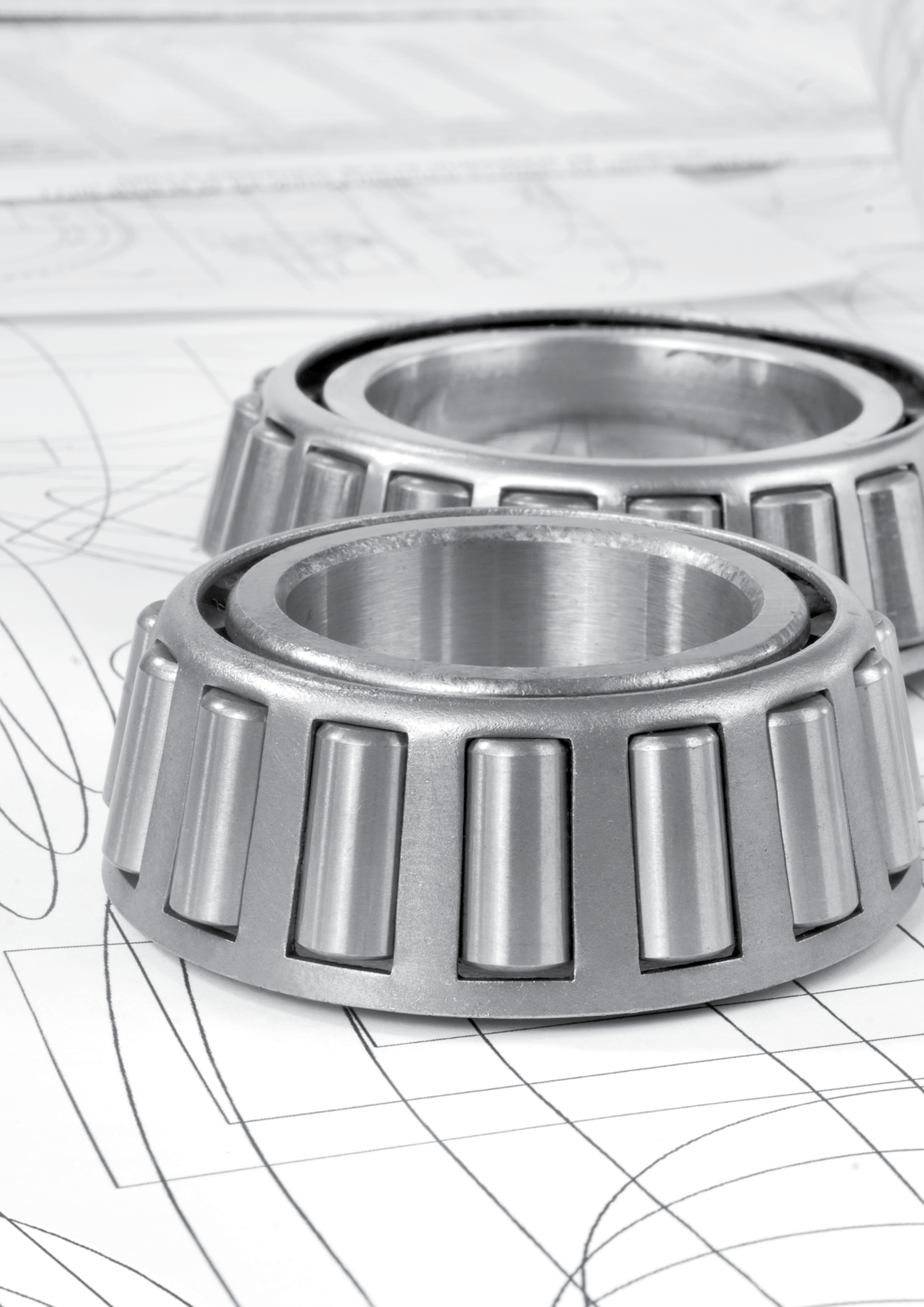
ATKINS

Technical Journal

10

Papers 150 - 161

Plan Des gn Enable



Welcome to the milestone tenth edition of the Atkins Technical Journal which reflects the core of what we excel at: we fix the problems of the past and we plan and design for a better future.

Of course, we do far more than fix the problems of the past; we ensure assets can function as required well into the future. For example, our reliability based assessment of the Severn Suspension Bridge in the UK has allowed a code-deficient structure to continue in operation with minimal intervention. And our asset management work on the M4 Elevated Viaducts has arrested reinforced concrete deterioration and reinstated full strength with minimal disruption to road users and no complaints from locals. Across our regions we are accomplished at improving safety and squeezing more capacity from our roads networks; papers describing enhanced roadside messaging with microwave technology in Orlando, USA and improved traffic controls to ease congestion with 3G data in Slough, UK demonstrate this.

The papers also show we have learnt from the past and have incorporated this learning into our current designs. Papers on cracking in prestressed beams and concretes with improved durability and reduced carbon footprint will help to ensure our future designs are sustainable and require considerably less maintenance. Learning from the Blackwater Valley Road project will also ensure we continue to build increasingly environmentally sound roads.

Looking to the future, our increasing capability in BIM and 3D modelling makes increasingly complex geometrical design possible, but we still need to consider the need for standardisation of components to facilitate construction, which is the subject of the paper Design Rationalization and Optimised Paneling for Architectural Freeform Surfaces. And as our projects become increasingly complex with more distributed design centres, we need to consider more sophisticated means of design management as reflected in the design management tool developed for HS2.

I hope you enjoy the selection of technical papers included in this edition. This tenth journal, and all previous editions, are available on our external website. If you wish to subscribe to our email alert service for future editions of this journal please visit: http://www.atkinsglobal.com/en-GB/about-the-group/our-publications/technical-journals;

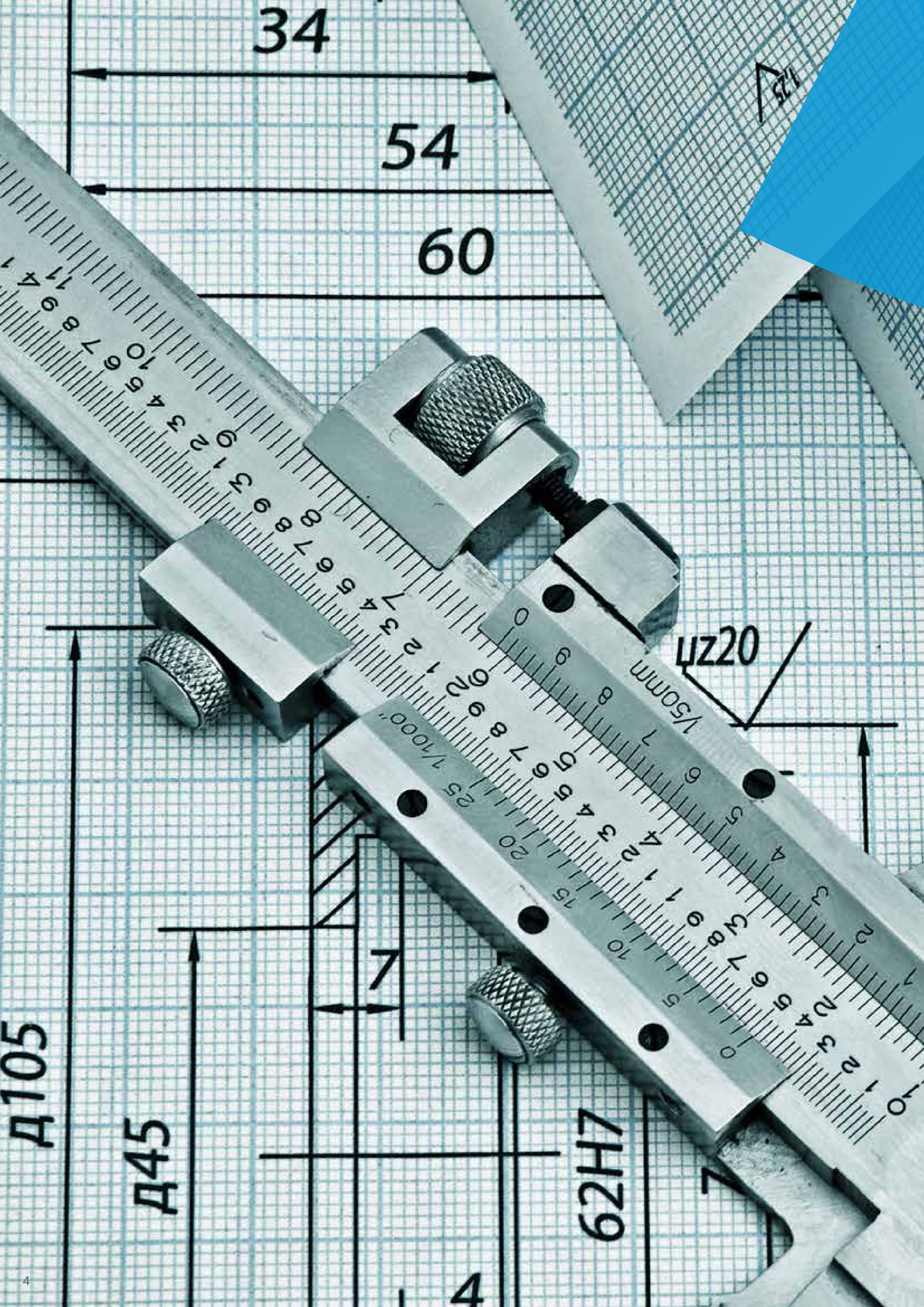
Chris Hendy

Network Chair for Bridge Engineering
Chair of H&T Technical Leaders' Group

Technical Journal 10

Papers 150 - 161

Don't forget you can view previous editions of the Technical Journal on our website, where you can also sign up for email alerts for future publications. Visit http://www.atkinsglobal.com/en-GB/about-the-group/our-publications/technical-journals to find out more.


34
54
60
uz20
7
д105
д45
62H7
4
1/50mm
1/1000"

HDM-4 adaptation for strategic analysis of UK local roads

Abstract

The UK Department for Transport is concerned to know that long-term investments in maintenance of local roads are contributing optimally to the UK economy, and that road maintenance funds are distributed equitably and provide value for money for the taxpayer. This can be accomplished by using decision support tools capable of predicting the long-term performance and costs of maintenance of the UK local road network. This paper describes the development of the World Bank's highway development and management model HDM-4 for use by the Department for Transport at a strategic level. This involved adaptation and calibration of HDM-4 to accurately model pavement performance and road user effects in England, linking HDM-4 with the existing database system used by the Department for Transport to facilitate strategic level analysis and investigation of road investment choices, and conducting a trial strategy analysis of the English local road network to quantify long-term maintenance needs and assess the effects of different maintenance funding levels on the condition of the network and costs to road users. The strategic analysis derived optimal capital and recurrent maintenance needs to clear existing maintenance backlogs and thereafter keep the road network in good condition on a sustainable basis.

Jennaro B. Odoki

Technical Director

Infrastructure Management and Engineering Services Ltd

Birmingham, UK

Michael Anyala

Senior Consultant

Highway Asset Management and Resilience Group

Atkins

Edward Bunting

Senior Policy Advisor

Department for Transport

London, UK

Introduction

There is a widespread recognition in the UK of the importance of highway maintenance and the value placed on the issue by users and the wider community. There is also an increasing understanding of the serious consequences of failure to invest adequately and effectively in maintaining the local highway network. Central government provides a significant proportion of local authority funding in England (outside London). In particular, the revenue support grant provides over £2 billion a year for highways-related revenue expenditure and the Department for Transport (DfT) provides more than £600 million a year from the highways maintenance element of the local transport plan capital settlement. The department is concerned to know that these funds are contributing optimally to the UK economy, are distributed equitably and provide value for money for the taxpayer. This can be accomplished by using decision support tools capable of predicting the long-term performance and costs of maintenance of the UK local road network.

The World Bank's highway development and management model HDM-4 (Kerali, 2000) is a powerful decision support tool that can be adapted and calibrated to correctly simulate the long-term behaviour of the UK local road network. Simulation of the road network involves prediction of road deterioration, the effects of maintenance works, the effects of road condition and standards on road users, and social and environmental effects (e.g. vehicle emissions and energy consumption).

This paper describes the work undertaken by the highways group at the University of Birmingham from 2006 to 2009 to adapt HDM-4 as a road investment decision support tool for use by the DfT at the strategic network level in England. In comparison with other countries such as Australia, Brazil, Czech Republic, India and Poland, there is a wider scope of HDM-4 usage as a support tool, including road policy formulation, strategy analysis, preparation of works programmes and project-level studies. The main users of this tool in the UK would be the DfT, local authorities, academic institutions and other agencies.

Table 1 gives a summary of the road network length in England by road class. Motorways were outside the scope of the study, although in principle the analytical framework of HDM-4 enables its application to any road class.

Road class	Length: km
Motorway	
Trunk	2904
Principal	45
Dual carriageway	
Trunk urban	561
Trunk rural	2825
Principal urban	2004
Principal rural	1132
Single carriageway	
Trunk urban	387
Trunk rural	3575
Principal urban	6740
Principal rural	14 966
B roads	19 871
C roads	64 720
Unclassified roads	179 436
Total	299 166

Table 1. The length of roads in England by road class

The DfT intends to use HDM-4 for the medium- to long-term planning of maintenance expenditure on principal and classified roads in England and for determining the effects of different levels of funding to be provided to local authorities. Furthermore, the department wants to analyse and examine the funding to be allocated to each local authority to improve the road network to a specified level of service, the time it would take to eliminate backlog, and the level of funding required to maintain the road network at the desired service level.

The novelty of this paper rests in the adaptation and calibration of the HDM-4 model to a country with with good roads and advanced asset management systems. Thus, the problems treated in this paper are twofold. First, the adaptation and calibration of the different components of the HDM-4 model to conditions in England and the linkage of HDM-4 to existing DfT systems to allow automatic exchange of data; and second, performing strategic analysis of the road network in England. The main issue considered is one of effectiveness and efficiency in road asset management, and specifically the determination of what allocation of money (both capital and revenue) would maximise the delivery of the Department's objectives. The concept used in this paper is timely and significant, particularly in the wake of the present global financial and economic crisis.

HDM-4 modelling logic

Overall logic sequence

The basic unit of analysis in HDM-4 is the homogeneous road section. Several investment options can be assigned to a road section for analysis. The vehicle types that use the road must also be defined together with the traffic volume specified in terms of the annual average daily traffic (AADT).

The analytical framework of HDM-4 is based on the concept of pavement life cycle analysis, which is typically 15–40 years depending on the pavement type. This is applied to predict road deterioration (RD), road works effects (WE), road user effects (RUE), and socio-economic and environmental effects (SEE) (Odoki and Kerali, 2000). The underlying operation of HDM-4 is common for the project, programme or strategy applications. In each case, HDM- 4 predicts the life cycle pavement performance and the resulting user costs under specified maintenance and/or road improvement scenarios. The agency and user costs are determined by first predicting physical quantities of resource consumption and then multiplying these by the corresponding unit costs.

Two or more options comprising different road maintenance and/ or improvement works should be specified for each candidate road section, with one option designated as the base case (usually representing minimal routine maintenance). The benefits derived from implementation of other options are calculated over a specified analysis period by comparing the predicted economic cost streams in each year against that for the respective year of the base case option. The discounted total economic cost difference is defined as the net present value (NPV). The average life cycle riding quality measured in terms of the international roughness index (IRI) is also calculated for each option.

The overall logic sequence for economic analysis and optimisation is illustrated in **Figure 1**. This figure shows the following (Odoki and Kerali, 2000) elements.

a. The outer analysis loop. This enables economic comparisons to be made for each pair of investment options, using the effects and costs calculated over the analysis period for each option, and it allows for variations in generated and diverted traffic levels depending on the investment option considered

b. Effects, costs and asset values. How annual effects and costs to the road agency and road users, and asset values are calculated for individual road section options

c. Optimisation procedures and budget scenario analysis. These are performed after economic benefits of all the section options have been determined

d. Multiple criteria analysis. This provides a means of comparing investment options using criteria that cannot easily be assigned an economic cost. Note that this capability has not been used in the present study.

