stations were placed such that the best possible intersection angles could be achieved within the limitations of the site. There were some instances where it was not possible to measure the photo control by intersection, in these cases the REDM measurements were relied upon to calculate the 3D coordinates.

During the photo control survey steel monitoring studs, placed previously by the English Heritage Property Curator, were measured. These points are to be used for future referencing and monitoring of the subject and were also used as check points for the block triangulation bundle adjustment.

The survey observations were processed in Atkins in-house survey calculation software in order to produce the 3D coordinates of the photo control points. The REDM measurements were used to calculate 3D coordinates independently of the values calculated by intersection

as a check. The results of the survey calculations were very good (as confirmed by the results of the triangulation and bundle adjustment) enabling very accurate measuring to be achievable from the resultant stereo models.

Triangulation and block adjustment

Before submitting the stereo photography and photo control to English Heritage for further processing, some of the project areas were triangulated using Intergraph ISDM to check the quality of the collected data and assess the accuracy potential for measurement from the resultant stereo models. The control points placed where measured and included in the bundle adjustment solution and the steel monitoring studs were measured as check points and hence not included in the bundle adjustment solution. The summary results of the bundle adjustment for area 6_1 of the project are given in Table 1.

Further processing by English Heritage

Following the photogrammetric survey, English Heritage has undertaken the production of various photogrammetric data products. The software package used to process the stereo photography supplied by Atkins was Topcon PI 3000. Topcon PI 3000 is a relatively low cost software package which enables stereo viewing and can be used to produce 3D models, VRML models and Orthophotos. The processing of selected areas commenced in January 2009 and at the time of writing is ongoing.

The images were imported into Topcon PI3000, the photo control was observed and stereo models were created. The next task was to create an overall Digital Terrain Model (DTM) of the subject area at a grid density of 5mm. From this DTM an overall VRML 3D model and an overall orthophoto was created for each subject area. Each of the 7 subject areas covered by the stereo photography consisted of many separate instances of historic graffiti.



Figure 7 - Camera ready



Figure 8 - Undertaking photo control survey



Figure 9 - 3D VRML model of an individual stone with graffiti.

In order for finer analysis to be undertaken, individual stones with graffiti, were recorded in the same manner, but at the higher resolution of 2mm. So for each subject area an overall DTM, VRML and Orthophoto at 5mm resolution was created as well as individual DTMs VRMLs and Orthophotos at 2mm for individual stones.

At the time of writing, the production stage has generated some 207 stereo pairs, from which 979 DTMs at 2mm and 979 DTMs at 5mm have been generated. A total of 979 individual stone VRMLs and 979 individual stone Orthos at 2mm have been created and 18 overall area VRMLs and 18 overall area Orthos at 5mm have also been created.

Table 1 - Bundle adjustment key statistics for Area 6_1 (figures given are in metres unless otherwise stated)

Parameter	X/Omega	Y/Phi	Z/Kappa	XY
RMS Control	0.000295	0.000239	0.000174	0.000268
RMS Check (steel mon. studs)	0.000278	0.000222	0.000586	0.000251
Max Ground Residual	0.000545	0.000505	0.000472	
Mean Std Dev Object	0.000092	0.000092	0.000201	
Mean Std Dev Photo Position	0.000426	0.000454	0.000173	
Mean Std Dev Photo Attitude	0.022076	0.021363	0.006284	

Conclusion

The results of the survey and bundle adjustment show that accurate millimetre or even sub-millimetre measurement is achievable utilising a well calibrated semi-metric digital camera, modern total stations and a robust technique.

Undertaking the stereo photography in a way which replicates the network design for an aerial survey creates the opportunity to triangulate each project area as a single block adjustment. This can be a very useful and time saving methodology when using high end digital photogrammetric workstations where many aspects of production can be automated, such as automatic DEM creation and orthophoto production. It also means that once triangulated, there will be seamless joins between individual models making the output from stereo plotting more reliable.

The quality and resolution of the 3D VRML and Orthophoto data products makes it possible for off site digital analysis of the graffiti to be undertaken in very fine detail without the need for high powered computers or expensive software. This 'close up' level of recording has provided a very accurate and detailed archaeological record of the graffiti at Dover Castle.

Acknowledgments

The author would like to thank Paul Bryan of English Heritage for contributing some text and editing this paper prior to submission, Steve Tovey of English Heritage for supplying information and images from the production stage of the project, Don Martindale for processing the survey data and Dr Stuart Robson for undertaking the calibration of the camera/lens combination using his calibration field at UCL and his camera calibration software.



Figure 10 - Orthophoto of individual stone with graffiti.

References

1. ATKINSON, K.B., 1968. The recording of some prehistoric carvings at Stonehenge. *Photogrammetric Record* 6(31): 24-31.
2. CHANDLER, J. H., FRYER, J. G. and KNIEST, H. T., 2005a. Non-invasive three-dimensional recording of Aboriginal rock art using cost-effective digital photogrammetry. *Rock Art Research*, 22(2): 119-130.
3. CHANDLER, J.H. and FRYER, J.G., 2005. Recording aboriginal rock art using cheap digital cameras and digital photogrammetry. *CIPA XX International Symposium, Torino*. Pages: 193-8.
4. CHANDLER, J.H., BRYAN, P. and FRYER, J.G., 2007. The development and application of a simple methodology for recording rock art using consumer-grade digital cameras, *The Photogrammetric Record*, 22(117): 10-21.
5. LUHMANN, T., ROBSON, S., KYLE, S., HARLEY, I. 2006. Chapter 7, Measurement concepts and solutions in practice. *Close Range Photogrammetry, Principles, Methods and applications*, page 441-442, ISBN 1-870325-50-8.
6. BRYAN, P. G., BLAKE, B., BEDFORD, J., ANDREWS, D., 2008. *Metric Survey Specifications for Cultural Heritage*. English Heritage. York, 197 pages.

Conjugate heat transfer study of a spin pit rig: Application to the lifing of HP turbine disc firtrees



**Naveen Prasad
Gopinathrao**

Engineer II
Energy



**Christophe
Mabilat**

Engineer I
Energy



**Dr Sohail
Alizadeh**

**Chief Engineer -
Thermofluid & CFD
Technologies**
Energy



**Dougal
Jackson**

**Aerothermal
Technologist**
Turbines, Rolls-Royce



**Dr Rory
Clarkson**

**Engineering Manager
- Thermals**
Combustion & Casings,
Rolls-Royce

Abstract

Structural life and durability of gas turbine engine discs and blade firtree have always been key design issues for designers and technologists. This is particularly important for gas turbine components which operate at elevated temperatures under high loads. The testing for these components needs to be undertaken at operating conditions, particularly when Low Cycle Fatigue (LCF) lives of components are evaluated. The spin pit and bi-axial rig tests, the former of which is the concern of this study, are used in the aero gas turbine industry for the lifing analysis of these critical engine parts.

In this study a three dimensional conjugate heat transfer analysis of the spin pit rig with the HP turbine disc and blade assembly for a large turbofan engine was performed. The scope of the Computational Fluid Dynamics (CFD) analysis has allowed a detailed understanding of the flow and thermal field within the domain to be obtained.

Flow within the rig was found to be complex, with multiple flow mechanisms simultaneously at play in several distinct areas of the solution domain. All heat transfer processes, radiation, convection and conduction were found to be important in the overall heat. Radiation heat transfer was found to be dominant in terms of temperature levels achieved in components, whilst convection heat transfer greatly influenced temperature gradients in solids. The appropriate capturing of flow leakages in the rig proved to have a significant impact on rig performance and component thermal fields. The component temperatures predicted by the tuned CFD model compared well against thermocouple measurements and predictions of the conjugate CFD analyses were directly used as boundary conditions in follow-on stress analyses for lifing purposes. The methodology adopted has resulted in a very satisfactory outcome on declared disc life.

It is understood that this is the first time where a full conjugate heat transfer CFD study of the spin pit rig without any separate thermal analyses, has been applied directly to the lifing of HP disc firtree and the eventual engine certification, in the aero gas turbine industry.

Introduction

It is understood that this is the first time where a full conjugate heat transfer CFD study of the spin pit rig without any separate thermal analyses, has been applied directly to the lifing of HP disc firtree and the eventual engine certification, in the aero gas turbine industry.

A spin pit is a test rig designed to spin turbo-machinery components up to design speeds in a controlled environment to evaluate the structural integrity of components. Spin pit tests are typically conducted in vacuum to reduce the power required to drive the test article and to reduce hazards when testing at high temperatures. Under these conditions inserted blade turbines often show shorter life than they do in engine operation. The source of the problem is accelerated fretting at the firtree caused by

the lack of air/oxidation. This problem is eliminated by operating the spin pit at absolute pressures of 1/3 of an atmosphere level.

Over the years, intense research efforts have been devoted to address this issue both experimentally and computationally. Mahorter et al.⁶ discussed the use of cyclic spin pit testing of gas turbine disks as a means of predicting the life and life predictions were made using statistical analysis of a limited number of tests.

Advancements in analytical and computational methods have brought significant developments in the life assessment of turbomachinery components. Matthew E. Melis and Erwin V. Zaretsky⁷ compared two series of LCF test data reported by Mahorter et al⁶ for two different

groups of gas turbine compressor disks using probabilistic methods. Predictions were made using finite element (FE) and analytical methods and the predicted failure location was correlated with the LCF tests. Similarly gas turbine disk life predictions were carried out by Sattar and Sundt⁵. They emphasised on analytical stress methods and related that to experimental testing of the specimen. Hernan et al.⁸ attempted to predict the fretting failure of gas turbine components using FE methods. They focused on identifying the driving factors for fretting failure and 2-dimensional FE analyses was carried out to quantify the influence of the key factors on stress and strain. The predictions correlated well with the experimental observations on the spin pit tests.

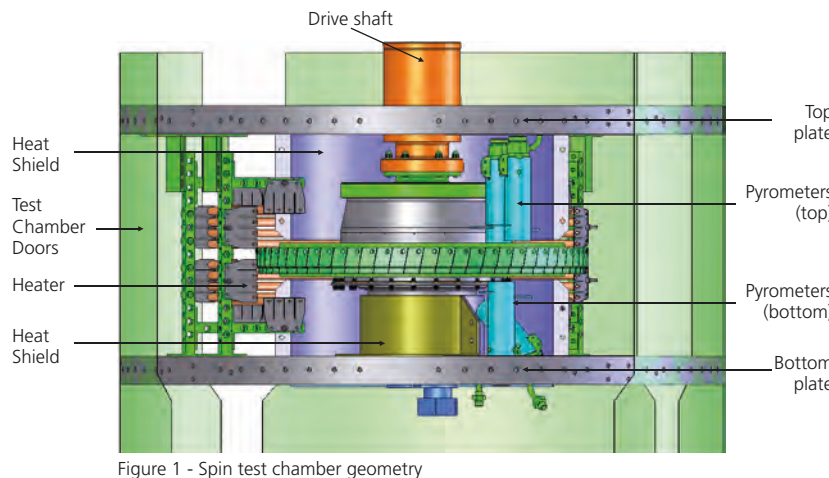


Figure 1 - Spin test chamber geometry

Gas turbine components are subjected to thermal and mechanical stresses during their operation which are transient in nature and these are generally modelled using FE for certain flight cycles. Management of the thermal loading in these components plays a key role in component lifing and hence predictions of detailed and accurate temperature profiles for disk and blade assemblies are highly important. Given the complex geometries involved and the challenging physics associated with engine operation, CFD is a good tool of choice for detailed component performance assessment. In recent developments, CFD is used more for deriving appropriate heat transfer boundary conditions to thermo-mechanical analyses. Approached vary from fully coupled⁹ to loosely coupled¹¹ approaches. In further developments, full conjugate heat transfer CFD applications for detailed disc temperatures predictions can be found for turbojet disc cavity¹⁰ and Bi-axial rig enclosure¹² studies.

In the present study, the 3-D conjugate heat transfer predictions of the spin pit have been used for the lifing assessment of the HP disc firtree for a large turbofan engine. CFD was used to get a better understanding of the overall rig behaviour and heat transfer mechanisms. Moreover it was intended to tune the CFD model to the rig thermocouples measurements and upon satisfactory model matching, the detailed disc-blade assembly temperature predictions could be exported and used directly for follow-on stress analysis, circumventing the need for separate thermal models.

This also ensured that the final component temperatures exported from the CFD model benefited from the detailed physics, taken into account in the conjugate heat transfer analysis, instead of making simplifying assumptions in thermal modelling about metal-fluid boundary conditions.

Test rig setup

Figure 1 and Figure 2 show the geometry of the spin-test chamber. This test chamber is placed inside a pit which is maintained at 1/3 atmospheres and the test disc and blade assembly was driven at 13,500rpm from a backing disc bolted to the drive shaft.

The test disc was fitted with 66 dummy blades connected to disc by firtree type attachment geometry. The dummies do not resemble the real airfoil geometry, but had a bulk mass equal to airfoil mass to exert the equivalent rim load on the disc.

The test disc and blade assembly were heated through three banks of heaters. The heater power levels were adjusted during the test along with the air bleed just sufficient to achieve the required operating conditions representative of engine operations. The rig was instrumented with thermocouples and pyrometers to measure the metal and air temperatures in the pit. Figure 2 shows the locations of the thermocouples (M01 –M06) around the test disc and blade assembly. This provided test data to validate CFD models at steady state conditions.

Modelling approach

Methodology

The 3D conjugate heat transfer of the HP turbine spin pit was carried out on a Linux cluster (3.2 GHz, 3GB, Intel Xeon) using Fluent 6.3.26 for running the simulation and ICEM-CFD for geometry healing, manipulation and mesh generation.

Three CFD models were generated namely 2D, 3D sector and a full 3D model to understand the rig behaviour as well as to analyse the effect of asymmetric blockage on flow field and heat transfer characteristics.

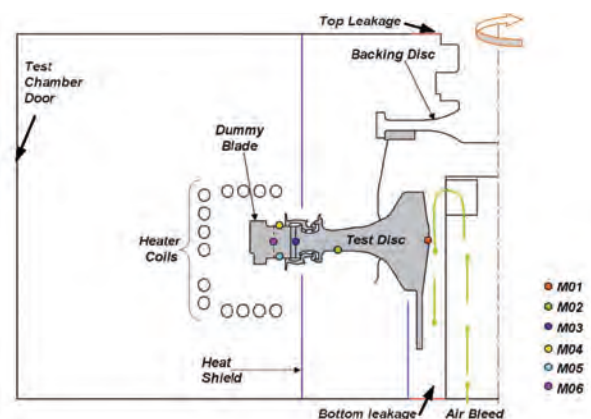


Figure 2 - 2D Schematic of Spin pit test rig chamber geometry

The 3D sector model was 1/66 in size and importantly it captured the blade and firtree geometry accurately. The modelling strategy adopted maintained a level of compromise between accuracy and model size. The solid domains were restricted only to test disc, dummy blades, seal plates and retaining rings which were of prime importance in this study.

The flow regime inside the rig is largely dominated by the rotational motion of the blade disc assembly and the entrainment of flow onto the spinning disc, which pumps the flow radially outwards. The temperature field in the domain is largely impacted by radiation heat transfer between the heaters and component surfaces, convection heat transfer between the components and the fluid and the conduction in the components. Additionally there is significant heat generated through the viscous heating generated by the shearing action of the disc. In order to predict the thermal environment and heat transfer mechanism at the walls accurately, appropriate wall y^+ values were maintained, particularly for the rotating surfaces and radiation exchange between surfaces were calculated. The standard turbulence model with wall functions was employed with viscous heating option enabled. The Discrete Ordinate radiation model was used to model the radiation.

By default, the equations of fluid flow and heat transfer are solved in a stationary frame of reference. However, for cases involving moving zones, it is advantageous to solve the equation of motion in moving

reference frames. With this approach the flow around the moving part can be modelled as steady state. For the spin pit there are many moving parts and stationary surfaces which were not surfaces of revolution. Such problems could be modelled using the Multiple Reference Frames (MRF), Mixing Plane or Sliding Mesh approaches. Owing to the limitations of the Mixing Plane model and the computational cost associated with the Sliding Mesh Model, it was deemed appropriate to use MRF for the present problem. A significant drawback of using the MRF approach is that it does not account for the relative motion of a moving zone with respect to adjacent zones, the grid remaining fixed for the computation. In full 3D model because of the above mentioned limitation non-uniform thermal gradients are setup in the solid components. This was overcome by averaging the temperature field in the solids circumferentially using a custom FORTRAN code.

Due to a lack of information on the full operating conditions of the rig, setup of the boundary conditions was difficult and assumptions had to be made on the flow rates and heater settings. In order to get fastest turn around times for the calculations the 2D axisymmetric model was used. This route was taken with the objective of gaining a quick understanding of the flow field and thermal behaviour of the model and the sensitivity of the solution to the operating variables. The amount of leakage present in the domain was estimated using simple heat balance calculations.

Later a 3D sector model of the rig was run with boundary conditions imposed from the 2D study, primarily to be able to transfer the thermal field for follow on FE analysis. The sector model was finely tuned to predict the metal temperatures to an accuracy of $\pm 10K$ with respect to thermocouple data. Finally the full 3D model was run with settings derived from the sector model to investigate the impacts of asymmetries in rig setups on its behaviour and operation. The temperature field in the solids were circumferentially averaged and compared with the sector model predictions.

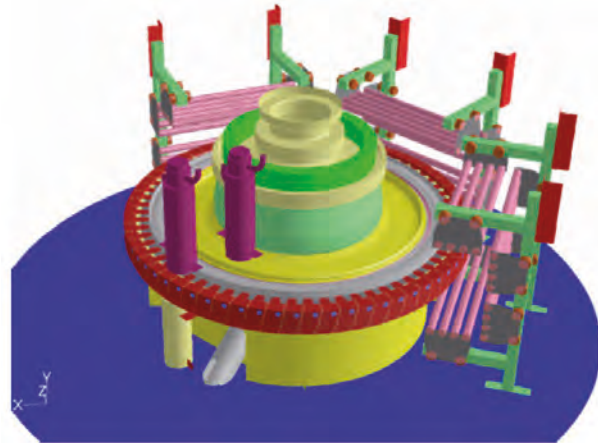


Figure 3 Simplified geometry (doors and top plate not shown)

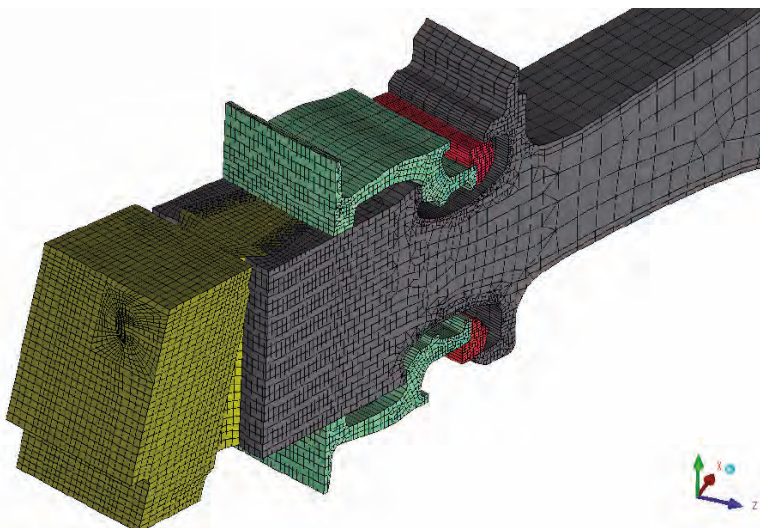


Figure 4 - Mesh-blade and disc assembly

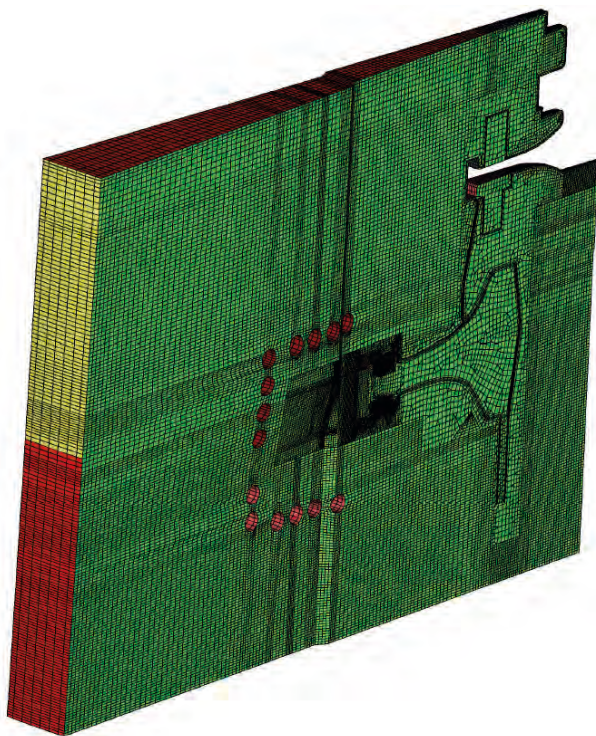


Figure 5 - Computational mesh of the 3D sector model

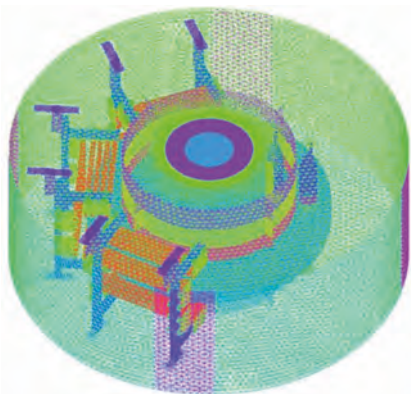


Figure 6 - Computational mesh of the full 3D model

The 3D sector model calculations were run on one compute node, whereas the full 3D model was run on 14 compute nodes over 7 machines.

Geometry and mesh

Figure 1 shows the three dimensional CAD model of the spin pit test rig. For CFD modelling, minor modifications and simplifications were applied to the CAD geometry, resulting in the domain presented in Figure 3. The simplifications adopted are as follows:

- The extent of the CFD domain was limited to test chamber assembly only
- The turbine disc drum, top and bottom lids and vacuum chamber doors were modelled as baffles with a thickness for conductive heat transfer
- Removal of features and components within the rig with a resolution deemed too small to resolve explicitly
- Cooling coils were not modelled explicitly.

In the MRF implementation the calculation domain is divided into sub-domains each of which may be rotating with respect to the inertial frame. Hence, the domain was divided into two zones namely rotating and stationary, with interface boundaries separating the zones. Different cell generation techniques were used to keep the overall cell count manageable. Moreover a significant degree of modularity was used for ease of model handling and manipulation. This approach allowed the meshes to be generated separately and any potential change in the geometry could be easily replicated. The mesh generated had the highest degree of cell resolution in the fluid about the disc and blade assembly and a reasonable core mesh for the remainder of the domain. This achieved an appropriate mesh resolution to capture the complex flow features around the disc and blade assembly whilst keeping the number of computational cells within reasonable and practical bounds. Cell counts for each of the meshes are summarised in Table 1.

Boundary conditions

The air bleed was defined to be a fixed mass flow rate boundary with prescribed total temperature and turbulence parameters. The heaters used to heat the disc and the blade assemblies to engine operating temperatures were defined to be a constant heat flux boundary with a prescribed output. Although the test chamber was well insulated, gaps existed in the domain. This was clear from the digital photographs of the rig. Some of the key gaps through which flow can either leave or enter the domain were those between the top plate and the drive shaft, annular gap at the bottom and the gap between the test chamber doors. The test chamber is placed inside a pit which was maintained at 1/3 atmospheres. Hence, a fixed pressure boundary was applied to these regions with respective backflow temperatures and turbulence parameters. In the full 3D model the gaps between the test chamber doors were discrete as in the test rig (see Figure 8), whereas in the 3D sector model it was approximated by defining the upper portion of the test chamber door. Cyclic boundary conditions were applied to the angular extents (Sides coloured in green, Figure 5) of the domain for the 3D sector model.

Table 1 - Cell counts

Mesh	Cell Count
3D Sector Model	300x103
Full 3D Model	20x106

Conjugate heat transfer study of a test pit rig:
Application to the lifing of HP turbine disc firtrees

Table 2 - Boundary conditions

Zone	Boundary condition
Air Bleed	Mass Flow Inlet
Top Leakage	Pressure Outlet
Bottom Leakage	Pressure Outlet
Leakage through test chamber doors	Pressure Outlet
Dummy Blades	Coupled
Blade-Disc Firtree	Coupled, $\epsilon = 0.9$
Disc surface	Coupled, $\epsilon = 0.9$
Heater assembly	Constant Heat flux
All other rotating surfaces	Adiabatic, $\epsilon = 0.5$
Top & Bottom plates	Constant Heat flux
Test chamber doors	Constant Heat flux
All other stationary surfaces	Adiabatic, $\epsilon = 0.5$

As conjugate heat transfer is of importance in the present study a wall will have a fluid or solid region on each side. In the present case the two sides of the wall were coupled hence no additional thermal boundary conditions were required. However, appropriate material type, wall thickness and radiation properties were specified. The disc and blade assembly was coated with constant emissivity paint with a value of 0.9. In the Multiple Reference Frames (MRF) model used in the present case definitions of velocity boundary conditions at the walls are important. For each fluid cell zone in the domain, the angular velocity (rpm) and the axis about which it rotates is specified and wall velocity is made relative to the velocity of the adjacent cell zone. Hence, all solid walls that are moving are given the speed of the moving frame and a relative angular velocity of zero was

specified. Likewise, all walls that are stationary in the moving frame of reference are given a velocity of zero in the absolute reference frame. Most surfaces away from the blade-disc assembly, where conjugate heat transfer was not of interest were considered adiabatic. Wall regions like test chamber doors, top and bottom lid of the test chamber were applied a constant heat flux boundary condition based on simple heat transfer correlations to account for the heat loss through these boundaries. All surfaces were also assumed to be machined with a significant degree of finishing, allowing the surface roughness to be specified using the default smooth roughness coefficient of 0.5.

Material properties

The analysis included 2D and 3D conjugate heat transfer calculations.

Hence the solution domain included both fluid and solids:

- The fluid inside the solution domain was defined as air. Density was varied as a function of temperature and pressure using the ideal gas law. Whereas specific heat capacity, dynamic viscosity and thermal-conductivity were varied as a function of temperature using the piecewise-linear option.
- The solids inside the solution domain were modelled based on their respective materials. Blades, HPT disc, seal plates and retaining ring conductivities and specific heats were modelled using data provided by Rolls-Royce.
- The small gaps between solid components were modelled as solid-air with thermal conductivity of air.

Convergence

In order to see the evolution of the thermal field and also to tune the CFD model against the temperatures obtained from thermal survey, monitoring points were defined at the exact locations of the thermocouples around the HPT disc and blade assembly. The assessment of run convergence was based on the magnitude and histories of the global residuals. The levels of convergence of residuals were of the order of 10^{-3} for continuity, 10^{-5} for the momentum, 10^{-6} for the energy equation and 10^{-6} for the radiation. Convergence was also based on the evolution of the temperature monitors/histories. The steady

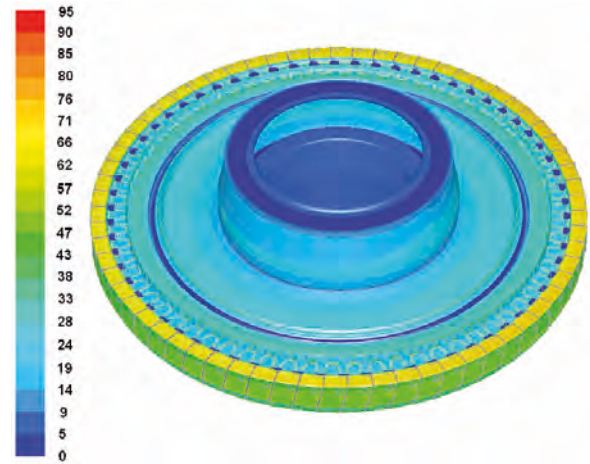


Figure 7 - Contours of y+ (disc & blade assembly)



Figure 8 - Gap between the test chamber doors

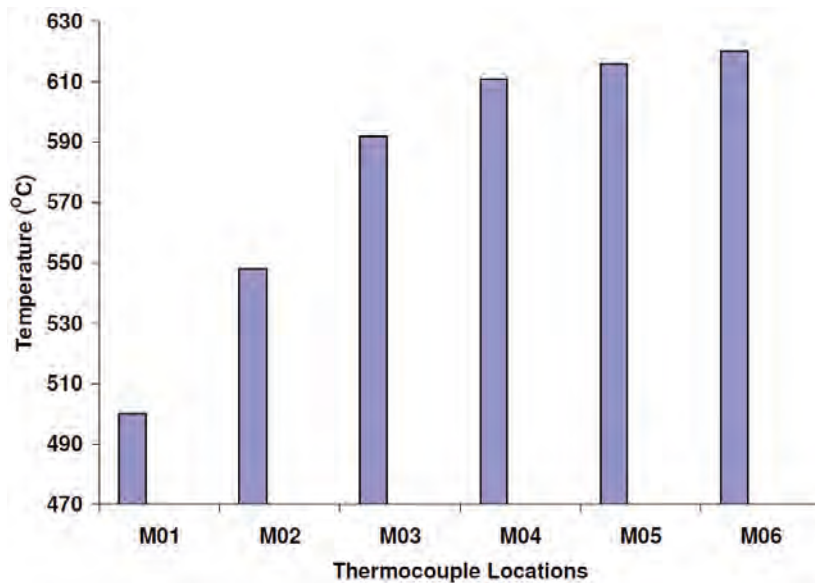


Figure 9 - Thermocouple data for steady-state test

solution was thought to be converged when the residuals were as low as indicated above and the monitor mean temperatures stable.

Spin pit tests

The test data 2 used for the validation of the CFD model was a steady state spin pit test performed on an HP turbine disc-blade assembly by Rolls-Royce. This test had been scheduled to demonstrate the life of the disc firrtree attachments. The test was conducted to optimise the thermal gradients within the HP turbine disc and achieve engine operating conditions.

The instrumented test rig provided metal and air temperatures in the pit. All thermocouples recorded metal temperatures to an accuracy

of $\pm 10K$. Along with the rotational speed, heater power settings were also measured during the test. The spin pit test was run for a variety of different running conditions and it included the steady state run. The steady state test was run at 9000rpm for approximately 2 hours and then for 15 minutes at 13,500rpm. During the test the top sets of heaters were operating at 100%, while the bottom ones were operating at 33% of the maximum. Figure 9 shows the thermocouple data obtained at the end of the spin pit test.

Results and discussion

This section of the paper contains results from the 3D sector model and the full 3D model. Results from the initial 2D model are not presented

here as it only served to provide a rapid understanding of the rig setup. In each case the results have been illustrated using contours and vector plots of velocity and temperature. The illustrated flow field within the spin pit rig corresponds to the results obtained at the end of the tuning process.

Flow features

The large-scale flow patterns inside the domain can be identified from the velocity vector and contour plots of the computations. The main flow features were observed to be the following:

- The disc performs work on the fluid. The intense shearing at the faces of the disc contributes to the heat rise in the fluid through viscous dissipation of turbulence
- There is a clear evidence of radial pumping of the flow over the rotor surface. The stationary enclosure walls around the disk and blade assembly promote recirculation fields above and below the blades
- In both the sector and the full 3D model the jet off the blade is at an inclination (Figure 10 location 4). This is due to the geometrical feature shown in Figure 12 (the line joining the two blades being slightly at an angle). With rotation, this 3D geometrical feature exerts an upwards force on the fluid, bringing about the observed jet inclination.

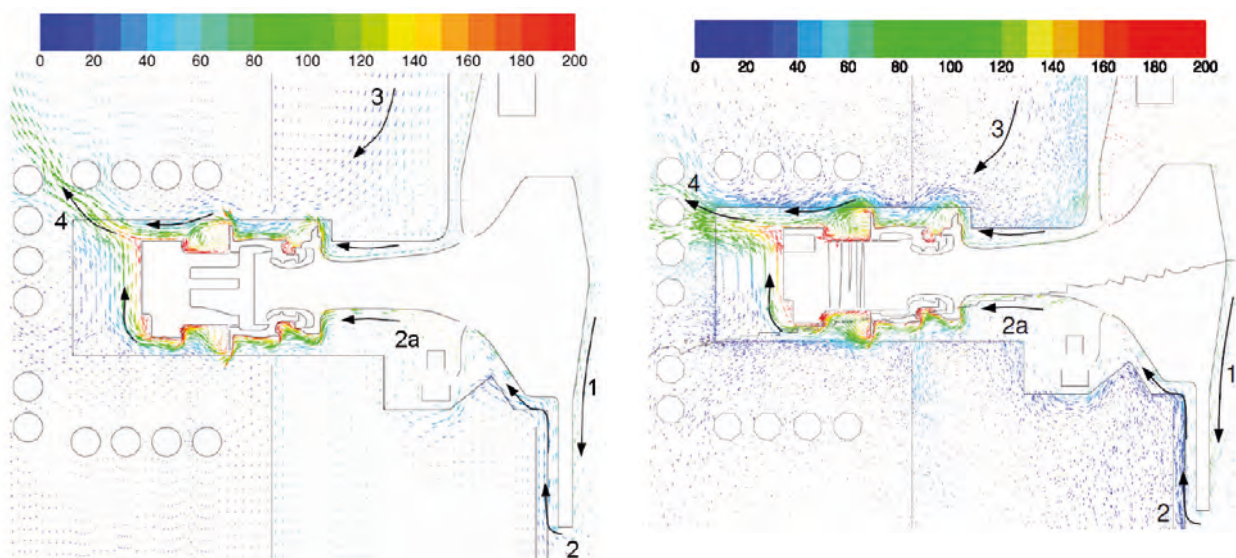


Figure 10 - Vectors of velocity magnitude (m/s) (full 3D)

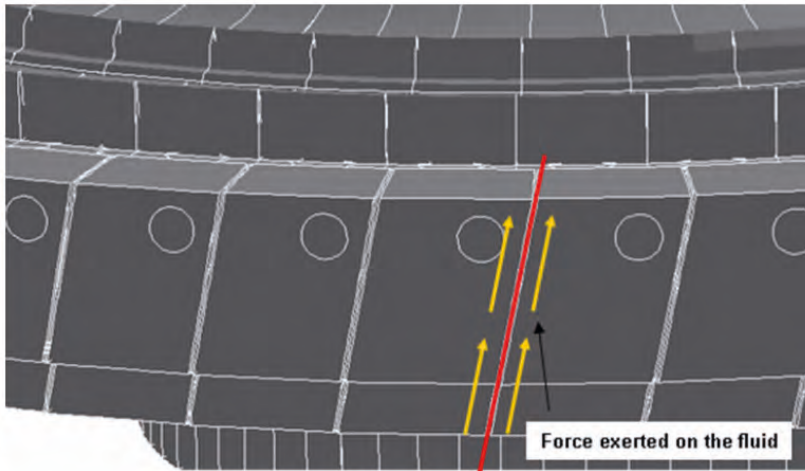


Figure 12 - Direction of force exerted on the fluid



Figure 13 - Contours of velocity magnitude (m/s) (3D Sector)

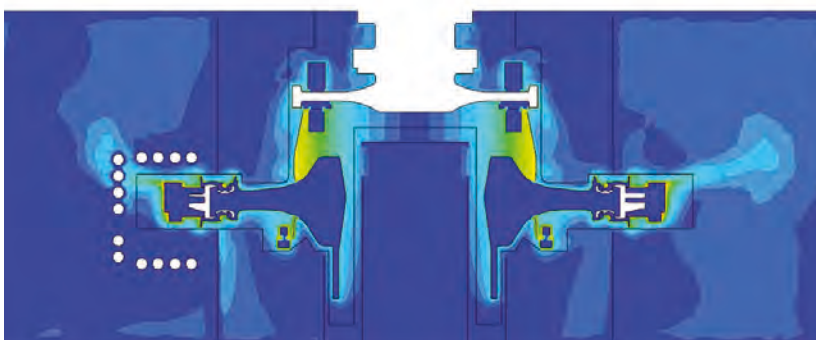


Figure 14 - Contours of velocity magnitude (m/s) (Full 3D)

Temperature (C)

- Cool inflow from the top pressure boundary (Figure 10 location 3) is largely entrained by the rotating disk surfaces
- Figure 10 shows the cool air entering from the air bleed at the bore (location 1). The flow entrained from the bottom pressure boundary mixes with the flow coming from the bore bleed at the bottom of the stub shaft (location 2) and leaks through the small gap between the stub shaft and the small enclosure at the bottom. This flow mixes with the flow entrained and pumped on the underside of the disc (location 2a)
- Significant difference can be noticed in the flow separating from the tip of the blade. In the full 3D model the flow leaves with less of an incline and impinges on the side set of heaters whereas, in the sector model the flow escapes through the gap between the top and side heaters (location 4, Figure 10 & Figure 11). It is thought that the difference is due to the restriction of the sector model domain to one blade (1/66).
- The variations in the flow field due to asymmetric blockages were captured in the full 3D model. The impact of such blockages on the flow field came from the heater components and pyrometers mainly. Due to the frozen rotor assumption these effects were not transported around the domain and averaging had to be performed using a custom code
- Comparison of flow features inside the disc drum and bore regions are very similar and the flow is mainly rotational (Figure 13 and Figure 14).