Input data requirements

The main data sets required as inputs for HDM-4 analyses are categorised as follows (Kerali et al., 2000).

a. Road network data, comprising inventory, geometry, pavement type, pavement strength and road condition defined by different distress modes

b. Vehicle fleet data, including vehicle physical and loading characteristics, utilisation and service life, performance characteristics such as driving power and braking power, and unit costs of vehicle resources

c. Traffic data, including details of composition, volumes and growth rates, speed-flow types and hourly traffic flow pattern on each road section

d. Road works data, comprising historical records of works performed on different road sections, a range of road maintenance activities practised in the country and their associated unit costs

e. Economic analysis parameters, including time values, discount rate and base year.

Predicting road deterioration

Road pavements deteriorate as a consequence of several factors, most notably traffic volume and loading, pavement design, material types, construction quality, environmental weathering, effect of inadequate drainage systems and works on

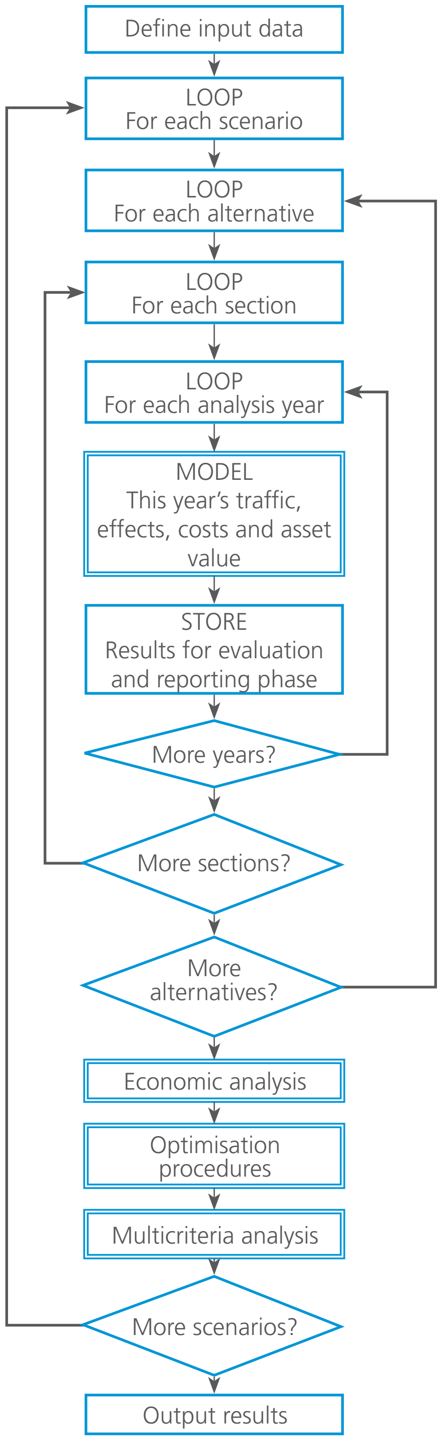


Figure 1. Overall analysis sequence

utilities. The HDM RD models are deterministic models which were developed using a structured-empirical approach (Paterson, 1987). This combines the advantage of both theoretical and experimental bases of mechanistic models with the behaviour observed in empirical studies. The types of models used for predictive purposes are incremental recursive and this gives the annual change in road condition from an initial state as a function of the independent variables.

Road deterioration is modelled in terms of cracking, ravelling, potholes, edge-break, rutting, roughness, friction and drainage. Roughness draws together the impacts of all other pavement distresses and maintenance. It is the dominant criterion of pavement performance in relation to both economics and quality of service as it gives most concern to road users. For each pavement type and each distress type there is a generic model which describes how the pavement deteriorates. To take account of the different behaviour of a particular pavement type constructed with different materials, the coefficients of the generic model depend on the different combinations of the materials. After maintenance treatments, the generic pavement type can change.

Choice and effects of maintenance actions

Standards refer to the levels of conditions and response that a road administration aims to achieve in relation to functional characteristics of the road network system. The choice of an appropriate standard is based on the road surface class, the characteristics of traffic on the road section, and the general operational practice in the study area based upon engineering, economic and environmental considerations. In HDM-4, a standard is defined by a set of works activities with definite intervention criteria to determine when to carry them out. In general terms, intervention levels define the minimum level of service that is allowed. Road agency resource needs for road maintenance are expressed in terms of the physical quantities and the monetary costs of works to be undertaken. The annual costs to the road agency incurred in the implementation of road works are calculated in economic and/or financial terms depending on the type of analysis being performed. The cost of each works activity is considered under the corresponding user-specified budget category (capital, revenue or special).

When a works activity is performed, the immediate effects on road characteristics and road use need to be specified in terms of the following: pavement strength, pavement condition, pavement history, road use patterns and asset value. The long-term effects of a works operation are considered through the relevant models, for example: rate of road deterioration, changes in road user costs, changes in energy use and environmental impacts. Thus, both the immediate and long-term effects are combined to determine the benefits of carrying out different sets of roadworks activities at different times over the analysis period.

Predicting road user effects

The impacts of the road condition, road design standards and traffic levels on road users are measured in terms of road user costs, and other social and environmental effects. Road user costs comprise vehicle operation costs (fuel consumption, tyre wear, oil, spare parts, depreciation, interest, crew hours and overheads), costs of travel time for both passengers and cargo holding, and costs to the economy of road accidents (i.e. loss of life, injury to road users, damage to vehicles and other roadside objects). The social and environmental effects modelled in HDM-4 comprise vehicle emissions and energy consumption (Odoki and Kerali, 2000).

Motorised vehicle speeds and operating resources are determined as functions of the characteristics of each type of vehicle and the geometry, surface type and current condition of the road, under both free flow and congested traffic conditions. The operating costs are obtained by multiplying the various resource quantities by the unit costs or prices. Thus, the annual road user costs are calculated for each vehicle type, for each traffic flow period and for each road section alternative.

Optimisation

The NPVs computed for the different section alternatives are used by the optimisation process to select the best alternative for each road section subject to the budget constraints not being exceeded. The optimisation problem therefore becomes one of searching for the combination of road investment alternatives that optimises the objective function (e.g. maximisation of economic benefits) under a budget constraint. The set of investment alternatives to be optimised is user-defined in such a way that they result in a different selection of treatments and it is not the set of all possible options for the network.

In this study, the ‘incremental benefit/cost’ ranking method was used to perform the optimisation. This involves searching through investment options on the basis of the incremental NPV/cost ratio of one alternative compared against the designated base alternative. The incremental NPV/cost is defined as

$$E_{ji} = \left[\frac{(V_{NP_j} - V_{NP_i})}{(C_j)}\right]$$

(Odoki and Kerali, 2000)

where E_{ji} is the incremental NPV/cost ratio, V_{NP_j} is the net present value of the selected investment alternative j, V_{NP_i} is the net present value of the designated base alternative i and C_j is the financial capital cost of the selected investment alternative j.

The objective of the incremental method is to select road sections successively starting with the largest NPV/cost ratio (E_{ji}), since this maximises the NPV for any given budget constraint. The algorithm starts by checking that there is sufficient budget available such that the minimal routine maintenance (base alternative) for all sections can be performed. An incremental search technique is then used to select the alternatives with successively lower incremental NPV/cost ratios, ensuring that at any time there is no more than one alternative per road section selected. Before selecting an alternative, the algorithm ensures that the works defined for that alternative do not exceed the remaining available budget for all budget periods defined. If the alternative exceeds the constraint in any of the budget periods, it is discarded. The process continues until the budget is exhausted for each budget period.

Reliability of results

The reliability of the results obtained from HDM-4 analysis is dependent upon two primary considerations (Bennett and Paterson, 2000). These are:

- How well the data provided to the model represent the reality of current conditions and influencing factors, in the terms understood by the model
- How well the predictions of the model fit the real behaviour and the interactions between various factors for the variety of conditions to which it is applied.

Application of the model thus involves two important initial steps.

a. Data input: interpreting correctly the data input requirements, and achieving a quality of input data that is appropriate to the desired reliability of the results

b. Calibration of outputs: adjusting the model parameters to enhance how well the forecast and outputs represent the changes and influences over time and under various interventions. Prior to using HDM-4 for the first time in any country, the system should be configured and calibrated for local use.

The accuracy required of the input data is dictated by the objectives of the analysis. For a very approximate analysis there is no need to quantify the input data to a very high degree of accuracy. Conversely, for a detailed analysis it is important to quantify the data as accurately as is practicable given the available resources.

Figure 2 illustrates the impact of the accuracy of input data on road deterioration predictions and the timing of future maintenance interventions (Bennett and Paterson, 2000). HDM-4 uses incremental-recursive models and the existing condition (denoted by point 1 or 2) is the start point for the modelling. The pavement will deteriorate and reach that condition, defined by a given set of criteria for maintenance intervention, in a certain period of time depending on the existing condition. The difference in the start point will have as great, if not greater, impact on when the treatments are triggered, as will the calibrated deterioration factor. **Figure 2** also illustrates a second point: That HDM-4 model predictions are based on the mean deterioration rate and therefore will have a certain time interval within which a particular treatment will be triggered by a given set of intervention criteria.

Typical values that define the slower and faster rates of deterioration into a band vary across the different distresses modelled. The further into the future one predicts the deterioration, the greater the spread in the trigger interval. Consequently, this will impact on the analysis results as costs incurred in the future are discounted to the base year value.

Configuring and calibrating HDM-4 to conditions in England

This activity involved two processes which are referred to as configuration and calibration of HDM-4. Since HDM-4 is designed to be used in a wide range of environments, it needed to be configured to reflect the norms that are customary in England. Calibration of HDM-4 is intended to improve the accuracy of predicted pavement performance and vehicle resource consumption.

Configuration of HDM-4

The configuration activity involved defining sets of default data and aggregate parameters that are subsequently used in HDM-4 strategic analyses. The parameters that were configured include traffic flow patterns to represent hourly distribution of flows on different road categories; speed–flow relationships

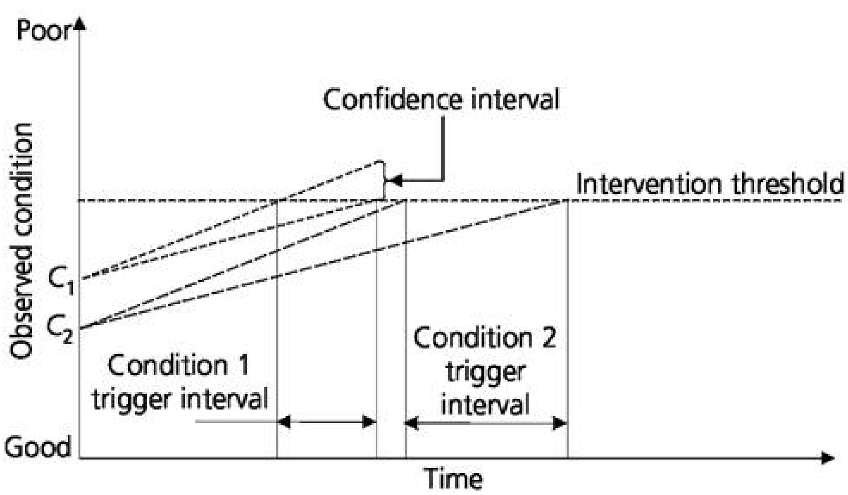


Figure 2. The impact of the accuracy of data on road deterioration predictions

for different road types, road classes, traffic bands, road geometry class and climate characteristics; and default data for road surface condition, structural condition, riding quality, surface texture and construction quality.

The speed–flow relationships used in the DfT (2008) COBA manual given in Design Manual for Roads and Bridges (DMRB), Volume 13, Section 1, Part 5, Chapter 9 (see www.leics.gov.uk/part_5.pdf) were transformed into the format used in HDM-4 as described in **Figure 3**, which shows that each road type has a separate speed–flow relationship.

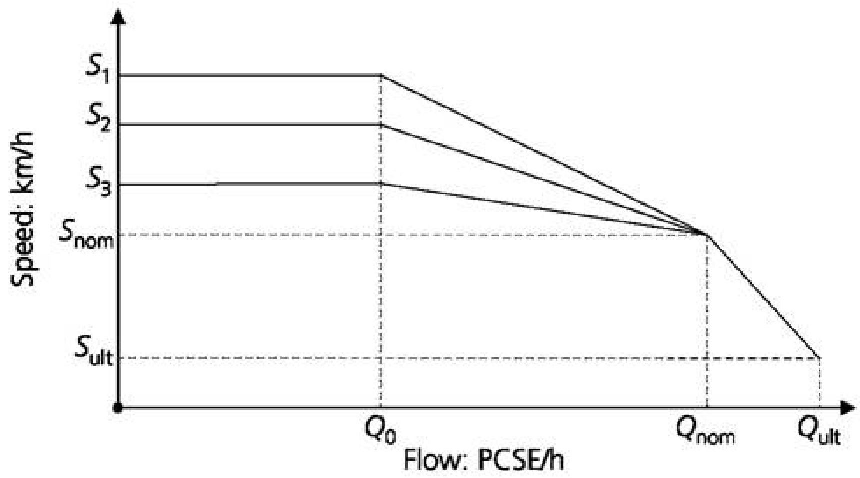


Figure 3. HDM-4 speed–flow model used to define the relationship for each road type. Q_0 is the flow level below which traffic interactions are negligible, in PCSE/h; Q_{nom} is the nominal capacity of the road (in PCSE/h); Q_{ult} is the ultimate capacity of the road for stable flow (in PCSE/h); S_{ult} is the speed at the ultimate capacity, also referred to as jam speed (in km/h); S_{nom} is the speed at the nominal capacity (in km/h); S_1 to S_3 are the free flow speeds of different vehicle types (in km/h); and PCSE/h is passenger car space equivalents per hour

Calibration of HDM-4 models

Calibration levels

There are three levels that define the extent of HDM calibration, which involve low, moderate and major levels of effort and resources as follows (Bennett and Paterson, 2000).

- Level 1: Basic application that determines the values of basic input parameters, adopts many default values and calibrates the most sensitive parameters with best estimates based on experience, desk studies and/or minimal field surveys
- Level 2: Verification, which involves moderate field surveys and measurement of additional input parameters required to calibrate key predictive relations to local conditions
- Level 3: Adaptation, in which major field surveys and controlled experiments are undertaken to enhance the existing predictive relationships or to develop new and locally specific relationships.

Calibration of HDM-4 to conditions in England was generally completed to Level 1 with a few models calibrated to Level 2, as described in the paragraphs below.

Road deterioration and works effects

Attention was focused on calibrating the key RD models for cracking, rutting and roughness-age-environmental effects since these have great impact on pavement serviceability life. The models were calibrated using historical pavement performance data gathered from two local authorities of Birmingham City Council and the London Borough of Hammersmith and Fulham. The goal of calibration was to reduce any bias of the predictions by the models to acceptable levels and increase precision. Bias is defined as the systematic difference that arises between the observed and predicted values, and precision is a measure of how closely the observed and predicted data are to each other. To augment this, expert knowledge of observed initiation and progression of distresses, and treatment life were obtained from discussions at a project workshop held at the University of Birmingham in 2006. This information is summarised in **Tables 2** and **3** and was used in the calibration of HDM-4 RD and WE models, respectively. **Table 2** indicates that the calibration was based on observed levels of deterioration before a maintenance treatment can be applied. **Table 3** gives the average treatment life by treatment type and group. Overall, the calibration of the cracking, rutting and roughness-age-environmental effects models was achieved at Level 2 calibration. The WE models and the other RD models were calibrated to Level 1 calibration.

Road user effects

Attention was focused on calibrating RUE models for vehicle speeds, fuel consumption, spare parts and capital costs. The calibration of RUE models required the collection of vehicle fleet data from the Driver and Vehicle Licensing Agency (DVLA) offices. The UK vehicle fleet characteristics are required in HDM-4 for the estimation of vehicle resource consumption, traffic flow and capacity, and average travel speeds, together with the unit costs. The definition of representative vehicle took into account the current traffic classification system defined in the DfT MRSF2003 database, as this allowed the use of existing traffic data. The overall aim was to select an appropriate number and mix of vehicles based primarily on factors such as vehicle operating costs (VOC), loading and vehicle occupancies. The Level 1 calibration of the VOC models involved a cross-check of the selected vehicle types against that defined in the COBA manual. Value of time figures during working and non-working hours used in the case study described in 'Strategy analyis case study' are given in **Tables 4** and **5** (see www.webtag.org.uk).

Data for validating HDM-4 analysis results

Historical data on traffic volume and loading, road condition, average annual maintenance expenditure, vehicle maintenance and fuel usage, and freight and passenger tariffs are the key data that need to be collected in order to validate HDM-4 analysis results, provide feedback and improve its performance. Bennett and Paterson (2000) provide guidance on how to collect these data.

Customised HDM-4 workspace for the Department for Transport

The workspace is the central hub of HDM-4. The results of the configuration and calibration activities were used to develop a customised master HDM-4 version 2 workspace for the DfT. The customised data included vehicle fleet, road network, work standards, configuration, project analysis, programme analysis and strategy analysis.

Linking HDM-4 to the Department for Transport database system

At the strategic level, HDM-4 requires that all road data be supplied in the form of representative and homogeneous sections. Homogeneity is defined in terms of physical attributes and traffic characteristics (e.g. pavement type, annual average daily traffic and road class). Secondly, it requires a relatively large number of data items for each road section.

Pavement defects	Initial values on typical new roads in UK	Expert knowledge of observed deterioration rate/ level
Cracking	0%	Initiation: 5% visible cracks at 5–8 years for asphalt surfaces and 4–5 years for surface dressings Progression: 5% wide structural cracks at 15 years for asphalt surfaces and 10 years for surface dressing
Rutting (structural)	0 mm	11 mm structural rutting at 15 years for asphalt surfaces and 10 years for surface dressing
Roughness	1.5 IRI	Depends on several factors, see Section on road deterioration and works effects
Texture depth	1–1.2 mm	Depends on road use and environmental factors

Table 2. Assumptions on pavement condition of UK local roads

Treatment group	Treatment type	Average treatment life: years
Surface improvement	Slurry seal (macro-asphalt) Single surface dressing	7
Resurfacing	Thin overlay Shallow mill and replace Medium mill and replace	10
Strengthening	Thick overlay Deep mill and replace Reconstruction	15

Table 3. Observed average treatment life

Vehicle occupant	2008 forecast prices and values: £/h
Car driver	21.86
Car passenger	15.66
LGV (driver or passenger)	8.42
OGV (driver or passenger)	8.42
PSV passenger	16.72

Table 4. Value of passenger travel time during working hours

Vehicle occupant	2008 forecast prices and values: £/h
Commuting	4.17
Other	3.68

Table 5. Value of passenger travel time during non-working hours

Many of these data items are model calibration parameters that are specific to HDM-4 and may not be readily available within the DfT databases.

A data transfer software tool was developed to simplify the transfer of road network data from the DfT's NRMCS system to HDM-4 to cater for the specific data input requirements. **Figure 4** illustrates the relationship between HDM-4 and NRMCS/MRSF databases.

The data transfer software tool can generate road network matrices and homogeneous sections in an efficient, effective and integrative manner for use in strategic level analyses in HDM-4. The automated process of data transfer significantly reduces the amount of time used for data preparation and analysis and consequently cost. **Table 6** provides an example of a typical road network matrix that may be produced using the NRMCS database.

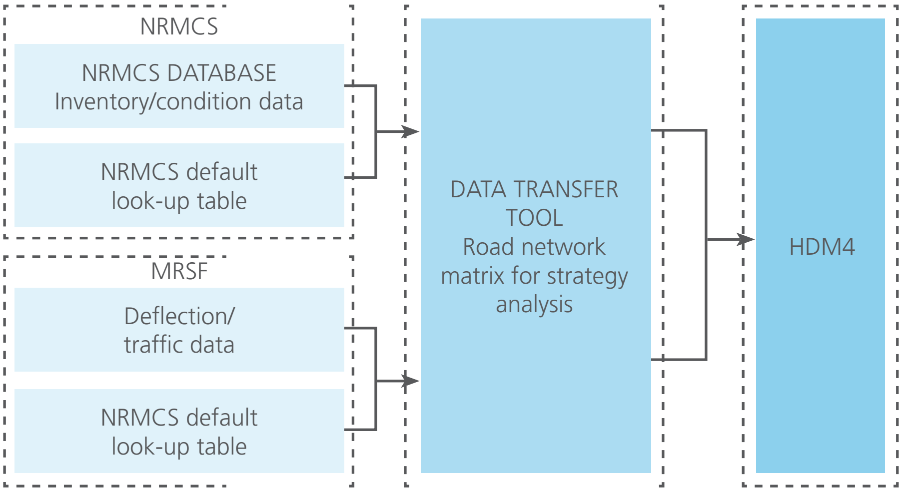


Figure 4. Relationship between HDM-4 and NRMCS/MRSF databases

Strategy analysis case study

Outline methodology

A strategic analysis of the principal DfT road network in England was undertaken to assess its medium- to long-term maintenance expenditure needs. The methodology adopted for the case study can be summarised by the following steps.

a. Generation of a road network matrix to represent the entire principal DfT road network and importing this matrix into the customised HDM-4 workspace

b. Definition of maintenance standards, intervention criteria and unit costs as applicable to different road classes and road types

c. Definition of budget scenarios

d. Performance of strategic analysis using the customised HDM-4 workspace and presentation and discussion of results.

Each of these steps is explained in the following sub-sections.

Generation of road network matrix

The entire principal road network in England was transformed into a matrix of homogeneous road sections according to key attributes that are likely to influence pavement performance and/ or road user costs. These attributes included the local authority, road type (urban or rural), road class (principal) and condition (good, fair or poor). The thresholds used to define limits for condition parameters (rut depth and cracking) between the matrix cells vary according to road class as shown in **Table 7**.

The road network matrix for the case study was generated using the data transfer tool described in on linking HDM-4 to the Department for Transport database system. Each of the matrix cells represents a larger number of real physical sections with similar characteristics, scattered around the road network. Instead of each of the short constituent sections being analysed separately, just the representative sections have been analysed. The length of each representative section is the sum of the lengths of the constituent real sections. The advantages associated with the concept of road network matrices are the quick analysis run time and flexibility of information processing of different solutions and strategies.

Using this approach, a representative matrix comprising 348 representative sections was generated from 6094 physical road sections. A part of the matrix is presented in **Table 6**.

Maintenance standards

The work activities performed on the DfT local road network have been considered under four maintenance standards, namely minimum maintenance, surface improvement, resurfacing and strengthening. The standards were developed from recommendations of the UK Pavement Management System (UKPMS) rules and parameters and outcomes of workshops attended by pavement experts familiar with the road maintenance practices in the UK. Works standards used for the case and associated unit costs of work items are given in **Table 8**.

Intervention levels for triggering maintenance and rehabilitation works were defined for each work activity. Minimum maintenance activities such as patching, edge break repair, crack sealing and other routine works are 'scheduled' and triggered annually. However, capital road works such as surface dressing, slurry seal, overlay, mill and replace

Sect_ID	Sect_Name	Length: m
U_P_B_4250G G G G V	Urban:Principal:Bituminous:Wigan:Good:Good:Good:Good:Very Good	91.5
R_P_B_2371G_G_G_G_V1	Rural:Principal:Bituminous:Lancashire:Good:Good:Good:Good:Very Good(1)	76.8
R_P_B_2275G_G_G_G_V3	Rural:Principal:Bituminous:Kent:Good:Good:Good:Good:Very Good(3)	66.5
U_P_B_4705G G G G V	Urban:Principal:Bituminous:Bradford:Good:Good:Good:Good:Very Good	96.0
U_P_B_2600G_G_G_G_V1	Urban:Principal:Bituminous:Norfolk:Good:Good:Good:Good:Very Good(1)	73.1
R_P_B_2275G_P G G V	Rural:Principal:Bituminous:Kent:Good:Poor:Good:Good:Very Good	84.0
R_P B 540 G G G G V	Rural:Principal:Bituminous:Peterborough:Good:Good:Good:Good:Very Good	45.5
R_P_B_2280F_G_G_F_G_	Rural:Principal:Bituminous:Medway Towns:Fair:Good:Good:Fair:Good	24.0
U_P_B_1350F_G_G_P_G_	Urban:Principal:Bituminous:Darlington:Fair:Good:Good:Poor:Good	15.4
U_P_B_5720F_G_F_F_F_	Urban:Principal:Bituminous:Merton:Fair:Good:Fair:Fair:Fair	28.8
R_P_B_2371G_G_G_G_V2	Rural:Principal:Bituminous:Lancashire:Good:Good:Good:Good:Very Good(2)	76.8
R_P_B_1155G_P_G_G_V2	Rural:Principal:Bituminous:Devon:Good:Poor:Good:Good:Very Good(2)	71.0
U_P_B_1155G_G_G_G_V2	Urban:Principal:Bituminous:Devon:Good:Good:Good:Good:Very Good(2)	63.3
U_P_B_1590G_F_G_G_V_	Urban:Principal:Bituminous:Southend-on-Sea:Good:Fair:Good:Good:Very Good	30.8
R_P_B_1780G G G G V	Rural:Principal:Bituminous:Southampton:Good:Good:Good:Good:Very Good	6.6
R_P_B_3600G_G_G_G_V3	Rural:Principal:Bituminous:Surrey:Good:Good:Good:Good:Very Good(3)	70.3
U_P_B_2600G_G_G_G_V2	Urban:Principal:Bituminous:Norfolk:Good:Good:Good:Good:Very Good(2)	73.1
U_P B 800 G G G G V2	Urban:Principal:Bituminous:Cornwall:Good:Good:Good:Good:Very Good(2)	69.8

Table 6. Sample local authority road network matrix

and reconstruction are 'responsive' to certain conditions and are triggered by criteria such as rutting, cracking, ravelling and roughness. The intervention levels given in **Table 9** were derived from recommendations of workshops attended by pavement experts including DfT representatives with extensive experience of road maintenance practice in the UK.

The four maintenance standards or investment alternatives given in **Table 8** and the associated work items are assigned to each of the 348 representative road sections. For each representative road section, the HDM-4 strategy analysis tool selects the investment alternative that provides the most economic and social benefits.

Section alternatives

The set of alternatives defined for each section contains a 'base alternative' which is defined to be the minimal routine maintenance, and the other alternatives are compared against this to derive the benefits. The other alternatives defined will typically be for different

Parameter	Road class	Good	Fair	Poor
Rut depth: mm	Principal classified	0–5	5–12	>12
	Non-principal classified	0–10	10–20	>20
Cracking: %	Principal classified	0–5	5–10	>10
	Non-principal classified	0–10	10–15	>15

Table 7. Threshold values for condition parameters

levels of intervention, for instance resurfacing, strengthening or reconstruction, as well as the effects of delaying the intervention of these works from the start year of the analysis (which is especially apt in the case of reducing the budget in the early years).

Budget scenarios and analysis

The trial strategy analysis was conducted over a 20-year analysis period under three budget scenarios: unconstrained budget scenario; unconstrained budget less 25%; and unconstrained budget less 50%.

- Budget scenario 1: Availability of unconstrained budget. The budget obtained from this scenario reflects the funding needs under the unconstrained budget scenario. Under this mode of analysis, HDM-4 calculates the required budget for a solution that most satisfies the objective function of maximising the economic benefits to the nation by considering both the DfT (road agency) costs and the road user costs
- Budget scenario 2: Availability of unconstrained budget less 25%. This scenario examines the impact of a 25% reduction in the unconstrained budget as calculated under budget scenario 1
- Budget scenario 3: Availability of unconstrained budget less 50%. This scenario examines the impact of a 50% reduction in the unconstrained budget as calculated under budget scenario 1.

Maintenance standard	Work items		Units	Units costs of works: £	
	Work type	Description		Economic	Financial
Minimum maintenance (Base alternative)	Patching potholes	Repair of surface distresses such as potholing, wide structural cracking and ravelling	m^2	17	20
	Edge break repair	Patching edge failures	m^2	21.3	25
	Crack sealing	Treatment of transverse thermal cracking and wide structural cracking	m^2	8.5	10
	Miscellaneous works	Includes shoulder repairs, vegetation control, road sign repairs and replacement, line marking, guardrail repair and replacement and so on	per km per year	2125	2500
Surface improvement	Patching potholes	As for minimum maintenance standard	m^2	17	20
	Edge break repair		m^2	21.3	25
	Crack sealing		m^2	8.5	10
	Miscellaneous works		per km per year	2125	2500
	Slurry seal (macro-asphalt)	A layer of an 8 mm macro-asphalt designed to seal the surface and to defer major maintenance intervention	m^2	3.4	4
	Single surface dressing	Single sealing of the carriageway in order to delay major intervention and to renew the skid resistance	m^2	2	2.3
Resurfacing	Patching potholes	As for minimum maintenance standard	m^2	17	20
	Edge break repair		m^2	21.3	25
	Crack sealing		m^2	8.5	10
	Miscellaneous works		per km per year	2125	2500
	Thin overlay	30 mm asphaltic overlay.	m^2	6	7
	Shallow mill and replace	Mill and replace 30 mm surface course	m^2	7.9	9.25
	Medium mill and replace	Mill and replace 35–50 mm by removing the old surfacing and replacing it with a new SMA (or HRA) material	m^2	11.9	14
Strengthening	Patching potholes	As for minimum maintenance standard	m^2	17	20
	Edge break repair		m^2	21.3	25
	Crack sealing		m^2	8.5	10
	Miscellaneous works		per km per year	2125	2500
	Thick overlay	150 mm HRA (or SMA) and DMB (or DBM) asphaltic overlay	m^2	14.5	17
	Deep mill and replace	Mill and replace 150 mm by removing the old surfacing and base and replacing them with SMA (35 mm) and DMB (115 mm)	m^2	17	20
	Reconstruction	Includes 35 mm SMA surface course, 100 mm DBM binder course, 115 mm base and 200 mm crushed stone sub-base	m^2	42.3	50

Table 8. Work standards and unit costs of treatment works

Work item	Maintenance intervention criteria
Patching potholes	Potholing ≥1 no./km
Edge break repair	Edge break ≥5 m2/km
Crack sealing	Wide structural cracking ≥1%
Miscellaneous works	Interval ≥1 year
Slurry seal (macro-asphalt)	(Roughness ≥3.5 IRI) AND (Interval ≥5 years) AND (Ravelling ≥10%) AND (Total carriageway cracked a5%)
Single surface dressing	(Roughness ≥3.5 IRI) AND (Interval ≥5 years) AND (Ravelling ≥15%) AND (Total carriageway cracked ≥5%)
Thin overlay	{(Rut depth mean ≥8 mm) OR (Total carriageway cracked ≥5%) OR (Ravelling ≥30%) OR (Roughness ≥3.8 IRI)) AND {Interval ≥8 years)
Shallow mill and replace	{(Rut depth mean ≥10 mm) OR (Ravelling ≥40%) OR (Roughness ≥3.8 IRI)) AND {Interval ≥8 years)
Medium mill and replace	{(Rut depth mean ≥12 mm) OR (Total carriageway cracked ≥6%) OR (Ravelling ≥30%) OR (Roughness ≥4 IRI)) AND {Interval ≥10 years)
Thick overlay	{(Rut depth mean ≥10 mm) AND (Total carriageway cracked ≥4%) AND (Interval ≥12 years)) OR {(Roughness ≥4 IRI) AND (Interval ≥12 years))
Deep mill and replace	{(Rut depth mean ≥14 mm) AND (Total carriageway cracked ≥9%) AND (Interval ≥12 years)) OR {(Roughness ≥4.5 IRI) AND (Interval ≥12 years))
Reconstruction	{(Rut depth mean ≥16 mm) AND (Total carriageway cracked ≥8%) AND (Interval ≥15 years)) OR {(Roughness ≥5 IRI) AND (Interval ≥15 years))

Table 9. Work items and maintenance intervention levels

Analysis of results

The results of the strategy analysis are presented under the following sub-headings: road condition trend, treatment works, financial costs and road user costs.