Thermal field

Figure 15 to Figure 17 show a selection of temperature plots through the domains for the sector and the full models. The hot fluid region to the left of the plot is completely separated from the other zones due to the presence of the heat shields. Because of this the transfer of heat to the disc is primarily via conduction through from the dummy blades.

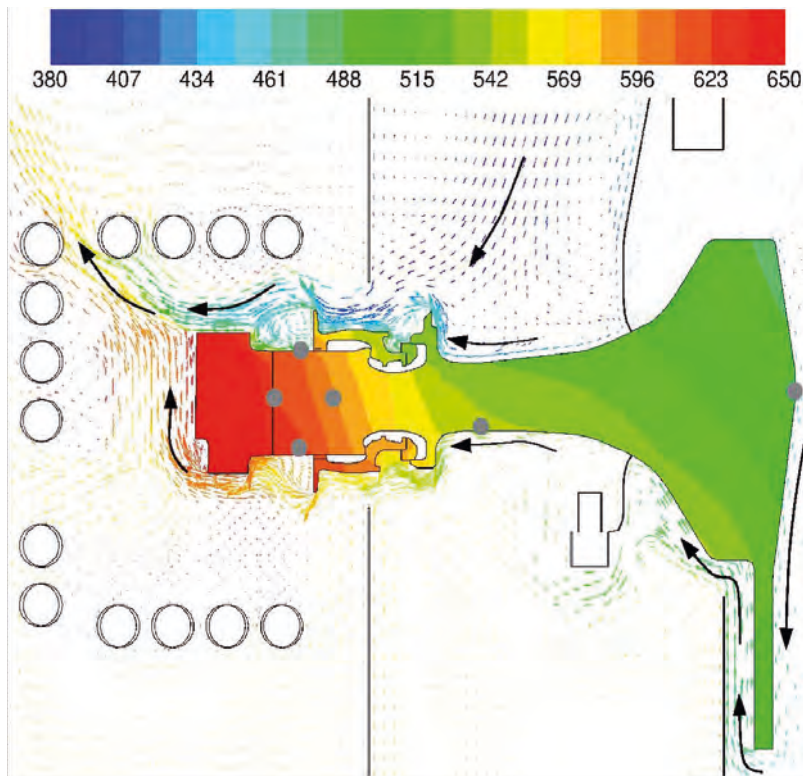


Figure 15 - Contours of temperature overlaid by velocity vectors coloured by temperature (OC) (3D Sector)

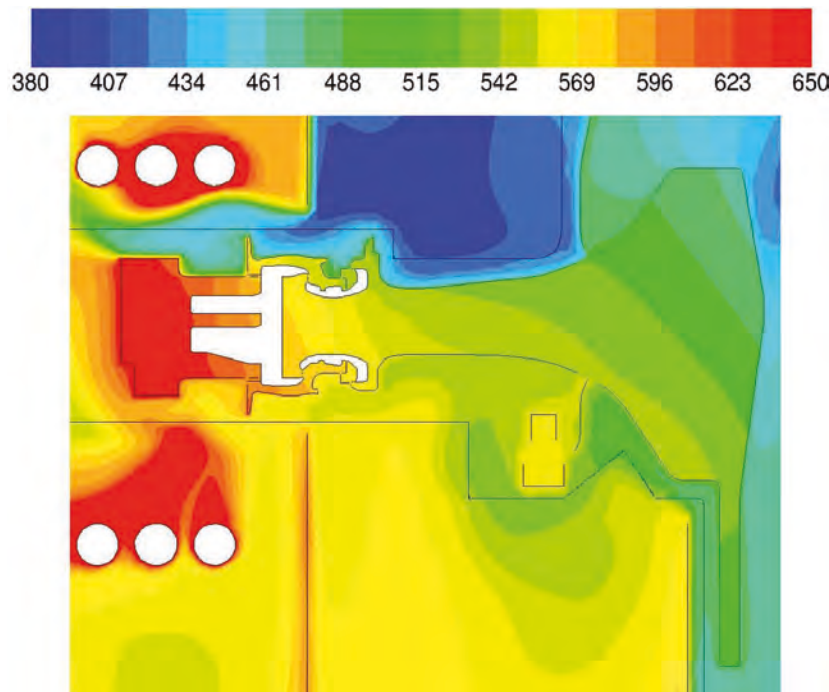


Figure 16 Contours of temperature (OC) (3D Sector)

The temperatures in the solid components reduce gradually from blades down into the disc and there is a significant temperature difference between the top and the bottom of the disc.

This is mainly due to the fact that large amounts of fresh cool air are entrained from the top opening, bringing about cooling on the disc topside. On the underside of the disc in the heat shield, there exists a warm recirculation field and therefore its impact on the cooling of the disc is less significant there.

This sets up a top to bottom temperature gradient across the disc as demonstrated in Figure 15 and Figure 16.

It is important to note that the cooling flow entrained through the top boundary bathes the top surfaces of the disc and blade, explaining the necessity for the high heater settings and the closer position of the top heaters. It is clear that for large parts of the disc blade assembly surfaces, the temperature of the fluid close to the solids is cooler when compared to the metal temperatures. This shows a clear evidence of radiation impact on the thermal field within the domain. The radiation was observed to be the main driver of the heating whilst convection heat transfer impacts top to bottom temperature gradients.

Figure 17(a) and (b) illustrates the results from the full 3D model before circumferential averaging. Figure 17(b) shows the temperature distribution opposite to the heater side. The temperature distribution in the disc and blade assembly is cooler and more inclined as the influence of top to bottom variation in convection heat transfer on this side in the absence of radiation from the heaters becomes more pronounced.

Averaged results

Use of the MRF model establishes non uniform thermal gradients in the solids which is not the case in reality. This problem was overcome by circumferentially averaging the temperature field in the solids using a custom written FORTRAN code. It is well appreciated that circumferential averaging is not the same as the rotor receiving effectively a time-averaged heat flux at its wall boundaries, which would be the case in reality. Given the time constraints on this work circumferential averaging of the thermal field was the best option available. The averaged results presented in this section are compared against the 3D sector model.

Identical plots of domains are generated from both full 3D and sector models to give a direct comparison. The temperature field from the 3D sector model was copied 66 times for comparison. Figure 18(a) and Figure 18(b) show the temperature distributions within the disc and blade assembly.

Conjugate heat transfer study of a spin pit rig: Application to the lifing of HP turbine disc firtrees

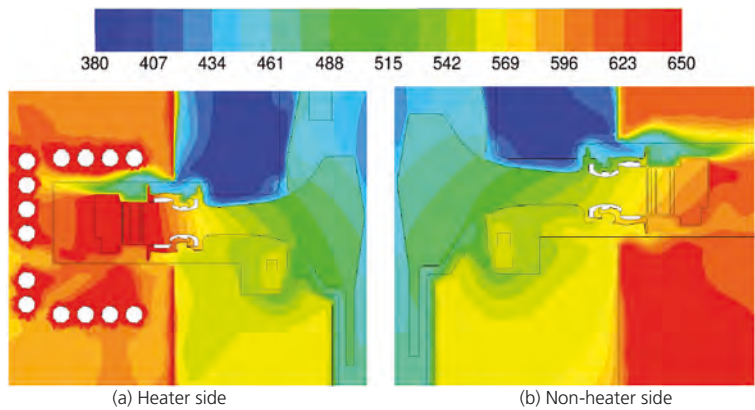


Figure 17 - Contours of temperature (OC) (Full 3D)

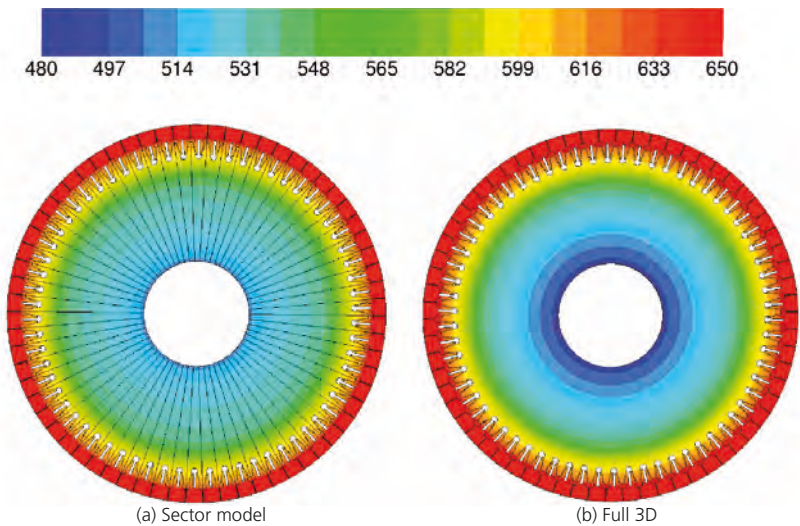


Figure 18 - Contours of temperature (OC) (Full 3D)

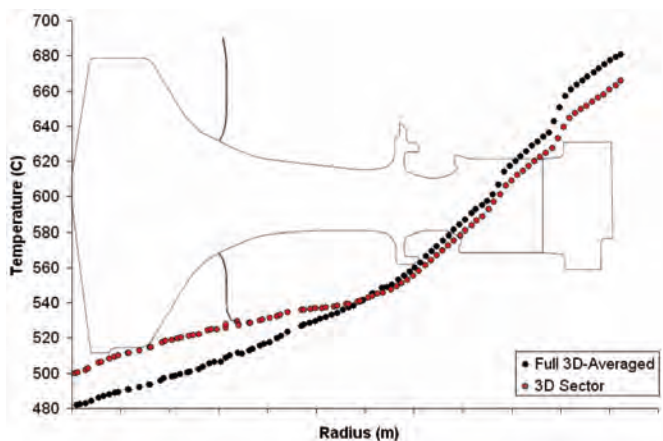


Figure 19 - Temperature differences between sector and full 3D model

Table 3 - Heater power settings

	Percentage of maximum power applied
Top	100%
Bottom	33%

From the plots it is clear that the circumferentially averaged temperatures within the disc and blade assemblies show generally similar distribution when compared against the 3D sector model.

The key findings of the comparisons were the following:

- The dummy blades are slightly hotter in the full 3D calculation
- The difference in temperature distributions around the firtree and the disc rim regions are less significant
- The temperature difference between the top and bottom of the disc firtree is slightly higher when compared to the sector model
- Temperature differences between the sector model and full 3D model are shown in Figure 19. Significant difference can be seen in the disc diaphragm and bore regions, where the sector model is appreciably hotter than the full model. The temperature gradient for the latter in this region is significantly steeper.

Tuning of the CFD model

There was very little information available concerning the spin pit rig operation conditions. In particular, there was no information on the heater settings and flow rates. At the end of the experiments the maximum power input used for each of the heaters was provided by Rolls Royce. These are shown in Table 3.

The 2D model was initially set-up with the estimates of the operating conditions and our best assessment of the radiation properties. A good first assessment of the rig behaviour was obtained from the 2D model. The boundary conditions derived from the initial running of the 2D axisymmetric model were applied to the 3D sector model. Further heater settings obtained from the tests were applied. Upon analysing the thermocouple data it was observed that the temperature on the top side of the disc firtree was higher compared to bottom. It was observed that there was a significant amount of cool flow being entrained from the top which was responsible for this feature. This also explained the necessity of high heater settings on the top set of heaters.

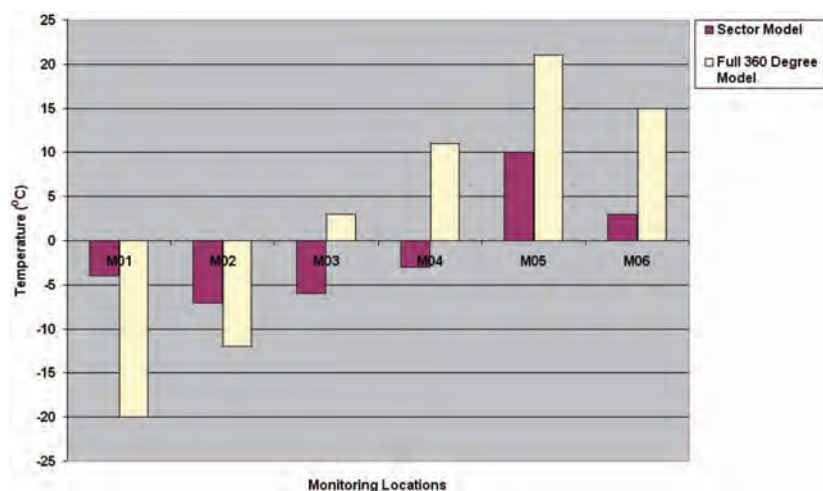


Figure 20 - Comparisons between tests and CFD

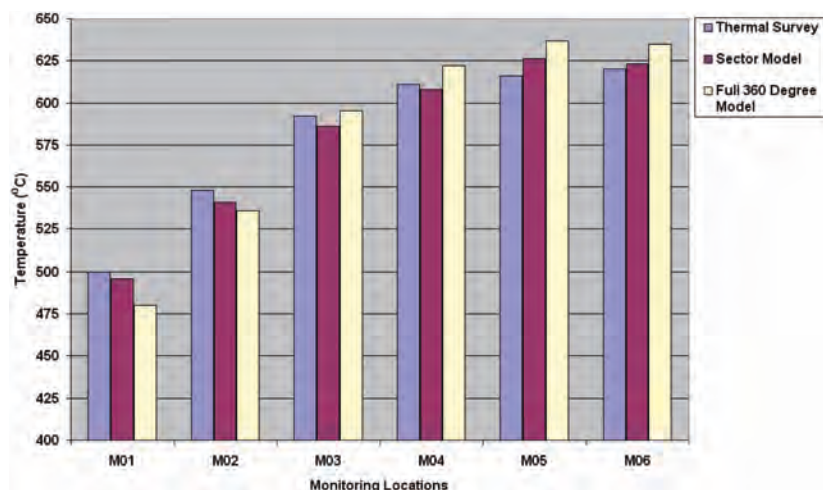


Figure 21 - Difference between CFD and tests

It was clear that the leakage flow temperatures (top and bottom) were important drivers and in the overall heat balance of the assembly and hence the leakage flow temperatures were fine tuned in order to get the best possible match to the thermocouples readings. Later, the final settings obtained from the sector model were applied to the full 3D model to examine impacts of asymmetries.

For the sector model it can be seen that a good match was obtained for all the thermocouples located around the test piece which was within the accuracy limits of the thermocouples. The largest difference is seen in M05 which is at the bottom of the disc firrtree. This region is in a relatively warm zone and hence received a large amount of heat. Further tuning was difficult as any small variation in settings had impacts on all thermocouples. Comparison between the experiment and the full 3D model is not within the accuracy limit of $\pm 10K$.

It is anticipated that this is due to the differences in the manner of leakage flow representations. Owing to the computational cost associated, the results obtained from the full 3D model were deemed appropriate and further tuning was not undertaken.

Conclusions and recommendations

A full 3D conjugate heat transfer analysis of the spin pit rig was performed. The scope of the CFD analysis has enabled a detailed understanding of the flow field inside the domain to be obtained. By conducting a steady-state calculation the variation in flow patterns and temperature distributions inside the domain have been successfully captured. The main findings from the analysis are as follows:

- Flow within the rig was found to be quite complex, with multiple flow mechanisms simultaneously at play in several distinct areas of the solution domain
- From both experiments and simulations it was clear that a large amount of flow leakage was present and this had a significant impact on the flow features and the thermal gradients
- Radiation heat transfer was found to be dominant in terms of the bulk temperature levels achieved in components, whilst convection heat transfer had a major influence on temperature gradients in solid components
- The finely tuned sector model predictions matched well with the thermocouple readings ($\pm 10K$ limit)
- A full 3D model simulation using the setup of the tuned sector model was run to assess the impact of blockage and any asymmetric flow features present. The steady state MRF attempted to capture wakes behind blockages such as pyrometers in the domain. However, due to the frozen rotor assumption these effects were static and not transported around the domain
- Comparison of the flow features and thermal field between the sector model and the full 3D model showed similar features with few exceptions
- Distributions in the fluid were less comparable, particularly in the outer annular enclosure, due to the more accurate treatment of the flow leakages in between enclosure door for the full model

- Temperature field from the steady-state MRF simulation was circumferentially averaged using a custom written FORTRAN program. The result from the averaging program was encouraging and showed a reasonable match with the sector model predictions
- The circumferentially averaged MRF temperatures used for the full model, carried with them the implicit limitation of MRF approach (frozen rotor assumption)
- A recommendation from this study would be to evaluate the changes between circumferential averaging and sliding mesh calculation, even if this is for a test case problem
- The asymmetric blockages seemed to have small impacts on the temperature distributions in the solid components
- The sector model captured the dominant physics for this task and achieved a large saving on the computational effort required
- The study carried out here demonstrates that CFD can be used as a tool for critical parts lifing, accounting for more realistic physics and in combination with testing circumventing the need for additional thermal models.

The predicted component temperature fields from the sector model were directly imported into in-house FE package in order to carry out the stress analysis of the HP turbine disc firtree and lifing of this component. The Low Cycle Fatigue (LCF) tests undertaken provided successful and satisfactory firtree life for the HPT of the engine. The results from the eddy current probe checks on the disc and blades showed no cracks. It was understood that the running of the rig was altogether very satisfactory.

It is understood that this has been the first instance, together with the Bi-Axial study¹² also undertaken concurrently where CFD conjugate heat transfer analysis predictions have been directly used for lifing of the aero engine critical parts leading to engine certification.

Acknowledgments

The authors acknowledge Rolls-Royce for permission to publish this paper.

References

1. Andrew Jopson and Dougal Jackson. HP Turbine Disc Firtree Thermal Survey Test. DNS111494. Rolls-Royce, 2006.
2. Dougal Jackson. HP Disc Spin Pit Thermal Survey Test Data Report. DNS132979. Rolls-Royce. June 2007.
3. Naveen Gopinathrao, Christophe Mabilat and Sohail Alizadeh, HP Turbine Disc Firtree – Conjugate Heat Transfer Analysis of Spin Pit Rig, DNS132072, Oct 2007.
4. Christophe Mabilat and Sohail Alizadeh, HP Turbine Disc Firtree – Conjugate Heat Transfer Analysis of the Bi Axial Rig Enclosure, May 2007.
5. Sattar, S. A., Sundt, C. V., Gas turbine engine disk cyclic life prediction, Journal of Aircraft, 0021-8669 vol.12 no.4 (360-365), 1975.
6. Mahorter, R.London, G. Fowler, Life prediction methodology for aircraft gas turbine engine disks, AIAA-1985-1141 SAE, ASME, and ASEE, Joint Propulsion Conference, 21st, Monterey, CA, July 8-10, 1985.
7. Matthew E. Melis; Erwin V. Zaretsky; Richard August, Probabilistic Analysis of Aircraft Gas Turbine Disk Life and Reliability, Journal of Propulsion and Power, 0748-4658 vol.15 no.5 (658-666), 1999.
8. Hernan V. Arrieta., Patrick Wackers., Ky Dang Van., Andrei Constantinescu., Habibou Maitournam, Modelling attempts to predict Fretting-Fatigue life on turbine components, RTO-MP-AVT-109. June 2004.
9. J A Dixon, J A Verdicchio, D Bentio, A Karl and K M Tham, Recent developments in gas turbine component temperature prediction methods, using computational fluid dynamics and optimization tools, in conjunction with more conventional finite element analysis techniques, Proceedings of the IMechE Part A Journal of Power and Energy, Volume 218, Number 4, pp. 241-255(15), Aug 2004.
10. Glen Snedden, Thomas Roos and Kavendra Naidoo, Detailed Disc Assembly Temperature Prediction: Comparison between CFD and Simplified Engineering Methods, ISABE-2005-1130, 2005.
11. K Saunders, S Alizadeh, L V Lewis and J Provins, The use of CFD to generate heat transfer boundary conditions for a rotor-stator cavity in a compressor drum thermal model, Proceedings of the GT2007, ASME Turbo Expo Power for Land, Sea and Air, Montreal, May 2007.
12. C Mabilat, S Alizadeh, D Jackson and R Clarkson, Conjugate Heat Transfer Study of a Biaxial Rig: Application to the Lifing of HP Turbine Disc Firtrees, ASME Turbo Expo 2008 Power for Land, Sea and Air, June9-13 Berlin, GT2008-51297.
13. John .H Lienhard IV and John .H Lienhard V, A Heat transfer Textbook, Phlogiston Press, Cambridge, Massachusetts, USA, 3rd Edition, 2007.
14. Frank .P Incopera and David .P .De Witt, Fundamentals of Heat and Mass Transfer, John Wiley and Sons, 3rd Edition, 1990.
15. Fluent 6.3.26 Help Manual.

Non-deterministic thermo-fluid analysis of a compressor rotor-stator cavity



Naveen Prasad
Gopinath

Engineer II
Energy



Dr. S. S. Sankar

Associate Professor
Energy



Dr. S. S. Sankar

Associate Professor
Energy

Introduction

In an engine, the compressor rotor-stator cavity is a complex thermo-mechanical system. The analysis of this system is undertaken, to predict the component temperatures and displacements which are ultimately used for material selection, component lifing and blade-tip clearance control. The latter, directly impacting component and engine efficiencies is a particularly challenging issue, influenced by both the thermal and mechanical loads of the rotating and stationary components. Of particular concern are the running clearances at engine high power condition. The blade centrifugal forces as well as the high thermal loading of discs and casing can induce undesirable blade/segment rubs, with the potential to adversely impact engine performance. The component thermal boundary conditions are usually obtained from a small number of standard flow field and heat transfer simulations which are usually undertaken using input parameters which are deterministic in nature. In practice, however, they are subject to inherent physical variability.

These physical input parameters and their associated uncertainty can lead to a significant deviation in the results. Therefore the quantification of such uncertainty becomes an important aspect in the modelling of the physical processes. A significant amount of work on uncertainty quantification in CFD applications applied to external aerodynamic problems has been published.^{5,6,7,10} However, as yet there have not been many publications on uncertainty quantification in gas turbine CFD applications.

In this paper we focus on applying an inexpensive Non-Intrusive Polynomial Chaos (PC) method for uncertainty propagation. The PC method is one among several methods to model and propagate uncertainty in stochastic CFD simulations. Xiu and Karniadakis² applied a generalized PC method, which is an extension of the original PC method first introduced by Wiener¹ to incompressible flow simulations around circular cylinder. This method was intrusive, meaning a coupled system of equations are obtained

and not be solved using existing CFD codes. This attracted researchers to develop an efficient method to model and propagate uncertainty which enables engineers to use their existing CFD codes for uncertainty analysis. Recently, several papers have appeared in the literature on implementation of PC methods that are non-intrusive in nature. Mathelin et al.⁴ investigated turbulent, compressible flow in a quasi 1D nozzle using a non-intrusive method called Stochastic Collocation and demonstrated a significant decrease in computational time when compared to intrusive methods. Walters et al.^{5,7} developed the Non-Intrusive Polynomial Chaos and applied it to three basic fluid dynamics problems. The method starts with replacing the uncertain variable of interest with their polynomial expansions. Then, for a given PC of given order, a random variable vector is chosen and the deterministic code is evaluated at these points. The method is non-intrusive however, and the polynomial coefficients estimated are not unique, since one can choose the random vectors arbitrarily.

In 2006, Loeven et al.⁶ introduced another non-intrusive method called Probabilistic Collocation for efficient propagation of arbitrarily distributed uncertainties. The method combines the idea of PC method and the Stochastic Collocation approach. Exponential convergence is obtained with respect to the polynomial order for arbitrary probability distributions since the optimal collocation points are computed based on the input distribution. The probabilistic collocation method computes the polynomial coefficients using Gaussian quadrature which results in a decoupled set of equations, making the method non-intrusive. The efficiency of Probabilistic Collocation method was demonstrated on a linear piston problem. Furthermore, the non-intrusiveness of the method was demonstrated by simulating steady flow around NACA0012 airfoil with uncertainty in free stream velocity using the FineTM/Hexa solver of Numeca Int.

In this paper, the Probabilistic Collocation method is applied to an industrial rotor-stator cavity in front of an IP compressor of a large bypass turbofan aero engine with uncertainty in the operating conditions.

The deterministic CFD simulations are carried out using FLUENTTM and thermo-mechanical modelling using the finite element program SC03. The background of the PC approach including the probabilistic collocation method is outlined, the results obtained from the investigation of the rotor stator cavity and the assessment of impacts of uncertainty on component temperatures and blade-tip clearances is presented and the conclusions drawn.

Probabilistic Collocation method

The PC method results in a spectral representation of the uncertainty parameters. An important concept of this approach is the division of uncertain variable into deterministic and stochastic parts. The polynomial chaos expansion of the solution is generally expressed as,

$$u(x, \theta) \approx \sum_{i=0}^P u_i(x) \Psi_i(\xi(\theta))$$

where, $u_i(x)$ is the deterministic part $\Psi_i(\xi(\theta))$ is the stochastic part of the random variable. The total number of expansion terms is $P + 1$, which is determined by \mathcal{N} the dimension of the vector of random $\xi(\theta)$ variables and the order of p of the polynomial Ψ_i .

$$P + 1 = \frac{(n + p)!}{n!p!}$$

In Probabilistic Collocation method the random trial basis $\{\Psi_i\}$ is a Lagrange interpolating polynomial denoted by h_i . Each variable depending on the uncertain input parameter is expanded as follows:

$$u(x, \theta) \approx \sum_{i=0}^P u_i(x) h_i(\xi(\theta))$$

The Lagrange interpolating polynomial is a function in terms of the random variable $\xi(\theta)$. It is a polynomial chaos $h_i(\xi(\theta))$ of order $p - 1$ that passes through the p collocation points. The collocation points are chosen such that they correspond to Gaussian quadrature points used to integrate the function $u(x, \theta)$ in the θ domain. Computation of Gaussian quadrature points and corresponding weights is detailed in Loeven et al.⁶ The Probabilistic Collocation method is a generalization of Stochastic Collocation method.

The inner product in the stochastic space is expressed as,

$$\left\langle \sum_{i=1}^P u_i(x) h_i, h_k \right\rangle \quad k = 1, \dots, P$$

This projection is approximated using Gaussian quadrature, with optimal collocation points and corresponding weights based on the input distribution. This results in a system of equations which is fully decoupled. The mean and variance of the solution are calculated using

$$\mu_u = \sum_{i=1}^P u_i(x) w_i$$

$$\sigma_u^2 = \sum_{i=1}^P (u_i(x))^2 w_i - \left(\sum_{i=1}^P u_i(x) w_i \right)^2$$

where w_i are the weights corresponding to the collocation points θ_i .

Probabilistic CFD analysis of a rotor-stator cavity

The non-intrusive polynomial chaos (Probabilistic Collocation) approach has been applied to a rotor stator cavity in front of an IP compressor of a large bypass turbofan aero engine.^{8,9} Figure 1 shows the arrangement of the IP Compressor (IPC) stage 1 and the Front Bearing Housing (FBH) structure at the front of a large civil aero engine.

Air is drawn forward through the compressor drum, and emerges from the front stub shaft, via a low radius labyrinth seal (RBS), into the large cavity in front of the stage 1 disc diaphragm. There it mixes with the cooler air emerging from the large holes in the rear panel of the FBH structure. The whole flow exits the cavity at the high radius labyrinth seal in the front of the IPC 1 rotor (FAS).

The deterministic CFD simulations are carried out using the FLUENT CFD code. The model was 2D axisymmetric, with swirl transport equations enabled. The flow inside the cavity is modelled by the Reynolds-Averaged Navier-Stokes equations using the 2-layer, realisable $k - \epsilon$ turbulence model. Figure 2 shows the grid resolution in the vicinity of the RBS and FAS seals. The mesh contained a structured boundary layer mesh and an unstructured core with a smooth transition in between.

The RBS and FBH holes were defined to be fixed mass flow rate boundaries with prescribed values of total temperature, hydraulic diameter and turbulent intensity. The FAS was prescribed using fixed pressure boundary condition. The flow properties at the RBS, FBH and FAS are summarised in Table 1 is for Max take Off (MTO) condition. Rotational speed of the IPC 1 disc was 7824rpm. The surface boundary of the CFD cavity was split-up into 10 different wall sections as shown in Figure 1.

The stochastic rotor-stator cavity problem was formulated by introducing uncertainties through the operating conditions (boundary conditions). The input uncertainties considered are The results obtained from the Probabilistic Collocation method for the various input uncertainties detailed in Table 2 are illustrated here. A fifth order approximation is used, with six collocation points for each uncertain parameter. The deterministic CFD simulation was performed at each collocation point.

The mean, variance and standard deviation of adiabatic wall temperature and heat transfer coefficients are computed. The propagation of uncertainty is also represented by plotting standard deviation of the temperature field in the cavity.

Figure 3 to Figure 5 illustrates the standard deviation of the temperature field obtained for the various input uncertainties using beta and uniform probability distributions. Furthermore, the influence of different input uncertainties on mean and standard deviation of wall temperature and heat transfer coefficient is also given.

From the standard deviation plots it is clear that,

- The inlet (T0)FBH-Holes results in maximum deviations close to walls 1 and 2. Significant deviations are seen in the core due to the entrainment of this flow in the core recirculation (Figure 3).
- The influence of inlet m' FBH-Holes uncertainty had similar levels of effect on the temperature deviations in the cavity (Figure 4).
- In case of multiple uncertain parameters (m' RBS and m' FBH-Holes) the standard deviations differ from the results obtained from the uncertainty parameters considered individually (Figure 5).
- The influence of inlet (T0)RBS and m' RBS uncertainties had a weak effect on the temperature deviations in the cavity.
- In all the cases it was observed that the uniform probability

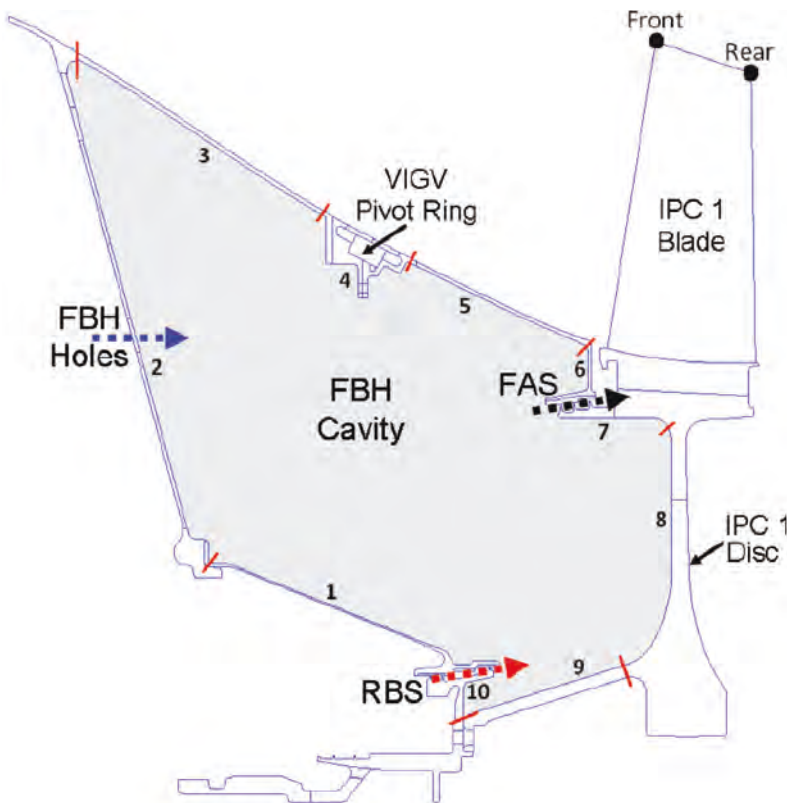


Figure 1 - IPC 1 and FBH Cavity

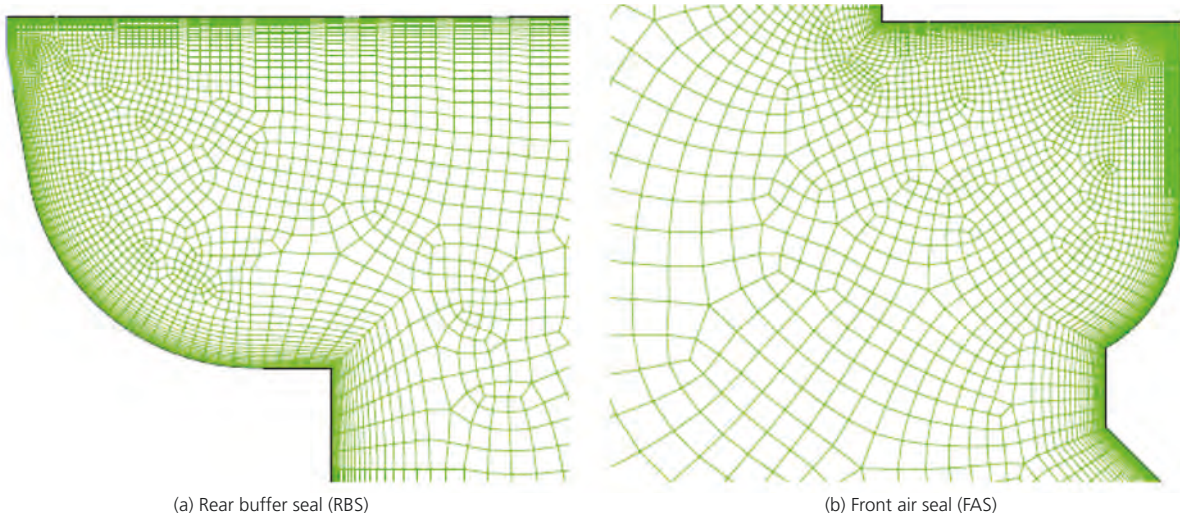


Figure 2 - Mesh resolution

Table 1 - Inflow and outflow boundary conditions

	Property	Value
RBS	\dot{m} (kg/s)	0.3
	T0(K)	578.9
	Whirl Fraction	0.5
	Hydraulic Diameter (mm)	0.8
	Turbulent Intensity	10%
FBH	\dot{m} (kg/s)	0.327
	T0(K)	483.0
	Holes Whirl Fraction	0.0
	Hydraulic Diameter (mm)	37.0
	Turbulent Intensity	10%
FAS	P(kPa)	401.5
	T0(K)	522.9
	Hydraulic Diameter (mm)	0.2
	Turbulent Intensity	10%

Table 2 - Uncertain parameters

		Beta PDF	Uniform PDF	
	Extent	Most Likely	Mean (μ)	SD (σ)
RBS Inlet Total Temperature (K)	[558.9, 598.9]	578.9	578.9	11.5
FBH-Holes Inlet Total Temperature (K)	[433.0, 533.0]	483.0	483.0	28.8
RBS Inlet \dot{m} (kg/s)	[0.06, 0.36]	0.3	0.210	0.08
FBH-Holes Inlet \dot{m} (kg/s)	[0.06, 0.39]	0.33	0.230	0.094
RBS and FBH Inlet \dot{m} (kg/s)	* See 3rd and 4th rows respectively.			

distribution brought about larger deviations to the temperatures in the cavity compared to the beta distribution.

The uncertainty on wall temperature and heat transfer coefficient due to the various uncertain parameters are illustrated in the charts of Figure 6 and 7. Mean wall temperatures obtained for uncertain parameters (T0)RBS and (T0)FBH-Holes were almost identical to deterministic values and hence not presented. However, \dot{m} FBH-Holes and multiple uncertain parameters with uniform distribution had significant influence on the mean wall temperatures. Comparison of standard deviations in Figure 6(b) shows significant deviations about the mean for (T0) FBH-Holes, \dot{m} FBH-Holes and multiple uncertain parameters with uniform input probability distribution.

For other cases the deviations in the wall temperatures were observed to be very modest. According to Lewis et al.³ experience a 15K change in the cavity temperature can require a change in the material from which the component is made.

Comparison of mean heat transfer coefficients obtained for the different input uncertain parameters showed no significant difference hence are not presented. The influence of different uncertain parameters on wall heat transfer coefficients is clearly evident from the standard deviations plots in Figure 7.

Probabilistic thermo-mechanical analysis of a rotor-stator cavity

The thermo-mechanical modelling was carried out using the Rolls-Royce proprietary finite element analysis code SC03. The code is capable of performing fully transient analysis of the engine running conditions.

The analysis program has the capability to represent a variety of thermal and mechanical boundary conditions found in gas turbine type applications. The capability to define the cooling flows to represent the heat exchange is of particular importance. The thermal boundary condition for example contains mass flow, flow direction, fluid temperature, heat transfer coefficient and pressure. Heat transfer

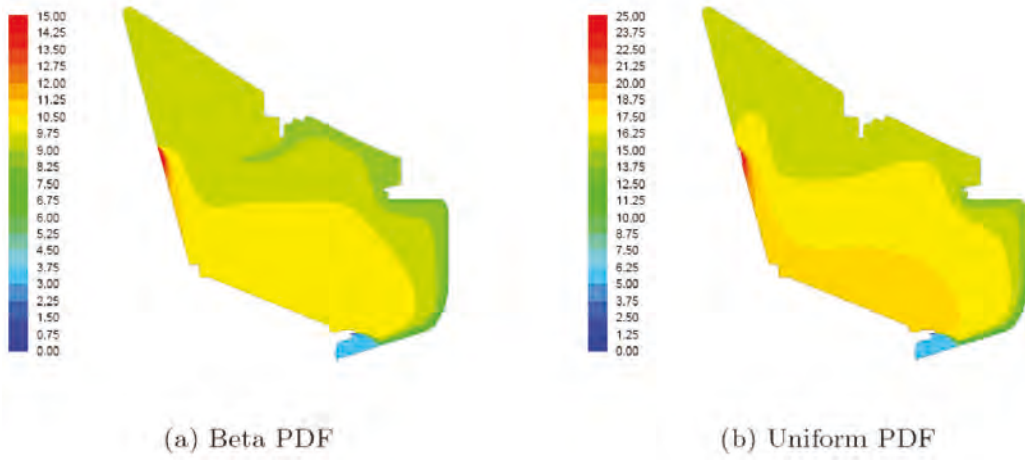


Figure 3 - Standard deviation of the temperature field (K) - Uncertain inlet $(T_0)_{FBH-Holes}$.

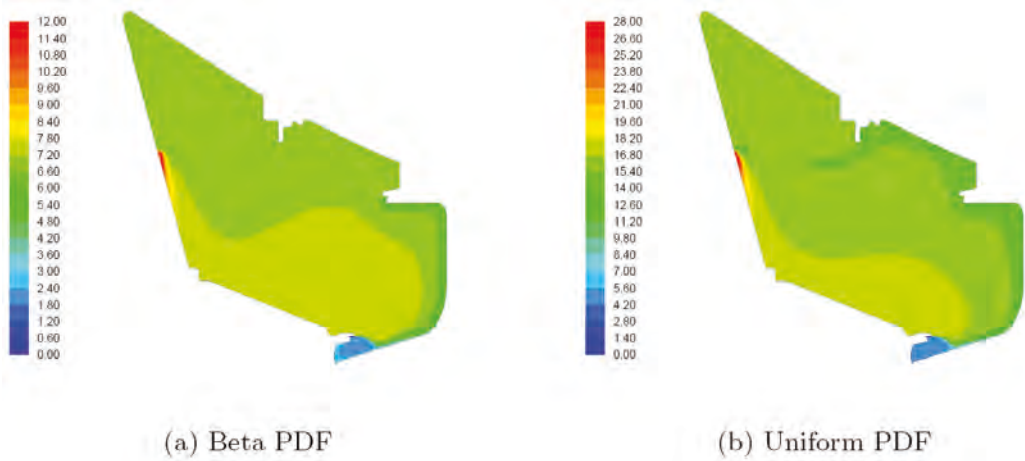


Figure 4 - Standard deviation of the temperature field (K) - Uncertain inlet $\dot{m}_{FBH-Holes}$.

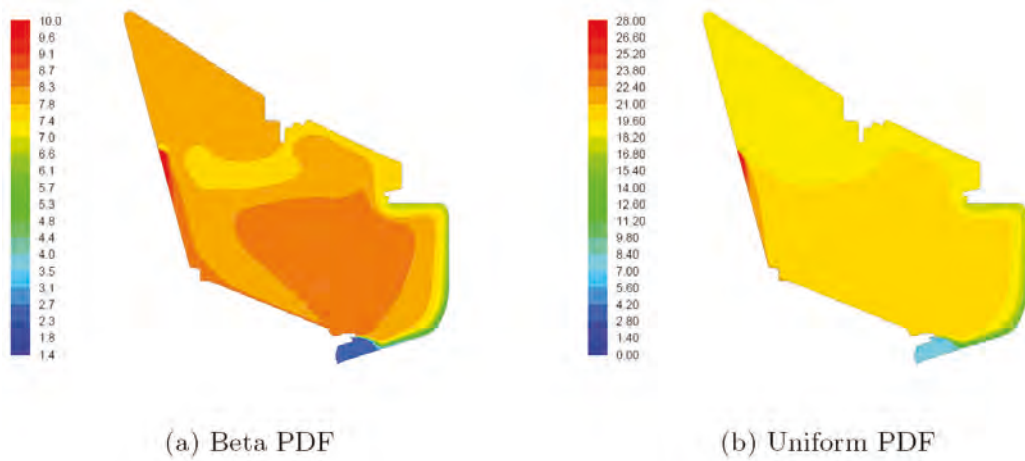
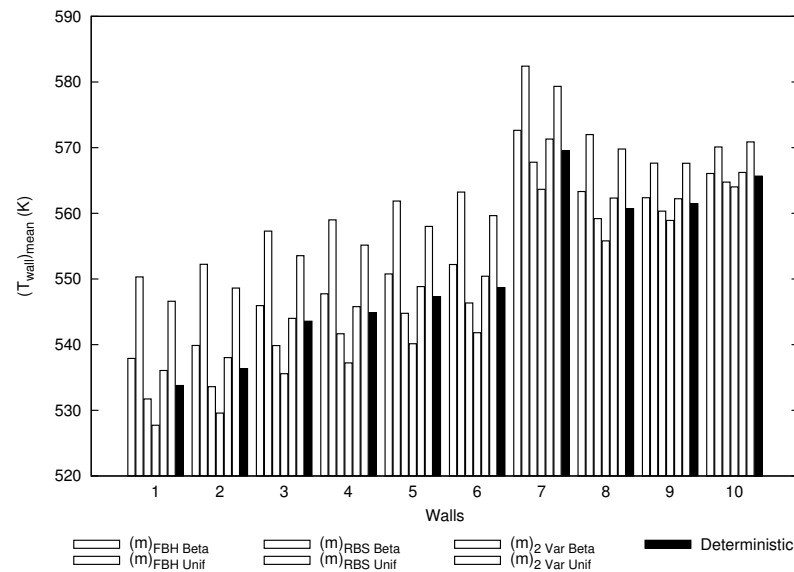
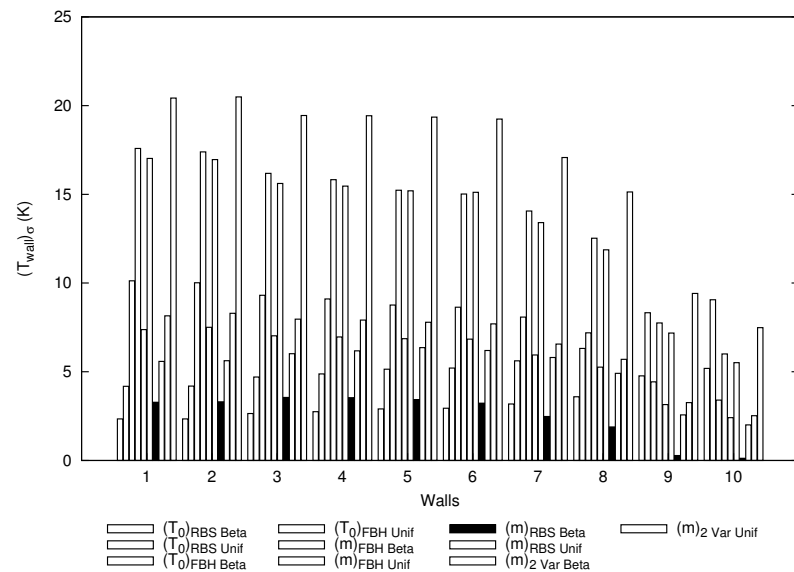


Figure 5 - Standard deviation of the temperature field (K) - Uncertain inlet \dot{m}_{RBS} and $\dot{m}_{FBH-Holes}$.



(a) Mean wall temperature



(b) Standard deviation wall temperature

Figure 6 - Influence of uncertainties on wall temperatures

coefficients are generally applied using standard convective heat transfer correlations. CFD is increasingly being seen as a tool to define the set of heat transfer boundary conditions in complex secondary air system cavities. In this study, the Temperature Influence Coefficient (TIC) method proposed by Lewis and Provins³ is used for transferring the heat transfer information from CFD to thermo-mechanical analysis.

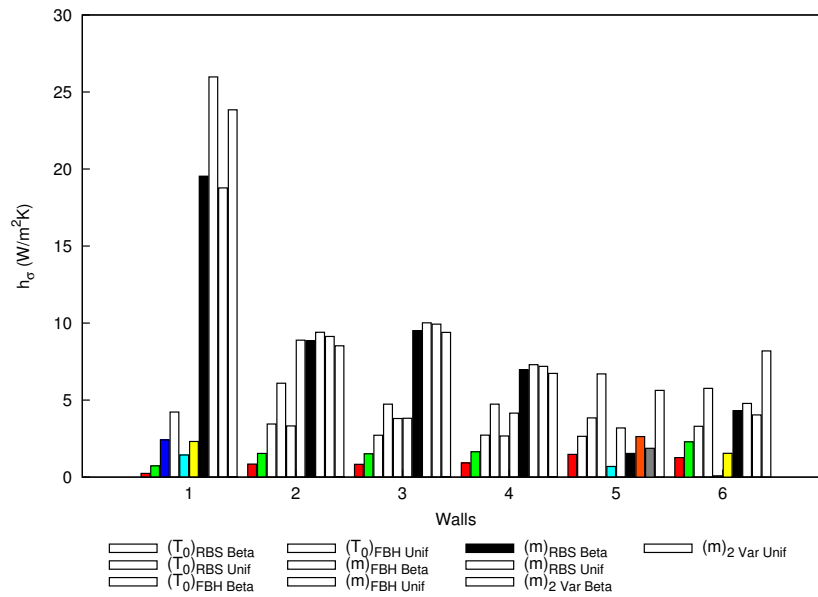
The thermo-mechanical models were 2D axisymmetric and comprised approximately 9000 six-noded triangular elements. The transient thermo-mechanical analyses were

carried out using a stabilised slam accel to MTO and a typical flight cycle as shown in Figure 8. The outputs from the thermo-mechanical analyses were transient temperatures and radial displacements.

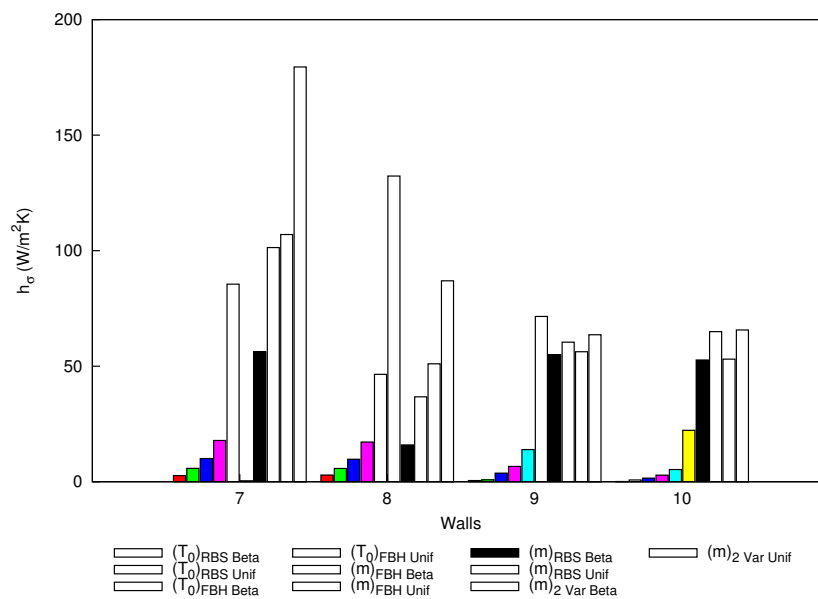
The heat transfer results obtained for the (T0)FBH—Holes with a uniform probability distribution exhibited the largest variations. Hence, results from this case was used to investigate the influence of input uncertainty on component temperatures and blade-tip clearances.

Figure 9 and 10 shows the mean and standard deviation of the temperature field obtained for the (T0)FBH—Holes

uncertain parameter with uniform probability distribution. Figure 9(a) shows temperature contours for the fully stabilised MTO case and Figure 9(b) at the end of the acceleration to MTO for a typical flight cycle (Figure 8). Plots clearly show the propagation of heat into the disc and blade assembly from the cavity fluid. Maximum temperatures and stresses occur during max take-off. At fully stabilised conditions the disc and blade regions are significantly hotter compared to end of accel due to thermal inertia.



(a) Standard deviation HTC, Stator walls



(b) Standard deviation HTC, Rotor walls

Figure 7 - Influence of uncertainties on heat transfer coefficients

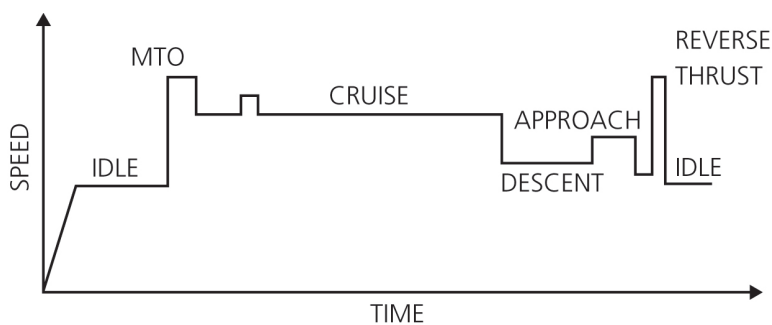


Figure 8 - Typical flight cycle

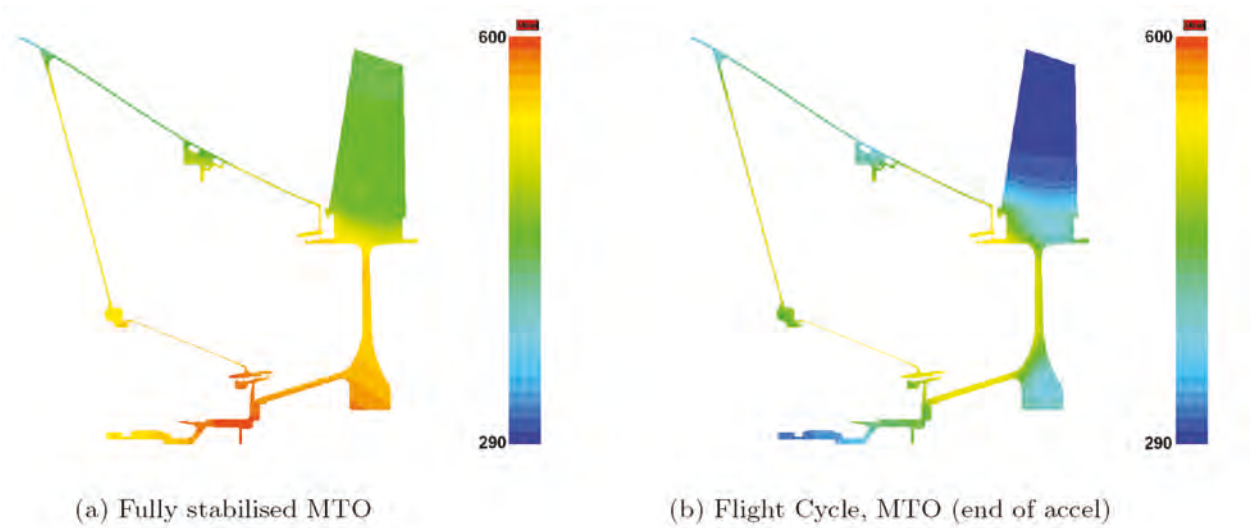


Figure 9 - Mean temperature contours (K)

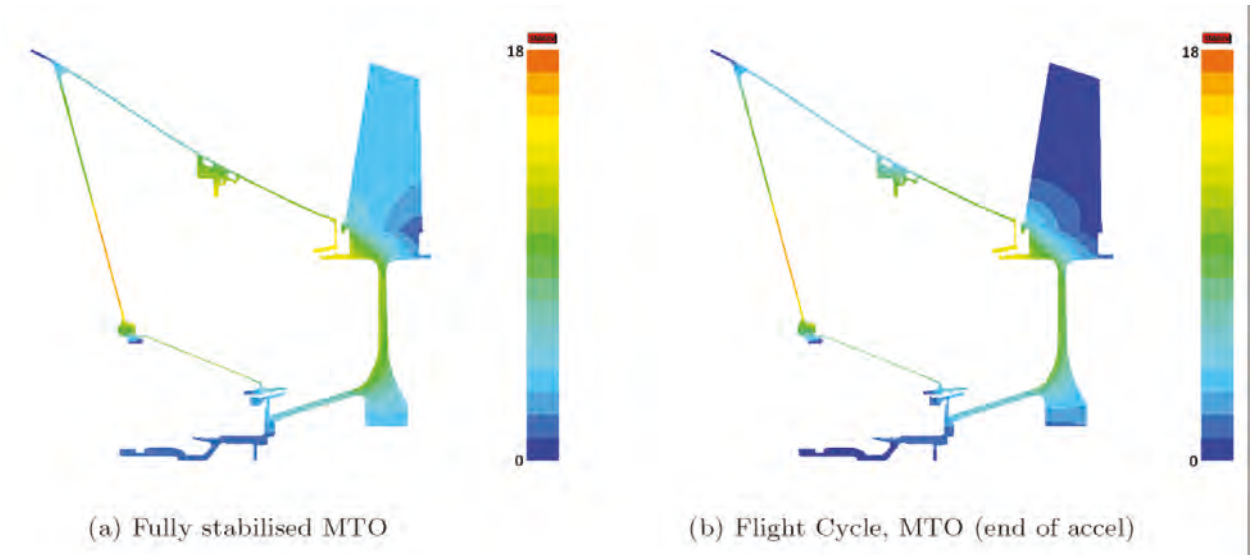


Figure 10 - Standard deviation of the temperature field (K)

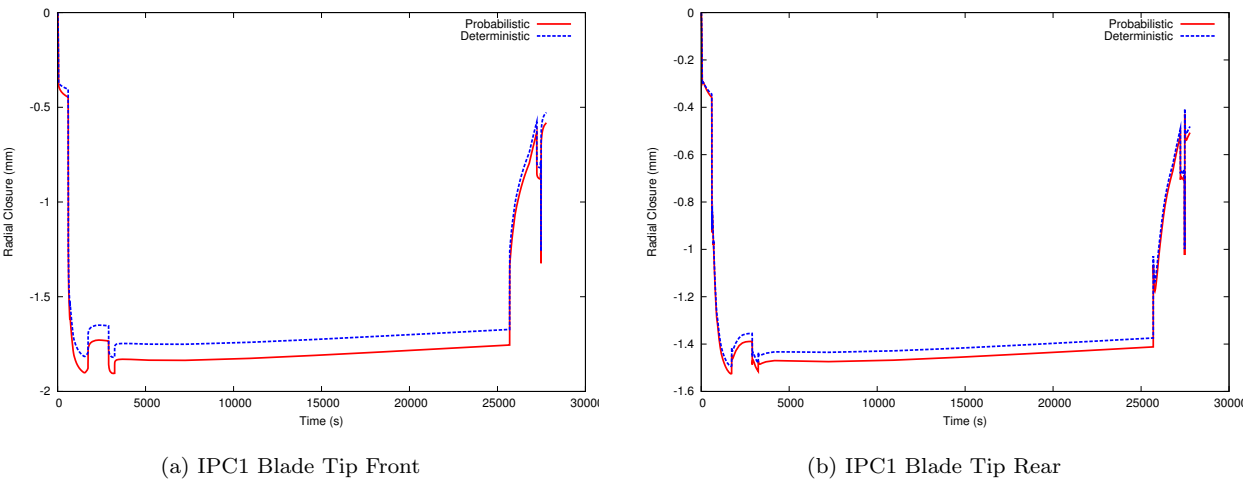


Figure 11 - Comparison of probabilistic and deterministic radial tip closures

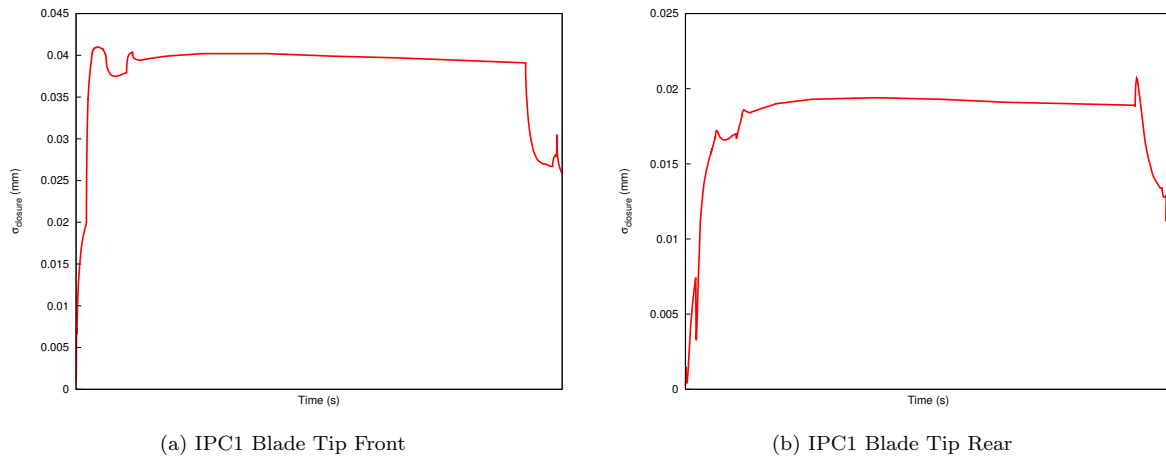


Figure 12 - Standard deviation - IPC1 blade radial tip closure

It is interesting to note that the standard deviations about the mean for both the stabilised and flight cycle are similar. Temperature deviations of about 14K and 10K are observed around the front air seal and disc diaphragm regions respectively, whilst, the lower portion of the static panel containing FBH holes shows the largest deviations. This latter observation is not surprising given the proximity of the source of the uncertainty and the jet flow bathing this panel below the FBH holes.

Blade tip clearance has been a challenging problem since the development of the gas turbine engine. They can have a significant influence on the component efficiency, surge margin and specific fuel consumption.

Better understanding of the variation at the earliest possible stage in the design cycle helps to improve engine efficiency by optimising both transient and steady-state tip clearances. The uncertainty on the blade tip clearances due to the (T0) FBH-Holes uncertain parameter with uniform probability distribution is illustrated in Figure 11 and 12. The comparison of mean radial tip closures with the deterministic results show a significant difference for the majority of the flight cycle. However, the standard deviations about the mean were observed to be modest. Such calculations can potentially increase the accuracy of tip-clearance calculation, reducing the "unknowns" and provide greater confidence in the setting of the cold-build clearances of the engine.

Conclusion

A non-deterministic thermo fluid simulation based on Probabilistic Collocation method has been applied to a rotor stator cavity in front of an IP compressor of a large bypass turbofan aero engine. The impacts on wall temperatures and heat transfer, have been assessed with different sources of uncertainty in operating conditions. The (T0)FBH-Holes and m' FBH-Holes as independent uncertain parameters showed higher deviations. A multivariable calculation with these two parameters could yield much higher deviations due to the combined effect. The results obtained from CFD are carried through a non-deterministic thermo-mechanical analysis to assess the impacts on component temperatures and displacements. Significant deviations are observed in key component parameters presented. The calculations presented demonstrate how uncertainty sources in engine operating conditions, can be managed in computational flow field and component thermo-mechanical calculations, providing greater confidence in the design and assessment of component performance.

Acknowledgments

This work was supported by the project NODESIM-CFD "Non-Deterministic Simulations for CFD-based Design Methodologies" funded by the European Community represented by CEC, Research Directorate-General, in the 6th Framework Program, under Contract No. AST5-CT-2006-030959.

The authors would also like to thank Rolls-Royce for the provision of the test case application and operating conditions data.

References

1. Wiener, N., "The Homogenous Chaos," Amer. J. Math., Vol. 60, pp. 897-936, 1938.
2. Xiu, D. and Karniadakis, G. E., "Modelling Uncertainty in Flow Simulations via Generalized Polynomial Chaos," Journal of Computational Physics, Vol. 187, No. 1, pp. 137-167, May 2003.
3. Lewis, L. V., Provins, J., "A Non-coupled CFD-FE Procedure to Evaluate Windage and Heat Transfer in Rotor-stator Cavities", " Proceedings of the ASME Turbo Expo 2004: Power for Land, Sea and Air, GT2004-53246, 2004.
4. Mathelin, L., Hussaini, M. Y., and Zang, T. A., "Stochastic approaches to uncertainty quantification in CFD simulations," Numerical Algorithms, Vol. 38, pp. 209-236, 2005.
5. Hosder, S., Walters, R. W., and Perez, R., "A Non-Intrusive Polynomial Chaos Method for Uncertainty Propagation in CFD Simulations," Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit, AIAA paper 2006-891, Reno, January 2006.
6. Loeven, G. J. A., Witteveen, J. A. S., and Bijl, H., "Probabilistic Collocation: An Efficient Non-Intrusive Approach For Arbitrarily Distributed Parametric Uncertainties," Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit, AIAA paper 2007-317, Reno, January 2007.
7. Hosder, S., Walters, R. W., and Balch, M., "Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables," Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, AIAA paper 2007-1939, Honolulu, April 2007.
8. Saunders, K., Alizadeh, S., Lewis, L. V. and Provins, J., "The Use of CFD to Generate Heat Transfer Boundary Conditions for a Rotor-Stator Cavity in a Compressor Drum Thermal Model, " Proceedings of the ASME Turbo Expo 2007: Power for Land, Sea and Air, GT2007-28333, Montreal, May 2007.
9. Naveen Prasad Gopinathrao, Christophe Mabilat, Sohail Alizadeh, "Use of Polynomial Chaos to Assess the Impact of Uncertainty on Heat transfer in a Gas Turbine Rotor-Stator cavity", Proceedings of the 10th AIAA Non-Deterministic Approaches Conference, AIAA paper 2008-2151, IL, April 2008.
10. Loeven, G. J. A., Witteveen, J. A. S., and Bijl, H., "Airfoil Analysis with Uncertain Geometry using the Probabilistic Collocation method", Proceedings of the 10th AIAA Non-Deterministic Approaches Conference, AIAA paper 2008-2151, IL, April 2008.

The optimum position for a tidal power barrage in the Severn Estuary



Rod Rainey
Chief Engineer
Energy

Introduction

Tidal power barrages in the Severn estuary were studied intensively 30 years ago, by a UK Government committee chaired by Bondi, see Bondi et al, 1981. It was concluded from computer models that the optimum position for a barrage from the power point of view was approximately half way down the estuary at Minehead. If the barrage was moved further downstream no more power was obtained because it was found that the barrage increasingly attenuated the incoming tides. Although tidal power barrages for the Severn have been studied on several more recent occasions, it appears that no more recent computer modelling has been undertaken on this point, see Burrows et al. 2009.

The problem can be investigated using G.I.Taylor's simple analytical model of the tidal flow in the Severn estuary (Taylor, 1921). This has the advantage of revealing the fundamental features of the problem more clearly than a computer model.

Taylor's model is described in Lamb's account of the "canal theory of the tides" (Lamb, 1932, pp 267-278), of which it is a special case. The canal theory considers tidal flow as a longitudinal gravity wave in a channel. Following Lamb's notation, if the width of the channel is $b(x)$ and its depth $h(x)$, both varying with position x along the channel, then the equation for the surface elevation $\eta(x, t)$ at time t is (Lamb, 1932, p.274):

$$\frac{\partial^2 \eta}{\partial t^2} = \frac{g}{b} \frac{\partial}{\partial x} \left(hb \frac{\partial \eta}{\partial x} \right) \quad (1)$$

where g is the acceleration due to gravity. In an estuary, high tide is assumed to occur at the same time $t=0$ everywhere, since the extent of the estuary, when measured in degrees of longitude, is small compared with the tidal cycle of approximately 180 degrees. A solution is therefore sought of the form:

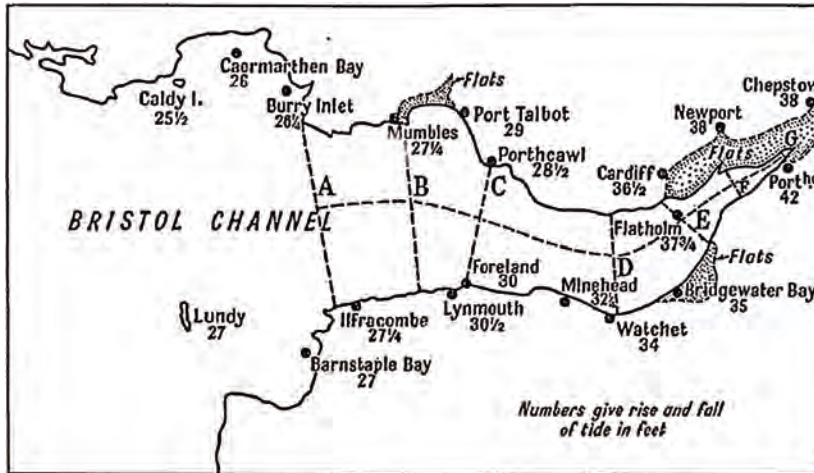
$$\eta(x, t) = \eta_0(x) \cos(\omega t) \quad (2)$$

where $2\pi/\omega$ is the tidal period of approximately 12 hours (half a lunar day). Thus (1) becomes:

$$\frac{g}{b} \frac{d}{dx} \left(hb \frac{d\eta_0}{dx} \right) + \omega^2 \eta_0 = 0 \quad (3)$$

In the case of the Severn estuary, Taylor observed that the width $b(x)$ and depth $h(x)$ both increase approximately linearly with distance x downstream (referred to henceforth as "west") of the head of the estuary at Portishead, see Figure 1 (originally Table 1 and Figure 1 in Taylor, 1921). He therefore took $x = 0$ at Portishead.

All the data in the table have been updated, from the latest Admiralty Charts. Breadth, depths and areas are at mean sea level. The area upstream of Sharpness (17km upstream of Chepstow) is excluded, since it is small and the tidal range is markedly reduced there. The time delays are mean values for neap tides (appropriate since we are considering a mean tidal range in Figure 3, which will be reduced by the barrage as in Figure 4) based on data for the north shore of the estuary (which appear more reliable than those from the south shore) in the 2009 Admiralty Tide Tables. Parameters for $n=8$ are defined to give the correct averages over the area upstream of Section G, when used in equation (9)



Taylor's Section	n	Distance x_n from Section G (km)	Mean depth (m)	Breadth (km)	Area S_n to next Section (sq.km)	Delay t_n of high tide, relative to Section A (min)	Loss angle (degrees) from equation (9)
A	1	114.3	36.9	40.6	800	0	3.8
B	2	92.10	28.7	37.7	585	2	3.9
C	3	77.83	24.4	30.0	695	2	5.2
D	4	46.33	16.3	22.7	383	6	6.1
E	5	28.72	16.3	13.2	220	14	5.2
F	6	14.82	9.5	15.2	166	19	
G	7	0	5.3	7.8	113	29	
	8	0				53	

Figure 1 - Taylor's model of the Bristol Channel

$$b = \beta x \text{ and } h = \gamma x \quad (4)$$

where β, γ are constants.

This reduces (3) to:

$$\frac{d}{dx} \left(x^2 \frac{d\eta_0}{dx} \right) + k\eta_0 x = 0 \quad (5)$$

with

$$k = \omega^2 / (g\gamma)$$

which can be solved exactly as a Bessel function:

$$\eta_0 = \frac{KJ_1 \{2\sqrt{kx}\}}{\sqrt{kx}} \quad (6)$$

where K is a constant. Taylor took $\gamma = \{25 \text{ fathoms}\} / \{80 \text{ UK nautical miles}\} = 0.0003084$ (β is immaterial) and the tidal period $2\pi/\omega$ as 12.4 hours, so that $k = 0.00655 \text{ km}^{-1}$, and found (6) to be a good approximation to the observed variation of tidal range in the Severn estuary, shown in Figure 1 (which are close to modern values).

This paper extends Taylor's analysis to the case of a tidal power barrage in the estuary.

Tidal power - the need for progressive waves

Considered as a function of time, the horizontal velocity in a tidal wave (and indeed in a water wave generally) is 90 degrees out of phase with the surface slope $\partial\eta/\partial x$, since the latter is in phase with the horizontal acceleration. And the pressure variations are in phase with the surface elevation η . Thus for a standing-wave solution of the form (2), where the surface slope is in phase with the surface elevation, the velocity and pressure are 90 degrees out of phase. Therefore the power flux (= velocity \times pressure) has a mean value of zero everywhere. This is of course to be expected, since the tidal energy is nowhere being dissipated in the estuary in potential flow, only stored.

When we extract tidal power with a barrage, however, we require an equal mean power flux inwards at the mouth of the estuary.

We thus reach the important conclusion that Taylor's solution (or any solution of the form (2)) is inadmissible west of the barrage, because it transmits no mean power. What is required west of the barrage is a progressive wave, in which there is a power flux, because the surface slope is 90 degrees out of phase with the surface elevation (and thus the velocity is in phase with the pressure). Rather than a solution of the form (2) we can seek a solution of the more general form:

$$\eta(x, t) = \text{Re}\{\eta_0(x)e^{i\omega t}\} \quad (7)$$

where $\eta_0(x)$ is now complex. This again leads to (5), which can be solved in the same way as:

$$\eta_0 = \frac{K_1 H_1^{(1)} \{2\sqrt{kx}\} + K_2 H_1^{(2)} \{2\sqrt{kx}\}}{\sqrt{kx}} \quad (8)$$

where $H_1^{(1)}$ and $H_1^{(2)}$ are a first and second Hankel functions of order 1, and we now have two constants K_1 and K_2 . The first term is a progressive wave travelling east, and the second is a progressive wave travelling west. Far to the west, both resemble tidal waves in open water of the same depth (since $H_1(x) \sim -\{\cos(x+\pi/4) \pm i\sin(x+\pi/4)\}/\sqrt{x}$, for large x).

East of the barrage, we can extend Taylor's solution empirically to include the observed delay-times of the tide. These are caused by the need to transport energy into the estuary, to overcome natural energy losses from turbulence, and may therefore be important in the context of a tidal power barrage (in fact they turn out to be of only minor importance, see Figure 3).

An equivalent electric circuit

In his account of waves in channels, Lighthill (1978, p.104) introduces the standard electrical analogy of voltage with pressure, and electric current with volume flow rate. If the level variation of a reservoir of area S is written $\text{Re}\{e^{i\omega t}\}$, then its pressure variation is $\text{Re}\{\rho g e^{i\omega t}\}$ (where ρ is the density of water), and the volume flow rate in and out of the reservoir is $S d/dt(\text{Re}\{e^{i\omega t}\}) = \text{Re}\{Si\omega e^{i\omega t}\}$. Thus on the electrical analogy its impedance is $\rho g / (Si\omega)$ so it is analogous to an electrical capacitance $S/\rho g$ (Lighthill, 1978, p.200 (3)).