Road condition trend

Figures 5, 6 and **7** show the progression of roughness, mean rut depth and whole carriageway cracking respectively, under the three budget scenarios. In all the three figures, budget scenario 1 generally gives the best condition profile and the predicted condition trend gets worse due to reduction in maintenance budget as defined in scenarios 2 and 3. The predicted average road condition trends illustrated in **Figures 5, 6** and **7** suggest that the effect of reducing the unconstrained maintenance budget by 25% (budget scenario 2) and 50% (budget scenario 3) only becomes significant from the seventh year of analysis.

In terms of international roughness index (IRI), the initial average condition of the road network is

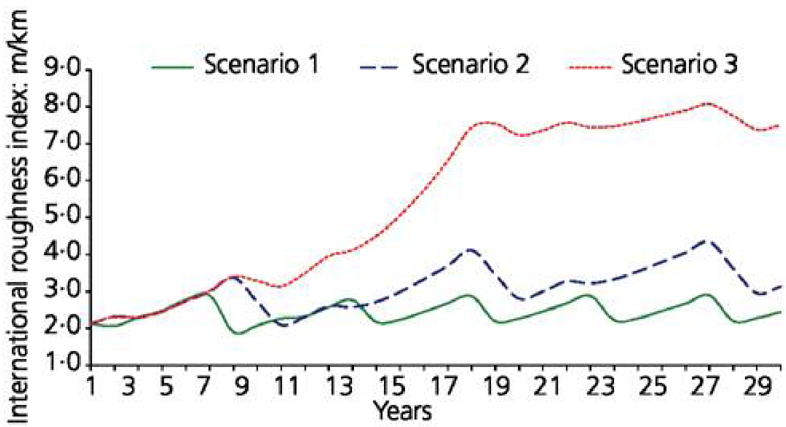


Figure 5. Average roughness progression by budget scenario

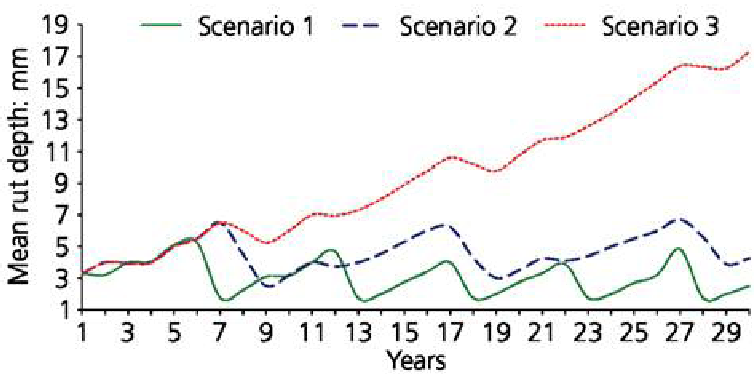


Figure 6. Mean rut depth progression by budget scenario

Asset Management

estimated at 2.0 IRI as shown in **Figure 5**. Analysis based on budget scenario 1 resulted in an average road network roughness of 2.4 IRI. Budget scenario 2 gives an average road network roughness of 3.7, while scenario 3 gives an average roughness of 5.5.

Outputs of HDM-4 analysis comprise several pavement deterioration parameters for analysed road sections including structural number, roughness, cracking, ravelling, potholes, edge break, rutting, texture depth and skid resistance. It is possible to use such outputs to estimate performance indicators such as the Scanner road condition indicator (RCI). The RCI has been developed to characterise the overall condition of the road carriageway, and provides a basis for reporting national indicators (NI) (Cartwright, 2009). If correctly established, the HDM-4 strategic analysis tool may be used as a what-if analysis tool to investigate the implication of various maintenance and rehabilitation options on performance indicators such as NI168.

Selected treatment works

The selected treatment works under different budget scenarios are summarised in **Table 10**. It can be observed that as the budget level is reduced, a higher proportion of lighter and cheaper types of treatment works are selected. Obviously the frequency of carrying out such treatment works will be higher with consequential increased delay costs due to work zone closures.

Financial results

The results of the total financial costs for capital and recurrent works by budget scenario are summarised in **Table 11**.

The results of the strategy analysis indicated an unconstrained budget (both capital and recurrent) for the preservation of DfT principal network to be around £17.6 billion over the 20-year analysis period. This requirement equates to an annual funding level of about £880 million per year.

The expenditure and condition profile for budget scenario 1 are shown in **Figure 8**. This shows a peak in the capital expenditure profile suggesting a total expenditure of about £2.6 billion in the sixth year of analysis.

A smoother expenditure profile may be achieved by limiting available annual budgets as illustrated in **Table 12** and **Figure 9**.

Road user costs

HDM-4 road user cost outputs considered in this study were vehicle operating costs (VOC) and travel time costs (TTC) (Bennett and Greenwood, 2004). A summary of average annual road user costs by vehicle types for the 20-year analysis period is given in **Table 13**. Vehicle operating costs are highest for OGV2 while the PSV

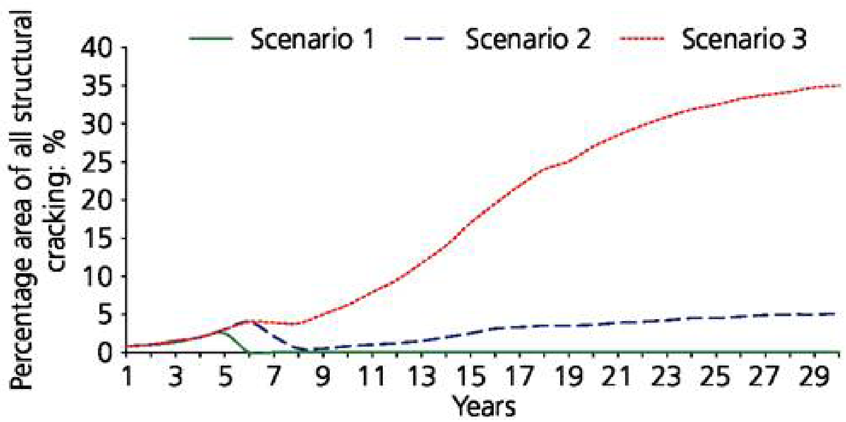


Figure 7. Average all carriageway cracking progression by budget scenario

Treatment category	Work type	Percentage of road length: %		
		Scenario 1	Scenario 2	Scenario 3
Surface improvement	Slurry seal (macro-asphalt)	0	10	25
	Single surface	0	13	31
Resurfacing	Shallow mill and replace	72	58	32
	Medium mill and replace	12	9	5
	Thin overlay	12	7	4
Strengthening	Reconstruction	4	3	3

Table 10. Summary of treatment proportions by road length for each budget scenario

Budget scenarios	HDM-4 derived inancial costs: million £			
Description	Capital budget limits	Capital costs	Recurrent costs	Total costs
Budget scenario 1 (unconstrained budget)	Unconstrained	15 358	2241	17 599
Budget scenario 2 (unconstrained budget less 25%)	11 519	11 519	2199	13 717
Budget scenario 3 (unconstrained budget less 50%)	7679	7679	1499	9178

Table 11. Capital and recurrent inancial costs by budget scenario

has the highest travel time costs. The average annual road user costs for OGV2 (**Table 13**) suggest that adopting budget scenario 2 over budget scenario 1 would result in an increase in average annual road user costs of £0.15 per vehicle kilometre. Similarly, adopting budget scenario 3 results in an increase in road user costs of £0.20 per vehicle kilometre.

A comparison of annual trend in vehicle operating costs and road condition is illustrated by budget scenario in **Figure 10**. The trend in VOC suggests that budget scenarios 2 and 3 generally result in higher operating costs than for budget scenario 1. However, the difference in VOC trend is less significant between budget scenarios 1 and 2, suggesting that fewer VOC savings would be achieved by choosing budget scenario 1 over scenario 2. From **Figure 10**, it can be observed that beyond analysis year 9 the road network condition trend under budget scenario 3 moves close to the poor condition threshold. A similar pattern is reflected in the VOC trends under scenario 3. **Figure 11** compares travel time costs trend over the 20-year analysis period. The differences in the effects of budget scenario 2 over budget scenario 1 are only significant beyond the 13th year of the analysis period.

Conclusion

The work presented in this paper has demonstrated that HDM-4 can be established as a comprehensive decision support tool for use by the DfT at the strategic level. This would enable the DfT to carry out medium- to long-term planning of maintenance expenditure and investigate investment choices on local roads. Prior to using HDM-4 in any country, the relevant prediction models should be adapted and calibrated to reflect local conditions. Owing to lack of appropriate time series data on road pavement performance at the time of the study, calibration of road deterioration models was largely based on expert knowledge and limited data on pavement performance of the DfT local road network. This calibration may be improved by using a minimum of four years' time series data of road condition such as those derived from the Scanner system, which provides more detailed and updated data.

The results of the strategy analysis case study indicated an unconstrained budget for

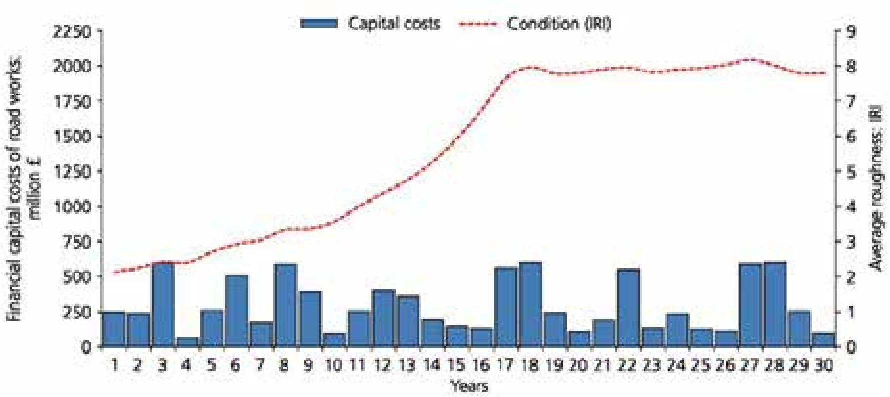


Figure 8. Comparison of expenditure and condition profile for budget scenario 1 without annual budget constraints

Start of year	End of year	Capital budget: million £
1	5	2560
6	6	512
7	11	2560
12	12	512
13	17	2560
18	18	512
19	23	2560
24	24	512
25	30	3072
Total		15 360

Table 12. Budget scenario details to reduce peaks in capital expenditure profile

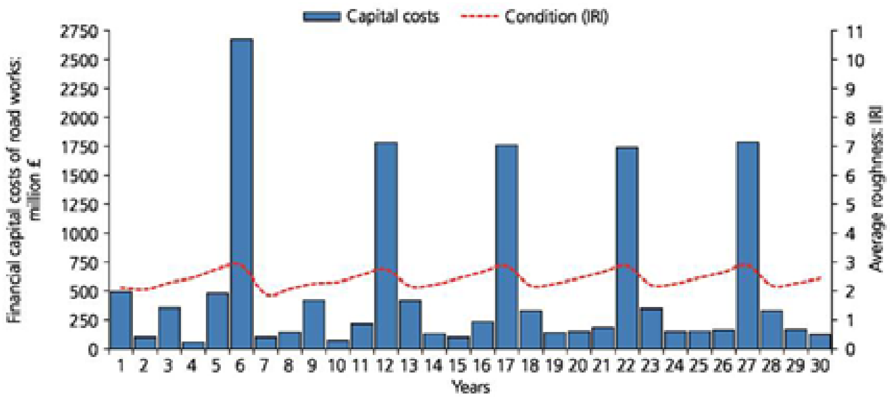


Figure 9. Comparison of expenditure and condition profile for budget scenario 1 with annual budget constraints

Budget scenario	Component	Average road user costs in £ per vehicle-km				
		Car	LGV	OGV1	OGV2	PSV
Scenario 1	VOC	0.42	0.44	1.20	2.66	0.54
	TTC	0.25	0.11	0.03	0.03	1.58
	RUC	0.68	0.55	1.23	2.69	2.12
Scenario 2	VOC	0.45	0.45	1.28	2.81	0.56
	TTC	0.26	0.11	0.03	0.03	1.59
	RUC	0.70	0.57	1.31	2.84	2.16
Scenario 3	VOC	0.45	0.46	1.30	2.85	0.57
	TTC	0.26	0.11	0.03	0.03	1.60
	RUC	0.71	0.57	1.33	2.89	2.17

VOC, vehicle operating costs; TTC, travel time costs; RUC, road user costs; LGV, light goods vehicle; OGV1/2, other goods vehicle catgegory 1/2; PSV, public service vehicle.

Table 13. Average road user costs by budget scenario over 30 year analysis period

preservation of DfT principal network to be around £17.6 billion over the 20-year analysis period. This equates to an annual average funding level of about £880 million. This funding level comprises both capital and recurrent maintenance needs to clear existing maintenance backlogs and thereafter keep the road network in good condition on a sustainable basis.

The strategy analysis case study demonstrated the capability of HDM-4 in determining the effects of various funding levels. The strategy analysis tool optimises investment options subject to available budget to minimise total transport costs by considering the costs to the road authority and road users. To that end, HDM-4 provides a good framework for ensuring that funds for maintenance of roads are distributed equitably among local authorities and provide value for money for the taxpayer.