A similar calculation applies in our case, for a reservoir formed by a barrage at one of Taylor's Sections A-E in Figure 1. The reservoir area can be discretized into the sub-areas S_n between the successive Sections, given in Figure 1. The level variation at the barrage is given by Taylor's formula (6) with his x-coordinate x_n given in Figure 1, and this formula can also be used to find the average amplitude of the level variations of each sub-area. The phases of these level variations is given by the average delay-times t_n in Figure 1. Thus the reservoir impedance Z_1 of a barrages at the n th of Taylor's Sections A-E can be written as:

$$Z_1 = \frac{\rho g J_1 \{2\sqrt{kx_n}\}}{\sqrt{kx_n}} \sum_{j=0}^{n-1} \left[\frac{J_1 \{2\sqrt{k(x_j + x_{j+1})}/2\}}{\sqrt{k(x_j + x_{j+1})}/2} S_j i\omega e^{-i\omega(t_j + t_{j+1})/2 - t_n} \right] \quad (9)$$

Evidently (9) is no longer purely imaginary, but has a real part analogous to a resistance R_L as well as an imaginary part analogous to a capacitance C . The resistance R_L gives the natural energy dissipation in the reservoir - to continue the electrical analogy, it can be expressed as a "loss angle" $\tan^{-1}(\omega C R_L)$, which is readily calculated from the argument of (9) and is given in Figure 1.

West of the barrage, it is convenient to consider the water pressure variation ($= \rho g \times$ level variation) as the sum of the pressure variation $Re\{P e^{i\omega t}\}$ which would be seen in the absence of the barrage, and the additional pressure variation $Re\{P' e^{i\omega t}\}$ caused, immediately west of it, by the presence of the barrage. The additional pressure $Re\{P' e^{i\omega t}\}$ at the barrage produces a tidal wave which propagates out to sea - as far as the flow to the west of the barrage is concerned, the barrage is acting like a wave-maker. We require its wave-making impedance Z_2 , i.e. the ratio of pressure to volume flow rate in the tidal wave it generates.

A unit wave propagating west is described by the second term in (8), with $K_2 = 1$.

The water acceleration in this wave, in the direction of propagation, is minus the surface slope times g , whence we can obtain the water velocity in a westward direction by integrating, as the real part of:

$$\frac{-g}{i\omega} \frac{d}{dx} \left(\frac{H_1^{(2)} \{2\sqrt{kx}\}}{\sqrt{kx}} \right) e^{i\omega t} \quad (10)$$

The volume flow rate in the direction of propagation is this velocity times bh , and the water pressure is $\rho g \eta$. We obtain the impedance Z_2 by dividing the latter by the former, which gives this impedance as:

$$Z_2 = \frac{-i\rho\omega H_1^{(2)} \{2\sqrt{kx}\}}{bh\sqrt{kx}} \left/ \frac{d}{dx} \left(\frac{H_1^{(2)} \{2\sqrt{kx}\}}{\sqrt{kx}} \right) \right. \quad (11)$$

which we can consider as a resistance R in series with an inductance L , giving a combined impedance of $R + i\omega L$. For large x , the wave resembles a tidal wave in open water, for which the impedance is known to be purely a resistance of $\rho c/(bh)$ (Lighthill, 1978, p.104), where c is the open-water wave speed \sqrt{gh} . This gives a useful cross-check, when (11) is evaluated numerically.

In the absence of the barrage, the (complex) volume flow rate at the barrage location is P/Z_1 , in an eastward direction. The additional wave-making volume flow immediately west of the barrage is P'/Z_2 , in a westward direction.

Thus the total (complex) volume flow rate at this location, in an eastward direction, can be written:

$$\frac{P}{Z_1} - \frac{P'}{Z_2} \quad (12)$$

If we write the total (complex) pressure at this location as $P'' = P + P'$, then (12) can be re-arranged to:

$$\frac{P \frac{Z_1 + Z_2}{Z_1} - P''}{Z_2} \quad (13)$$

On the electrical analogy, this the same current as would be produced by a voltage generator $P(Z_1 + Z_2)/Z_1$ with a source impedance of Z_2 . The flow in an eastward direction produced by this voltage generator passes first through the barrage, and then into the reservoir beyond it. The impedance seen by the flow is thus the flow-resistance of the turbines in the barrage, in series with the reservoir impedance Z_1 . The turbines can be taken for simplicity as allowing flow in both directions. This is the most common arrangement, see Baker, 1991, p.31, and also the most efficient, before turbine losses, see Prandle, 1984. Also for simplicity, the flow-resistance of the turbines can be taken as a constant R_B , because very similar results have been obtained in simpler cases with linear and quadratic turbine characteristics (Garrett & Cummins 2004). Thus the equivalent circuit of the complete system is as shown in Figure 2.

When $R_B = 0$, it may be seen that the pressure at the barrage is its undisturbed value P , as it should be.

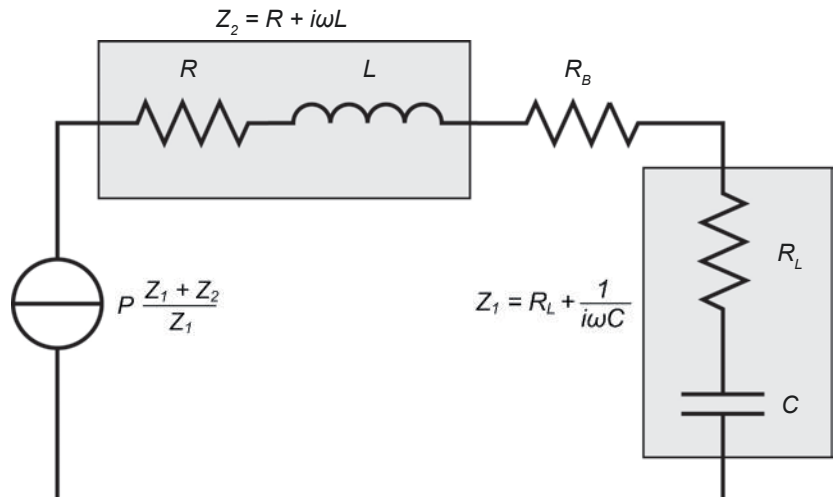


Figure 2 - Equivalent electric circuit of barrage

Similarity to wave power

At first sight it may seem curious that to provide the inward power flux needed to power the barrage, we introduce an additional tidal wave travelling in an outward direction. The reason is that from (8) Taylor's standing-wave solution (6) can be seen (by putting $K_1 = K_2 = K$ in (8) and noting that $H1^{(1)} + H1^{(2)} = 2J_1$) as the superposition of a tidal wave travelling east and an equal one travelling west. Our additional wave travelling west is cancelling part of his, giving a net inward wave.

This situation is familiar in wave power, see for example Mei 1989 Section 7.9. Two-dimensional wave power devices likewise need to radiate waves out to sea, to cancel out wave reflections.

Power available at various locations in the Severn estuary

We can now calculate the power from the equivalent circuit of Figure 2. The argument does not rely on the approximations above, but applies equally if accurate values for Z_1 and Z_2 are available. The (complex) volume flow rate through the barrage is:

$$P \frac{Z_1 + Z_2}{Z_1(Z_1 + Z_2 + R_B)} \quad (14)$$

and thus the average power is:

$$\frac{1}{2} |P|^2 \left| \frac{Z_1 + Z_2}{Z_1(Z_1 + Z_2 + R_B)} \right|^2 R_B \quad (15)$$

This is readily calculated as a function of R_B , using the expression (9) and (11) above for Z_1 and Z_2 . It is given in Figure 3 for Taylor's sections A to E of Figure 1. The (complex) tidal pressure P in the absence of the barrage is taken as $4\rho g$ at Watchet, or 8m tidal range, which is the approximate root-mean-square value between the mean spring range of 10m, and the mean neap range of 5m, and thus gives the annual-average power. The values elsewhere are extrapolated from this 8m figure, using Taylor's formula (6).

Rather than plotting against R_B , Figure 3 is plotted against the pressure difference across the barrage (i.e (14) times R_B), expressed as a fraction of the tidal pressure variation $|P|$ in the absence of the barrage.

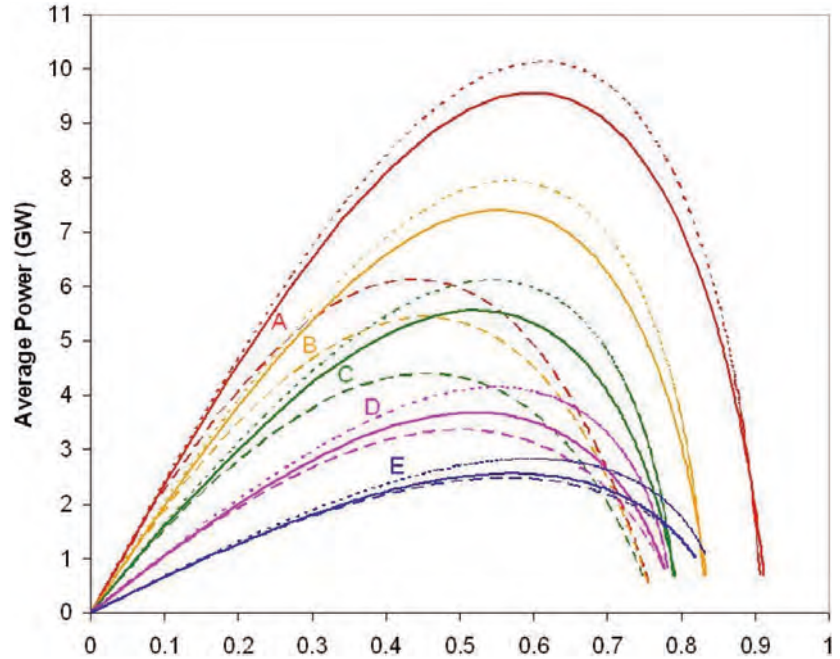


Figure 3 - Average power (GW) for barrages at locations A - E of Figure 1

The horizontal axis is the water level difference across the barrage, divided by the tidal amplitude (= range/2) in the absence of the barrage. Solid lines are with the outer estuary model included. Dashed lines are without it. Dotted lines are with it included, but with the delay times t_n in Figure 1 set to zero, to remove natural energy losses.

Evidently the optimum value for this fraction is between 0.4 and 0.6, and the power increases steadily as the barrage is moved west. This is of course to be expected - as we move west, the reservoir area increases much more than the tidal range reduces, see Figure 1.

Effect of the shape of the estuary west of Taylor's model

Taylor observed that the shape of the Severn estuary changes abruptly west of his estuary boundaries (section A in Figure 1), and ceases to follow his formulae (4), even approximately. The width of the estuary approximately doubles immediately west of section A, and thereafter follows another of Taylor's linearly-tapering profiles, with both depth and width increasing approximately linearly with distance from a notional apex at Abergavenny, 100km east of section A. The depth of 40m at section A gives a new value of $\gamma^* = 40\text{m}/100\text{km} = 0.0004$ for γ , and thus a new value $k^* = 0.00505\text{km}^{-1}$ for k . We wish to find the effect of this transition to a new profile on the barrage wave-making impedance Z_2 .

The effect of the abrupt transition

will be to reflect some of the wave travelling west considered in section 3, back up the channel. This reflection will be re-reflected from the barrage, and then again from the abrupt transition after section A, in an infinite sequence. We can sum all the waves travelling west into a single wave travelling west between the barrage and Taylor's section A, and likewise sum all the waves travelling east into a single wave travelling east in this region. We can write the (complex) volume flow rates in the direction of wave propagation as:

- V_o and V_b for the wave travelling west, at respectively the outer end of the region at section A, and at the barrage.
- V'_o and V'_b for the wave travelling east, at respectively the outer end of the region at section A, and at the barrage.

We can first find the ratio of V'_o to V_o , which we can express as a reflection coefficient r where $V'_o = rV_o$.

In the wave travelling west, the

impedances at the two locations just considered are given by (11), we can write them as Z_o and Z_B . In the wave travelling east the impedances can be seen from (11) to be the complex conjugates of Z_o and Z_B (the Hankel function $H_1^{(2)}$ from (8) becomes

$H_1^{(1)} = \overline{H_1^{(2)}}$ and the $-i$ from (10) becomes $+i$ because the acceleration in the direction of wave propagation is now plus the surface slope times g). In the region west of section A, we have only a wave travelling west, and the impedance is given by (11) with the new parameter k^* instead of k , and with $x = 100\text{km}$. We can write this impedance as Z^* .

The sum of the pressures in the two waves immediately east of the transition at section A can now be equated to that in the single wave immediately west of it. The latter is obtained from the volume flow rate $V_o - V'_o$ in the westward direction:

$$V_o Z_o + V'_o \overline{Z_o} = (V_o - V'_o) Z^*$$

i.e.

$$V'_o = \frac{Z^* - Z_o}{Z^* + Z_o} V_o$$

so that

$$r = \frac{Z^* - Z_o}{Z^* + Z_o} \quad (16)$$

When $Z^* = Z_o$ there is no reflection from the outer boundary, and (16) accordingly predicts that $V'_o = 0$, as expected.

We can now find the required wave-making impedance Z_2 of the barrage, in terms of the reflection coefficient r given by (16). From (8):

$$\frac{V_B Z_B}{V_o Z_o} = \frac{H_1^{(2)}(2\sqrt{kx_B})/\sqrt{kx_B}}{H_1^{(2)}(2\sqrt{kx_o})/\sqrt{kx_o}}$$

and

$$\frac{V'_B \overline{Z_B}}{r V_o \overline{Z_o}} = \frac{H_1^{(1)}(2\sqrt{kx_B})/\sqrt{kx_B}}{H_1^{(1)}(2\sqrt{kx_o})/\sqrt{kx_o}} \quad (17)$$

where x_o and x_B are the x -coordinates of section A and the barrage. Since

$H_1^{(1)} = \overline{H_1^{(2)}}$ the RHS of these two equations are complex conjugates of each other.

Thus:

$$\left(\frac{V_B Z_B}{V_o Z_o} \right) = \frac{V'_B \overline{Z_B}}{r V_o \overline{Z_o}}$$

i.e.

$$V'_B = V_B r \frac{V_o \overline{Z_B}}{(V_o \overline{Z_o})} = V_B r e^{-i2\omega T} \quad (18)$$

where we are noting that the argument of $V_o \overline{Z_B}$ is $-\omega T$, where T is the wave transit time between the barrage and section A (readily calculated from (8)). We can thus obtain the wave-making impedance at the barrage, as the sum of the pressures divided by the sum of the volume flow rates:

$$\frac{V_B Z_B + V_B r e^{-i2\omega T} \overline{Z_B}}{V_B - V_B r e^{-i2\omega T}} = \frac{Z_B + \overline{Z_B} e^{-i2\omega T}}{1 - r e^{-i2\omega T}} \quad (19)$$

When $r = 0$, there is no reflection at the outer boundary and (19) then predicts that the wave-making impedance of the barrage is Z_B , as expected.

The barrage powers can be re-calculated using this new wave-making barrage impedance Z_2 - the results are shown in Figure 3. Evidently the changed shape of the estuary west of Taylor's original model increases the power considerably, which is to be expected since the increased width of the estuary will lower Z_2 and thus from Figure 2 increase the power. The effect is more pronounced the closer the barrage is to this increased width. Thus the conclusion remains that the power increases steadily as the barrage is moved west - indeed it now increases more.

The question thus arises of the boundary condition even further out, where the second Taylor profile stops abruptly at the western extremities of England and Wales. This transition can be treated exactly like the transition at section A. If the impedance is assumed to halve at this transition, for example, and the calculations repeated, the maximum powers in Figure 3 all increase, by between 1% (barrage at section E) and 7% (barrage at section A). So again the effect is more pronounced for barrages closer to the transition - it appears that features beyond the UK are relevant to the barrages furthest down the Severn estuary. This supports the practice in the

most recent studies (e.g. Burrows et al. 2009) of extending computer models out to the limits of the continental shelf, although the type of boundary conditions applied there are very important (recent studies appear to be subject to the criticism that the boundary conditions are zero impedance, see next Section).

The calculations can also be repeated with the delay-times t_n in Figure 1 set to zero, which will remove natural energy dissipation. This is done in Figure 3, and reveals that natural energy dissipation is only of minor importance.

Finally, the changes in tidal range produced by the barrage are important. They are readily calculated from the equivalent circuit in Figure 2, using the full expression (19) for Z_2 , and are shown in Figure 4, on the same horizontal axis as Figure 3. Taking into account the fact that the power peak in Figure 3 is further to the left for section C, the changes to the tidal range are very similar for all barrage locations. With barrages operated at maximum power, the tidal range is cut to 70% of its former value east of the barrage, and 90% of its former value immediately west of the barrage. A very simple view of the barrage is that (from (9) and (11)) Z_2 is small compared with Z_o , and R_L is small compared with C . From Figure 2, the optimum power, as a matter of elementary electrical engineering, is when R_B has the same impedance as C . This is an existing result in the tidal power literature, due to Garrett and Cummins (2004). It would reduce the tidal range east of the barrage by a factor $\sqrt{2}$, and leave the range immediately west of it unaffected, because Z_2 is small.

The horizontal axis and colour coding are the same as Figure 3. Dashed lines are east of the barrage, solid lines are just west of it.

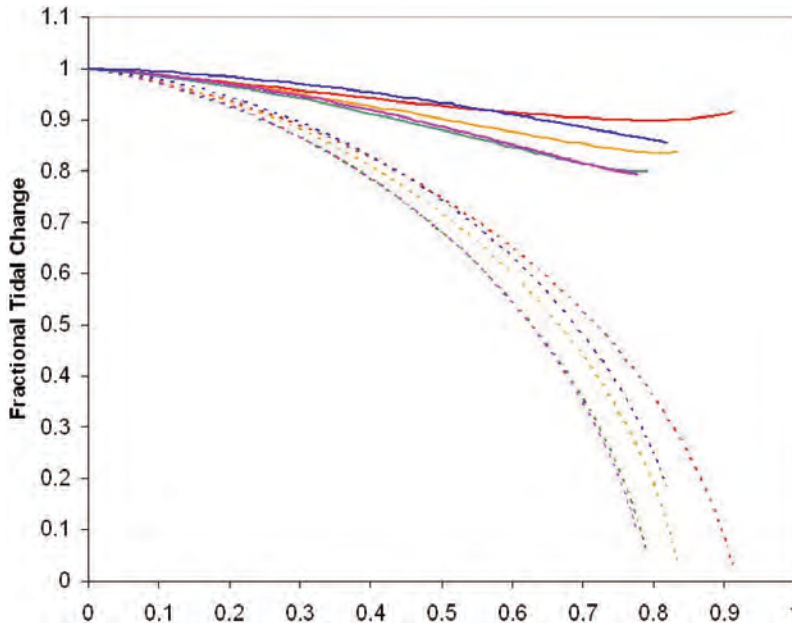


Figure 4 - Fractional tidal change for barrages at locations A – E of Figure 1.

Previous computations

The question of the optimum position for a barrage in the Severn estuary, from the power point of view, was studied 30 years ago, see Bondi et al. (1981). The power was computed with various finite-difference numerical models, some of which extended out into the Irish Sea. They showed the average power rising strongly from 0.5 GW to 2.3 GW as the barrage was moved west from Taylor's section F to D (Bondi et al. 1981, Vol1 p.18). This is similar to the results in Figure 3, allowing for conversion losses. However, very little increase was found for positions further west. By Taylor's section C, the power was starting to decline, in marked contrast to the increase seen in Figure 3 – albeit that significant discrepancies were found between computer models (Bondi et al. 1981, vol 2 p.57).

We now explore a possible reason for this decline, which is that all the models simply held the tidal range fixed on the model boundary, at the same value it would have were the barrage absent. This was then, and apparently still is, the usual assumption in tidal modelling (see e.g. Prandle, 1980), although it has been recognised as wrong in principle (Garrett and Greenberg, 1977).

It will produce a total reflection of the outgoing tidal wave – it is equivalent to setting Z^* in (16) equal to zero. This leads to:

$$V_o' = \frac{-Z_o}{Z_o} V_o$$

i.e.

$$r = \frac{-Z_o}{Z_o} = -e^{i2\varphi} \quad (20)$$

where $\varphi = \arg(Z_o)$ is the phase advance of pressure over volume flow rate, in an outwards-propagating tidal wave, at the model boundary. For a model boundary at section A, for example, it can be calculated from (8) as 45.3 degrees. If we similarly write $\theta = \arg(Z_B)$ then $Z_B = \zeta e^{i\theta}$ where ζ is real and θ is the phase advance of pressure over volume flow rate, in an outwards-propagating tidal wave, at the barrage. For a barrage at section E, for example, it can be calculated from (8) as 68.9 degrees. The wave-making impedance Z_2 of the barrage (19) thus becomes:

$$\frac{\zeta e^{i\theta} - \zeta e^{-i\theta} e^{i2\varphi} e^{-i2\omega T}}{1 + e^{i2\varphi} e^{-i2\omega T}} = \frac{\zeta e^{i(\pi-\theta+2\varphi-2\omega T)} + \zeta e^{i\theta}}{e^{i(2\varphi-2\omega T)} + 1} \quad (21)$$

Since $e^{i\chi} + e^{i\psi} = \{e^{i(\chi+\psi)/2} + e^{-i(\chi+\psi)/2}\} e^{i(\chi-\psi)/2}$
 $e^{i(\chi+\psi)/2} = 2\cos\{(\chi+\psi)/2\} e^{i(\chi-\psi)/2}$ this impedance can be written:

$$\frac{\zeta \cos\{\pi/2 + (\varphi - \theta - \omega T)\} e^{i(\pi/2 - \varphi - \omega T)}}{\cos(\varphi - \omega T) e^{i(\varphi - \omega T)}} = i \zeta \frac{\sin(\omega T + \theta - \varphi)}{\cos(\omega T - \varphi)} \quad (22)$$

Thus the wave-making impedance at the barrage is purely imaginary (i.e. reactive), as we would expect – the barrage can radiate no wave power, because the waves it sends west are perfectly reflected back by the model boundary. Its amplitude is small if the model boundary is close to the barrage, because then θ and φ are nearly equal, and the phase delay ωT of a tidal wave between the barrage and the model boundary is then also small. Thus the change in the results will be small, because Z_2 is small anyway, as noted at the end of the previous Section.

However, when the model boundary is a long way from the barrage, φ will be small because the tidal wave at the model boundary will resemble an open-water wave. Thus when the phase delay ωT reaches 90 degrees, the denominator in (22) will drop to zero, and the wave-making impedance of the barrage will become very large. The power from the barrage will accordingly drop. This condition requires the transit time T of a tidal wave between the barrage and the model boundary to be a quarter of the tidal period, or $12.4/4 = 3.1$ hours. This is a resonant condition, with the natural sloshing period of the basin between the barrage and the outer boundary equal to the tidal period. With a mean tidal wave speed of 25m/s, say, it corresponds to a distance from the barrage to the outer boundary of $25 \times 3600 \times 4 = 360$ km. This is comparable with the size of the larger models used by Bondi et al. (1981).

It is thus possible that the models used by Bondi et al. (1981) were giving spurious results due to internal resonances, caused by the incorrect model boundary condition, in which the tidal range was held at the same value it would have were the barrage absent. The appropriate boundary condition is an “absorbing” one, which does not reflect waves – these are standard in naval architecture, and familiar in physical model testing too, as the beach in a wave tank.

Acknowledgement

This work was performed under contract to the RSPB, WWF, WWT, the National Trust and the Wye and Usk Foundation. Equations (7) and (8) are due to F.J.M.Farley, who kindly reviewed the manuscript. The author had formerly used an exponential-horn approximation downstream of the barrage, which fits the geometry of the Severn estuary much less well.

References

1. Baker, A.C. 1991 Tidal Power Peter Peregrinus Ltd. on behalf of the Institution of Electrical Engineers.
2. Bondi, Sir Hermann et al. 1981 Tidal Power from the Severn Estuary. Dept. Energy, Energy Paper No.46. HMSO.
3. Burrows, R. et al. 2009. Tidal Energy Potential in UK Waters Proc. ICE - Maritime Engineering, 162, Issue 4 (in press).
4. Garrett, C. & Greenberg, D. 1977 Predicting Changes in Tidal Regime: The Open Boundary Problem. J. Physical Oceanography 7 171-181
5. Garrett, C. & Cummins, P. 2004 Generating Power from Tidal Currents ASCE J. Waterway, Port, Coastal & Ocean Eng. 130 114-118
6. Lamb, Sir Horace 1932 Hydrodynamics 6th Ed. Cambridge University Press.
7. Lighthill, Sir James 1978 Waves in Fluids Cambridge University Press.
8. Mei, C.C. 1989 The Applied Dynamics of Ocean Surface Waves. World Scientific.
9. Prandle, D. 1980 Modelling of Tidal Barrier Schemes: An Analysis of the Open-Boundary Problem by Reference to AC Circuit Theory Estuarine & Coastal Marine Science 11 53-71
10. Prandle, D. 1984 Simple Theory for Designing Tidal Power Schemes Adv. Water Resources 7 21-27
11. Taylor, Sir Geoffrey 1921 Tides in the Bristol Channel in The Scientific Papers of Sir Geoffrey Ingram Taylor, Vol 2. pp 185-192. Cambridge University Press 1960

Greater Plutonio project - Subsea flowline design and performance



David Bruton
Senior Advisor
Energy



Forbes Sinclair
Head of Pipeline
Engineering,
Atkins Boreas
Energy

Abstract

Greater Plutonio is a large subsea development of five fields (BP 50%, Sonangol/Sinopec 50%) in deepwater offshore Angola Block 18. The development comprises a Floating Production Storage and Offloading (FPSO) connected to some 43 subsea wells by some 130km of infield flowlines, which were commissioned and delivered first oil in October 2007. An EPCIC contract was awarded in 2004 to a consortium of Acergy and Technip for the design and installation of the complex subsea systems. The design had to consider not only the flow assurance issues of long distance tie backs, but also the potential for lateral buckling and pipe walking under the anticipated operational cycles. As a starting point, the Atkins Boreas SAFEBUCK design principles were used, supported by specific test programmes that included fatigue testing and detailed pipe/soil interaction testing. Following the successful completion of the installation programme, detailed monitoring of the performance of the flowlines during hydrotest and early operation was carried out to confirm system integrity and validate the assumptions used in design. This paper will cover several aspects of the challenging design for the flowline systems, the development of the test programmes, the installation and survey techniques used to establish the operating performance as well as an assessment of the actual performance against design predictions.

Field layout and overview

Field layout

The development is located in 1400m of water some 165km North West of Luanda off the coast of Angola. First oil was achieved on 1st October 2007.

The development comprises the following flowline systems which serve 43 subsea wells:

- 2 Off oil gathering networks (North and South totalling approximately 23km and 20km of insulated 12inch flowline)
- 2 Off Water Injection (WI) networks (North and South totalling approximately 25km and 30km of 14inch and 12inch flowline respectively)
- 1 off gas injection system (South – totalling approximately 13km of 12inch flowline)

Hydrocarbons are processed onboard an FPSO and offloaded via a CALM buoy. The field produces approximately 250,000 bpd. Gas is currently being re-injected into the reservoir, and will continue until such time as a dedicated exported line to an onshore LNG terminal is completed. WI rates will average 750,000 bpd.

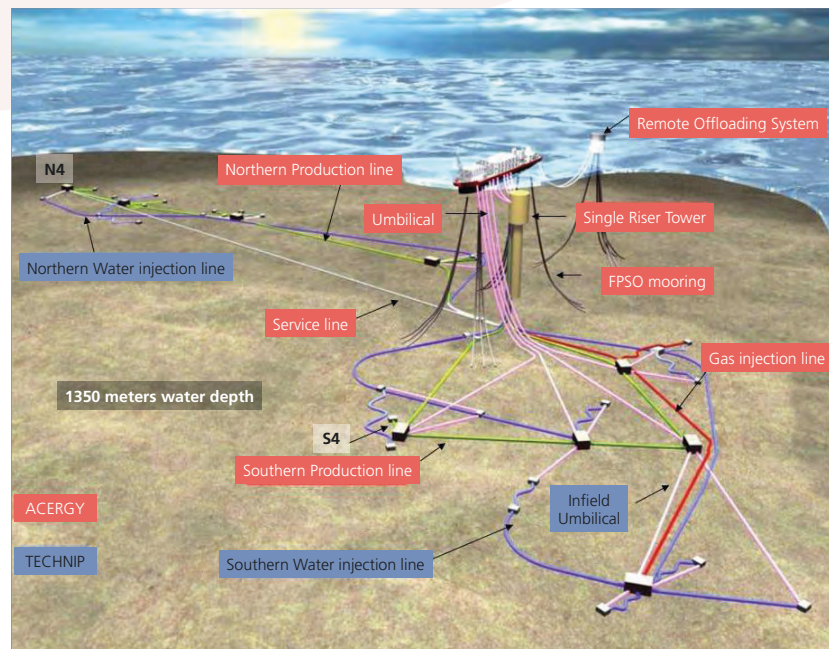


Figure 1 - Development schematic

Each of the flowlines utilise a number of flexible and rigid jumpers/spools to communicate with the subsea wells via In Line Tee (ILT) or Flowline Termination Assemblies (FTA).

Schematics of the development are shown in Figure 1 and Figure 2.

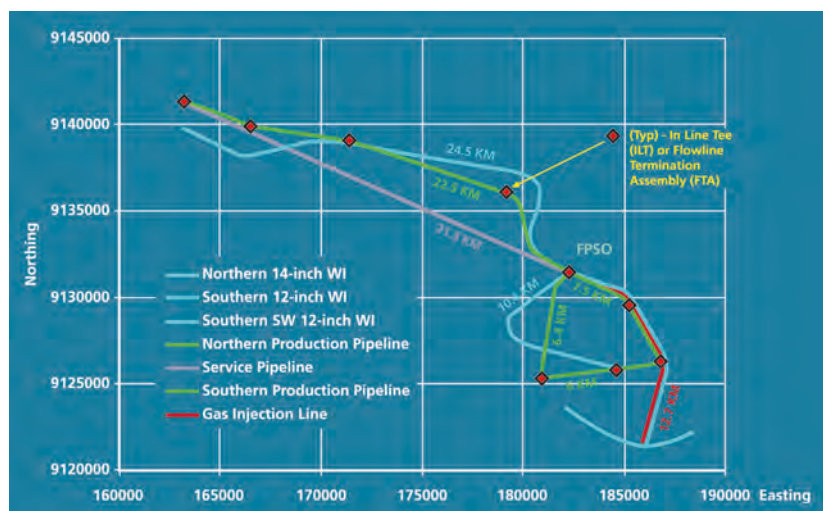


Figure 2 - Flowlines schematic plan

Execution plan and strategy

The issue of thermal expansion and the impact on the development was realised in August 2004. First oil was scheduled for 2Q 2007, and the market (supply chain) environment at this time was overheated. The aforementioned challenge was further compounded by the lack of available data on key input parameters such as soil friction and process operating performance data, which led to a wide range of mitigation possibilities. Finally, it was also realised that the analysis itself was going to require complex FE models and significant engineering input.

To overcome these challenges an execution strategy was established that encompassed seven key elements:

- (1) A new contractual framework
- (2) A fully integrated team would be established – this would include BP/Technip/Aceryg and Atkins-Boreas personnel

- (3) Special “integration” focus would target specific disciplines (geotechnical and materials and welding) that were considered key input drivers
- (4) The soils, materials and welding test programmes would be managed by a third party and funded by BP separately
- (5) All FE analysis work would be 100% independently verified
- (6) All mitigations would be scalable, simple, and not require new technologies
- (7) A joint (Acergy/Atkins Boreas/ BP/Technip) Peer Team would be established and a robust peer review programme implemented.

The execution plan to deliver the programme within the available schedule was structured as detailed in Table 1.

Table 1 - Execution plan

Phase	Primary	Parallel Activity
1	Agree commercial framework	Analytic work (Atkins boreas)
2	Develop robust mitigation elements	Procurement/soil data acquisition/fatigue testing
3	Develop robust design	Procurement/fatigue/soils/materials testing/ third party verification
4	Optimise using QRA	Testing/fabrication
5	Final design	Testing/third party verification/fabrication

Basis of design for thermal expansion

A key challenge for the subsea flowline design was the likelihood of lateral buckling, pipe walking or route curve instability of the flowlines between riser base and the drill centres. The analytic engineering studies (Phase 1 of the Execution Plan) indicated that all of the flowlines were susceptible to lateral buckling and several were susceptible to flowline walking.

Uncontrolled lateral buckling can lead to high strains and cyclic loads at the buckle crown and seriously compromise the integrity of the flowline. In addition, the flowlines have ILT structures which require controlled lateral buckling to prevent unplanned “rogue” buckles forming close to the ILT structures and compromising their integrity.

Flowline walking can occur when a flowline is subjected to steep thermal transients during start-up/shut-down cycles or when the flowline is laid on a seabed slope and subjected to load cycling¹. Over a number of cycles, this movement can lead to very large global axial displacement, with associated overload of the FTA or ILT connections. The production flowlines are subject to steep thermal transients and are on moderate slopes.

Flowline design in accordance to API1111 was supplemented by the SAFEBUCK design guideline² on lateral buckling and flowline walking issues. The guideline was used to perform the key lateral buckling limit state checks, including local buckling, fatigue and fracture.

Key design data

Key design data to address lateral buckling and walking were reviewed and established early in the detailed design, including:

- Field layout routes, bathymetry and flowline configurations
- Design operating conditions and operating cycles through life
- Thermal transients during shutdown/restart operations
- Fatigue and fracture design parameters, SN-curves, environment effects and Stress Concentration Factor (SCF)
- Pipe material response

- Pipe-soil interaction, including non-linear axial and lateral friction response.

Flowline configurations

The configurations are summarised in Table 2, all flowlines are single pipes. The production lines are wet insulated with 5 layer PolyPropylene (PP) coating. All other lines have 3 layer PP coating.

Operating conditions

Operating conditions for each flowline are summarised in Table 3, the most onerous co-incident pressures and temperature combinations were used for lateral buckling and walking design checks. The table also includes the maximum seabed slope along each flowline used in the walking analyses.

Thermal transients

On the production flowline thermal transients occur during shutdown/restart cycles, a full shutdown this includes cold displacement, hot-oiling and well restart. Thermal transient profiles were determined for all production, service and gas injection flowlines. Figure 3 shows an example of hot-oiling thermal transients.

Soil friction determination

Pipe-soil response is the largest uncertainty in the design of pipelines susceptible to lateral buckling and pipeline walking. Guidance was provided from the force-displacement models developed for the Safebuck Joint Industry Project (JIP)³, which included project-specific test data donated to the JIP by BP and other participants.

Improving the understanding of pipe-soil response provided the greatest scope for optimising the pipeline design because of the extreme sensitivity of design solutions to the axial and lateral resistance imposed by the soil. The Greater Plutonio Project therefore decided to carry out project-specific pipe-soil testing programmes that dramatically improved understanding, enhancing the pipeline design process and significantly reducing installation costs.

The T-bar test data provided

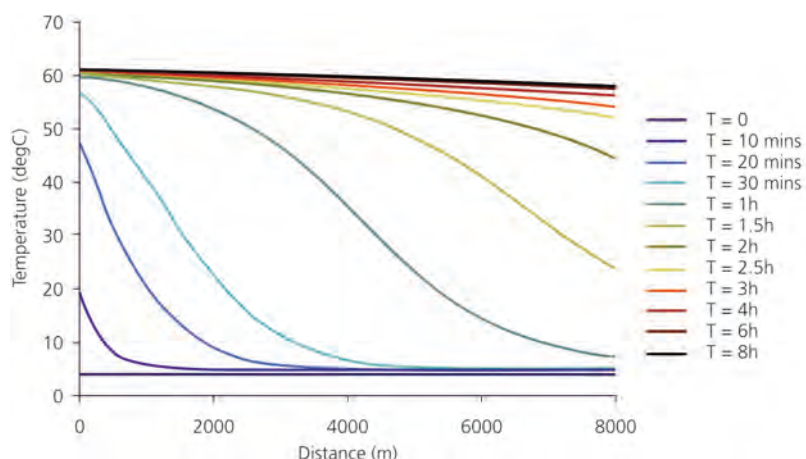


Figure 3 - Hot oiling thermal transient profiles

Table 2 - Flowline data

Flowline	Total Length	OD (mm)	WT (mm)	Coatings
Production	22.6km (North)	323.9	20.6	90.5mm 5LPP
	19.5km (South, 3 lines)	323.9	19.1	91.5mm 5LPP
Service	21.2km	323.9	19.1	2.5mm 3LPP
Water Injection	24.4km (North)	355.6	15.9	2.5mm 3LPP
	28.8km (South, 3 lines)	323.9	14.3	2.5mm 3LPP
Gas Injection	12.6km	355.6	15.9	2.5mm 3LPP

Table 3 - Design data

Flowline	Temperature	Pressure	Design Cycles	Seabed Slope
Production	95°C (North)	160 bar	92 major, 1656 minor	1.0° 0.96°
	88°C (South)	135 bar		
Service	51°C	265 bar	184	0.22°
Water Injection	25°C (North)	225 bar	1000	1.05° 0.92°
	23.5°C (South)	225 bar		
Gas Injection	40°C	365 bar	1650	0.81°

excellent data on near-surface shear strength profiles, see Figure 4, and soil sensitivity, both fundamental to the design of surface-laid pipelines in very soft clays.

As laid pipeline embedment is fundamental to the prediction of axial and lateral friction. The values of as-laid embedment predicted using published models^{4,5,6} is generally found to be substantially exceeded in practice, due to the load concentration at touchdown and the dynamic displacement of the pipe catenary during installation causing vertical and lateral oscillations of the pipe at touchdown. Calculated levels of embedment were therefore compared with observed levels of embedment from similar pipelines in similar soils to calibrate embedment predictions.

The Greater Plutonio soils testing programme included laboratory tests to assess axial resistance at the low contact effective stresses that occurs in the soil under typical flowline weights (<10kPa). Specialist equipment including the tilt table device, see Figure 5, at the University of Texas at Austin⁷ and the Camshear device, see Figure 6, at the University of Cambridge were used to measure pipe-soil interface friction at a range of loads and pipe velocities.

Large-scale axial and lateral cyclic tests were also undertaken at NGI, see Figure 7, using 4.5m³ of soil collected from the field and re-consolidated in a large tank⁸. The lateral soils testing at NGI measured lateral first load response and lateral cyclic response, including the soil resistance at span touchdown points to each side of a sleeper.

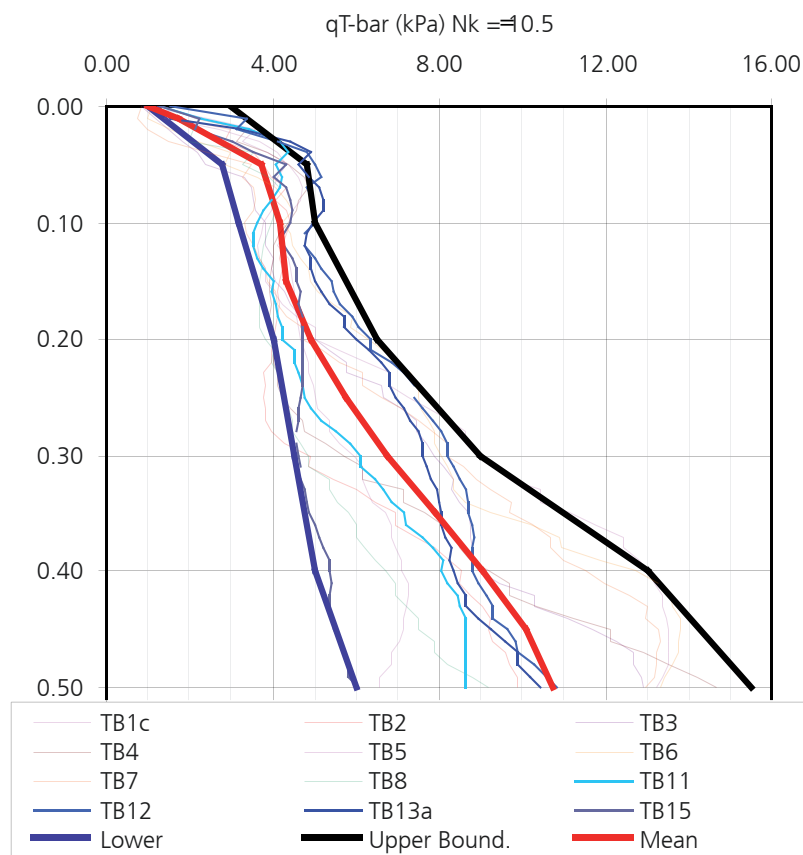


Figure 4 - Undrained shear strength profiles defined by in situ T-bar tests

These tests confirmed that the light production flowline lateral response fitted well with SAFEBUCK predictions³, as shown in Figure 8.

However, the tests also demonstrated that significantly higher levels of lateral friction could be expected for the heavier WI flowlines.

Cyclic tests showed that soil berms are quickly established under cyclic loading and provide significantly increased resistance to lateral displacement after the first cycle, see Figure 9.

This test programme greatly improved the industry's understanding of pipe-soil response and the significant resistance provided by the soil berms.

If walking predictions are unacceptably high then flowline anchoring would be required. The design had already shown that axial friction was hugely significant to the pipe walking response because, while low levels of friction could be expected, a high level of axial resistance can reduce walking significantly or prevent walking altogether.

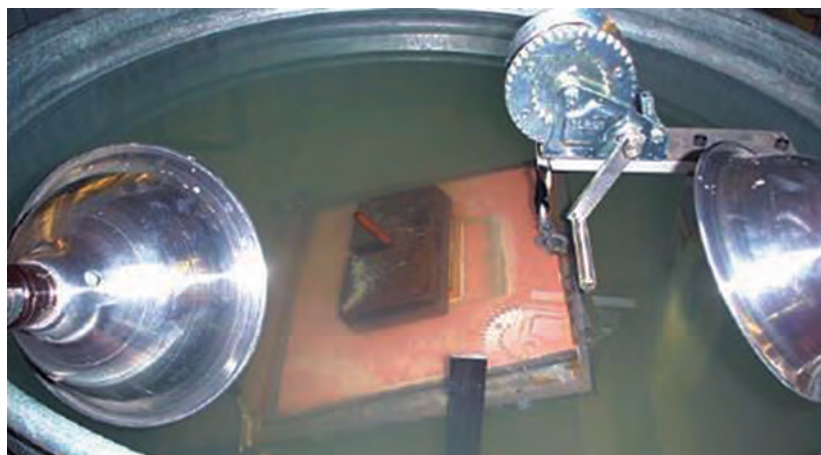


Figure 5 - Tilt-table device at the University of Texas

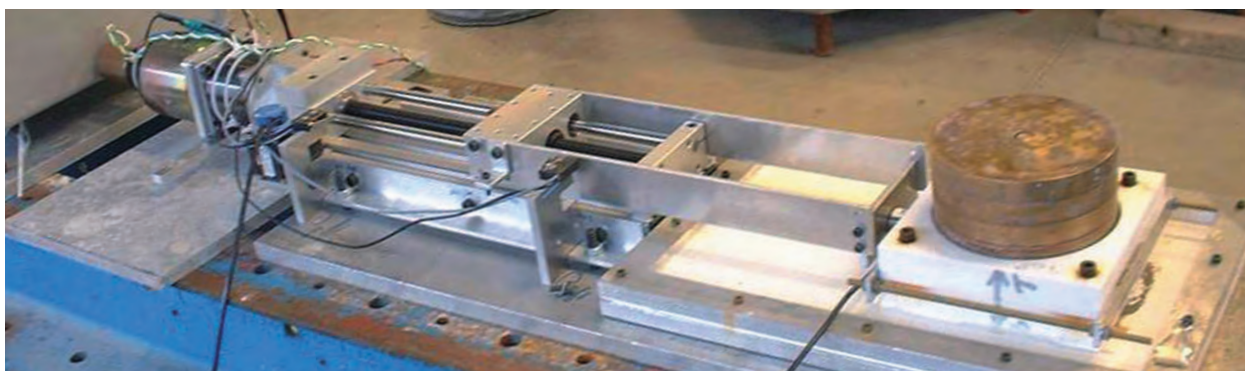


Figure 6 - Cam shear device at the University of Cambridge

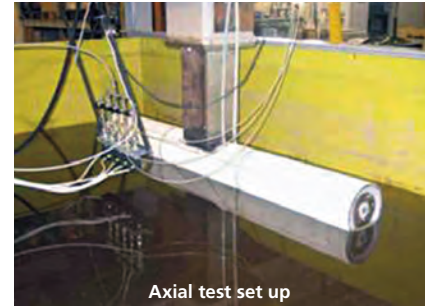


Figure 7 - Large-scale axial and lateral cyclic tests at NGI

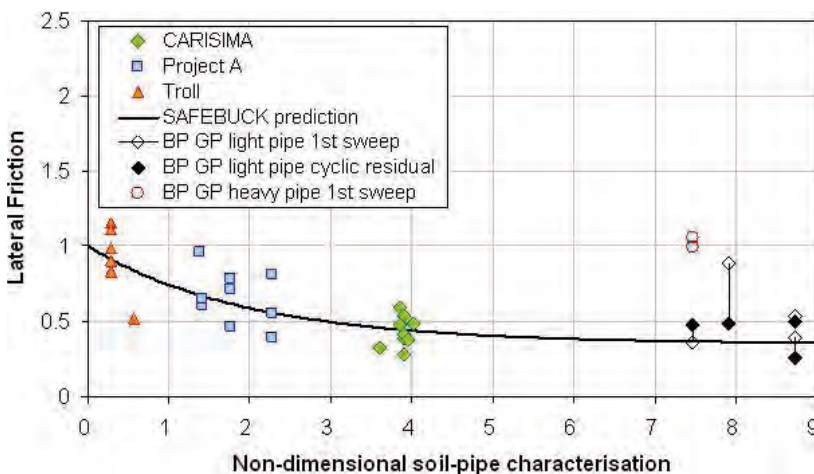


Figure 8 - Lateral sweep resistance - Greater Plutonio flowline tests at NGI compared with SAFEBUCK residual friction model and data

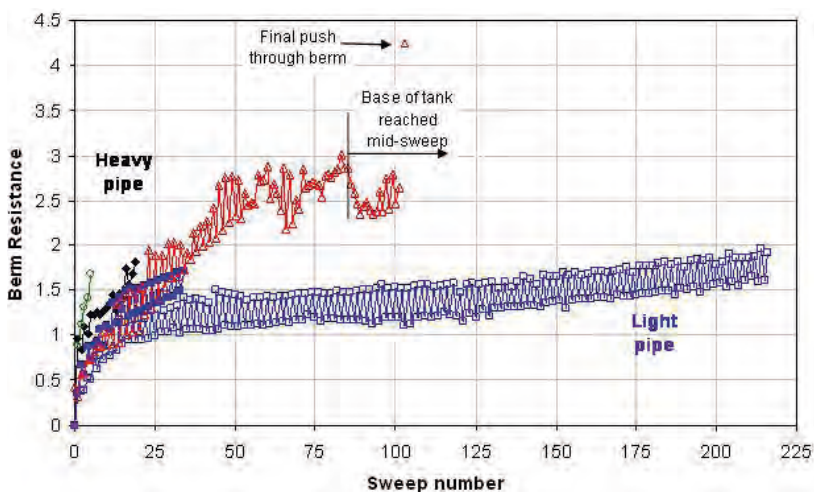


Figure 9 - Lateral resistance provided by soil berms - NGI tests

So as testing progressed BP co-ordinated a high level geotechnics workshop team, including key individuals from within BP, Cambridge University, NGI, Technip, Acergy and Atkins Boreas. This team was instrumental in agreeing an unconventional approach, which relied on the long-term “drained” soil response for slow cyclic axial loading. This approach was supported by test results which demonstrated:

- (1) Higher levels of axial friction at low contact stresses associated with very light flowlines (particularly wet-insulated flowlines)
- (2) A roughened PP pipeline coating adopted for all flowlines was also shown to benefit the design by increasing axial friction
- (3) Low pipe displacement velocities are more likely to result in a “drained” soil response.

The increased level of axial friction resulted in significant project savings by avoiding the need for anchoring on a number of flowlines.

Overview of mitigations

Lateral buckling mitigation

If uncontrolled lateral buckling were to occur, the flowlines would have been subjected, at the buckle locations, to extreme loading conditions that could seriously compromise the integrity of the flowline, as outlined earlier.

It was decided at the early stages of detailed design to engineer and install mitigation measures to control the locations and shapes of these lateral buckles.

The main driving parameters for an efficient mitigation are:

- Determination of the buckle initiation location:**
 The actual spacing between lateral buckles has a significant impact on the loads induced to the flowline. This parameter is linked to the virtual anchor spacing which characterises the load in each buckle. As such, the first engineering step was to determine the planned lateral buckle locations so that the corresponding virtual anchor spacings would enable a sufficient reduction of the expansion forces in the flowline
- Buckle initiation under hydrotest/operation induced loads:** It is necessary to create an artificial imperfection along the pipelay route to efficiently trigger buckle formation. This was achieved with the use of sleepers pre-laid perpendicularly to the flowline route, these supports creating vertical imperfections large enough to enable buckle firing as illustrated on Figure 10

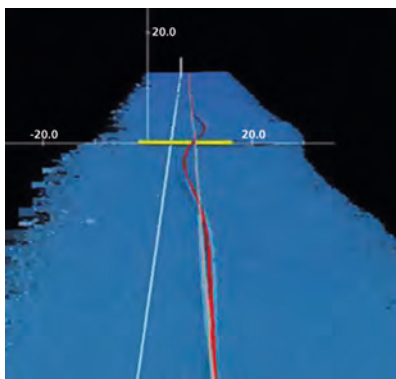


Figure 10 - Buckle firing (Initiation) on a sleeper

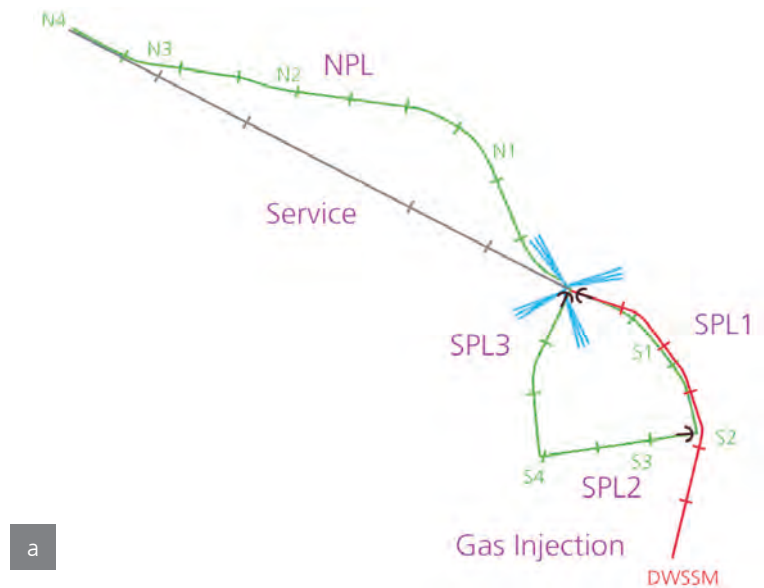


Figure 11 - Sleeper location plan (a) and sleeper assembly (b)

- Flowline integrity over field life needs to be ensured:** The loads in the buckle are impacted by the actual pipe/sleeper friction coefficient; the cyclic loading induced by field operations; and the touchdown conditions within the soils. For all planned buckle locations the flowlines were welded using fatigue class weld parameters over a 200m distance. For the production flowlines local additional buoyancy was used to increase the buckle length so as to achieve acceptable stresses at the buckle crown. This was achieved through extra thickness of the insulation coating over 100m length either side of the sleepers.

The solutions engineered during the detailed phase of the project led to an intensive testing programme which encompassed friction behaviour of materials, welding and fatigue testing described in subsequent sections. The design of the sleeper assemblies relied on the use of 32inch diameter pipe sections to achieve the required

vertical imperfection. The length of each sleeper was determined based on the predicted maximum lateral displacement of the flowline.

The ranges of lateral displacement were up to 15m and this could occur to either side of the sleeper centre. In view of installation tolerances and safety margin this led to a maximum sleeper length of 42m.

The final location of planned buckling sites, along with the buckle parameters such as amplitude, was also determined by the need to minimise the expansion levels at in-line structures and at crossing locations. Sleeper locations are presented on Figure 11(a).

The coating on top of the sleepers was selected based on tests to ensure low friction levels and durability throughout the flowline life. Figure 11(b) shows a typical sleeper assembly used for the production flowlines.

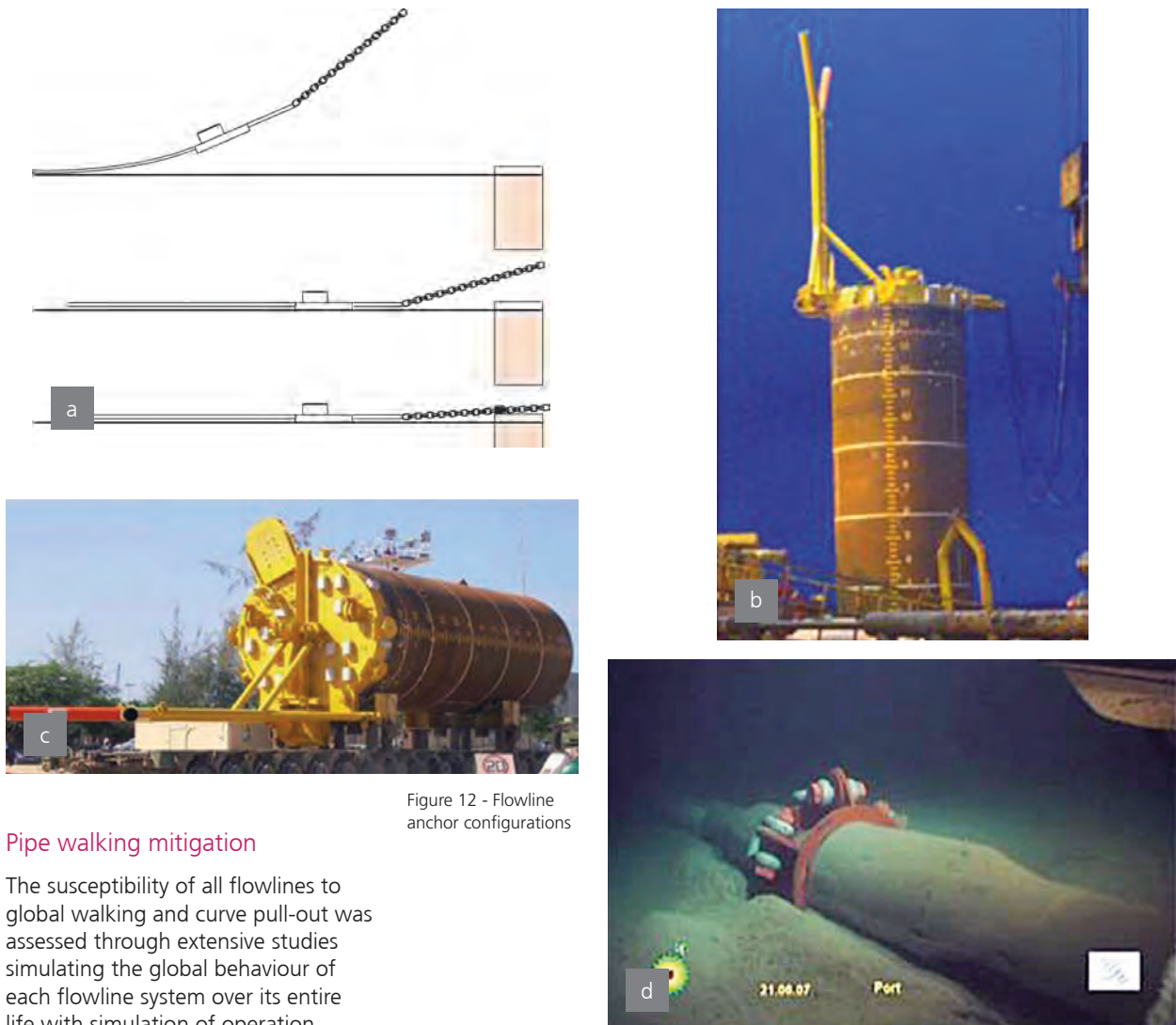


Figure 12 - Flowline
anchor configurations

Pipe walking mitigation

The susceptibility of all flowlines to global walking and curve pull-out was assessed through extensive studies simulating the global behaviour of each flowline system over its entire life with simulation of operation shutdown and restart cycles.

With the aid of rough coating, applied to the external coating to increase axial friction, it was shown that the water and gas injection and service lines were not impacted by pipe walking issues.

However, it was shown that all three branches of the South Production system were affected and needed to be anchored at one end to prevent walking. The potential retaining loads calculated were in the range of 100tonnes which dictated the use of suction anchors. In view of the field layout and due to operational reasons, it was not possible to anchor the flowline ends during flowline lay initiation which could have allowed the use of the traditional and relatively simple stab and hinge over method. Anchoring of the flowline ends needed to be done during lay-down with a length of chain connecting the pipe end to the suction anchor as illustrated

on Figure 12(a). The difficulty was to ensure very precise alignment of the pipe end section, including the FTA structure with the locking device installed on top of the suction pile to avoid side loads that the chain may have induced on the FTA if not properly aligned. This necessitated the design of the innovative chain capture solution shown on Figure 12(b) and Figure 12(c).

Using quantified risk assessments it was decided that the northern production system would not need to be anchored to prevent walking. However, the residual uncertainty in axial friction predictions meant that the risk was significant, enough to warrant the implementation of an in-line anchor attachment arrangement to allow later retrofitting of pile anchors Figure 12(d).

Sliding structures for axial displacements

In order to accommodate the anticipated levels of expansion at flowline ends or at in-line structures, all piping modules were designed so as to be capable of sliding relative to the mud mat. This was achieved through low friction rails on either side of the structure frame. In view of the weight of up to 70tonnes and the size of up to 14m length by 6.5m width for these structures, this was quite a challenge. Selection of the low friction rail and pad materials was made through extensive laboratory testing, as presented in Section 5. For the production flowlines the maximum sliding distance required was found to be 4.5m for FTA structures and 4m for ILT structures. The range of expansion was determined for each of the structures. For the ILT structures of the WI flowlines 1m sliding was required.



Figure 13 - Sliding structures assemblies

Pictures of a typical structure assembly and sliding rail details are presented in Figure 13.

Markings on the rails and expansion recording pads (shown in Figure 13) will allow future monitoring of axial displacements throughout the field operational life.

Crossing design

There are 3 locations where the production flowlines cross the WI flowlines. In view of the water depth

and soft soil conditions it was decided to achieve the separation between both lines with sleepers similar to those used for lateral buckle initiation. The vertical imperfections created made it necessary to consider the presence of those crossings in the lateral buckling studies. It was then required to ensure that no lateral or vertical displacement of the flowline could occur at the crossing location in order not to make the crossing design further complicated. Finally, the design should ensure free longitudinal

displacement of the flowline above the sleeper. The crossing location was adjusted as much as possible in view of the layout constraints to correspond to areas of low pipe axial compression force to minimise the risk of buckling. However, the risk of having an undesired buckle at the crossing could not be totally eliminated. It was then decided to restrain lateral and vertical movements by installing a locking beam after laying the pipe on top of the sleeper, as illustrated in Figure 14.

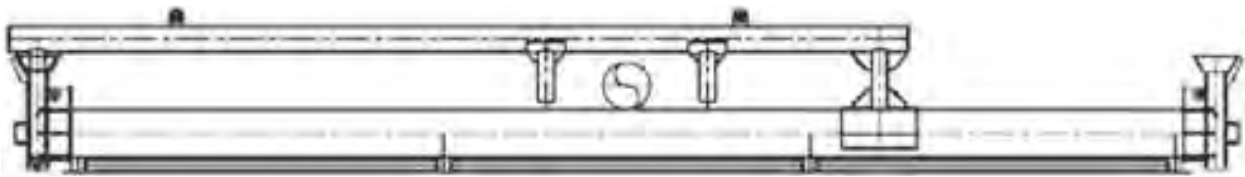


Figure 14 (a) - Elevation on Crossing Design

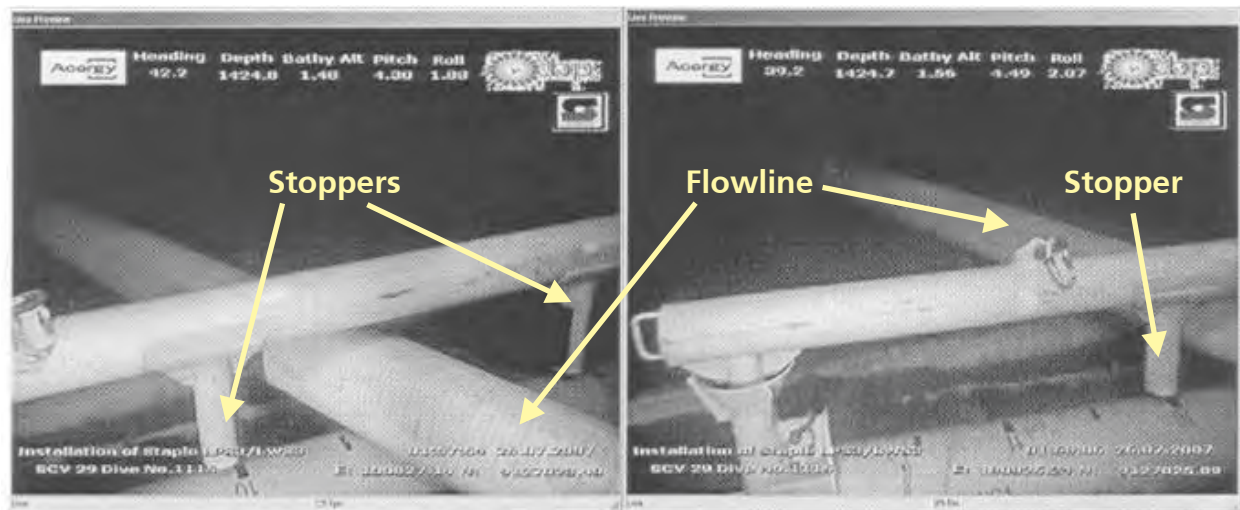


Figure 14 (b) - Installed crossing

Testing of sliding materials

Lateral buckling of flowlines on the Greater Plutonio development is controlled by the use of sleepers to raise the pipe off the seabed at intervals along the length of each flowline. The design relies on the flowlines sliding across the top of the sleepers. The friction and wear performance of the pipe-sleeper interface is important to system performance for the life of the field.

The ILTs along the flowlines and FTAs are supported on mudmat foundations with the flowline components sliding on rails with low friction support pads. The foundation loads and the flowline end loads are affected by the frictional resistance of the pad-rail interface, which is important to system performance for the life of the field.

To understand the behaviour and long-term performance of these sliding interfaces a test programme managed by Atkins Boreas on behalf of BP and the Greater Plutonio Contractors was carried at Doosan Babcock (Scotland).

The aims of the test programme were:

- (1) To measure the static and dynamic friction coefficient of the various material combinations under various design loads and displacement speeds
- (2) To confirm material wear and creep rate over the equivalent design life, to ensure that the loss of thickness would not compromise the wear surface performance
- (3) To confirm the full scale performance of the pad design, including the attachment method.

The test programme proved of great value in eliminating unsuitable materials and enabling the project to confirm suitable material combinations with acceptable friction and wear properties for all required applications.

FTA and ILT sliding pads and rails - friction testing

Suitable FTA/ILT pad material combinations tested, see Figure 15, included Polytetrafluoroethylene (PTFE) on PTFE or Epoxy coating, which achieved a very low friction coefficient (up to 0.04) and Ultra-High Molecular Weight Polyethylene (UHMWPE) on Epoxy (up to 0.08). The PTFE/PTFE combination exhibited the best wear and friction properties but field experience has shown it is difficult to manufacture for this application. These tests also showed that the increase in friction that occurs after a long dwell period or with seabed soil at the interface was marginal.

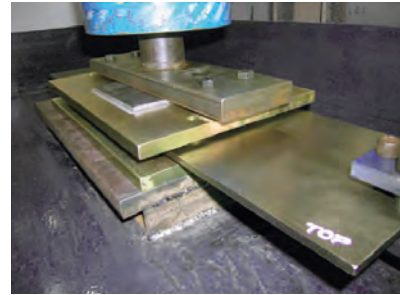
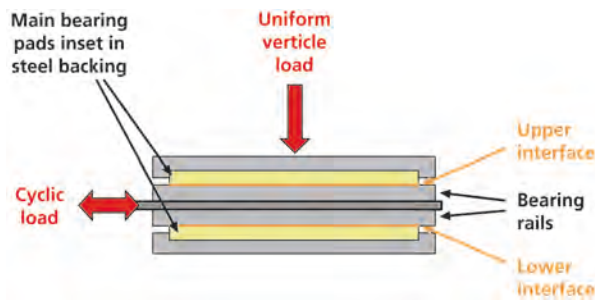


Figure 15 - Back-to-back bearing pad/rail configuration prior to seawater submersion

Flowlines on sleepers - friction testing

The pipe-sleeper interface tests, see Figure 16, measured consistent friction coefficients 0.14 to 0.36 under operational loads for insulated flowlines with PP coating, and 0.09 to 0.17 for the heavier WI flowlines with PP coating. These were lower than expected and within design limits. The tests demonstrated the expected reduction in friction coefficient with increasing load and showed that increasing dwell times between sweeps did not significantly influence results.

However, friction coefficients for the Polyurethane (PU) field joint were higher than expected at 0.54 to 0.76. Although not considered a long-term concern, some wear of the field joint coating was observed. The Glass-Flake Epoxy (GFE) sleeper coating was shown to be susceptible to wear over many cycles under high vertical loads. While this is not expected to be a concern for the Greater Plutonio flowlines, this coating is unlikely to be acceptable for heavier pipelines.



Figure 16 - Back-to-back pipe-sleeper test arrangement prior to seawater submersion

Fatigue testing and welding

In-air tests of welds confirmed acceptable in-air fatigue performance for all welds. Specimens tested for the WI flowlines were subjected to prestraining representative of the reeling cycle before testing. These tests confirmed that there was no clear influence of prior reeling strain on the subsequent fatigue performance of the welds.

A desk study carried out by TWI confirmed that little data existed on the fatigue performance of pipeline girth welds in a corrosive environment, under the high stress/low-cycle fatigue conditions applicable to the Greater Plutonio production flowlines.

A programme of sour service endurance fatigue testing in simulated pipeline operating conditions was therefore initiated at TWI to support the flowline design process. Testing compared the fatigue life of welds in air with that in a sour environment. A particular concern was the increase in fatigue damage due to low frequency cycling, associated with cyclic lateral buckle loading. Results of tests carried out at 0.2Hz suggested that the fatigue performance, for the three pipeline welding procedures tested, is similar and that the fatigue life reduction factor relative to air is approximately 10. The use in design of an 'E' curve divided by ten encompasses the 0.2Hz and 0.01Hz results, shown in Figure 17.

In parallel testing, the effect of low frequency Fatigue Crack Growth (FCG) rate in the same simulated sour environment was studied using frequency scanning at constant Stress Intensity Range (DK) as well as the more usual variable ΔK approach. The FCG rate was shown to increase with decreasing fatigue load frequency, to reach a plateau at around 0.1Hz, see Figure 18.

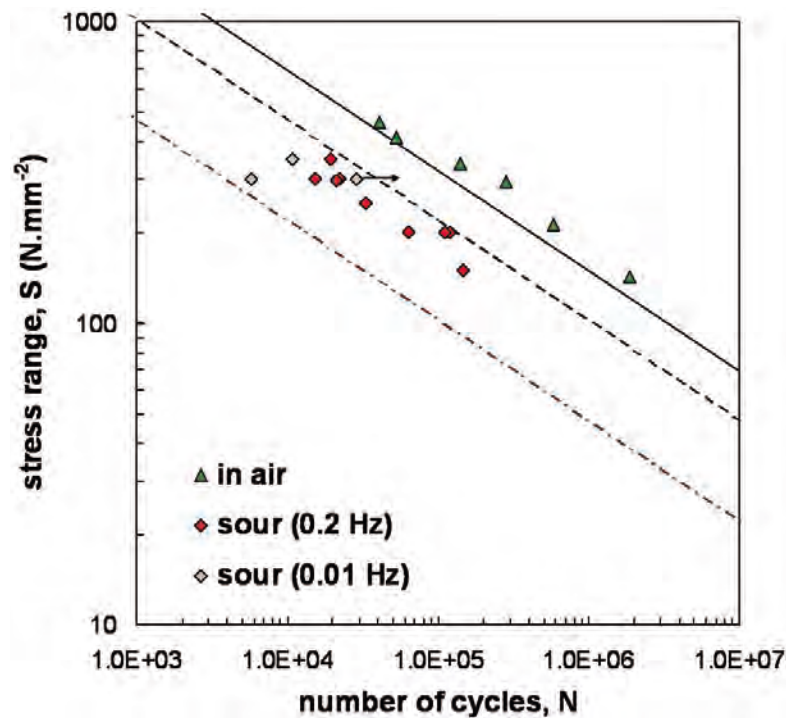


Figure 17 - Fatigue endurance test results in sour environment

These results proved useful in guiding the test frequencies employed in the endurance testing, discussed earlier. The crack growth rate data obtained was used as one of the required inputs to Engineering Critical Assessment (ECA) calculations from which weld acceptance criteria were derived for flowline girth welds, prepared onshore and offshore.

A further test programme by TWI, established the fracture toughness of welds in the simulated sour Plutonio environment under operating and shutdown conditions.

Lower bound data on toughness was measured in terms of the stress intensity at which crack extension occurred by stress corrosion cracking (K_{ISCC}) at test temperatures in the range 20°C to 60°C. Other tests conducted on hydrogen charged samples were used to measure the effect of hydrogen embrittlement on the overall toughness of pipeline welds at ambient temperature. These results were then used in the flowline ECA, to calculate the tolerable flaw sizes for surface breaking and embedded flaws.

Experience from low-frequency testing in seawater for the SAFEBUCK JIP had highlighted a similar concern about a significantly increased "fatigue life reduction factor" under low frequency loading in seawater with Cathodic Protection (CP). As a result, particular attention was given to field joint integrity on all flowlines to ensure that welds would not be exposed to seawater.

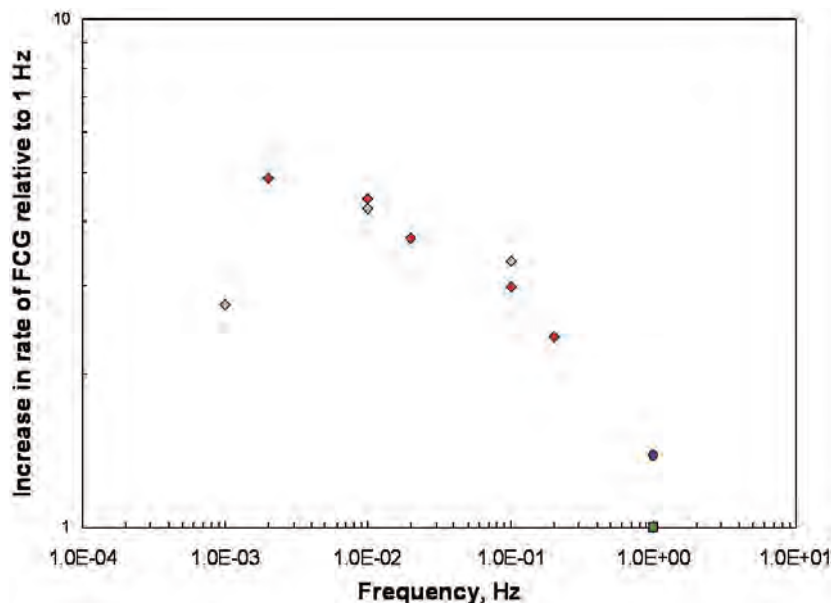


Figure 18 - Increase in fatigue crack growth with reducing frequency

Fabrication and installation

An overview of the fabrication and installation scope which was split between Technip and Acergy is presented in this section.