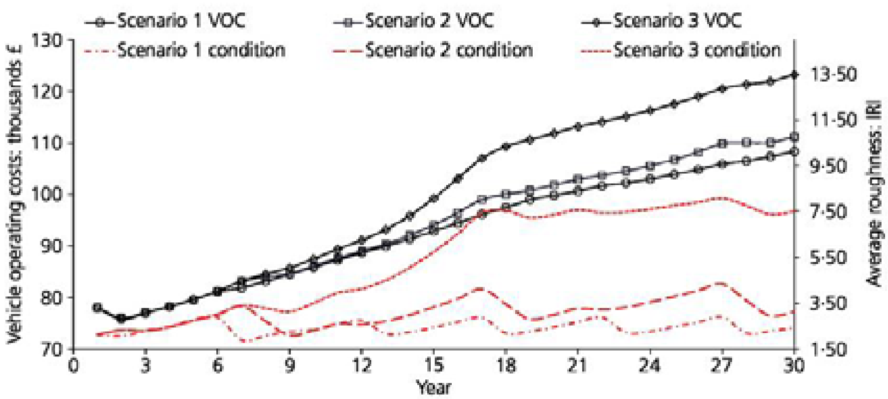


Figure 10. Average annual vehicle operating costs (VOC) and roughness progression by budget scenario

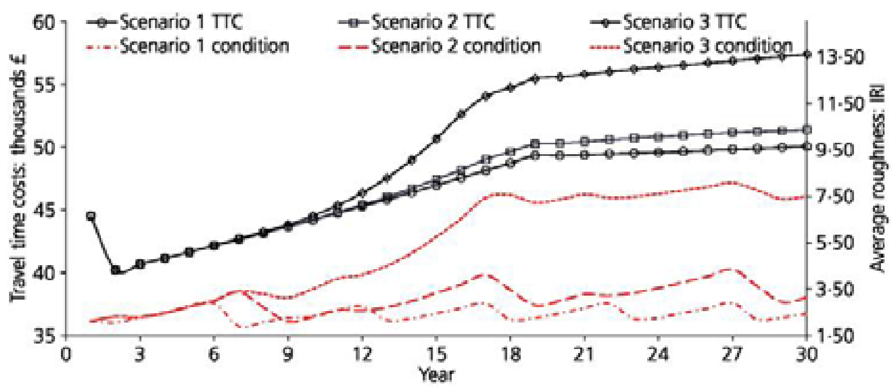


Figure 11. Average annual travel time costs (TTC) and roughness progression by budget scenario

References

1. Bennett CR and Greenwood I (2004) Modelling Road User and Environmental Effects in HDM-4 – Volume 7. International Study of Highway Development and Management Series. World Road Association (PIARC), Paris, France
2. Bennett CR and Paterson WDO (2000) A Guide to Calibration and Adaptation – Volume 5. International Study of Highway Development and Management Series. World Road Association (PIARC), Paris, France
3. Cartwright R (2009) Technical Note 44. NI168 (2008/2009) Guidance Notes for UKPMS Development. TRL, Wokingham, UK
4. DfT (Department for Transport) (2008) COBA Manual. Transport Analysis Guidance (TAG), Unit 3.5.6. See http://www.dft. gov.uk/webtag/documents/expert/unit3.5.6.php (accessed 20/08/2012)
5. Kerali HGR (2000) Overview of HDM-4 – Volume 1. International Study of Highway Development and Management Series. World Road Association (PIARC), Paris, France
6. Kerali HGR, McMullen D and Odoki JB (2000) Applications Guide – Volume 2. International Study of Highway Development and Management Series. World Road Association (PIARC), Paris, France
7. Odoki JB and Kerali HGR (2000) Analytical Framework and Model Descriptions – Volume 4. International Study of Highway Development and Management Series. World Road Association (PIARC), Paris, France
8. Paterson WDO (1987) Road Deterioration and Maintenance Effects, Models for Planning and Management. A World Bank Publication. The John Hopkins University Press, Baltimore, MD, USA.

Intelligently adding intelligence

Abstract

All decisions in winter service are based on input information. Decision makers trust this information whilst recognising its limitations. A weather forecast is generated using highly complex algorithms; however, human intervention is often required. Although highly reliant on model outputs, quality weather forecasts require the injection of further intelligence.

Decision making in winter service is also increasingly reliant on algorithms and the quest for higher levels of refinement comes at a cost. In the current economic climate it may not always be affordable to achieve high accuracy, therefore it is important to determine what can be done at lower cost. The fundamental questions are "How much accuracy do we need for different decision making activities?" and "Where is the intelligence best added?".

This paper considers data use in winter service and explores how the injection of intelligence into uncertain data sets can yield useful, lower cost results. It includes an outline of an experimental, long-range forecasting decision support tool that shows promise for salt stock management, resource planning and other benefits outside of winter service.

Paul Williams

Principal Consultant

Highways & Transportation

Atkins

Introduction

Winter service remains a high profile and politically important part of highway maintenance. This is to be expected given that transport disruption resulting from a typical winter in England was valued by the 2010 Winter Resilience Review at approximately £1bn. This is compared to the average annual spend on winter service in England of £160m.[1]

The delivery of winter service in the UK is heavily reliant upon data feeds and network based information. Detailed road weather forecasts, designated weather stations providing roadside observations and GPS tracked vehicles all contribute to a well informed operation.

Decision makers trust this information whilst recognising its limitations. A weather forecast is generated using highly complex algorithms requiring super-computing power, however, human intervention is often required. Although highly reliant on computer model outputs, high quality weather forecasts often require the addition of further human intelligence.

This paper considers data use in winter service and explores how the injection of intelligence into uncertain data sets can yield useful, low cost results. The findings to date show potential benefits to winter service managers responsible for both local and strategic road networks. In addition there is scope for benefits to be gained outside of the core winter service operation.

Data proliferation

The past 10-15 years has seen an explosion in the availability and use of data in most areas of life. Highway winter service is no different with dramatic increases in the collection, storage and dissemination of data. A state of the art winter service is likely to be reliant upon highly sophisticated route based weather forecasts. There will be numerous weather stations logging and reporting actual weather conditions every few minutes, 24 hours a day. The salt spreading vehicles will be tracked and reporting back not only their position but also their actions as they proceed along their route. Decisions to treat may even be made with semi-automated decision support tools which log the what, why and when of the decision making process.

Data is becoming absolutely critical to the delivery of winter service. As more data becomes available, more sophisticated tools are created to make good use of it. This creates an ever increasing cycle driving for increased accuracy and certainty. Ultimately this is driven by the desire to deliver a safe and appropriate winter service for as little financial cost as possible. However, as sophistication increases there is a risk that we lose sight of the limitations of the data source and what we are trying to achieve.

Decision support tools are developed to help to manage a defined problem. The recent proliferation of data provides the developer with more choice regarding the data set(s) they adopt and analyse.

Weather forecasting

As a deliberately controversial statement, weather forecasts are highly educated guesses. They are a prediction of the future based upon past recorded information and some highly complex modelling. The models are so complex that they require ever increasing supercomputing power to run. However, even with all the computing power and some of the finest minds, the weather forecasters still struggle to provide accurate long range forecasts. This is to be expected as there are so many variables, dependencies and interlinks within the modelled system. The UK Met Office state in their Long Range Forecast User Guide that:

"Long-range predictions are unlike weather forecasts for the next few days. The nature of our atmosphere means it is not possible to predict the weather on a particular day months to years ahead. At this range we have to acknowledge that many outcomes remain possible, even though only one can eventually happen. Over the course of a whole season, year or decade, however, factors in the global weather system may act to make some outcomes more likely than others."[2]

This does not mean that the outputs are not useful, it is just they have in-built uncertainty which needs to be understood and accommodated.

It is important to understand how the weather forecasts employed to make decisions are constructed. The following description is a generic summary of the type of process. There will be variation between different forecast organisations, however, it illustrates that there is uncertainty within every forecast.

The forecast model(s) rely on millions of data readings of current and past events from across the globe. These are fed into the computer and the complex algorithms run to predict the future. However there is no single 'answer'. The model may run hundreds of times with subtly different inputs resulting in a collection of 'answers'. These 'answers' can then be laid side by side and the most probable 'answer' identified. This is then output as 'the' current prediction of future weather events according to the computer. However this computer output is purely algorithm based. For short term, detailed forecasting a human weather forecaster will often use their knowledge and experience to refine the forecast. Despite the huge computing power and complexity of the algorithms there is still a need to apply some human intelligence.

Uncertainty

All data comes with some form of uncertainty. For example, the weather conditions reported by weather stations are not 100% accurate; they are the sensor's interpretation of conditions. The accuracy varies with what is being measured and the quality of the sensor. Winter service in the UK relies upon a fleet of weather stations with good quality sensors resulting in small amounts of error. However, this may not be the case elsewhere in the world.

As has been described above, weather forecasting has uncertainty built into it. Moreover because it is a prediction of events yet to occur it is not possible to confirm if the forecast is correct at the time of production.

Decision making has to embrace the various uncertainties and define a course of action. This is something that engineers are generally well equipped to do, however, it is important to understand uncertainty when making decisions. Making a decision without this understanding is potentially high risk as the decision has to be appropriate for not just the predicted conditions but also other credible scenarios. This is why a decision to monitor through the night may be made on marginal nights. The forecast says it will not freeze, however, the decision maker realises that there is a chance that the conditions will vary slightly and preparations have to be made.

Dennis Lindley is a British statistician who stated in his 2006 book "Understanding Uncertainty":

"There are some things that you know to be true, and others that you know to be false; yet, despite this extensive knowledge that you have, there remain many things whose truth or falsity is not known to you. We say that you are uncertain about them. You are uncertain, to varying degrees, about everything in the future; much of the past is hidden from you; and there is a lot of the present about which you do not have full information. Uncertainty is everywhere and you cannot escape from it."[3]

On this basis it is important to understand the uncertainty and if possible make use of it.

Decision support

Accompanied by the proliferation of data is a proliferation of decision support tools. These process the data in useful ways to provide guidance to decision makers. Some tools go as far as presenting an answer to the decision maker for acceptance. These tools are beneficial to the industry but they are not risk free.

As decision support tools become increasingly complex and widespread there is a risk that the ability to manually make the decision is lost. This would be a significant issue if the tool were to be unavailable or if the tool provided 'wrong' answers. Structural engineers are trained to carry out analysis manually so they do not have to rely on finite element analysis software and also so they can spot when the results do not look right. Winter decision making is no different.

There is also a risk that the decision support tools promote false confidence. The tools provide an answer to a problem which is helpful as long as the decision maker understands the limitations and accuracy of the information provided. Winter decision making involves analogue systems where there are ranges of possibilities. A decision support tool may only provide a yes / no answer to the most credible scenario. There will, however, be a range of credible scenarios which may need to be understood.

Understanding how decisions are reached, the limitations of the inputs and the 'status' of the outputs is critical to the safe use of decision support tools.

Drivers for thinking differently

In the UK, the severe winters of 2008-09, 2009-10 and 2010-11 highlighted that salt was not an endlessly available commodity. Salt is the primary de-icer in the UK and therefore any disruption to the supply during the winter season is a significant challenge to the safe operation of the road network. There is plenty of salt available on Earth, however, in the UK, the issue in those winters was producing it rapidly in a form suitable for road treatment. There were numerous reviews and lesson learning exercises but one of the key themes around salt supply is being able to better predict medium / long term demand within a winter season. Strategic salt stocks work well to safeguard against future national crises, however, better management of operational salt stocks is an important part of the jigsaw. The answer is not simply to hold ever increasing amounts of salt in stock as this is a significant drain on available cash. Therefore a mechanism is needed to help better predict salt demand over a number of months to enable timely restocking.

Traditionally winter service decision making has focused on the 'here and now'. Most winter service forecasts only looks five to ten days ahead. This is simply not enough time for salt suppliers to react to a sudden surge in demand.

The biggest problem faced by winter service managers was available input data. Relying on past salt use and using rolling averages has recently been proven as inadequate. Typically the calculations were done on a season by season basis so it was not refined enough to predict spikes in demand during a season. Given the variability of winter patterns in the UK using more detailed salt use figures and applying more complex statistical analyses still could not predict potential future spikes. Climate change or even the phenomenon some call “Global Weirding” make the reliance on past information an illogical solution.

The only option was to use long range weather forecasting. This could provide a salt forecast model of the coming months. However, it is well known that long range forecasting contains a high degree of uncertainty and there is no immediate prospect of significant improvements in accuracy. So the question was how can we use this ‘dirty’ data set to create a ‘clean’ answer?

Thinking in systems

The solution came as a result of a basic systems thinking approach. Systems thinking is an increasingly popular way of approaching problems. It accepts that the problem may not be a ‘component’ but is as likely to be between ‘components’. The quote below from Ballé in 1994 sums this up:

“To understand things we take them apart and study the pieces. To improve things we try to improve pieces individually. It is rather like trying to get a horse to run faster by teaching each of the legs to perform a more efficient movement ….. The systems approach focuses on the inter-relationships, how the horse’s legs relate to each other and back to the horse.”[4]

The concept that when ‘A’ happens it has an impact on ‘B’ which in turn impacts on ‘C’, ‘D’ and ‘E’ is important. This defined the approach to allow the uncertain forecast data to provide useful information. Donella Meadows provides a very relevant insight in her 2012 “Thinking in Systems: A Primer” book.

“As our world continues to change rapidly and become more complex, systems thinking will help us manage, adapt, and see the wide range of choices we have before us. It is a way of thinking that gives us the freedom to identify root causes of problems and see new opportunities.”[5]

We understand that salt stocks diminish when road treatments occur and there is no restocking. Treatments only occur when the road surface is cold and there is moisture available to form snow, frost or ice. The road surface is cold at the same approximate times that the air is cold. Given we could obtain forecasts of likely conditions we therefore could calculate a potential salting scenario. This salting scenario will outline the likely future treatments which in turn would allow the generation of a forecast salt stock profile. This may sound trivial, however, it allowed a link between the available forecasts and the desired answer. It also highlighted that the tool could have potential to contribute to much wider decisions than just salt stock management.

Taking the example of the 2008-09 winter, if numerous decision makers had access to information predicting salt usage then the impending crisis could have been foreseen earlier and as orders increased the salt producers would have been able to increase production. Everyone had access to long range weather forecasts but there is little evidence to suggest that many made the link and acted upon it.

There is a direct relationship between salt usage and number of treatments. This means that the same ‘answer’ would allow forecasting of labour and plant demands allowing better management of the winter service. It also has wider potential as a forecast spike in salting activity is also likely to mean an increased number of calls to contact centres, more demand on council support services and an increased demand on doctors due to respiratory illness. It is all part of ‘the system’.

It was apparent that there were likely to be significant benefits if the problem could be solved. The various pieces of the jigsaw existed but it now had to be put together.

Developing the idea

Continuing the ‘systems’ based approach the first step was to identify inputs, constraints, linkages and the desired outcomes. This allowed the problem to be approached as efficiently as possible. A key point was the aim to produce a relatively low-tech output to ensure it was applicable and understandable to as wide a range of people as possible.

Forecast data input

The first activity was to examine the commercially available forecast information. This is where the recent proliferation of data helped. It is possible to purchase raw weather forecast model data over the internet for modest sums of money. There was an option to purchase higher cost, forecaster intervened data, however, the project was an experiment to identify the potential use of raw un-intervened model data. The option remains for the future to consider the costs and potential benefits of a move to intervened data.

Having reviewed the readily available forecast data the Global Forecasting System (GFS) model and a post processed Climate Forecast System (CFS) model were chosen. By combining these two forecast data sets we gained an accessible total forecast period of 60 days in XML format.

Constraints

The ideal data would have provided road surface temperature forecasts and an indication of the likely hazards in terms of frost, ice or snow. However, this was not possible so the first constraint was the available data. A mechanism to convert the available data into more useful information was required.

The forecast data did not supply the overnight minimum temperature which is the standard decision making reference point. The forecast minimums at different times of day and night were available, therefore the worst forecast condition was adopted as the minimum overnight temperature. This clearly inserts further uncertainty.

It was understood that there was uncertainty around the accuracy of the forecast data. The longer the range, the more inaccurate; however, there was no ready way to determine how big the error would be. It was not possible to change the accuracy levels available and therefore a way to embrace this uncertainty and drive benefit from the available data was needed.

The forecasts only contained atmospheric data and therefore could not be directly translated into winter treatments. The decision making guidance in the UK relies upon road surface temperature. It was understood that there is no direct relationship between road and air temperature therefore an assumption would be required.

The weather varies with location and at times can vary significantly within a small geographic area. Forecast information was available for numerous points around the UK, however, each additional forecast point added to the cost of the data and also the complexity of the tool. Ideally hundreds of forecast points would have been ingested and processed, however, this was not practical in initial development. Seven points were chosen to provide a spread of forecasting across the UK. These were in ‘average’ locations as opposed to the more unusual locations such as tops of hills or adjacent to the coast.