Technip scope

Technip installation scope was as follows:

- 55km of WI flowlines (plastic lined pipe)
- Sleepers for buckle initiation
- 16 WI FTAs and ILTs
- 12 WI flexible spurs (20km length in total).

The subsea structures piping was fabricated in Europe, and then assembled to the structural beams and mud mats in South Africa. It is worth noting that all piping within these structures was internally clad with Inconel 625 alloy.

The rigid pipelines were fabricated at the Dande Spoolbase jointly owned by Technip and Sonangol. Over 500 000 man hours were spent on the following activities:

- Rigid line pipe jointing into 1000m stalks: circa 2000 automated welds were performed, with no recordable time loss, see Figure 19(a)

- Rigid line pipe SCR type welding for fatigue sensitive areas: more than 270 manual GTAW with only 1-off repair
- HDPE liner fusion welding into 1000m stalks, see Figure 19(b)
- HDPE insertion into line pipe stalks, see Figure 19(c)
- HDPE liner energization, see Figure 19(d)
- Stalks assembly into circa 3km sections, see Figure 20
- Pipeline spooling onto the installation vessel. See Figure 21.

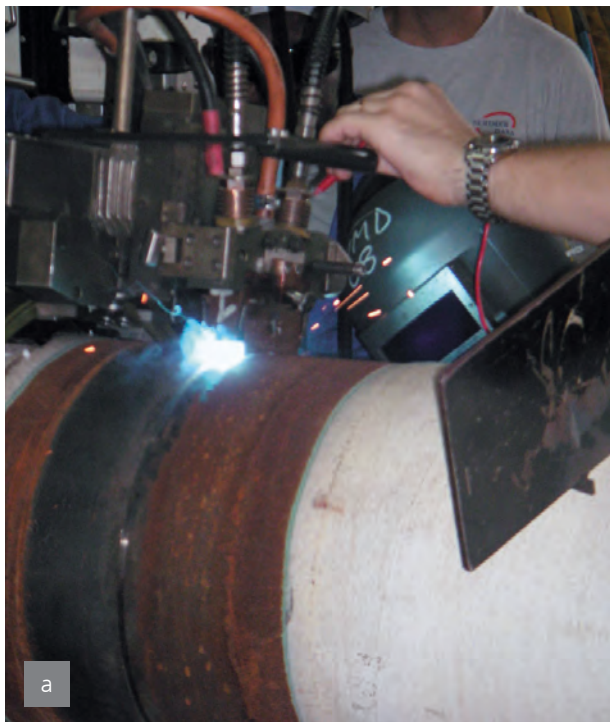


Figure 19 - Rigid pipeline fabrication at the Dande Spoolbase

The sleepers were assembled at the Dande Spoolbase, and then loaded out on the Geofjord vessel, to be installed prior to the pipelay activities, see Figure 22.

The 56km of WI rigid flowlines were spooled onto the Deep Blue, and then reel-laid. The 15-off subsea structures inserted in the rigid WI Flowlines were also installed by the Deep Blue, while the flexible spurs and the associated subsea structure were installed using the Constructor, see Figure 23. The Geofjord also providing field assistance to the Deep Blue, retrieving the buoyancy modules used to land the subsea structures.



Figure 20 - Stalk assembly



Figure 21 - Pipeline spooling onto the Deep Blue installation vessel



Figure 22 - Sleepers loaded out at the Dande Spoolbase



Figure 23 - Installation of subsea structures (a) and flexible spurs (b)

Acergy flowline installation scope

Acergy installed the production, service and gas injection flowline systems using the Acergy Polaris (flowline and large spools) and with the Acergy Eagle for the anchor piles. The flowline lay and structure installation was carried out using the new J-lay tower installed on the Acergy Polaris as illustrated on Figure 24.

The new tower enabled a maximum tension capacity 950tonnes and Gimbaling ability of 15° pitch and roll, see Figure 25. Two work stations allowed for welding and field joint coating.

The Acergy installation scope was as follows:

- 34km of Production flowlines (with insulation coating)
- 42km of Gas Injection and Service flowlines (non insulated)
- 24 sleepers for buckle initiation
- 3 flowline anchors and 2 devices for retrofit anchors to control pipe walking
- 18 FTAs and ILTs
- 17 rigid spools (including spools in excess of 100m length).

Flowline initiation was performed as follows:

- (1) FTA inserted in tower from bottom by crane, see Figure 26
- (2) Top end collar landed on bushing
- (3) First weld performed, see Figure 27
- (4) Initiation structure linked to initiation wire/chain for seabed connection.

Two welding procedures were qualified for the production flowlines to cover the welding of double joints onshore and the offshore installation by the J-lay method.



Figure 24 - Acergy Polaris installation vessel

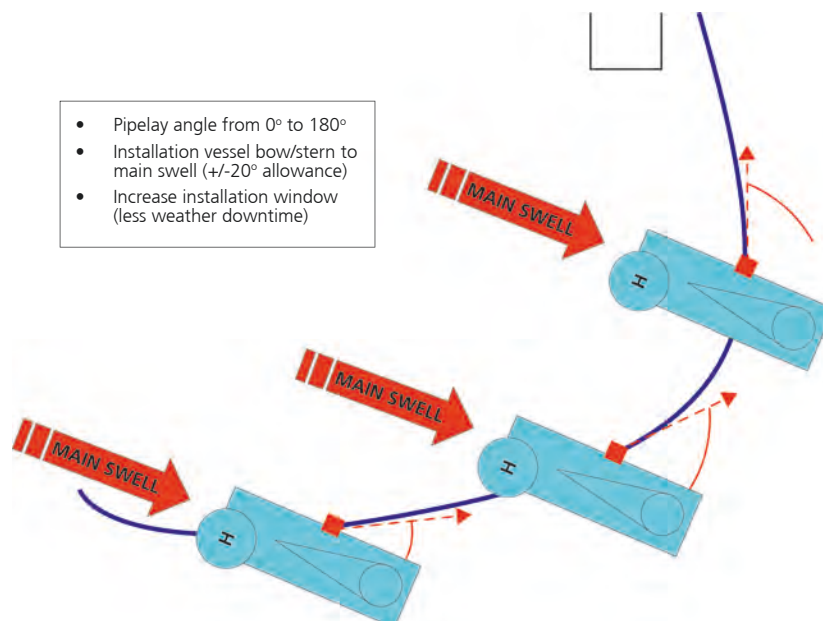


Figure 25 - J-Lay Tower Gimbaling

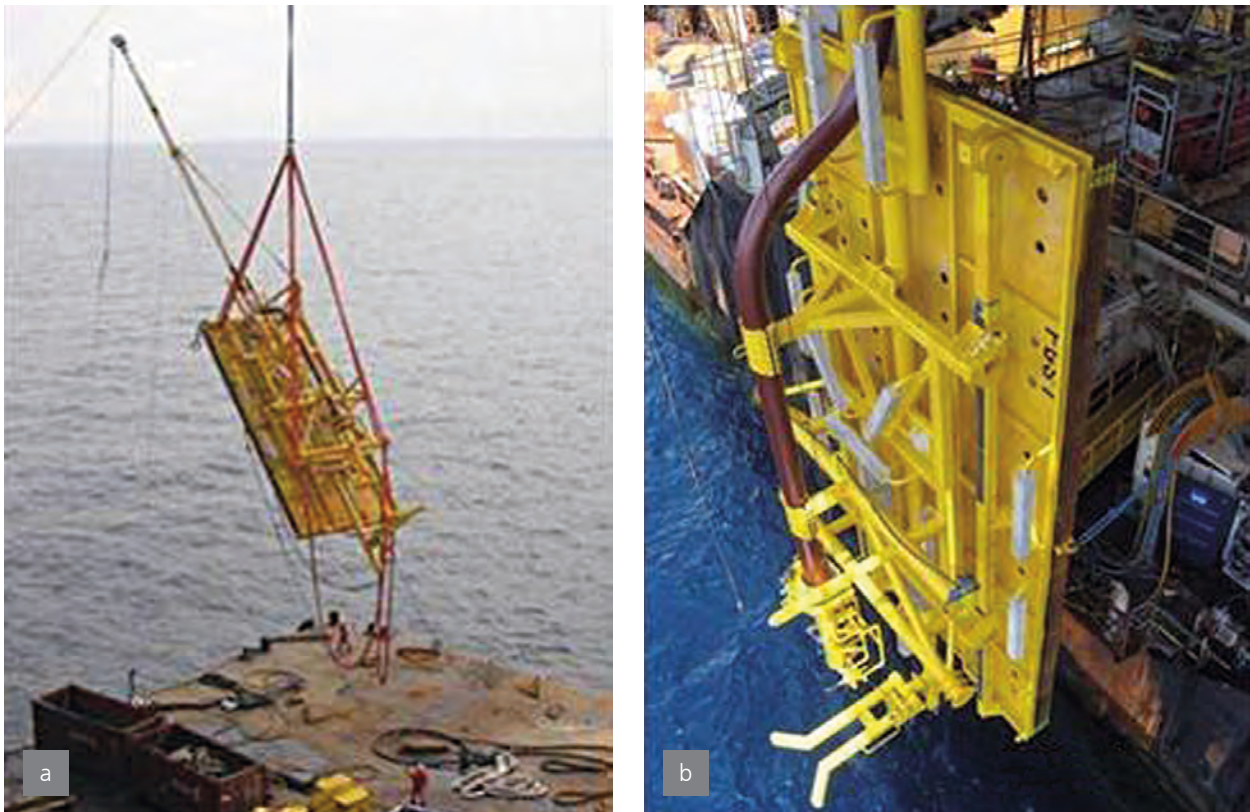


Figure 26 - Flowline initiation - FTA installation

The welding technique qualified by the double joint contractor involved a 60° included angle, single “V” preparation with manual GMAW root and intermediate runs, sufficient to support the subsequent fill and capping by SAW. This work was performed at a double jointing facility set up for the project in Brazil. The same welding processes were used for double joints destined for both fatigue sensitive buckle zones and other areas in the flowline, although

the allowable root misalignment on fatigue welds was more restrictive. The welding procedure qualified for the J-lay installation used a standard narrow gap weld preparation. The Surface Tension Transfer (STT) GMAW technique was used for the root run, as an alternative to the more traditional copper backing technique. Fill and cap runs utilised mechanised pulsed GMAW.

The fatigue sensitivity of the buckle zone welds meant that weld acceptance criteria based on workmanship standards were not appropriate. Automated ultrasonic testing (AUT) was essential to permit the through thickness extent of flaws to be determined and achieve the necessary reliability of flaw detection and sizing. The selected AUT Contractor was able to provide limited previous qualification data for the double joint weld configuration, hence a more extensive AUT validation programme was undertaken. This involved the preparation of test welds with artificial defects and metallographic sizing of defects to compare with the flaw sizes given by the AUT system. For the offshore J lay welds, previously qualified data was available, but again was supported by a validation test programme.

For both the double joint and J-lay welds the degree of AUT under-sizing was less <1.0mm and a probability of detection in accordance with the requirements of DNV OS-F101 was achieved.



Figure 27 - First weld performed on flowline



Figure 28 - Installation of FTA and ILT structures

The sizing accuracy values from the validation testing were used to modify the tolerable flaw sizes obtained in the flowline ECAs, allowing the development of the final weld acceptance criteria applied to production welds.

Laydown with FTA and ILT structure installation were performed by inserting the structures in the J-lay tower according to similar steps as initiation structures, see Figure 28.

On the south production system 3 FTAs required to be connected to anchor piles for pipe walking prevention. Suction anchors needed to be installed prior to laydown with the Acergy Polaris. The top of pile, as installed, is shown in Figure 29 (a).

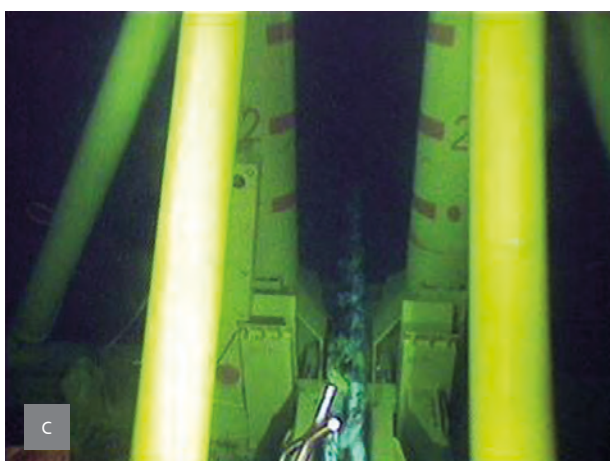


Figure 29 - FTA connection to anchor piles



Figure 30 - Sleeper deployment and pipelay over a sleeper

The steps for connection of the pipe end to the anchor pile with a chain during laydown are shown in Figure 29 (b) to (d).

Sleepers were deployed with the barge crane while pipe-laying, since the touch down point was significantly behind the barge location, see Figure 30 (a).

Transponders installed on the sleepers and Remote Operated Vehicle (ROV) observation ensured that the flowline was later laid as close as possible to the centre of the sleeper, see Figure 30 (b), but with an offset on lead in.

Rigid spools were installed with the Acergy Polaris. A specific spreader frame was fabricated allowing deployment of the whole range of shapes and sizes as illustrated on Figure 31. The longest spools measured 80 x 40m.

All spools between FTAs or ILTs and the manifolds were connected using FMC horizontal connectors with the CAT tools whereas on the riser tower side, the spools were connected using the Matis flanged connection tool illustrated in Figure 32 (a) while Figures 32 (b) and (c) show the subsea connection steps.

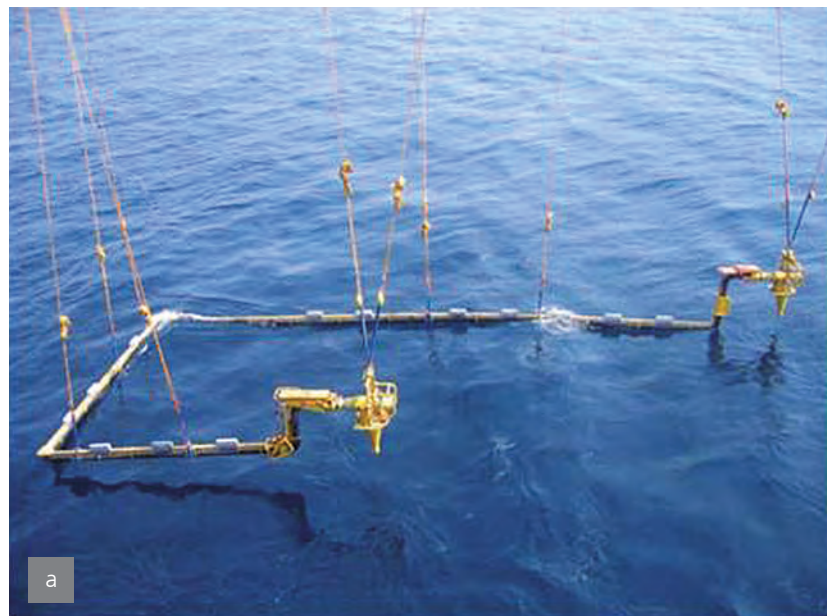


Figure 31 - Rigid spool installation



Operational monitoring

Overview

Operational monitoring for each of the systems relies on an integrated process that re-evaluates the remaining fatigue and allowable expansion available within the structures on a periodic basis. A highly accurate, as-laid survey was carried out to establish the initial condition of the flowlines including out-of-straightness (OOS) along the full length of each flowline. Additional surveys are being undertaken during hydrotesting and post-startup, to verify buckle formation identify rogue (unintended) buckles and confirm operational loading at each lateral buckle site.

Deflected position survey requirements

The survey requirements for successful monitoring required accuracies of better than 0.2m deflection over 50m length in a water depth of 1400m without using an LBL array. To enable this, the following survey equipment was integrated to the ROV:

- (a) Kongsberg HiPaP 500 transponder/responder in Ultra Short Base Line (USBL) mode
- (b) Kongsberg Hydroacoustic Aided Inertial Navigation System (HAIN), fitted with CDL miniRLG 2 or Ixsea IMU90 subsea unit, and NavLab post processing software
- (c) Doppler Velocity Log (DVL) RDI WorkHorse Navigator 600KHz or 1200KHz
- (d) Subsea gyrocompass CDL miniRLG2 and/or Ixsea Octans
- (e) Paroscientic Digiquartz Depth Sensor (DQ)
- (f) ROV mounted Valeport or SAIV CTD and mini SVS
- (g) Dual Heads Multibeam Echosounder (MBE) Reson Seabat 7125
- (h) Starboard and portside boom cameras, with central camera. All three videos were logged using VisualSoft digital video system
- (i) Wheeled undercarriage system, adjustable for 300mm to 500mm flowline OD.

Figure 32 - Matis flanged spool connection

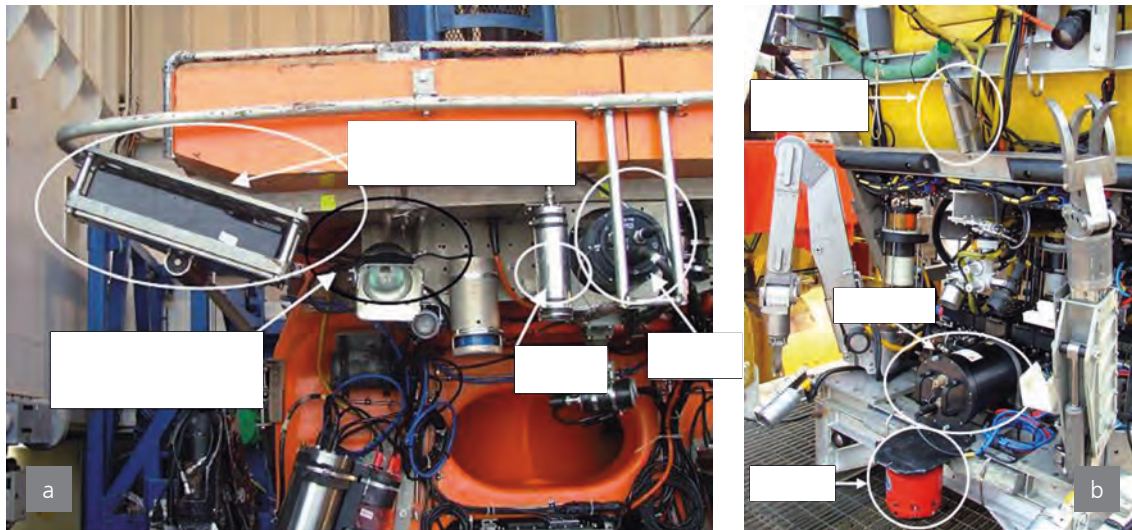


Figure 33 - HiRov 3000 (a) and SCV 32 (b)

The above configuration enabled two OOS survey options:

Contact Method: With the ROV continuously rolling along the flowlines. This was the primary method as it allows direct derivation of the XYZ position. The HAIN INS results are directly used as flowline reference. They can be compared with ROV tracks computed using alternative processing methods such as Gyro+DVL and/or HiPaP+DVL, for quality control and removal of possible positioning artefacts.

Non Contact method: With the ROV flying 0.5m to 1.5m over the flowlines. The XY values of the Top of Pipe (TOP) are then determined from the processing of the swath bathymetry data in relation with the ROV Track (HAIN INS aided USBL HiPaP). The Z values of the TOP are determined from the Digiquartz combined with the MBE. This method was used when the flowlines were buried or too deeply embedded and also for spools.

A very detailed HAIN, DVL and MBE calibration procedure was performed offshore prior to the survey start up. This involved an onshore analysis at Kongsberg Horten for fine tuning of the Inertial Measurement Unit (IMU), to be reset within the Navlab post-processing parameters. The following deliverables were produced onboard for immediate engineering analysis:

- (a) Processed ASCII Flowline 5 points files at 1m spacing
- (b) Cross Profiles at 1m spacing
- (c) Digital video logs, combined with flowline events

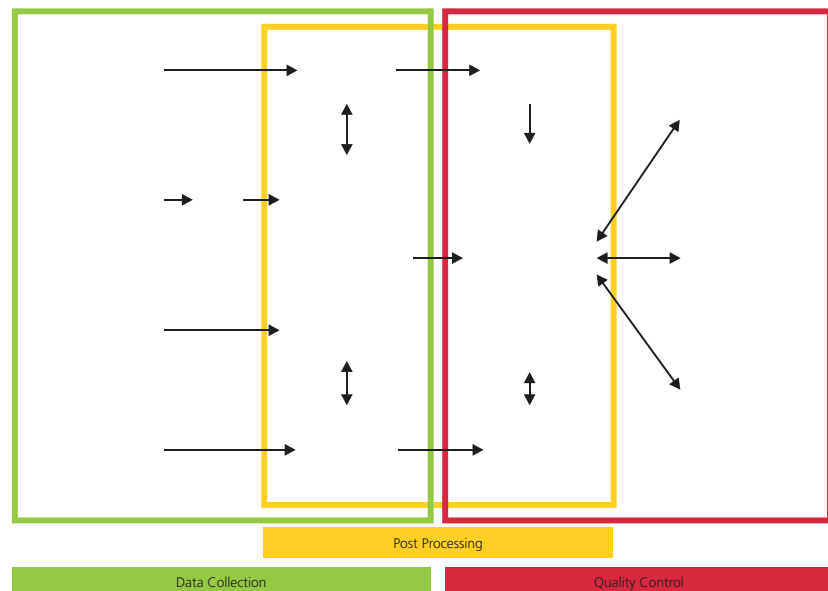


Figure 34 - Data flow/quality control schematic

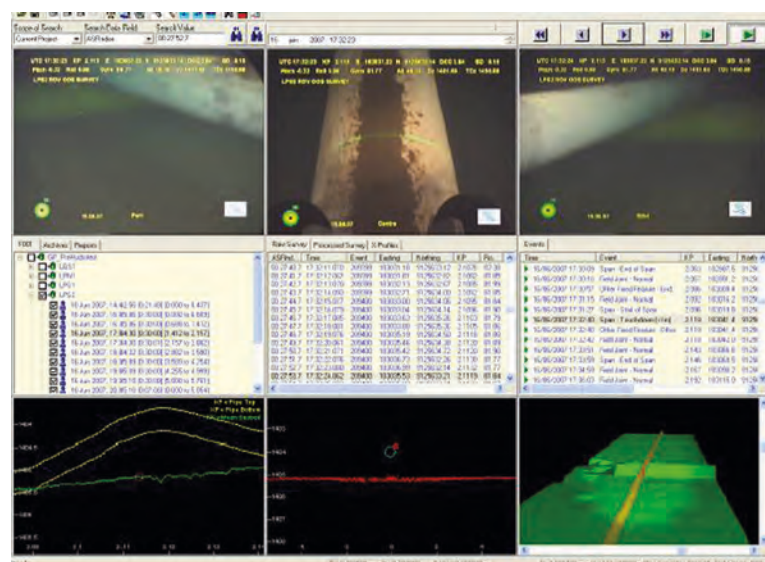


Figure 35 (a) - Thermal expansion

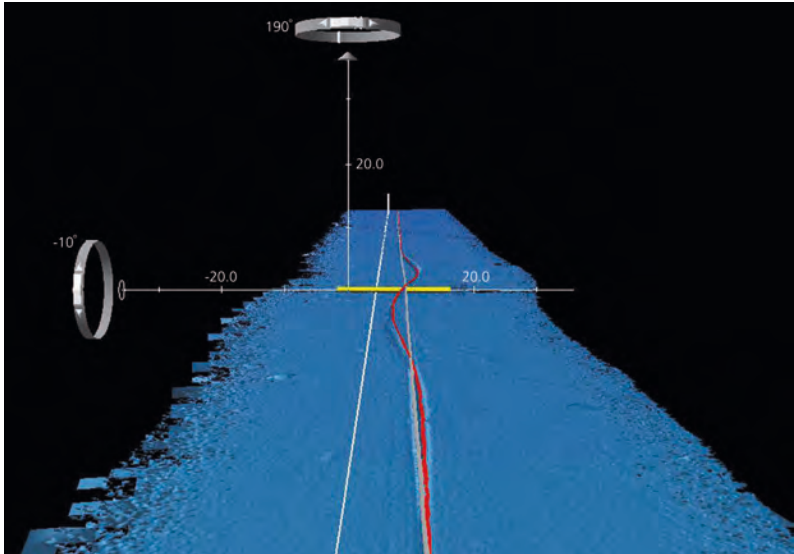


Figure 35 (b) - OOS survey/project data flow

- (d) OOS check list
 - (e) Raw and processed MBE data for Fledermaus import and 3D visualisation by NetSurvey (0.15m grid for detailed buckles areas and 0.25m grid over the complete flowline length).
- All available information was combined during analysis of the OOS survey results. In particular, very small rogue buckles identified during hydrotest with an amplitude of 0.2m to 0.5m, over a 50m long section, were confirmed from positioning track comparison, and differences in the left/right differential embedment level, combined with

visual observations of the soil berm created to one side of the line.

Comparison made between pre-hydrotest, hydrotest and post start up surveys have allowed high quality data to be extracted and used within engineering FE analyses. Data is currently being imported into Geographical Information System (GIS) for use during the operational monitoring of the field.

The majority of the surveys conducted to date were undertaken using a DeepOcean WorkRov HiRov 3000, see Figure 33 (a), deployed from the M/V Normand Tonjer, which is a DP2 ROV/Survey vessel.

Occasionally, some lines were also surveyed using SCV 32, see Figure 33 (b), deployed from M/V Acergy Legend, for example during hydrotest OOS surveys along the south production system. During the remaining field life, ongoing surveys will be performed using a dedicated infield support vessel.

Use of GIS data base as an IMR tool

Operational monitoring uses geographically referenced information that enables engineering analysis to be undertaken on a holistic basis. The use of GIS also allows ease of recognition of apparent areas of concern, and rapid prioritisation of ongoing inspection or intervention activities. Table 4 provides an example of some of the data that is contained and regularly updated within GIS for the flowline system. Integrated data display examples are shown in Figure 36 and Figure 37. The red dot on the field map in Figure 37, indicates position of ROV when acquiring these images.

Table 4 - Example of GIS flowline data

Item	Description	When	Comments
1	Deflected position of flowline/ Spool (geo referenced video/ MBES/ 5 point cross profiles/ embedment etc)	At installation	Provides key input data for fatigue life calculations and monitors issues such as curve pull out / rogue buckle growth etc.
		During hydrotest	
		Post start up	
		Intervals during field life	
2	Linear expansions on ILT/FTA assemblies	As for 1 above	Provides data on walking
3	External coating integrity on deflected flowline areas	Post start up	Provides data on whether knock down factors require adjusting (i.e. welds exposed to seawater)
		Intervals during field life	
4	Internal crack detection on deflected flowline welds	Intervals during field life	Provides data on fatigue weld integrity.
5	CP field gradient	Intervals during field life	Data on coating integrity
6	Internal and external corrosion on deflected flowline areas	Intervals during field life	As above
7	Start up and shutdown history	Continuous	Allows subsea and topsides teams to evaluate possible contingency plans

Greater Plutonio project - Subsea flowline design and performance

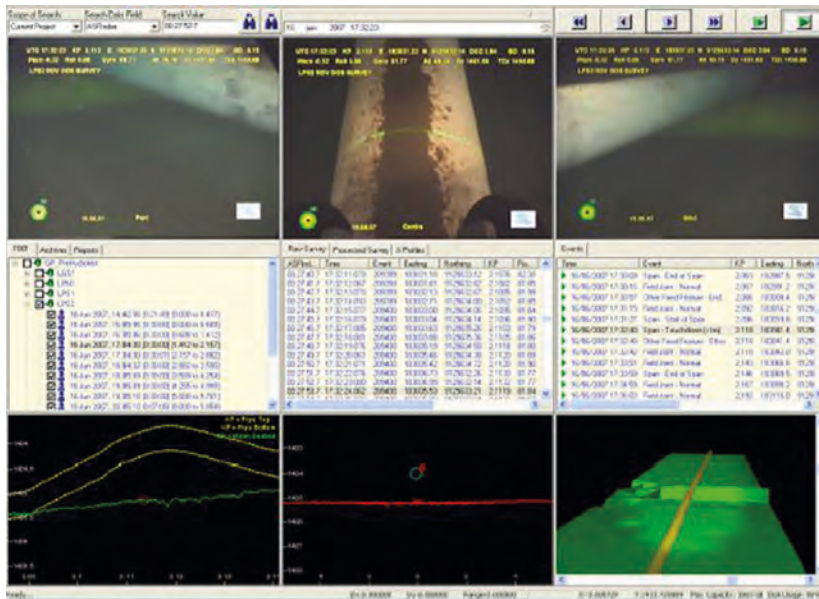


Figure 36 - Cross profiles and combined MBES 3-D imaging

AS-BUILT / DIGITAL VIDEO

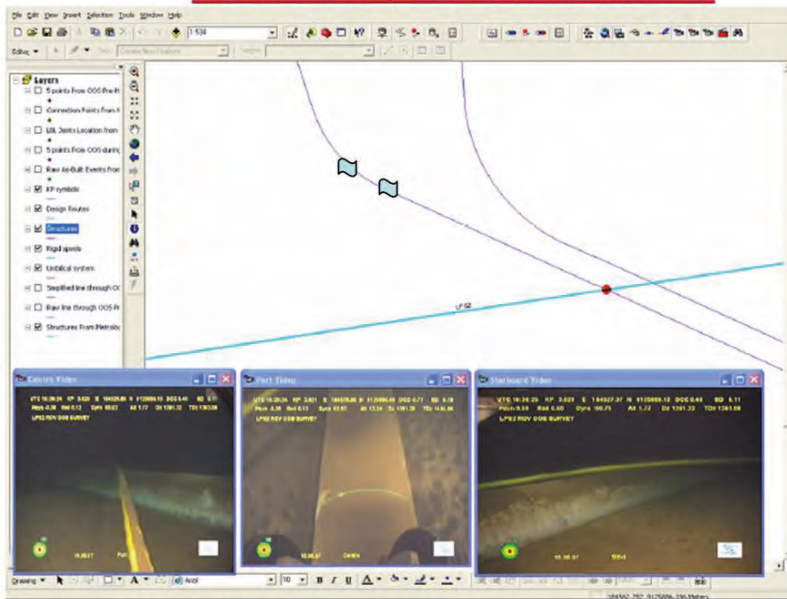


Figure 37 - As built digital video images

Performance

Detailed Out of Straightness (OOS) and end expansion surveys were carried out along each flowline post installation, during hydrotest and after first operation. OOS surveys were also performed for the spools. The data from these surveys was then used to determine the performance of each flowline and allow it to be compared to design. The hydrotest and first operation surveys were also used to determine whether any unplanned (rogue) buckles had formed.

Buckle formation

The primary function of the OOS surveys during hydrotest and first operation was to determine whether buckles had formed at the planned locations. Detailed design analyses concluded that the hydrotest would generate enough axial force to initiate planned buckles on all lines. This effectively meant that the hydrotest could be used as a "dummy run" for buckle initiation before flowlines were subjected to the more onerous operating conditions. It also gave the project scope to force buckle initiation during start-up, if required.

A summary of the buckle formation at hydrotest and during first operation is included in Table 5.

Buckle formation was very reliable on the production, service and gas injection flowlines during hydrotest, with all 24 planned buckles forming; in addition only 4 unplanned "rogue" buckles formed. Each rogue buckle formed between planned initiators and did not hinder planned buckle formation. The flowline design accounts for on-bottom rogue buckles forming between initiators.

A number of the WI Buckle Initiators (BI) did not trigger buckles during hydrotest. Several of the initiators are expected to form during normal operation. The reason that some WI buckles did not initiate during hydrotest has yet to be fully explained, although there are a number of potential causes, including potentially high residual tension in the flowline, low out-of-straightness level, the designed pipe-sleeper configuration being less conservative than those on the production or gas Injection flowlines, or possibly the fact that the WI lines are heavier than the production flowlines.

Table 5 - Buckle formation

Flowline	Planned Buckles	Hydrotest buckles formed		Operation buckles formed	
		Planned	Rogue	Planned	Rogue
Production	North	9	1	TBC	TBC
	South	6	0	6	0
Service	4	4	2	TBC	TBC
Water Injection	North	7	3	TBC	TBC
	South	11	5		
Gas Injection	5	5	1	TBC	TBC

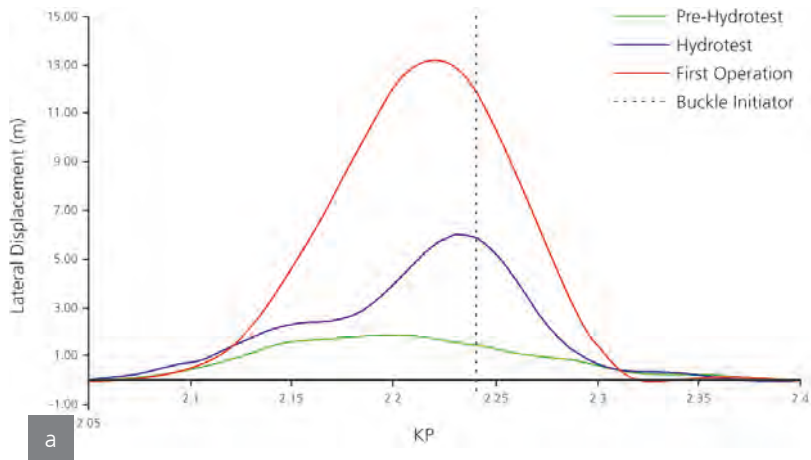


Figure 38 - OOS surveys – (a) Survey Data comparison, (b) Pipe movement on initiator

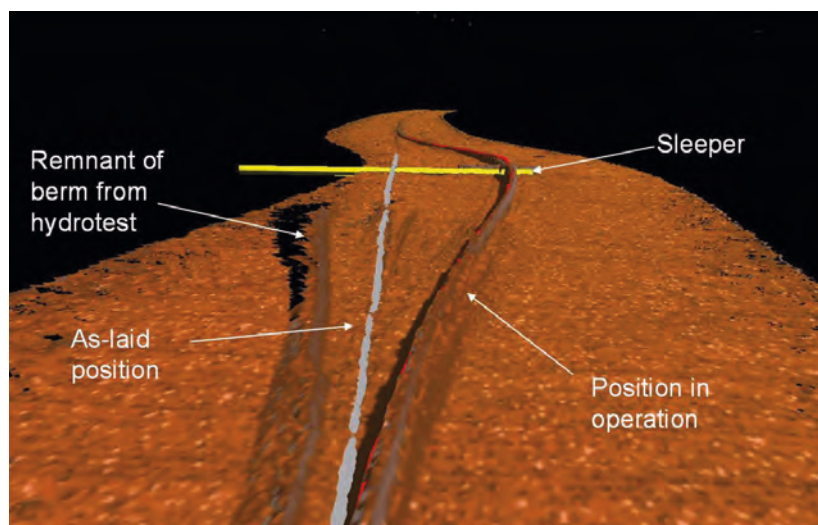


Figure 39 - Buckle shape in operation

Buckle shapes

Detailed OOS surveys were carried out along each flowline prior to hydrotest, during hydrotest and after first operation. Each survey provided a very accurate understanding of the behaviour of each buckle under different loading regimes.

An example of typical survey data is presented in Figure 38 (a) includes an overlay of all three flowline OOS surveys. Survey video was also used to cross check the buckle amplitudes on each sleeper; lateral measurement markings on each sleeper allow estimation of buckle movement. An example pipe on sleeper is presented in Figure 38 (b).

The survey data clearly shows the formation of a buckle on the initiator during hydrotest which grows in amplitude at first operation. For each flowline the surveyed buckle shapes were used in FEA analyses to determine the loads in each buckle and also estimate the actual pipe-soil response.

Buckle shapes from first operation, see Figure 39, were within the bounds expected in design. As part of the ongoing integrity monitoring programme the buckle shapes were fitted using finite element models, the majority of buckles achieved excellent shape fits with lateral friction at large displacements well within the design range and close to the mean defined by pipe-soil interaction testing.