Winter treatment decision making varies with different local authorities and agencies therefore the amount of salt used for a given set of weather conditions will vary across the UK. Whilst it may be feasible to utilise multiple different treatment guides within the decision making

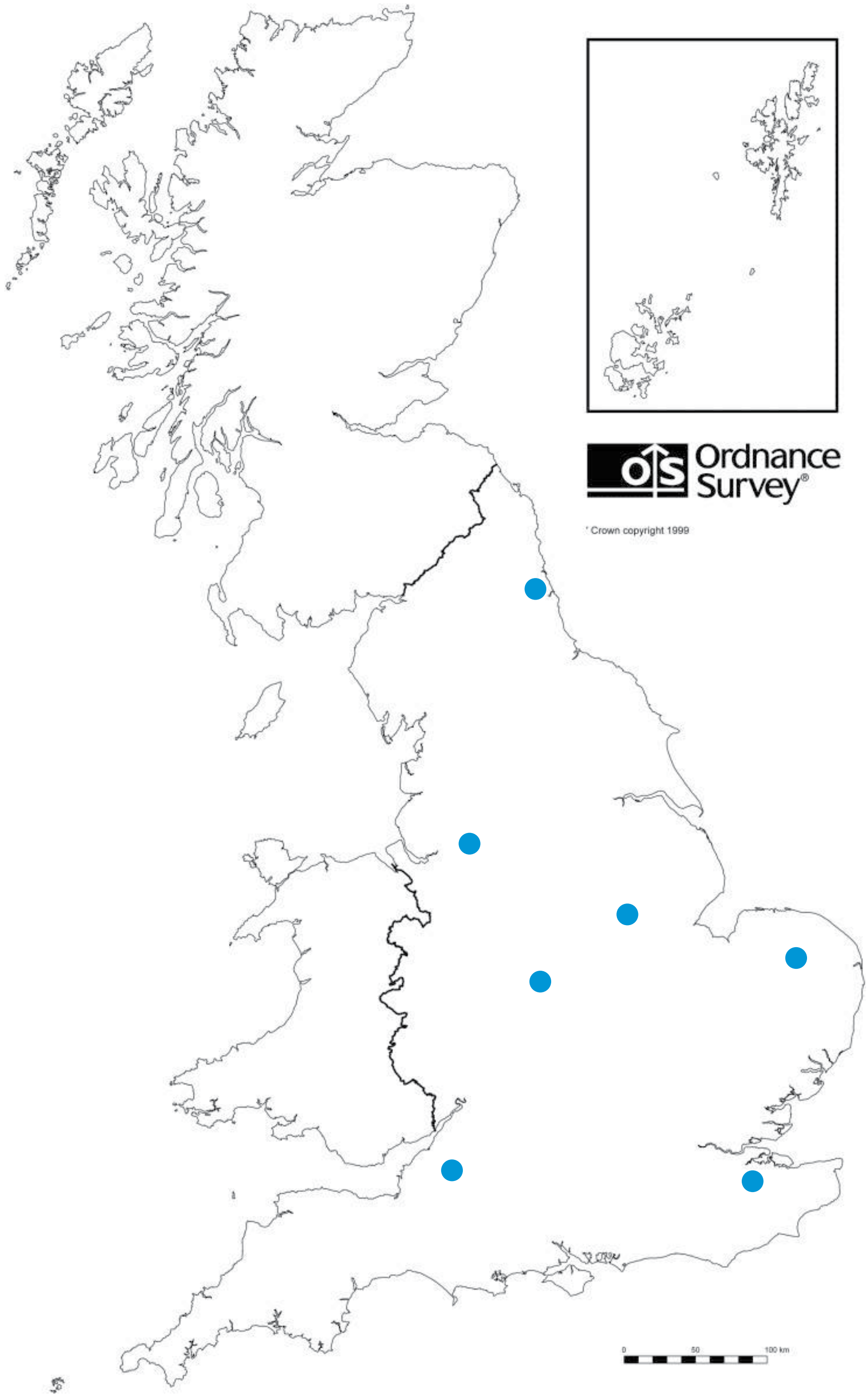


Figure 1. Map of forecast points

algorithm it would significantly increase the complexity and it is not clear if it would yield any significant benefits. A typical treatment decision guide would be needed to maintain a simple algorithm.

This pilot project was highly innovative and there was little to benchmark against. It was important to calibrate and test the output against actual data. Since the 2008-09 winter, the UK records its salt usage in more detail and therefore there was the potential of a data set to test against.

Linkages

The intention of building a tool to help predict salt usage opened up the opportunity to provide information to a wider audience than just the winter decision maker. As mentioned previously a forecast spike in salting activity is also likely to mean an increased demand on other public services.

The outputs could influence a significant range of a local authority's services. The Winter Service Manager is obviously interested in being able to manage their salt stocks, fleet and drivers. The output will assist with all of this but will impact on other areas. For example if salt suppliers see an influx of early customer orders they are likely to increase production to meet the need. Therefore the output at a local authority level could help the salt producers. Obviously the producers could adopt a national model to aid their planning, however, it does not tell them when the orders will actually be placed by customers.

The plant and labour needs of winter service will often impact on the ability of a local authority to deliver other activities. Therefore if a spike in winter treatments is expected then it is likely that other services will have to reduce or put in place plans to bring in additional resources.

Taking this one step further, there are numerous local authority services which are not directly impacted by the winter service activities, however, they rely on the highway network to deliver their service. For example, carers visiting vulnerable people may struggle to get to them all if the road network is not fully available.

There are also local authority services that are impacted by the weather and can have a significant impact on the local economy. For example businesses may struggle to operate normally if their staff have childcare issues following school closures due to preventable issues such as heating failures. Forward planning may be able to prevent some of the school closures. Snow is obviously a different issue.

Finally there is the public. It would not be appropriate to disseminate the outputs of any tool based on such uncertain data to a 'layman'. However, it would allow planning of media campaigns such that as a cold spell approaches the media managers are well prepared to engage, warn and inform as is required by the UK Civil Contingencies Act: 2004.

It is therefore clear that there are numerous potential interested parties. The challenge is how to deliver usable output to meet as many needs as possible.

Outcomes

The desired outcome is simply a tool which provides guidance on the potential future winter service demands. This can then be employed in different ways to yield benefit.

It is clear from the number of linkages that there are numerous potential interested parties. The challenge is how to deliver usable output to meet as many needs as possible. Some of the intelligence from the forecasting inputs had been consciously omitted so a decision was needed to determine the best places to inject more intelligence.

The amount of uncertainty within the forecast data and the number of assumptions needed to process the data mean that a 'single answer' output is unlikely to be useful. The outputs will have to be potential scenarios which can then be used by a competent person to make an informed decision. In doing this some of the uncertainty is absorbed. Indeed a 'range answer' is only possible where there is uncertainty therefore this approach is capitalising on the uncertainty. It paves the way for techniques such as sensitivity analysis to provide confidence levels for the decision maker.

It was realised that this could only be a decision support tool and not a decision making tool.

Pilot implementation

Having undertaken the initial research it was clear that an experimental pilot should be developed. This was against a backdrop of the second severe winter in the UK and another potential salt crisis. It was realised the tool would not be suitable for operational decision making but it may yield some useful results.

Assumptions

The first step was to consider the assumptions needed. This was probably the single most significant injection of intelligence into the tool. Using experienced winter decision makers who had received advanced meteorology training it was possible to translate a standard treatment decision guide into a set of rules which could be used against the available forecast data.

There were actually three sets of rules established to provide a high, medium and low scenario. These scenarios effectively took optimistic, pessimistic and neutral assumptions based on the weather forecast data (assumption variance). This assumed that the forecast was accurate and the 'rule' contained all the error.

However, these scenarios could also be seen as translating the single forecast into an optimistic, pessimistic and neutral forecast (forecast variance). This assumes the 'rules' are accurate and creates an artificial variance in the forecast.

The truth is likely to be somewhere in the middle with the forecast and assumptions both being incorrect (balanced output). This muddies the waters of uncertainty somewhat but **Figure 2** shows these concepts diagrammatically. It requires the user to take a step away from pursuing the definitive answer and accepting that there is a range of answers.

It does assume that the forecast and assumptions will not both be very wrong in the same direction. However, given the 'neutral forecast' and 'neutral assumptions' are the most likely results it is likely that the true output will be close to the centre of the output range.

This approach is not overly scientific, however, it does only aim to give the most probable range of results. It is conceivable that the true answer will lie outside the results range, however, this needs to be accepted by the user. It is a salt forecast and therefore may not be the perfect answer.

By varying the scenario settings it would be possible to carry out a rudimentary sensitivity analysis. This would provide an indication of the confidence the decision maker should have in the output.

The assumptions used air temperature, snow risk and precipitation amount as the main indicators to the conditions. The treatment guides were reverse engineered such that a set of conditions were matched to each treatment option. It was even possible to identify where multiple treatments could be required. This allowed, for an assumed network area, the calculation of the amount of salt required to deliver the treatments.

Data handling

Having obtained the two forecast data feeds these were initially parsed into a daily single forecast data set. This was the combined GFS and CFS model data. This data feed was processed against the rules to provide three different treatments and therefore salt use scenarios.

The processing in the initial pilot was carried out via a multi-sheet Microsoft Excel workbook. It used standard functions and can be performed without the need for Visual Basic programming. However future versions may benefit from a more sophisticated computing solution. Once the forecast data had been loaded the processing time was typically under 10 seconds.

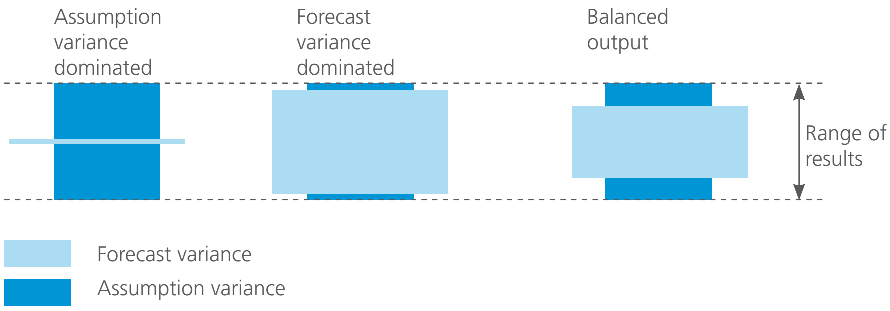


Figure 2. Uncertainty in the model

The three scenarios are then plotted on a graph against time to provide both instantaneous and cumulative salt use. The instantaneous use graph allows ready identification of spikes of winter service activity. Given a known starting stock, the cumulative graph can be used to model the salt stock diminishing with use. This can then trigger thresholds for restocking.

Calibration

The first working edition of the tool required calibration. Assumptions had been made and it was not certain if these were correct. It is akin to manufacturing a clock, setting the time for the first time and simply hoping it kept the correct time. The output required baselining against known results.

It was possible to access some Highways Agency salt usage figures which allowed calibration. The model's results were reviewed against the now known outcomes. By varying the assumptions it was possible to generate salt use forecasts that aligned to the actual results. The real test was then the tool's ability to forecast salt use into the future using the newly calibrated assumptions.

Refinement

The tool produced some interesting results over a period of time. For example during the severe weather event of December 2010 the model predicted that the UK salt use would dramatically reduce prior to salt levels falling to critical levels. The actual salt stocks did move as the tool had predicted with a sudden stop to the severe conditions that had put such strain on the winter service across the UK.

Over an extended period the settings were refined to what is thought to be a good 'all-round' set of rules for the UK as a whole. However there is further refinement to do and it is anticipated that the rules will need validating for each different network prior to any operational use.

When re-running models with slightly different rule settings it was possible to carry out a basic sensitivity analysis. Through prudent use by competent people it is possible to refine the output to provide the decision makers with a degree of confidence in the results. However, there is an element of 'gut' feel required. It reverts back to the concept of forecaster intervened weather forecasts. The computer models suffice the majority of the time, however, sometimes you can't beat the human touch.

This is as far as the research has gone to date. There is a working model in existence which appears to offer promising results. Taking a national view it has produced workable results which, at times, have been shared with those responsible for supervising reserve salt stocks in the UK. The results may not have been acted directly upon but they were available if required.

Example outputs

A number of output graphs have been included in the paper to provide an impression of the output that can be generated. The visualisation in the final edition is likely to change as this is currently based on a Microsoft Excel spreadsheet and charting function.

Figure 3 is a good example of the weather forecast driving likely salt use over the Christmas 2012 period. From the graph a winter service manager would be able to see that there is some activity throughout the period however the majority of the activity falls after the Christmas period. This would be an important piece of information for both salt and resource management.

Going into a little more detail the 'neutral scenario' line sits closer to the 'low use scenario' than the 'high use scenario'. Without any further analysis this indicates that the true situation is likely to be closer to 'low' than 'high'. This is very basic but it highlights the potential benefits sensitivity analysis could bring.

Long term forecasting generally gets the 'shape' of the temperature over time graph correct. The precise timing is not necessarily correct and potentially the precise temperature is also inaccurate. However, getting the shape correct is important. **Figure 3** also shows that the three scenarios accentuate the same shape of curve. The 'high use scenario' amplifies the peaks more than the other lines.

Using some Highways Agency actual treatment information it was possible to draw an approximate comparison to the forecast model.

Figure 4 shows the average number of treatments delivered and it is clear that the model is capable of following the overall 'shape' of the service. At times the timing is slightly incorrect, however, considering the level of data being input and the range of assumptions being made

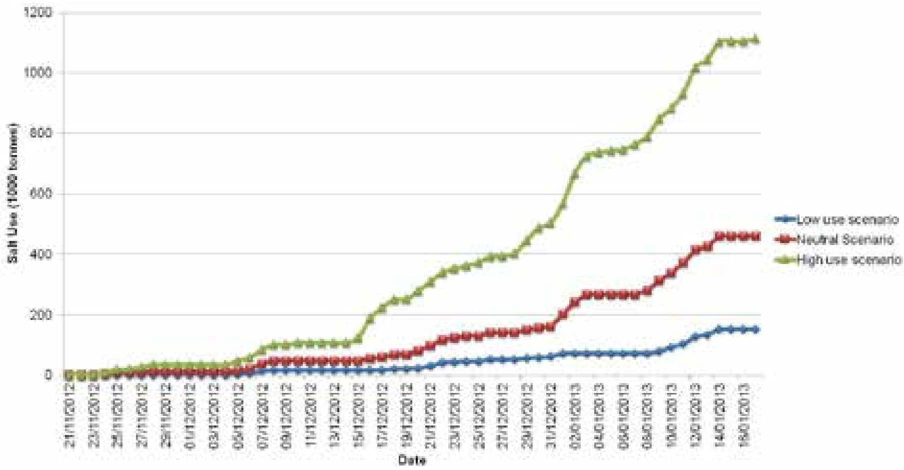


Figure 3. Results graph from the Christmas period 2012

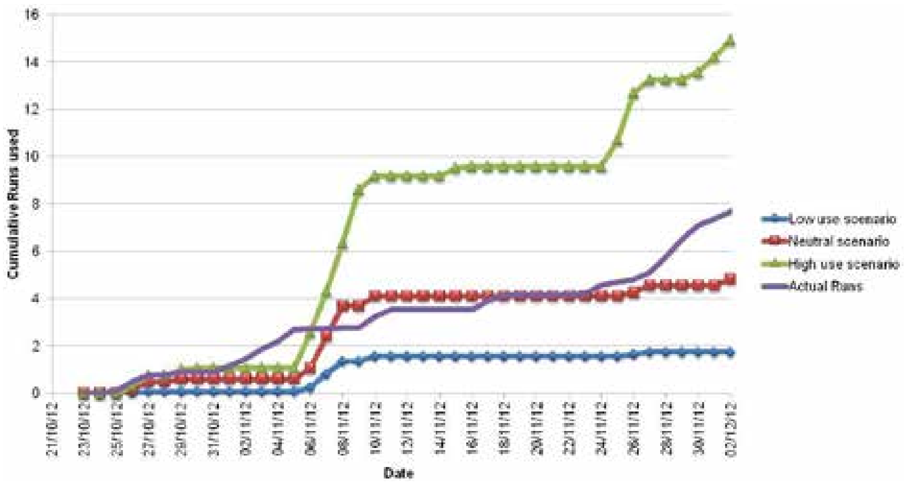


Figure 4. Results comparison

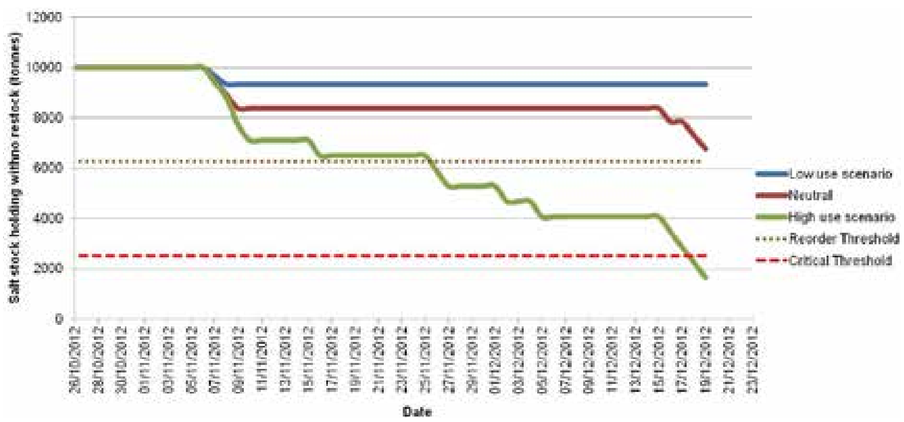


Figure 5. Salt stock model

the results are encouraging.