End expansion

Each FTA and ILT was designed with measurement bars on the sliding frame so that accurate readings of axial movement could be taken during each survey. Readings were taken at each FTA during hydrotest and during first operation. Figure 40 presents typical video-grabs of slider movement on an FTA for hydrotest and operation.

The observed end expansions in operation for each of the southern production flowlines were within the bounds set in detailed design, most of the expansion values lay reasonably close to the mean design predictions, typically within 0.1m.

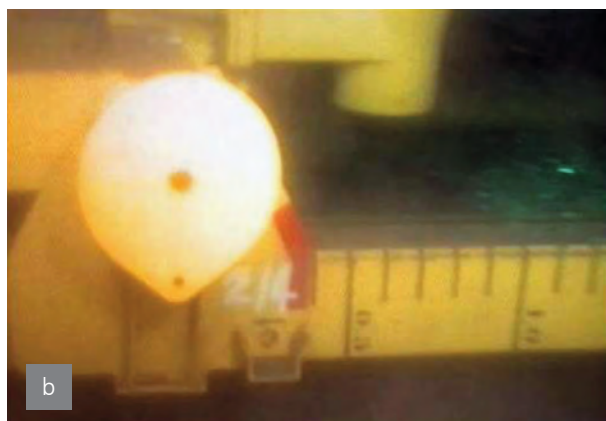
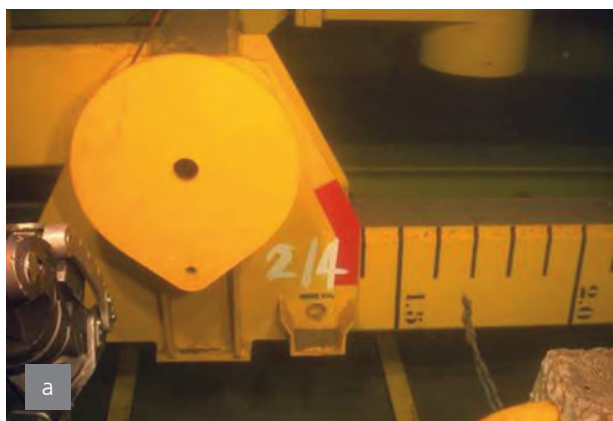


Figure 40 - End Expansion Readings – (a) Hydrotest, (b) First Operation

Ongoing flowline monitoring

The flowline integrity management plan calls for regular flowline surveys to check buckle shapes and end expansion readings. The survey data will be combined with regular monitoring of pressure and temperature conditions in each flowline so that models can be updated regularly to monitor buckle shapes, predict fatigue life consumption and also determine whether flowline walking is occurring.

Conclusions

The flowline system was installed off the critical path with first oil being achieved on October 1st 2007. The performance of the system was within design limits and no remedial works have been required.

Several world firsts were accomplished along the way. These include the soil testing and analysis, sour service low frequency fatigue testing, the use of integrated buoyancy and sleepers to control buckles, the use of rough coated pipe to increase axial friction, a novel installation scenario for connection and tensioning of piles to flowlines, and the provision of a future (contingency) mid-line anchor pile connection.

In addition, there were several design details which should receive mention: these include the use of high integrity field joints, machine profiled J-lay collars, the elimination of anodes on the flowlines to assist with overall coating integrity, and the use of wet lay to enhance embedment and reduce operational loads (less effective force in the operation).

Key lessons learnt include:

- (i) The necessity to perform analytic modelling at the Front End Engineering Design (FEED) stage of any design
- (ii) Early integration with topsides design teams
- (iii) Early implementation of GIS
- (iv) Recognition that soils and fatigue behaviours will require an integrated test programme
- (v) The realisation that the modelling complexity would benefit from independent third party analysis
- (vi) Evaluation of the effects of pipelay systems and methodology on embedment, residual lay tension, and the provision of initial offsets to assist buckle formation.

Acknowledgements

The authors would like to thank all those individuals and companies that have contributed to the success of the Greater Plutonio Project and give a special mention to those individuals who have also contributed to the writing of this paper:

Peter Bowden (BP) – Fatigue Testing & Welding

Andrew Cosham (Atkins Boreas) – Fatigue Testing & Welding

Abbreviations and nomenclature

AUT	Automated Ultrasonic Testing
BI	Buckle initiator
COF	Coefficient of Friction
CP	Cathodic Protection
DK	Stress intensity range
FCG	Fatigue Crack Growth
FEED	Frontend Engineering Design
FPSO	Floating Production Storage and Offloading
FTA	Flowline Termination Assembly
ECA	Engineering Critical Assessment
GFE	Glass-Flake Epoxy
GIS	Geographic Information System
ILT	In-line Tee
JIP	Joint Industry Project
OOS	Out of straightness
PP	Polypropylene
PTFE	Polytetrafluoroethylene
PU	Polyurethane
R	Ratio of minimum load to maximum load
ROV	Remote Operated Vehicle
SCF	Stress concentration factor
UHMWPE	Ultra-high molecular weight polyethylene
WI	Water Injection

References

1. Carr, M., Sinclair, F., Bruton, D., "Flowline Walking – Understanding the Field Layout Challenges, and Analytical Solutions developed for the SAFEBUCK JIP", OTC-17945 Offshore Technology Conference, Houston (2006).
2. "Safe Design of Flowlines with Lateral Buckling - Design Guideline" Report BR02050/SAFEBUCK/ Rev B, SAFEBUCK JIP (2004). Confidential to JIP Participants
3. Bruton, D., White, D., Cheuk, C., Bolton, M., Carr, M., "Pipe-Soil Interaction Behaviour during Lateral Buckling, Including Large Amplitude Cyclic Displacement Tests by the Safebuck JIP" OTC 17944 (2006).
4. Murff, J.D., Wagner, D.A., & Randolph, M.F. "Pipe penetration in cohesive soil". *Géotechnique*, 39(2):213-229 (1989).
5. Lund, K.M.. Effect of Increase in Pipeline Penetration from Installation. Proc. of the Conference on Offshore, Marine and Arctic Engineering. New Orleans, USA, (February 2000)
6. Cheuk, C.Y., White, D.J., "Centrifuge Modelling of Pipe Penetration due to Dynamic Lay Effects" Proc. Conf. on Offshore Mechanics and Arctic Engineering (2008).
7. Najjar, S.N., Gilbert, R.B., Liedtke, E.A., McCarron, W. Tilt Table Test for Interface Shear Resistance Between Flowlines and Soils. Proc. Conf. on Offshore Mechanics and Arctic Engineering (2003).
8. Langford, T.E., Dyvik, R. & Cleave, R., Offshore pipeline and riser geotechnical model testing: practice and interpretation OMAE2007-29458 (June 2007)

Development of an armour mass estimation tool for land vehicles



Matt Geary
Graduate Engineer
Defence

Abstract

The Future Rapid Effects System (FRES) is the British army's largest procurement programme; aiming to deliver a family of medium weight armoured vehicles. This paper describes the Atkins survivability team's development of a spreadsheet tool to model the weight and performance of conceptual armour systems for FRES vehicles. The selection of armour materials is complicated by the wide variety of impact mechanisms, which have conflicting demands on armour material properties. This paper first discusses some of the more common types of impact mechanism and armour type. The general armour concepts of Whittaker weighting and mass efficiency are then described. The development of the 'Weight stack' spreadsheet tool is split into the three main stages of; describing vehicle geometry, storing material information and calculating armour weight.

Background

The Future Rapid Effects System project covers a wide variety of vehicles, including protected troop carriers, reconnaissance and direct fire roles. Despite their different purposes, the vehicles all have the common goals of providing high levels of protection and mobility. In addition, the vehicles are all to be air transportable so as to provide the British army with a rapid reaction capability.

Atkins has worked alongside the Ministry of Defence (MoD) Defence Equipment and Support (DE&S) group as a customer friend and provides a variety of systems engineering and technical services.

Statement of problem

In armoured vehicle design there has always been a tension between the requirements for protection and for mobility. In order to provide higher levels of protection, additional armour is inevitably added, decreasing the power to weight ratio of the vehicle. In order to effectively balance the protection requirements for an armour vehicle, the Survivability team requires a way of assessing:

- (1) The protection levels which are achievable within a given weight budget, to ensure that protection requirements given to vehicle providers are achievable.
- (2) The likely weight of a defined protection level, in order to assess a vehicle provider's armour proposals
- (3) The problem of assessing armour weight and performance is complicated by the variety of threats, impact mechanisms, armour types and other variables.

Threat types

There a huge variety of threats in existence, but the type of impact can largely be broken down into four categories:

Low kinetic energy projectiles

This is a broad category encompassing any projectile travelling below approximately 1200m/s, with the threats ranging from blast fragments to heavy machine gun projectiles. Even within this category there are many different impact mechanisms at work, several of which can occur at different stages of the impact. The specific impact mechanisms of the projectile will be dependent on many factors including the velocity, material properties, geometry and angle of incidence.



Figure 1 - Sample range of low KE projectiles

To provide effective protection against a projectile, an armour material must be capable of both resisting penetration and absorbing the projectile's energy.

To resist the initial penetration of the projectile, it is desirable for the material to have high hardness and high yield strength. A particularly hard material is also likely to deform the nose of a projectile, or in some cases shatter it. This distributes the impact and reduces the kinetic energy density of the projectile, so reducing the penetration.

Obviously, for a projectile to pass through a surface it must move material out of its path. This deformation is particularly dependant on the geometry of the projectile. The impact from a blunt projectile will cause the armour to form a dish, while a sharp nosed projectile will force material to flow radially from the point of impact. This movement of material through plastic deformation absorbs a greater amount of energy than if brittle failure of the armour material occurs¹. An armour material is therefore also required to have high ductility to effectively absorb energy.

Thin sheets of ductile armour material, particularly when impacted by a sharp projectile, are likely to form radial cracks from the site of impact as the material is less constrained. This reduces the amount of energy absorbed in the impact and goes some way to explaining why a single thick block of armour is more efficient than several thin layers.

Particularly blunt projectiles, with low kinetic energy (K.E.) density, are known to produce a 'plugging' effect. In these cases large shear stresses form at the edges of the blunt impact, leading to localised heating and causing thermal-shear instability². This leads to the separation of a plug of material from the armour, forming a secondary projectile. It has been suggested that this occurs when the thermal softening effect generated by impact overcomes the heat capacity and work hardening of the material.

Therefore an armour material should be specified to have high shear yield strength and ductility, while also having high resistance to shear instabilities through having a high work hardening index and low thermal softening rate.

Taking low K.E. threats alone, the ideal armour material will therefore have high ductility, high hardness, high yield strength, high shear yield strength, high work hardening index and low thermal softening rate. Clearly there are many competing and in some case mutually exclusive demands. In practice a balance must be struck in order to ensure an armour material provides protection against a variety of projectiles. The competing demands have lead to the use of a composite of material types in specific layers. The most common of these is the placing of a hard ceramic onto a ductile composites backing.

High kinetic energy projectiles

The impact load of a projectile is proportional to the square of its velocity. As the projectile approaches approximately 1200m/s the impact load becomes so great that the materials begin to behave in a hydro-dynamic fashion. These high levels of K.E. are achieved in ballistic projectiles through the use of a casing known as a sabot. As the projectile leaves the barrel the sabot is discarded leaving a projectile with an extremely high K.E. density. In order to maximise the K.E. density a large length to diameter ratio is required, hence the collective name of long-rod penetrators. This has led to the use of fins to stabilise the projectile in flight once the sabot has been discarded.

At such high speed the penetration is dominated by material density and so the strength of materials can be largely ignored. However, while the penetration is inversely proportional to the square root of the armour density, the armour thickness is linearly proportional to the density. A low density material will therefore be more effective per kg of material at preventing complete penetration by long-rod penetrators.

Shaped charge

Shaped charge devices go by a variety of names, but broadly speaking a shaped charge is any device using an explosive lined cavity to deform a material into a fast moving projectile. The devices make use of two related effects; the Munroe effect and the Mises-Schardin effect. The Munroe effect, discovered in the 1880s by Charles Munroe at the U.S. Naval Torpedo Station at Newport³, describes the focussing of a blast by the shape of an explosive. Munroe famously demonstrated the effect by noting that the serial number written on the explosive was etched onto the surface of the target by the explosion. The Mises-Schardin effect states that when a sheet of explosive is detonated, the majority of the blast energy is dissipated normal to the larger surfaces.



Figure 2 – Armour Piercing Fin Stabilised Discarding Sabot (APFSDS) projectile in flight

A precision manufactured shaped charge, such as that found in a Rocket Propelled Grenade (RPG), consists of an explosive with a conical cavity lined with a ductile material. The explosion collapses the liner and focuses the material into a fine fast moving jet. Typical velocities are in the order of 8500m/s at the jet tip and 1500m/s at the jet rear³. This velocity difference stretches the jet, eventually destabilising it into a particulate stream. The relationship between penetration and armour density for this threat is largely the same as for a high K.E. round.

Not all shaped charges are made to the same precision as an RPG. Many of the Improvised Explosive Devices (IEDs) found in current operations make use of the shaped charge concept in a more basic way.

Blast

While the initial pressure wave from a blast is capable of causing rapid buckling of a vehicle's facets, blast can cause damage to a vehicle through a variety of other mechanisms. For a start, the blast itself is also generally accompanied by fragments and hot gases. If the integrity of the vehicle fails due to either fracture or fragment impact, the hot gasses cause significant damage to the interior of the vehicle. The impulse from the blast can result in the vehicle overturning, particularly in cases of partially buried explosives, where the blast will be focussed upwards.

A blast in very close proximity to the vehicle will form a compressive shock wave travelling through the vehicle hull material. Due to the lower acoustic impedance of the air at the back face the wave is reflected to form a tensile wave. The tensile stress can cause a scab of material to be thrown from the back face, producing a secondary projectile within the vehicle without actual

penetration. This effect is seen in projectile impacts as well as blast.

By far the most effective method of protecting against blast is to increase the stand-off distance from the point of detonation. This is seen in vehicles purchased for recent conflicts, which raise the floor of the vehicle off the ground as far as possible. The recent vehicles also tend to incorporate a V shaped hull to direct the blast away from the vehicle. In addition there is great potential for the use of false floors to prevent damage to the crew through buckling of the belly plate.

Armour types

Metal

Rolled Homogenous Armour (RHA) is the base-line steel armour for land vehicle. As the name suggests the material has uniform properties through its section. While armour materials are often capable of superior protection levels, their effectiveness is still judged against RHA. The obvious advantage of RHA is that it is a relatively cheap and well understood material. As a steel it also has the advantage of experiencing work hardening around the site of an impact. This can result in a multi-hit performance better than other materials which outperform it in single hit scenarios. Aluminium is another commonly used monolithic metallic armour. Against most K.E. threats it has a better performance level than RHA, however, due to its lower density a greater thickness is required, increasing the size of the armoured vehicle. The use of aluminium armour has been held back by instances of stress corrosion cracking until improvements in the metallurgy and welding were developed.

Aluminium can also be more susceptible to plugging, due to its lower heat capacity per unit volume than steel. Titanium is also occasionally used as an armour material, but its cost has prohibited its widespread use.

All of the metal armour solutions so far are based on an energy absorption approach. In practice absorbers are often coupled with a disruptive outer layer. This aims to disperse the energy of the projectile either by disrupting its motion so that it impacts with a greater surface area, or deforming the projectile.

The most commonly used metallic disruptor/absorber system uses a sheet of High-Hardness Steel (HHS) perforated at regular intervals, backed after an air gap by a block of aluminium. The perforations in the steel add extra edge effects and encourage the projectile to tumble. The high hardness of the disruptor deforms or shatters the projectile, reducing its K.E. density. This armour system is aimed at addressing the competing material property demands described in the previous section.

Ceramic

Ceramics are known to make effective disruptor materials due to their high hardness and low density. Their brittle nature means that they require a backing absorber material, with a composite often used to maintain the low mass of the system. The fragmentation of the ceramic due to an impact can significantly reduce its ability to withstand multiple hits. This has led to the design of smaller ceramic segments, in order to limit the damage to the surrounding material. Unfortunately ceramic solutions tend to be expensive, in the region of 2-12 times the cost of an equivalent protection using RHA, depending on the grade of ceramic⁴.



Figure 3 – Anti-tank weapon showing shaped charge in nose



Figure 4 – Warrior armoured vehicle fitted with a mixture of bar armour and ERA

Bar armour

Bar armour consists of a cage like set of bars with defined spacing, surrounding the vehicle at an offset. The armour disrupts the correct functioning of the rocket propelled grenade (RPG) rather than simply putting more material in its path. This has the advantage of being an extremely lightweight and cheap method of protecting against a highly penetrative threat, whilst being highly resilient to multiple strikes.

Unfortunately the bar armour requires a relatively large offset from the vehicle hull. This greatly increases both the space envelope and visibility of the vehicle. Additionally, there are RPGs in existence specifically designed to bypass the bar armour's effectiveness.

Explosive Reactive Armour (ERA)

As a concept ERA is a counterintuitive method of protecting a vehicle. The armour consists of two metal plates sandwiching a layer of explosive, with the whole cartridge positioned at an angle to the likely direction of impact. When a jet of copper from an RPG impacts the cartridge, the explosive is detonated driving the plates apart. The plates' movements at an angle to the jet cause the jet to cut a slot where before it would only have had to make a small hole. This increases the effective thickness of material through which the jet must pass. The opposite directions of travel of the plates also cause the jet to destabilise, greatly reducing its penetration capability.

Against a standard RPG an ERA cartridge can perform as well as a steel plate many times heavier. However, ERA is still too heavy for many light or medium weight vehicles to use it extensively.

The threat defeat mechanism described above does mean that the ERA performance is very dependent on the angle of incidence and the position of impact on the cartridge. A jet impacting the cartridge at normal will find the ERA little more effective than an equivalent mass of RHA. This is due to the plates' directions of motion now being in line with the jet, no longer causing a slot to be cut or having a destabilising effect.

To counter ERA, RPGs have been developed which incorporate two shaped charges. The device is carefully timed so that the first charge detonates the ERA, giving the second charge a clear path to the vehicle hull. Fortunately these threats are currently not particularly prevalent.

Armour concepts

Whittaker weighting

It is unsurprising that as the threats to armoured vehicles have evolved, the amount of armour a vehicle is expected to carry has increased dramatically. Unfortunately, there is a limit to the weight a vehicle can carry before its functionality is impaired. This is illustrated well by the German World War 2 effort

to build an impenetrable tank.

The Panzerkampfwagen VIII Maus weighed approximately 180 tonnes, but unfortunately the power to weight ratio was such that it could only reach a top speed of 13 km/h. As it has been shown impractical to fully protect a vehicle, an effective distribution of armour must be sought. This has led to the use of Directional Probability Variation (DPV) to determine the most likely directions of attack. The most famous of the DPVs is that developed by Lt Col. J M Whittaker in 1943⁴, which assumed a vehicle travelling towards a line of enemy vehicles at a constant velocity. Whittaker's assessment showed the majority of shots would impact on the front segments of the vehicle, up to an angle of approximately 60° from normal. Such an approach is useful in conventional warfare, where the direction of the enemy is fairly certain. However, its applicability in more modern scenarios is less certain.

Areal density

Areal density, defined as the mass of material per square metre, is the term generally used to describe an amount of armour. This is found to be more appropriate than using the overall mass, as it allows the protection levels of different sized facets to be compared, while giving a greater appreciation of the mass than using thicknesses.

Mass efficiency

Armour effectiveness is most commonly considered using the general term of 'mass efficiency'. This is defined as the ratio of the areal density of RHA (A_{RHA}) required to stop a particular threat, to the areal density of another armour material (A_X) required to stop the same threat.

$$em = \frac{A_{RHA}}{A_X}$$

While it is convenient to use RHA as a baseline, the performance of RHA and other armour materials depends on an assortment of variables including; angle of incidence, material thickness and threat type.

Taking angle of incidence as an example; for low K.E threats when a critical angle is reached, the geometry of the round will play a large part in the impact. As the tip of the projectile

contacts the surface a moment will be imparted causing the projectile to tumble and so impact with a larger surface area. Obviously this leads to a reduced penetration level but by different amounts according to the armour material and threat geometry.

Armour performance is also known to vary with the thickness of material, so that penetration through a number of layers of material is not equal to the penetration which would be achieved into a semi-infinite block. Put another way, a bullet with 15mm RHA penetration capability might take 4 closely spaced layers of 5mm RHA to stop it. Even taking this into account, the change in mass efficiency may still be hard to predict; when a thin plate is hit by a low K.E. projectile at a high angle of incidence, the projectile would not be expected to take a linear path through the material. At some point the projectile would pass too close to the back face for it to be contained. At this point the projectile would rotate towards the back face and burst through.

As discussed previously, there are a wide variety of penetration mechanisms which vary with the threat and so the mass efficiency of a material is particularly threat specific. Low K.E. projectiles can be expected to ricochet or gouge a scoop of material when impacting metal at a steep angle. However, a copper jet from an RPG is travelling at such speed when it hits the metal, that its impact has been likened to a hot knitting needle passing through butter. Clearly the angle is of no consequence in the latter case.

Modelling

When attempting to predict the penetration of a threat, the approaches generally fall into the categories of numerical simulation or empirical/analytical assessments.

Numerical simulation is undertaken using a dynamic analysis software package such as Autodyne. While able to give a great understanding of the impact mechanism and provide accurate results, the process is computationally intensive and extremely time consuming both to set-up and run. Such an approach would take far too long to assess the number of combinations of armour, threat and impact angle

required by the Survivability team.

As described previously, the specific projectile impact mechanism is dependent on a great many variables. However, empirical relationships, being derived from test results, are specific to particular impact types. As such they will only function over a limited range of variables. For instance Recht produced the separate equations for the residual velocity of blast fragments and sharp penetrators⁵. For the Survivability team's purposes, many different empirical relationships would have to be used according to the projectile type, impact angle, material type, and so on. This would require the precise detail of the impact to be understood before a new projectile type or armour material could be modelled. In this way empirical relationships were considered too case specific for the team's purposes. Recht's equations are by no means the only relationships produced. There are a great many conflicting empirical equations, each with benefits and disadvantages.

As with the empirical relationships there are a great many case specific competing analytical methods for assessing impacts. Analytical methods are generally used to understand the specifics of the impact mechanism at work and identify the important variables when numerical simulation is unavailable. For the purposes of assessing the weight or performance of concept armours on a vehicle, standard empirical and analytical methods were considered too case specific to be of use.

Weight stack

The solution arrived at by the Survivability team was to base the analysis of armour solely on the experimental results determined by the Defence Science & Technology Laboratory (DSTL). The team developed a 'weight stack' spreadsheet tool to make use of DSTL's extensive knowledge of mass efficiencies for armour materials against the required threats; effectively a fundamental form of empirical relationship.

Division of problem

The problem of developing the weight stack tool was divided into the three main areas of:

- (1) Defining the geometry of a complex vehicle shape
- (2) Storing and accessing the mass efficiency information
- (3) Combining the geometry and mass efficiency data to determine the mass of a selected armour system.

Geometry

To define the vehicle geometry the user of the tool starts by entering the 3D co-ordinates of the corner points of the vehicle in an alphabetical list. The geometry of the vehicle is then defined in a facet by a facet basis, by selecting the pre-defined coordinates of the corner points from drop-down lists. Each facet is defined by its corner points, assuming that straight lines connect the points and that the facet is flat, as tends to be the case with military vehicles. The normal vector of the facet is determined by taking the cross product of two vectors formed from the first three co-ordinates of the facet. From the angle of this normal vector, the elevation and azimuth of the vehicle facet is then determined. By applying a rotation matrix to the facet coordinates; the facet is rotated so that it is only in one plane. Now a 2D problem, Green's theorem, which relates a closed area to the path surrounding it, is applied to determine the area of the shape.

$$Area = \frac{1}{2} \sum (x\Delta y - y\Delta x)$$

This method does not limit the number of sides which a facet can have, allowing complex shapes to be entered and analysed using their corner point coordinates. While this method determines the area, azimuth and elevation of a facet it does not indicate which side of the facet faces the inside or outside of the vehicle. In other words the method does not distinguish between the roof and floor of a vehicle. This poses a problem when threats are defined as only impacting over specific angles, as is the case with mine blast, which should not be considered as impacting the roof. To address this, the tool estimates

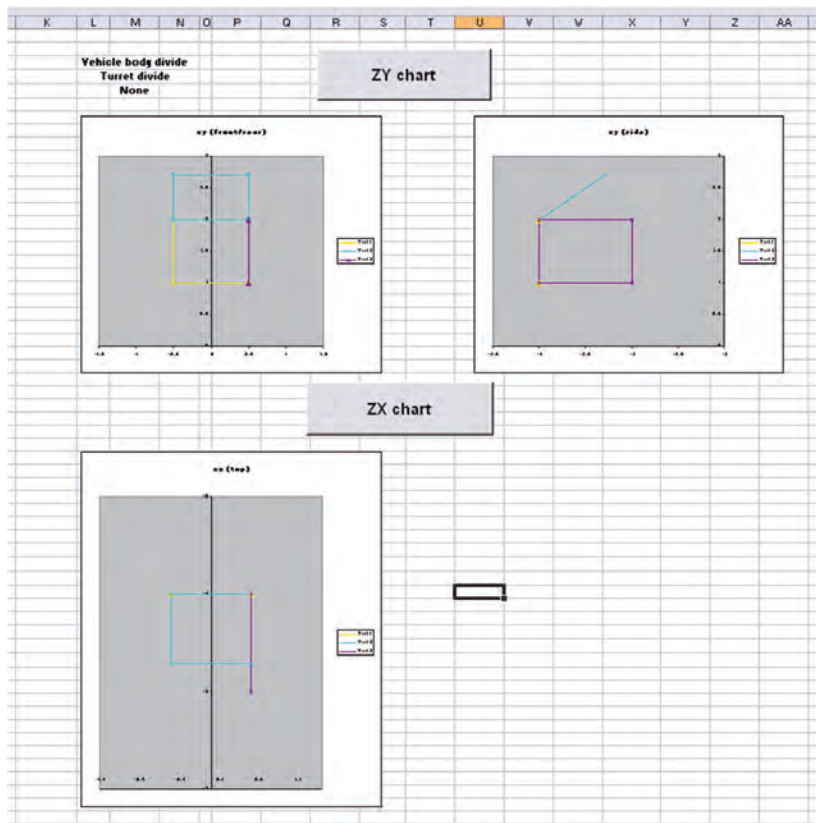


Figure 5 – Complex vehicle geometry is built up from simple shapes

the centre point of the vehicle as the midpoint of the maximum and minimum coordinates in each plane. The position of the facet relative to the centre point then defines which side faces outwards. This works adequately as long as the vehicle is a standard shape. A manual override function is included for particularly unusual vehicle shapes to be analysed. The importance of the Whittaker weighting system was discussed earlier. In assessing concept armour systems it is desirable to be able to move the line defining the frontal arc division of the vehicle. For example, the question might be posed as to how much weight is saved by moving the crew section forward. A feature was included to allow the user to define a moveable front/rear dividing line, rather than force the user to specify two sets of facets for front and rear. The geometry sheet automatically forms separate front and rear facets from each facet intersected by the dividing line. This is done simply by linear interpolation in three dimensions, allowing the front and rear of a facet to be armoured independently.

DSTL provided mass efficiency values dependent on threat type, armour type, impact angle and material thickness based on experimental results. The information is stored in the tool as a series of lookup tables displaying mass efficiency values varying with angle of attack and thickness of material. Each combination of threat and material type has an individual lookup table. Bilinear interpolation is used to estimate the mass efficiency at a particular angle and thickness. While taking a relatively large space to store the information, this method offers a great deal of flexibility. New materials can be included with rough estimates of effectiveness and as more information becomes available through testing the mass efficiency data can be updated accordingly.

As the mass efficiency values are taken from experimental results they automatically take account of a variety of physical effects whether or not the precise mechanism is fully understood. For instance a bullet impacting steel at a steep angle may be subject to ricochet, may break through or may deform enough to cause plugging. It need not be understood which the dominant mechanism is, as the mass efficiency values are experimentally derived and so take account of all factors.

Applying information

The first step in the process of determining the mass of an armour system obviously requires the user to specify the armour material. The vehicle geometry data is imported into the calculation sheet and the user selects a single armour material and a threat to be applied for each facet. With the threat selected, the calculation sheet imports the associated angles of incidence and RHA penetration levels from the threat data sheet. The worst case angle of impact is determined as the angle closest to normal at which the threat is capable of impacting a facet. The worst case impact elevation (χ) is therefore defined as:

Where:

$$\chi = \begin{cases} 0, & \theta \geq \alpha \\ \theta - \alpha, & \theta < \alpha \end{cases}$$

$\alpha = \text{facet elevation}$

$\theta = \text{Largest elevation of threat incidence}$

With the worst case impact azimuth (β) similarly defined, the vector of the impact at the worst case elevation and azimuth is determined by trigonometry as:

$$V = \left[\frac{1}{\tan\left(\frac{\pi}{2} - \chi\right)}, \frac{1}{\tan\left(\frac{\pi}{2} - \beta\right)}, 1 \right]$$

From this the angle from normal of the worst case impact, combining elevation and azimuth, is determined as:

$$\psi = \cos^{-1}\left(\frac{1}{|V|}\right)$$

This impact angle is used to determine the thickness of RHA (t_{RHA}) required to stop a threat with penetration P(mm of RHA) using the cosine rule.

$$t_{RHA} = P \cos(\psi)$$

In practice, the cosine rule does not apply exactly to calculating penetration at an angle, due to the effects, such as ricochet, mentioned previously. However, the mass efficiency values are set relative to the performance of RHA at normal and so these effects are inherently accounted for. The calculated thickness of RHA and the impact angle are then cross referenced with the material type and threat type to determine an effective mass efficiency value from the stored mass efficiency data. The thickness of RHA is converted into an areal density of RHA by multiplying by the density. Finally the areal density of armour required to stop the projectile impacting a facet is determined by dividing the Areal density of RHA by the determined mass efficiency. An overall value of armour mass can be determined by multiplying the areal density by the facet area.

$$\text{Mass} = \frac{(t_{RHA} \times \rho_{RHA})}{em} \text{Area}$$

Once this calculation has been performed for each facet subject to the defined threat, a new threat is selected by the user and the calculations repeated. The tool compares the mass of material required for each threat on a facet by facet basis, showing the increase in armour mass required for protection against increasing levels of threat severity and giving a cumulative armour mass.

However, the description above is only covering the most basic case of single module, single material solutions. In this case it would not be possible to alter the level of protection offered to the vehicle unless the entire armour pack was removed from the facet and replaced with a new one. For a modular approach where the protection level can be altered, the armour must consist of a stacked set of layers of material. The tool therefore allows the user to define at which threat levels there are divisions between armour layers. To enable this, the contribution of the lower armour layers to reducing the penetration of a more severe threat needs to be assessed. Unfortunately the existing thickness of the lower level cannot simply be subtracted from the total thickness of armour required to stop the more severe threat, as the mass efficiency is known to vary with layer thickness.

The mass efficiency of each pre-existing layer must therefore be recalculated for subsequent threats. Once the penetration reduction due to the pre-existing layers is subtracted from the penetration of the threat, the mass of armour for the new layer can be calculated as before. An adjustment factor is added to the penetration capability of the threat according to the number of extra layers in place. This accounts for uncertainty in the residual penetration capability, once a threat has passed through a layer. The thickness (t_x) of the new layer of armour is therefore found by:

$$t_x = \frac{(t_{RHA} - (\sum_{i=1}^n \frac{t_i \rho_i em_i}{\rho_{RHA}}) + R) \rho_{RHA}}{em_x \rho_x}$$

Where:

t_i = thickness of pre-existing armour layer i

ρ_i = density of pre-existing armour layer i

em_i = mass efficiency of pre-existing armour layer i

R = Residual penetration correction factor

The capabilities required in order to assess modular armour, reassessing mass efficiency for each layer against each threat, also provide the capability to use layers of different material types. This is a useful function when defining 'top-up' armour, where a different material which is more efficient against a particular threat is placed on top of the standard armour.

Despite the flexibility of this model, some complex forms of armour do not lend themselves to the mass efficiency approach. This includes bar armour, which due to its very threat specific nature provides no protection against other threat types. ERA also poses a problem due to its varying angles and thicknesses. Fortunately these armours are generally only available in standard forms, where their capabilities are well understood. They are therefore applied as Areal densities without assessing their effectiveness and so the tool requires a certain amount of armour experience from the user. By using a spreadsheet as a basis, the tool can be easily modified or expanded and has an inbuilt capacity for growth, as the armour material and threat data sheets can be either added to or updated with more recent experimental results.

Conclusion

The weight stack tool described in this paper provides the Survivability team with a rapid method of assessing the mass of complex concept armour systems and evaluating survivability requirements, incorporating a variety of impact mechanisms through the use of experimental data and a layer function to compare modular armour to monolithic systems.

The tool has been used to explore how the FRES protection requirements could be achieved and what the impact of the necessary weight of armour would have on the vehicles mobility. It has been concluded that it is impractical to provide an armour solution that could defeat all the identified threats simultaneously. This has subsequently driven a great deal of analysis to see how the threat distribution changes with the type of operation. In turn it has resulted in "mission tailored" protection requirements which reflect the changing balance of threats and better represent the user's needs. Recently the tool has been used in conjunction with armour cost information to investigate options for "value for money" protection solutions within the constraints of the vehicle.

References

1. Energy and momentum changes during ballistic perforation. Hetherington, J. G. 3, s.l. : International Journal of Impact Engineering, 1996, Vol. 18.
2. Penetration: Mechanical aspects and mathematical modelling (review). Aptukov, V. N. 2, s.l. : Strength of Materials, 1990, Vol. 22.
3. Payne, C. M. Principles of Naval Weapon Systems. s.l. : Naval Institute Press, 2006 .
4. Hazell, P. J. Ceramic Armour: Design, and Defeat Mechanisms. s.l. : Argos Press Pty, Limited, 2006.
5. The effect of projectile nose shape upon ballistic limit velocity, residual velocity and ricochet obliquity. Recht, R. F. s.l. : Weapons Development Department, Naval Weapons Center, China Lake, California, 1973.



Chris Marriott

Senior Safety
Reliability Engineer
Defence

Abstract

A history of the development of explosives, and related accidents, provides a background to the practical application of safety assessment techniques for projects involving explosives systems. The analysis techniques include 'fault tree' and 'event tree' analysis, which have been used qualitatively to help develop 'lines of defence' arguments. These have been subsequently developed using a graphical technique termed 'bow tie' analysis. The development of 'safety functional requirements' and 'safety integrity levels' is also described for equipment which will be used in explosives processing. Finally, the difficulty in determining the consequences of an insult to explosives is discussed, together with the implications for 'quantitative' risk assessments. This paper aims to demystify the terms used above by demonstrating how they have been used in practical applications.

The opinions herein are those of the author and not necessarily those of Atkins Limited.

Introduction

If a material is said to be 'explosive' then it can be considered to be chemically or otherwise energetically unstable and liable to release energy in a sudden and often violent manner. This is likely to be accompanied with the generation of high temperature, large changes of pressure, a flash and loud noise. The level of such an energy release can be quite staggering: In 1917 the City of Halifax, in Nova Scotia, was devastated when a French cargo ship exploded in its harbour. Over 1600 people were killed instantly and all buildings and structures within a vast area were devastated¹.

Safety, on the other hand, is typically defined as "the condition of being safe; freedom from danger, risk or injury". The term 'explosives safety' therefore fits quite neatly into the definition of an oxymoron, i.e. "a figure of speech in which incongruous or contradictory terms are combined". However, the need for safety when handling explosives is far from contradictory and it is rare for the word 'explosive' to appear without some reference to the word 'safety' in close proximity. This paper provides a brief history of the development of explosives and their attendant lessons in safety, and then goes on to show how modern safety cases are developed for processes and equipment which, by their very nature are designed to be explosive.

A history of explosives and related accidents

The history of explosives is closely linked with tales of tragedy, either from the misery it can cause in its intentional application or from accidents during its development, manufacture, transportation and storage. Ironically, the discovery of the ingredients for the earliest known forms of explosive originated from a Chinese attempt to find an elixir of life! In 850 AD, a Chinese book warns of a particular elixir: "Some have heated together sulphur, realgar and saltpetre with honey; smoke & flames result, so that their hands and faces have been burnt, and even the whole house where they were working burnt down"². The ingredients of this potent mixture include two of the three ingredients of what is now known as 'gunpowder', namely sulphur and saltpetre, with the third ingredient being charcoal. The earliest known weapons employing this mixture were probably Chinese gunpowder fire arrows reported in 969 AD³, but it took a couple of centuries before knowledge of gunpowder was transferred to Europe with a description by Franciscan monk Roger Bacon appearing in 1242.