A visualisation was developed to help Winter Service Managers to forecast at county level. **Figure 5** is an example output showing how salt stocks may vary over the coming period.

Figure 5 shows how the previous two outputs could be translated into a forecast stock holding. The re-order threshold is purely as an example which shows that the point where re-ordering should occur could be as early as mid November but is more likely to be mid December. This should aid decision making in terms of timing of salt ordering. This example also shows that there is a possibility that the service could become compromised as early as mid December if restocking does not occur.

Lessons and the future

The pilot project was created out of a need for longer term decision guidance and a curiosity to see what can be achieved at low cost. Along the route a great deal has been learnt not only about weather but also about how data can be used. It has enabled the identification of ways to take the pilot forward.

Lessons learned

The concept of uncertainty is an important one in long term planning. It is simply not possible to create certainty over an extended period, especially if natural systems are implicitly linked to the activity. Rather than expend significant efforts failing to achieve the impossible the pilot embraced the uncertainty.

The decision not to use forecaster intervened weather forecasts did increase uncertainty however the amount of benefit did not outweigh the costs of the more detailed forecasts. The key to the decision tool was translating weather data into treatment information and ultimately salt use forecasts. By injecting the intelligence into the right part of the system it has shown that usable results can be achieved even using low-cost 'dirty' data. However, currently it does require an intelligent user to ensure the results are appropriate.

Decision support tools are a useful and inevitable part of winter service. However, it is important to question the inbuilt uncertainty around the inputs and the algorithms. If basing decisions wholly on the outputs of tools it is important that the output is truly understood.

When considering decisions and decision support it is important to consider not only the activity that the tool directly supports but the subsequent activities it influences. The systems thinking approach helps identify linkages and drive benefit from the tools you have available.

Planning for the future

The pilot project has used coarse data inputs and tested against national figures. It appears possible to forecast salt use down to a regional level, however, with an increased data set it is thought that this could be further refined to provide more detailed salt forecasting. This would then benefit individual local authorities in their salt stock management.

The forecast data we ingested works well, however, this is low cost data and with the advances over the recent years in long term forecasting a new and improved data set may now be available.

By reviewing historic long term weather forecasts against an authority's salt use of the same period, the model could be calibrated and validated to a specific network.

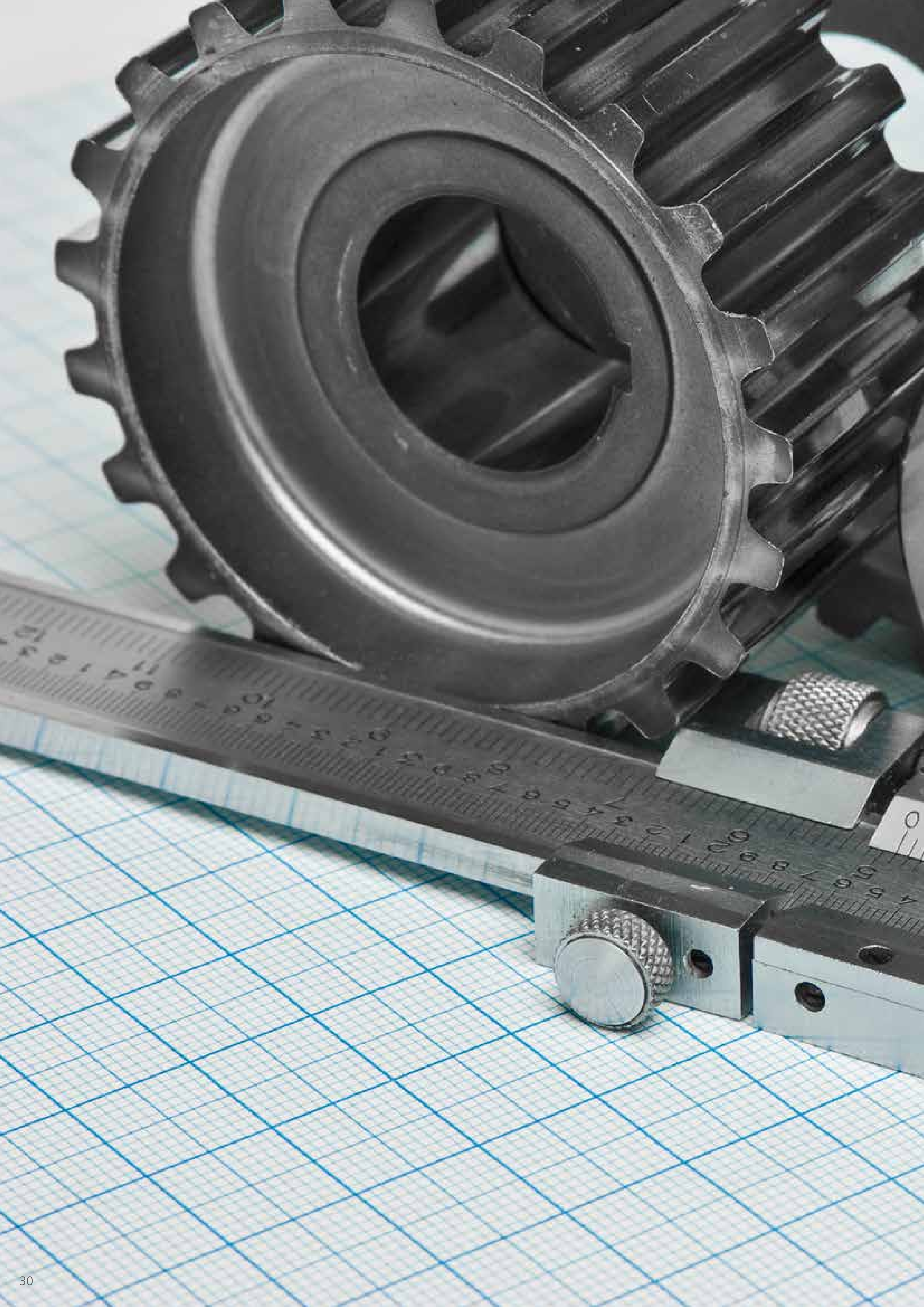
It is hoped that operational trials with highway authorities can be established to test the usefulness of the tool.

Further work is also needed to identify scope to integrate the tool, or similar tools, into the wider operational planning of local authority services. The tool currently serves tactical level managers in the highway department but it has potential to provide guidance at a more strategic level to other departments. Emergency planning, social services, education and health are all areas where long term identification of severe weather is important. They are also areas of budgetary constraint.

It is important in times of economic restraint to maximise the benefits from any investment you make. A business case is needed to identify the core and peripheral benefits, attempt to monetise these and then offset against the cost of implementing the tool. It is hoped that this business case can be developed with 'live' operational partners in the near future.

References

1. Quarmby, D (2010) The Resilience of England's Transport Systems in Winter, Para 109
2. http://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/long-range/user-guide
3. Lindley, D. (2006). Understanding Uncertainty, Preface pp xi
4. Ballé, M (1994) Managing with Systems Thinking. pp35
5. Meadows, D (2009). Thinking in Systems: A Primer. pp17.



152

Cyber Security

The use of static analysis to detect malware in embedded systems

Abstract

The use of formal static analysis techniques to ensure that software-based safety systems are free from compiler introduced errors is well established (Pavey, Winsborrow, 1995)[1]. This technique ensures that the executable binary code created by the compiler is mathematically equivalent to the original source code. This paper reports on extending this technique to detect malware inserted into executable code. The source-code comparison process was originally developed by British Energy for the verification of the Primary Reactor Protection System software of the Sizewell 'B' Nuclear Power Plant. The process takes the executable binary file that is resident on the target computer and recreates the equivalent assembler code using disassembler tools. This is then formally compared to the original source code using the MALPAS Compliance Analysis tool, and any discrepancies are revealed. The process has the ability to detect any executable binary code that cannot be traced back to the source code, and may therefore be used to detect the presence of malware in the executable. The paper will report on experiments conducted by Atkins to determine whether malware that has been deliberately inserted into the executable can be detected using formal proof. The applicability of the process to software developed for general purpose operating systems (e.g. Windows) will also be evaluated.

Chris Sampson

Principal Consultant

Defence, Aerospace & Communications

Atkins

John Drever

Innovation Director

Defence, Aerospace & Communications

Atkins

Brian Third

Senior Security Consultant

Defence, Aerospace & Communications

Atkins

Introduction

Many of our most trusted devices, critical to our everyday life and wellbeing, contain embedded software. From our cars, to our phones, heating and cooling, lighting, domestic appliances, and televisions all contain embedded microprocessors running software. Further, our national infrastructure: communications systems, water supply, food production, electricity and gas, petrol pumps, road traffic signalling and trains – things we rely on to keep our economy moving - all rely on software. We also rely on embedded software to keep us safe – for example in nuclear reactors, medical devices, avionics and air traffic control, police and defence systems. Embedded software is prolific and vital to our everyday lives.

Traditional malware detection and prevention relies on two key techniques:

1. Detecting when the embedded software is doing something that it probably shouldn't be doing (eg transmitting data, accessing information)
2. Detecting patterns of software known as being malware.

Preliminary results from Symantec published in 2008 suggested that "the release rate of malicious code and other unwanted programs may be exceeding that of legitimate software applications."[2] According to F-Secure, "As much malware [was] produced in 2007 as in the previous 20 years altogether[3]. malware is prolific and threatens our everyday lives.

This paper presents another detection and prevention method using a technique previously employed to detect errors in compilers – formal proof that embedded software is mathematically identical to the source code that generated it.

Malicious hackers use malware to:

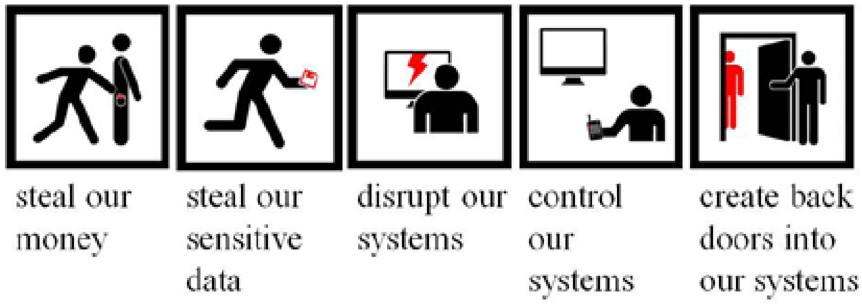


Figure 1. Malware attacks

How does malware enter embedded systems and what do we currently do about it?

The software lifecycle

malware attacks on our desktops and laptops are well understood; backdoor code, Trojan horses, worms and viruses are all well-known techniques. However attacks on embedded software are more complex; many embedded applications are not open to the internet and communicate only across segregated and tightly controlled networks.

To understand the vulnerabilities, it is worth looking at a typical embedded software development lifecycle (**Figure 2**).

Here are a few methods of inserting malware into an embedded system:

1. Develop malicious code (or pay a developer to write some code) during the development lifecycle. This could be mitigated during a standard peer review process or independent test, but there is a residual risk that peer review and module test could miss something very subtle
2. Create malicious code and hide it from other developers – then include the malicious code at compilation time. Independent testing might find this, however most testing is based on positive affirmation that functions are correctly coded. There might be a possibility of finding this code if there is MCDC and full path coverage at assembler level
3. Include malicious code at target cross-compilation time. System testing might find this if you are very lucky. If it is common malware, then malware detection could pick up (if you run the malware scanner on the target code!). If you have exceptional defensive code, you might detect something strange happening during run time
4. Develop code which modifies your code during patches. Patches can be verified using hashing algorithms and comparing the hash output to a fixed length signature. However, there is a risk that patches are not always checked for integrity
5. Include malicious code while the binary file is being transferred to the target hardware. This technique is very hard to detect and can have a high impact if successful
6. Write code which modifies your code during run-time. Real-time analysis would be required for this to be detected. However, the real-time analysis could affect the system and in itself compromise system integrity.

Current methods for detecting malware

Currently detecting embedded malware is a complex process involving advanced forensic techniques and cryptographic hashing functions. The main issues surrounding these techniques are the high cost of maintenance and the technical issues which can affect the control systems which they are designed to protect.

There are three main techniques of detecting malware:

1. Anomaly based detection: after the detection engine has learned what forms a clean/safe environment, the detection engine can then detect malware which forces the system outside normal operation. The drawback of this type of detection is the high false alarm rate
2. Specification based detection: in order to reduce the high false alarm rate of anomaly based detection, specification based detection attempts to approximate the requirements of a system as opposed to the behaviour of a system
3. Signature based detection: signature detection uses known attack patterns to detect attacks on a system. This is a good approach, however, if an unknown attack is launched against a system then it is unlikely to be detected as the signature has not been added to the detection database.

Each technique has its uses, mainly in a corporate environment. The main issue with each of the techniques is the level in which they interact with the system. These techniques are great at detecting network based attacks however the systems have to be taken off line in order to check for malware on each device. This approach may be unsatisfactory if the system has long periods of run-time and little maintenance down-time.