The first recorded use of gunpowder by English soldiers was at the battle of Crecy in 1346. However, even by the 15th Century, early weapons could cause as much harm to the user as the intended target, as James II of Scotland discovered when he was blown up by his own cannon whilst besieging Roxburgh castle in 1460⁴.

The manufacture of such a lethal substance was also a risky business. The three components had to be ground together intimately to be effective and at first this was done with a mortar and hand-operated pestle. Later, 'pestle mills' were used, the mixture being pounded mechanically in wooden mortars by wooden stamps. Although the use of wood avoided the generation of sparks, to which the powder is very sensitive, this was still a hazardous operation because the impact and friction from the pestle often caused the gunpowder to explode. In Britain, stamp mills started to be replaced by 'incorporating' (or rolling) mills in around 1740. An incorporating mill consisted of two heavy wheels which ran over a bed of partially mixed gunpowder and ground it together for a period of up to 8 hours, without subjecting it to any severe shocks⁵.

The requirement for such relatively large scale, 'industrial' processes, was driven not only by military requirements but also more peaceful applications in the field of civil engineering.

The earliest record of the use of gunpowder for blasting occurred in 1627 and it was being used in the Derbyshire copper mines in 1638⁶. Its first use in large scale engineering was the construction of the French Languedoc Canal in 1681 and over the next couple of centuries many of the canals, road and railways were all constructed with the aid of explosives. This was not without the attendant risks associated with explosives: by the time Isambard Brunel had completed the "monstrous and extraordinary, most dangerous and impracticable tunnel at Box"⁷ it had cost over a hundred lives. Even more perilously, explosives started to be used for underground blasting in around 1820. This was short-lived however, due to numerous catastrophes, and a Royal Commission on Accidents in Mines recommended that from 1887 it be replaced by specially blended and tested explosives.

The pressure for special types of gunpowder to meet the ever increasing demands of miners, civil engineers and the military led in the 19th century to replacements for gunpowder as an explosive. Christian Schönbein was credited with the invention of 'guncotton' (made from nitrocellulose and cotton wool) in 1846. It was more powerful than gun powder and was one of the first of the so called 'smokeless powders'. Also, in the words of Schönbein, "The manufacture is not attended with the least danger..."⁸. However, in 1847 the first ever guncotton factory blew up and 21 people were killed, possibly due to overheating in the drying process. An even more powerful explosive, nitroglycerine (NG), was first made by Ascanio Sobrero in 1847 by treating glycerine with sulphuric and nitric acids. He discovered its physiological effects when he tasted a sample and it gave him palpitations and weakness in his limbs, he discovered its explosive effects when a small sample exploded glass fragments into his hands and face. Its exploitation as an important explosive only arose when Alfred Nobel found a way to bring it under control. In

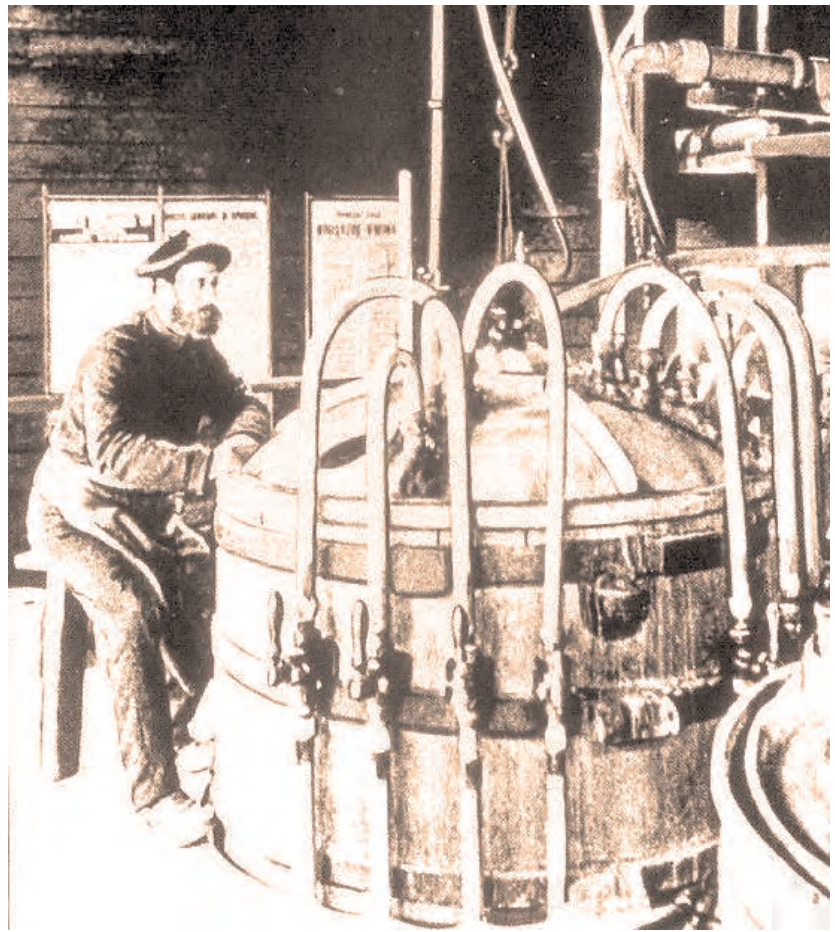


Figure 1: Nitroglycerine manufacture – the operator sat on a one-legged stool so that he would not fall asleep and let the mixture overheat.

1867 he patented the idea of taming liquid NG by absorbing it into a solid (diatomaceous earth or 'Keiselguhr') to make a dough. He named the substance 'dynamite', from the Greek word for 'power'. Production of the NG itself remained a hazardous operation, as shown by the operator's seating arrangements in Figure 1⁹.

Whilst dynamite was a major step forward, it suffered from the dangerous problem of 'sweating', i.e. the leaking of unstable nitroglycerine from the solid matrix. To resolve this, Nobel invented 'gelignite' (or 'blasting gelatin') in 1875. Gelignite is an explosive material consisting of 'collodion-cotton' (a type of nitrocellulose or gun cotton) dissolved in nitroglycerine and mixed with wood pulp and sodium or potassium nitrate. Its composition makes it easily moldable, and safe to handle without protection, as long as it is not near anything capable of detonating it. One of the cheapest explosives, it burns slowly and cannot explode without a detonator, so it can be stored safely.

The next major development in what were to be termed 'high explosives' was the introduction of 'Lyddite' (based on picric acid) as a filling for explosive shells in the British Army, first used at Omdurman in 1898. It was used up until 1916, despite its relatively poor performance and the many accidents associated with its manufacture, including: 16 deaths at Woolwich Arsenal in 1903 following the forcible insertion of a former in 9.2" shells; 10 deaths in Heckmondwike due to a foreign body in the grinding mill; and 38 deaths in 1916 at Low Moor munitions factory, when a fire in drum of picric acid spread to a packing building¹⁰. These deficiencies accelerated the introduction of the eponymous 'TNT' (trinitrotoluene), a substance considered so stable during its manufacture and storage that it was exempted from the 1875 Explosives Act⁶. TNT was, however, very poisonous and led to the deaths (from jaundice) of many female workers in the munitions factories of the Great War.

Between the wars, more powerful explosives were developed such as RDX, which was first used in the 1890's as a medicine by Hans Henning. When combined with TNT it provided a stable, relatively 'insensitive' munition. When combined with jellies, waxes and plasticizers the RDX forms what are known as 'plastic explosives', which are malleable, very stable and store well (for example, the notorious Czech material known as 'Semtex'). The most 'powerful' explosive currently in large scale use is HMX, which was discovered in the 1930's as a by-product from the manufacture of RDX. HMX is a common ingredient in Plastic Bonded Explosives (PBXs), which are powdered explosives to which plastic binders have been added. The manufacturing process for a typical PBX explosive in a modern weapon system requires the powder to be oven dried at temperatures in excess of 100°C (the powders are wetted for transport and storage to prevent explosive dusts clouds from forming), then isostatically pressed (i.e. subject to equal pressure from every side) at pressures up to 140MPa¹¹, before the final shape is machined using standard cutting, milling and drilling machines. Thus the explosive is exposed to every type of hazard and is expected to survive with the least sign of reaction and yet will detonate reliably on demand.

More than 8000 people are reported to have been killed in the 20th Century due to the result of explosives responding to unplanned stimuli such as heat or shock. As a result of this, and recent advances in explosives chemistry, there is an ever increasing drive for the employment of 'Insensitive Munitions', defined as: "Those munitions which reliably fulfil their performance, readiness and operational requirements on demand but will minimise the violence of reaction and subsequent collateral damage when subjected to unplanned stimuli"¹². It is to be hoped that this particular directive is pursued with due vigour.

Explosives - definition of terms

When explosive materials explode they produce a large volume of hot gases and a consequent large increase in pressure. However, the effects of a 'low explosive', such as gunpowder compared to a 'high explosive' such as TNT are on different scales: Low explosive can produce pressures up to 600MPa in milliseconds; a high explosive can produce nearly 50 times the pressure in a thousandth of the time (i.e. at supersonic speeds). Technically, a low explosive is said to 'deflagrate', or burn rapidly, whereas a high explosive is said to 'detonate' (from the Latin 'to thunder'). This rapid shattering effect, for which the term 'brisance' is employed, makes high explosives useful in mines, bombs and shells, whereas low explosives are more suitable where a steady thrust, or 'heave', is required as in gun propellants or mine blasting. Explosives are also commonly divided into 'primary' and 'secondary' types, where 'primary' refers to a detonating explosive which is extremely sensitive to stimuli (e.g. heat, friction, shock), and 'secondary', which refers to a detonating explosive with relatively low sensitivity. Generally, an 'explosive train' would consist of a primary explosive as initiator (i.e. a 'detonator'), a 'booster' and a secondary explosive as the main charge. The 'booster' translates the weak detonating wave from the initiator and amplifies it to initiate the main charge. This is important for modern munitions, where the main charge is designed to be as insensitive as possible to increase safety in manufacturing, transport and use.

Low explosives are generally referred to as 'pyrotechnics' or 'propellants.' The former is derived from two Greek words: 'fire' and 'art' and fireworks are a popular example. A pyrotechnic produces a self-sustaining combustion reaction which may or may not be accompanied by a visible flame and the production of gas. Compositions are mixtures of a fuel (e.g. carbon) and an oxidising agent (e.g. potassium nitrate). Propellants can be pure materials (e.g. nitrocellulose) or mixtures of chemicals designed to burn at a controlled rate and provide a predefined thrust to the system containing them.

Low explosives, such as those used in fireworks, can also detonate under certain circumstances, sometimes in tragic circumstances: In May 2000, a fireworks factory in Enschede exploded, destroying an entire neighbourhood and killing 23 people. This was considered to be due to the self-confining effect of a large mass of material. In the same month, an accident at a public fireworks display in Australia led to the death of a young girl when a roman candle detonated in the steel tube in which it was standing, again due to the effects of confinement. This led to the introduction of the Queensland Code of Practice in 2002, which had specific aims to minimise risks at all public firework displays.

Explosives regulation

In Britain, the first Act of Parliament to deal with public safety in relation to gunpowder was that of 1719, which regulated the carriage of gunpowder in London. This was followed in 1772 by the Gunpowder Act, which covered manufacturing processes (prohibiting pestle mills and limiting the quantities of powder)⁵. In the mid 1800's, amidst the discovery of new explosive chemicals, almost anyone could make and sell explosives and this inevitably meant that there were increasing numbers of accidents; a particularly severe explosion occurred at a Birmingham factory in 1875, killing 52 people. This incident, amongst the many others, led to the passing of the Explosive Act in 1875 (extended in 1923) and established the Explosives Inspectorate. From 1847 to 1974, every explosives related incident was published in the Annual Report of Her Majesty's Inspectorate of Explosives. The Health and Safety Executive maintain a database (EIDAS -Explosives Incidents Database Advisory Service) which includes these records, as well as those reported to HSE since 1974 or drawn from the public domain. There are currently nearly 9500 incidents recorded on the database¹⁰. Figure 2 shows particular information drawn from this database, relating to major explosives incidents in which 5 or more people died. Note the series of catastrophic events in the Great War period due to explosions in factories producing NG, TNT or picric acid, and the reduction in such incidents in modern times.

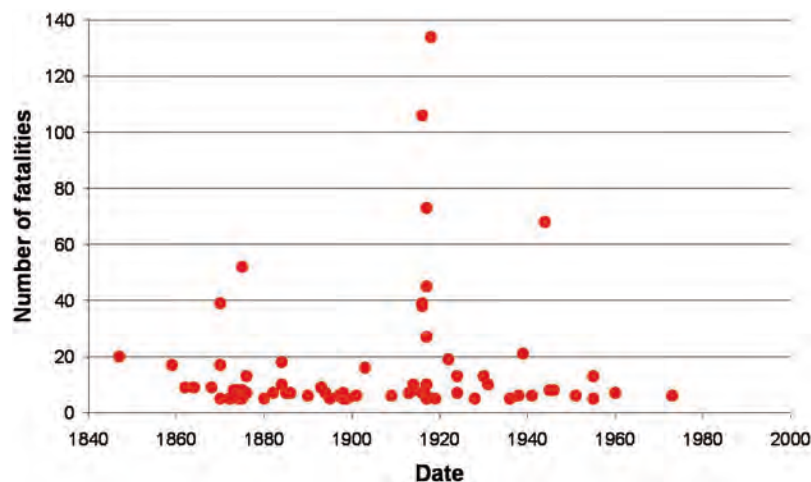


Figure 2 - Explosives related incidences with >5 fatalities.

Explosives - safety testing

The database contains such tragedies as the incident in 1909 where attempts to demolish a wreck by explosives means led to a tremendous explosion and caused 6 deaths. It later transpired that the wreck was that of a ketch loaded with gelignite! The report by the inspector suggested that "...before commencing the work of destroying a wreck by means of explosive it is desirable to ascertain, so far as may be, the nature of the cargo."¹³

Legislation such as the Gunpowder Act & Explosives Act have been replaced by the Manufacture and Storage of Explosives Regulations (MSER) 2005 and, for UK military applications, by the Joint Service Publications (JSPs). A key requirement of MSER is that premises must be licenced and that separation distances between explosive sites (also known as 'Quantity Distances') must be maintained, according to the 'hazard type' (ranging from hazard type 1 for mass explosive hazard to hazard type 4 for fire or slight explosion). This aligns with the UN recommended system for the transportation of dangerous goods, i.e. class 1.1 is equivalent to hazard type 1. Other key regulations include the Control of Major Accident Hazards regulations (COMAH) 1999, made under the Health and Safety at Work Act (HWSA) 1974, which aims to prevent major accidents involving dangerous substances and limit the consequences should they occur. The military equivalent is the Major Accident Control Regulations (MACR).

In 1887 the Coal Mines Act established a series of tests for explosives to be used in coal mines. These tests included the 'Trauzl test' (later the ballistic pendulum) to measure the power of the explosive and a test to measure the explosive effect on typical coal mine gases. The Sensitiveness Collaboration Committee (SCC) was established at Waltham Abbey in 1962 to produce procedures for the safety tests applied to energetic materials, and this resulted in the SCC Manual of Tests. Ownership was transferred to the Defence Ordnance Safety Group (DOSG) in 2005 and the manual was re-issued under the Energetic Materials Testing and Assessment Policy Committee (EMTAP)¹⁴.

The manual provides a compilation of those sensitiveness and explosiveness tests that are used to evaluate the risks from military explosives during development, manufacture, handling, filling, transport, storage and use, and to provide definitive methods for the tests. Traditionally it has been assumed that explosives intended to fulfil a specific function possessed broadly similar safety characteristics, i.e. the designations 'propellant', 'high explosive', 'pyrotechnic', 'initiator' or 'booster' could imply various degrees of potential hazard. However, the search for ever improved performance has blurred the distinction, in hazard terms, between explosive materials. This is shown by the overlapping regions in the figure below, which shows typical regions in the sensitiveness/explosiveness plane occupied by the various types of explosives.

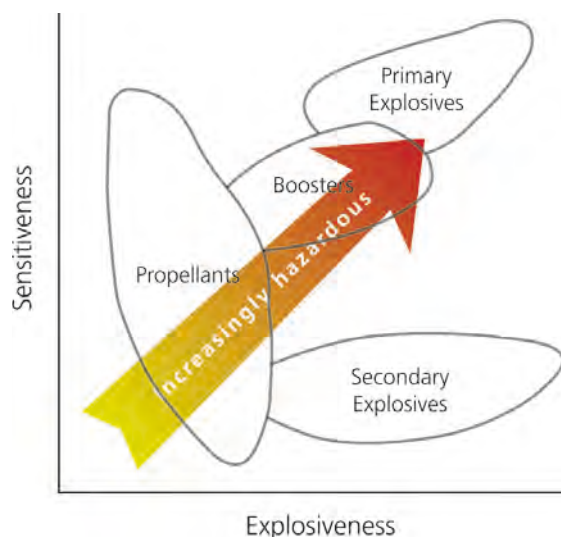


Figure 3 - Explosiveness and sensitiveness

Explosives risk assessment

It may appear that explosive safety is concerned with solely technical issues and is achieved by compliance with rules and regulations. Lessons from the many disasters involving explosives, however, suggest this is not the whole picture. Prescriptive regulations are accepted for routine situations and well-established practices; it is very difficult to write rules for unpredictable explosive materials and impracticable to be able to cover every eventuality. Thus, in the formation of a risk assessment, some basic questions need to be answered:

- (1) What can go wrong?
- (2) What is the likelihood & who can be harmed?
- (3) How can we reduce this?

The results of this analysis are then tested by assessing whether all practicable control measures have been taken (and that industry practice has been followed) and that further action is grossly disproportionate to any benefits gained from the risk reduction. This test is generally referred to in the UK as 'As Low As Reasonable Practicable' (ALARP).

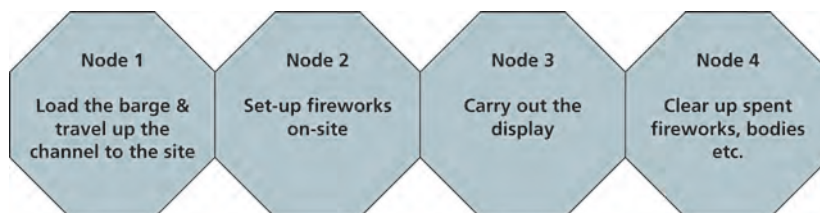


Figure 4 - Possible nodes for a firework display

What can go wrong?

The start point for any assessment of risks is the hazard identification process, or 'HAZID', which will also define responsibilities (duty of care) and interfaces with other safety assessments. There are many ways of conducting the identification process, including 'Structured What-If Techniques' (SWIFT) and 'Hazard and Operability Studies' (HAZOP), to name two of the more recognised methods. However, essentially they rely on the same principle of 'structured brainstorming' with an appropriate team of knowledgeable people. Regardless of the method, the schedule of hazards identified should be as comprehensive as possible. Structuring consists of laying out the 'nodes' or key break points in a process and working through a list of guidewords to prompt the team to identify any hazards associated with that node and guideword. This data is then recorded in a 'hazard log', which may be in a spreadsheet or database format.

Thus for a major public fireworks display on a barge in Bristol docks, the nodes may look something like that shown in Figure 4.

The associated guidewords may be generic (e.g. 'more of', 'less than', etc.) as would be used in a HAZOP, or more topic specific, as in a SWIFT analysis. A typical hazard log will consist of a unique hazard reference number, a description of the cause and consequence and a list of any existing safeguards. Safeguards relevant to fireworks include 'engineering safeguards' (e.g. separation distances, racks for mortars, and lightning arrestors) and 'procedural safeguards' (e.g. firework quality assurances and operator training). An example spreadsheet is shown in Figure 5.

Node 1: Barges Traveling to Display Site (Bristol Harbour)							
Hazard Ref.	Step / Equipment	What If	Cause	Hazard	Consequence	Existing Safeguards	
						Engineer	Procedural
N1-1	Fireworks product	Adverse Weather	Gust wind	Fireworks drop into sea	Fireworks soaked with seawater, potential risk for recovering the fireworks but no immediate threat	1. No metallic powder is allowed to be used in fireworks for display in Hong Kong	1. Fireworks display will be postponed or cancelled when wind speed is higher than 8m/s and the decision will be made before the barges leave.
N1-2	Barge	Adverse Weather	Fog	Barges crashed with other vessel causing engine fire	Fireworks exposed to heat and explode, affect nearby ships and ferries		1. The barges stop moving under low visibility condition 2. Marine department keeps updating the barges about the weather condition in
N1-3	Barge	External Fire	Engine Fire	Fireworks on deck exposed to heat	Fireworks exposed to heat and explode, affect nearby ships and ferries	1. Fire-fighting equipment such as hosereel and fire extinguishers are provided on board	1. Police vessels escort the barges and enforce a safety distance of 250m

Figure 5 - Example hazard log sheet

What is the likelihood & who can be harmed?

The next stage is to carry out risk estimation, which entails assessing the severity and frequency of the hazardous event. For a qualitative assessment, a risk matrix is a useful way of ranking and displaying the results. This is generally a grid with frequency down one side and severity along the top, ranked from high to low.

For a more quantitative risk assessment (termed 'QRA'), other techniques are applied to provide a more systematic analysis of the risks. These include: Failure Modes Effects and Consequences Analysis (FMECA), which is generally more suitable to the analysis of complicated component failures; Fault Tree Analysis (FTA) and Event Tree Analysis (ETA).

FTA is used to graphically display the means by which faults can lead to a failure of the system of interest (the 'top event'). Thus, in the development of an explosives manufacturing plant, a series of trees would be developed for each of the different operations (e.g. receipt, storage, drying, pressing, machining and inspection). A fault tree describing explosive pressing could look like Figure 6.

Taking fault 3.4 'over heating' further leads to the fault tree shown in Figure 7, where the bottom level events derive from the hazard log.

In this example, the gate with descriptor 'N3_F2' represents a 'Line of Defence' (LOD). To qualify as a LOD, robust justification is required to show that the safeguards (engineered features and/or managerial controls) provide adequate protection against the event under consideration from occurring¹⁵. These safeguards then become formal Safety Functional Requirements (SFRs). The SFRs identify specific requirements on systems and structures which make up the LOD, e.g. "The thermal trip shall isolate the heating medium should the heat exceed 102°C". There may be one or more LODs required to terminate the fault sequence arising from a hazard, depending on its severity and the level of protection afforded by the SFRs within the LOD.

If these systems are programmable, then a 'Safety Integrity Level' (SIL) may also need to be assigned. A SIL is a relative level of risk-reduction provided by a safety function. Four SIL levels are defined, with SIL 4 being the most dependable and SIL 1 the least. The international standard IEC 61508 defines SIL using requirements and IEC 61511 is an application specific adaptation for the process industry sector (including explosives manufacturing). This defines the 'Layer of Protection Analysis' (LOPA) methodology to be used in the assessments. This is a quantitative method (as opposed to a LOD, which is descriptive) which requires an understanding of the probability of failure of a system and, crucially, an understanding of the potential consequences in the event of failure.

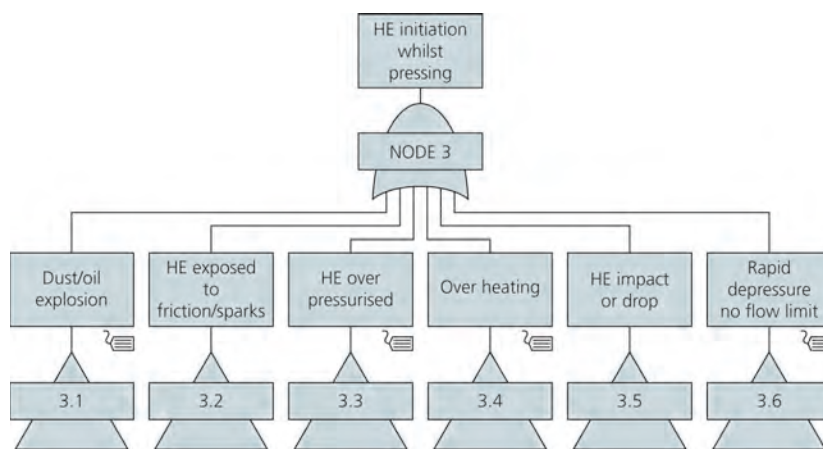


Figure 6 - Fault tree for explosive pressing

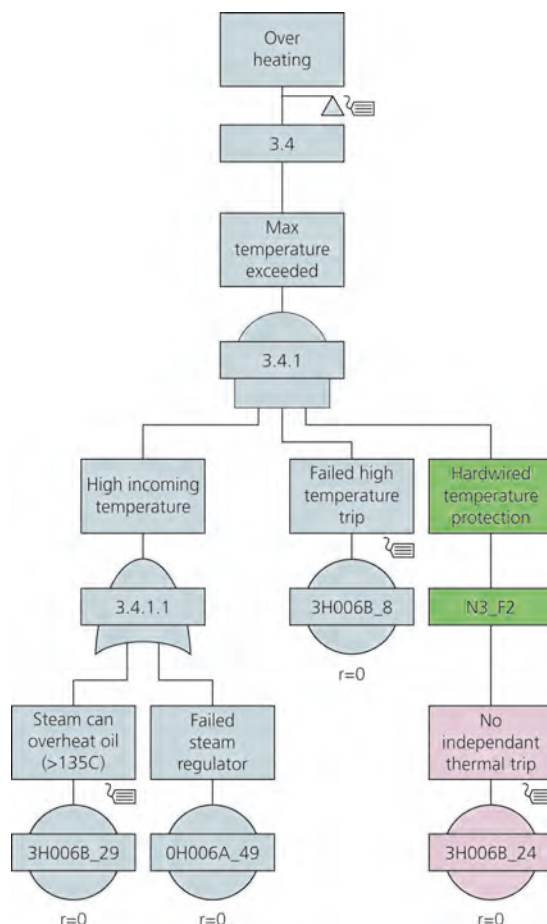


Figure 7 - Fault tree for overheating during explosive pressing



Explosive pressing is deemed to be sufficiently hazardous that it is required to be carried out remotely¹⁷, thus the consequences for other worker of an explosive event in the cell are related to the effectiveness of interlocks, as shown in Figure 8.

Finally, the results from FTA and ETA can be usefully displayed on a 'bow-tie' diagram, which is a representation of all the initiators and consequences of a particular scenario and the LODs in place to prevent or mitigate them¹⁶. Figure 9 considers the controls in place to prevent hazards identified in the press cell leading to fatalities.



Once the risks have been ranked and prioritised, they can be systematically assessed and risk reduction measures put in place, with the objective of meeting the previously set safety level criteria. Risk reduction is generally approached in a logical sequence, as shown in Figure 10.

Clearly a duty of care is owed to the public in the siting and management of any explosives facility and this is achieved through the enforcement of Quantity-Distance (Q-D) rules. These relate the quantity of explosives and the distance necessary to limit the severity of effects, in the event of an accidental ignition. Considering workers on-site or actually involved in explosives manufacturing, elimination of the hazard may only be achieved by a complete change to the process and may require some lateral thinking (e.g. remove the need for explosive machining by pressing directly to the final shape). Reducing the likelihood of an explosive event could be achieved by reducing process temperatures or pressures, but these are generally already at a minimum. Reducing the effect can be achieved by reducing quantities processed, but this will have an effect on throughput. Isolation can be achieved by removing the operator to a safe place, as is recommended for machining and pressing operations¹⁷. The design of controls (making up the LODs) will depend on an analysis of the consequences of ignition⁸ or initiation⁹. A clear understanding of the nature of the explosive hazard in terms of its likelihood and its potential for harm to people is required so that any engineering provisions are robust and can tolerate a wide range of faults.

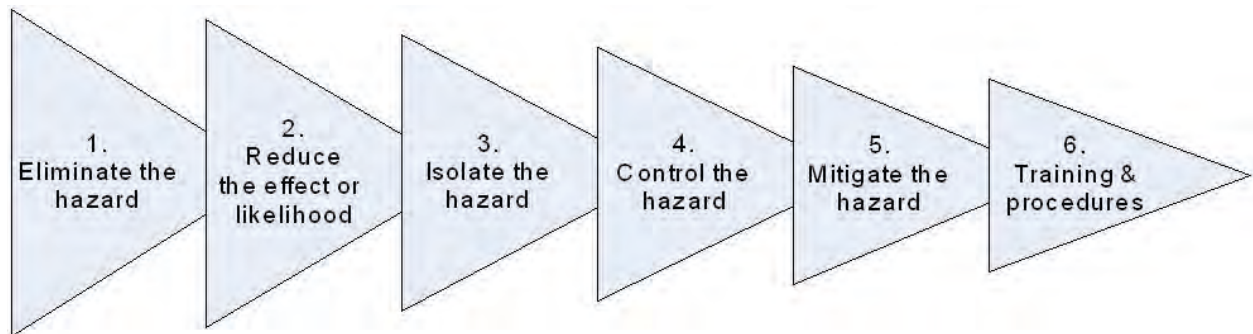


Figure 10 - Risk reduction hierarchy

Consequence analysis

By the very nature of explosives, the effects of accidental ignition or initiation are largely immediate and entail severe consequences. The magnitude is largely determined by the quantity of explosives material involved, while the type of effect is largely determined by the hazard type. A quantitative assessment of the probabilities of reactions of explosive to various stimuli is very difficult to achieve. Usually there are insufficient test results (for the particular stimuli/ explosive combination) to give statistically valid estimates, due to the cost of testing.

However, an understanding of the consequences of the response to an insult is necessary so that proportionate effort is applied in the risk assessment. For example, impacts to an explosive can occur at any time during processing operations, either intentionally as when machining, or accidentally because of drops or mechanical/ control system failures. Depending on the severity of the impact, this may lead to a localised temperature rise and the production of 'hot spots' as the mechanical energy is converted into heat.

Intrusion into the explosive by 'spigots' or pinching between hard surfaces can exacerbate this heating by providing additional confinement, enabling the hot spots to coalesce and initiate a runaway reaction within the explosive. The violence of the response depends on:

- Severity of impact (i.e. energy available).
- Transient confinement (i.e. as a result of the accident, such as crushing from a blast door).
- Degree of damage of the explosive.
- Reactivity of explosive.
- Mechanical strength of explosive.
- Close fitting case materials if present (providing confinement and a source of fragments).

Thus, for hazard assessments involving conventional explosives there is little benefit in carrying out probabilistic risk assessments when such great uncertainties exist in the data. It is considered that a rigorous assessment based on strong qualitative arguments will yield a better understanding of the hazards and their control¹⁵. It is acceptable to use a 'relative ranking' rather than a specific probability figure, grouping consequences into, for example, 'partial reactions', 'violent reactions' and detonations, against which the consequences for operators can be assessed. Some guidance is provided in the form of 'typical' effects from given quantities of explosive, as shown in Figure 11¹⁹.

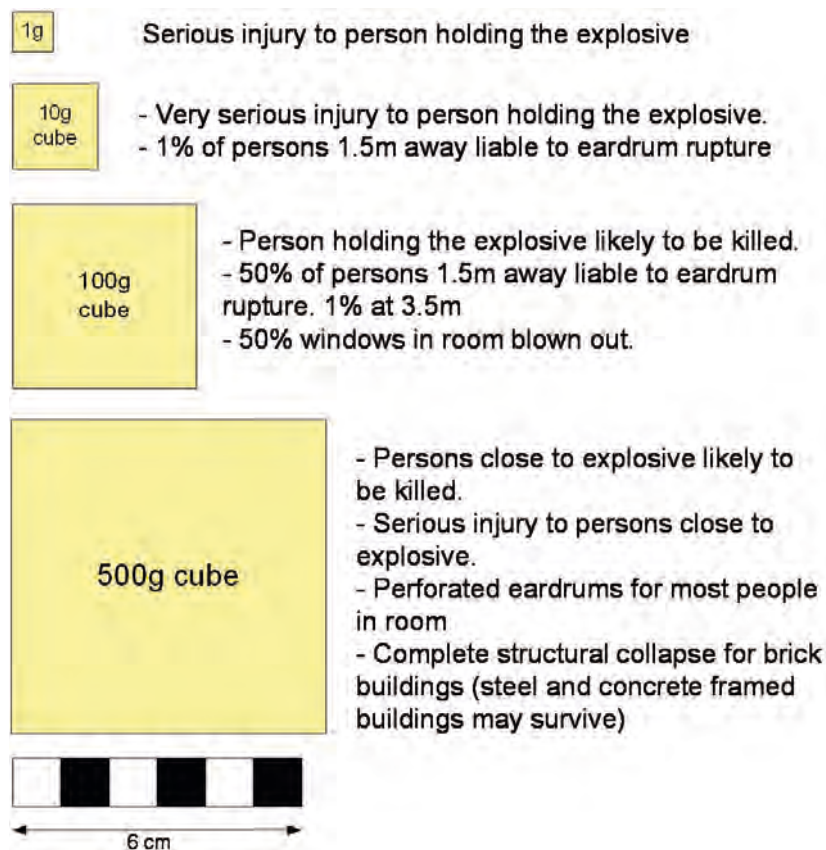


Figure 11 - Explosive effects within 6x6m building

Conclusions

This paper has shown how the standard risk assessment methodologies developed in the process industries are applied to those processes involving explosives. A number of the terms used, both in risk assessments and in explosives processing, have been demystified by demonstrating their application in a number of practical applications. The effect of applying such a systematic process has had a significant effect on the safety of the industry, as evidenced by the reduction in incidences involving explosives. Thus it is concluded that whilst 'explosives safety' appears to be an oxymoron, it is nevertheless an achievable goal.

Acknowledgements

Thanks are due to Richard Gough for proof reading this paper and providing invaluable comments.

References

1. Nova Scotia Archives & Record Management, "Halifax explosion 1917 online resources", <http://www.gov.ns.ca/nsarm/virtual/explosion>
2. Kelly J, "Gunpowder, A history of the explosive that changed the world", Atlantic Books, 2004
3. Davies N, "Pyrotechnics handbook", Cranfield University, (2001).
4. Brown GI, "The Big Bang: A History of Explosives", Sutton Publishing Ltd, (2005)
5. Crocker G, "The Gunpowder Industry", Shire Publications, (2002)
6. Rolt L, "Isambard Kingdom Brunel", Longmans, Green & Co., p. 136, (1957).
7. Holmyard E, "Makers of chemistry", Oxford University Press, p. 189, (1931)
8. Copyright ICI Archives
9. HSE, "Explosives Incidents Database Advisory Service", <http://webcommunities.hse.gov.uk>
10. Cooper PW, "Explosives Engineering", Wiley-VCH Inc., p.53, (1996)
11. Izod D, "Ammunition Technology Volume 1: Introduction", Cranfield University, (1996)
12. Cooper-Key A, "Circumstances attending an accident which occurred on February 1st 1909, in connection with the removal of a wreck...near Caister", Home Office Report, (1909)
13. EMTAP, "Manual of Tests", Defence Ordnance Safety Group, (2005)
14. Gilbert SG, "Assessing hazards from explosives as part of explosives facility safety cases", 27th DOD
15. HSE Information sheet, "Guidance on risk assessment for offshore installations", OIS 3/2006, (2006)
16. Explosives Industry Forum, "Requirements for remote explosives manufacturing facilities", HSE/CBI Explosives Industry Group, (2005)
17. Merrifield R, "Safe handling requirements during explosive, propellant and pyrotechnic manufacture", HSE Specialist Inspector Report No. 31, (1991)

ATKINS

© ATKINS Ltd except where stated otherwise. The ATKINS logo, the "open A" device and the strapline "Plan Design Enable" are trademarks of ATKINS Ltd.



Printed on 9Lives Offset, part of the family of 'born again' papers, manufactured with National Association of Paper Merchants Certification (NAPM) using approved 100% recycled fibre pulp, 9lives Offset offers the ultimate in environmental management. The manufacturing mill and printer have also been accredited with the internationally recognised standard ISO 14001.