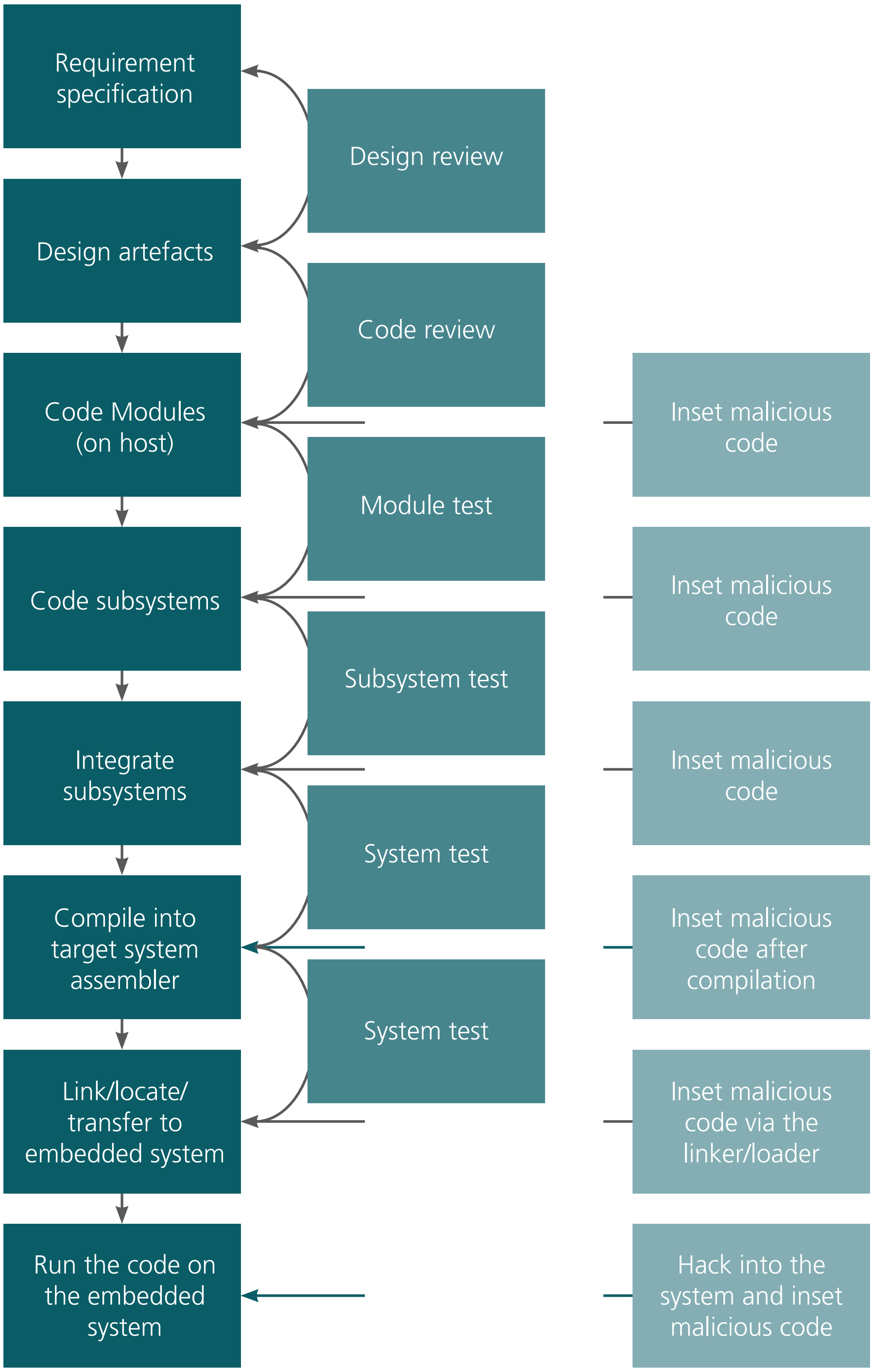


Figure 2. Standard software lifecycle and vulnerabilities

In addition to the limitations to detection, it is still possible that the malware has been designed to evade the malware detection engine by learning how the system operates and adapting accordingly. This is a technique used by some of the newest pieces of malware available.

Formal proof that the executable is the same as the code

Introduction

During the development of the Sizewell B nuclear power station, a software verification method was devised to identify any defects introduced into the Reactor Protection System software due to compiler errors. There was a worry that the compilation process used to generate the code for the target system may include bugs that might cause safety issues. The process devised made use of the MALPAS toolset to mathematically prove that the binary image on the embedded system was identical to the high level language code (PLM 86). The vital part of MALPAS that is unique in this context is the ability to mix third generation language and assembler in a single mathematical model through the use of the MALPAS Intermediate Language (IL).

Further to this, Atkins recently performed a feasibility study for EDF and Areva that proved a similar approach for Teleperm XS using the C language and Intel Assembler.

From a safety perspective we were looking to prove:

1. Program state. The process checks that all writes to memory and the stack are semantically consistent between the source code and target code
2. Constant initialisation. Note that as a special case of item a), the process will need to check that any initial values loaded directly into memory by the compiler are consistent between the target and source
3. Procedure calls. The process checks that the address of a procedure called in the ASM and the equivalent C function call in the source are consistent
4. Stack consistency. The process should check that the value of the stack pointer at the end of the procedure is equal to that at the start of the procedure
5. No added code. The process should ensure that no code has been added to the executable that is not traceable to the C source code.

From a security perspective we would want to check number 5 – there is no added code.

Process to check for no added code

The process for checking that there is no added code is simplified below.

Step 1: Disassemble and translate

a. Convert the binary file on the target into Hexadecimal
b. Convert the Hexadecimal File into Assembler
c. Translate the Assembler into MALPAS Mathematical Language (IL)

Step 2: Translate the source code (in this case "C") into IL

Step 3: Perform MALPAS compliance analysis

a. Convert the C IL into malpas_check proof assertions and insert those into the correct places in the ASM IL.
b. Run the MALPAS compliance analyser on the ASM IL.
c. Check the output for discrepancies (there might be compiler errors as well as malware).

Step 4: Report

Experimentation with MALPAS

Method of Experiment

The example below is of a single C statement (**Figure 3**). The Assembler produced by the compiler is given (**Figure 4**), and finally the IL Translation of the assembler (**Figure 5**). Note that the code assertion is given by the "malpas_check()" statement. It is necessary to replace the variable references in the C Translation by stack accesses (as this is how the variables are located by the compiler).

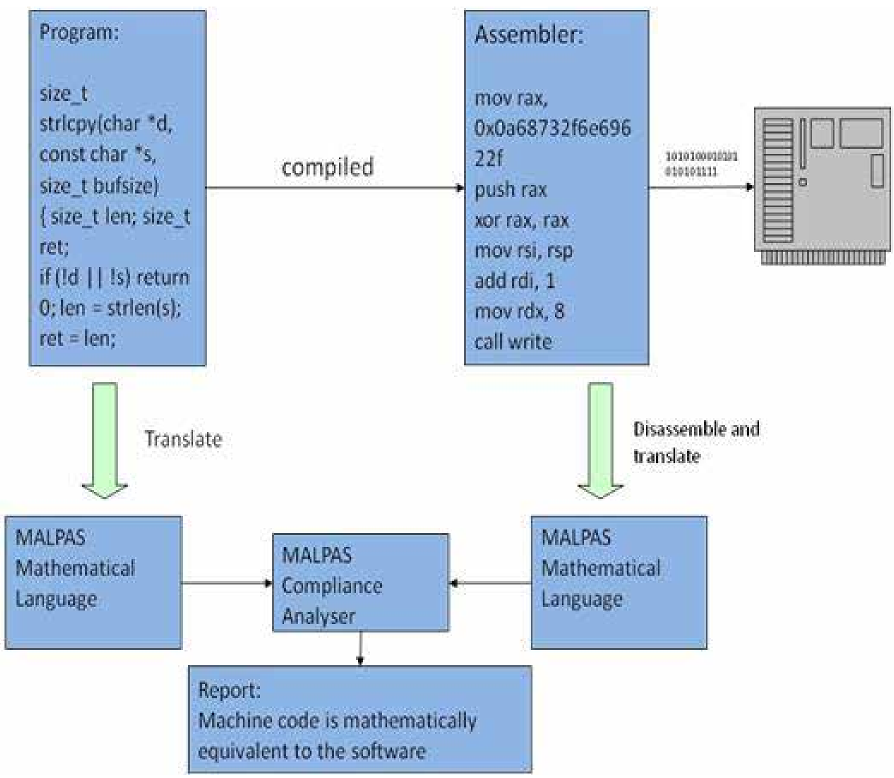


Figure 3. Simplified process for proving "no added code"

Sum = param1.v + param2.v + param3.v;

Figure 4. C Source code

```
                      @2:
006F   C45E08     LES   BX,[BP+args]
0072   26C47704   LES   SI,ES:[BX+4]
0076   26D904     FLD   ES:[SI]
0079   C45E08     LES   BX,[BP+args]
007C   26C47708   LES   SI,ES:[BX+8]
0080   26D804     FADD  ES:[SI]
0083   C45E08     LES   BX,[BP+args]
0086   26C45F0C   LES   BX,ES:[BX+12]
008A   26D807     FADD  ES:[BX]
008D   DD56F2     FST   [BP+Sum]
```

Figure 5. Intel Assembler from disassembled executable image

```
00220:   ; [ @2: ]
00230:   MAP [ LES    BX,[BP+ARGS] ]
             bx := stack !! (bp+args);
             es := stack !! (bp+args + 2);
         ENDMAP;
         bl := low(bx);
         bh := high(bx);

00240:   MAP [ LES    SI,ES:[BX+4] ]
             si := mem !! adr(es, bx+4);
             es := mem !! adr(es, bx+4
             + 2);
         ENDMAP;

00250:   [ FLD    ES:[SI] ]
         fsp := fsp + 1;
         fstack := update (fstack,
                          fsp,
             real_value (mem!!!adr(es,
             si)));

00260:   MAP [ LES    BX,[BP+ARGS] ]
             bx := stack !! (bp+args);
             es := stack !! (bp+args + 2);
         ENDMAP;
         bl := low(bx);
         bh := high(bx);

00270:   MAP [ LES    SI,ES:[BX+8] ]
             si := mem !! adr(es, bx+8);
             es := mem !! adr(es, bx+8
             + 2);
         ENDMAP;

00280:   [ FADD   ES:[SI] ]
         fstack := update (fstack, fsp,
fstack ! (fsp) + real_value (mem!!!adr(es,
si)));

00290:   MAP [ LES    BX,[BP+ARGS] ]
             bx := stack !! (bp+args);
             es := stack !! (bp+args + 2);
         ENDMAP;
         bl := low(bx);
         bh := high(bx);

00300:   MAP [ LES    BX,ES:[BX+12] ]
             bx := mem !! adr(es,
             bx+12);
             es := mem !! adr(es, bx+12
             + 2);
         ENDMAP;
         bl := low(bx);
         bh := high(bx);

00310:   [ FADD   ES:[BX] ]
         fstack := update (fstack, fsp,
fstack ! (fsp) + real_value (mem!!!adr(es,
bx)));

00320:   [ FST    [BP+SUM] ]
         stack := updated(stack,
         bp+dlsum, store_real (fstack !
         fsp));
```

[ASSERTION to check that the C functionality is preserved in the Assembler]

[Sum is replaced with real_value(stack!!!(bp+dlsum)), derived from the memory map produced by the compiler

Note that most of the time C is dealing with the objects whereas ASM is dealing with pointers to the objects which are located in mem or stack. The C objects are therefore replaced by the equivalent stack operations.

]

```
$malpas_check ((real_
value(stack!!!(bp+sum)))
EQ
real_value(deref_dword_pointer
($cmem, deref_dword_
pointer($cmem, $cstack!!!
(bp+args),4),0))
 +real_value(deref_dword_
pointer($cmem, deref_
dword_pointer($cmem,
$cstack!!!(bp+args),8),0))
+real_value(deref_dword_
pointer($cmem, deref_
dword_pointer($cmem,
$cstack!!!(bp+args),12),0)));
```

Figure 6. IL Translation of Assembler, with proof assertion shown in red

A compliance analysis result of threat := false; indicates that the assertion has been proved, and no Malware has been inserted.

Further examples will be developed to demonstrate that Malware inserted into the executable results is a threat at the compliance analysis stage.

Use in practice

This process can be applied to any translatable embedded software. We would envisage that once the target image is analysed once, then it could be saved and checked against the executable image on a regular basis to prove that no changes have been made. This could be performed using a simple checksum utility.

The initial process should only take two to three weeks to execute once the tools have been created to disassemble the code and translate the assembler.

Types of Vulnerable System

We believe that this process would be useful for the following applications:

- Secure communications systems including crypto
- Industrial control systems (civil nuclear, utilities, high value manufacturing)
- High BIL systems (cooling, life support systems)
- Weapon systems

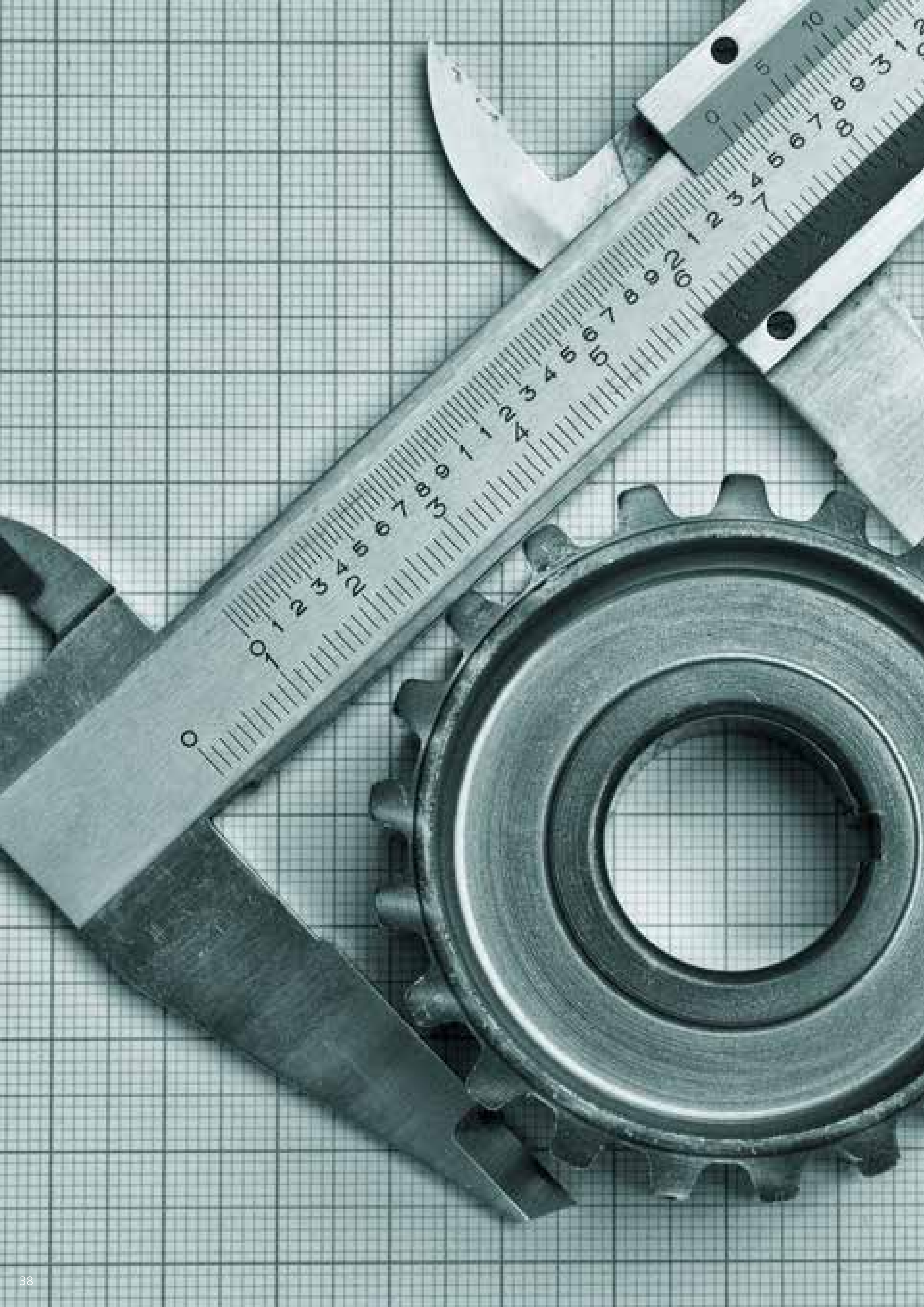
- Air traffic control and avionics
- Financial transaction systems
- EPOS
- Automotive systems
- Robotics.

Acknowledgement

This paper was previously presented at the 8th IET System Safety Conference, 16-17 October 2013, Cardiff, UK.

References

1. Formal demonstration of equivalence of source code and PROM contents, Proceedings of the IMA Conference on Mathematics of Dependable Systems, Oxford University Press, 1995, pp225-248 D J Pavey and L A Winsborrow
2. Symantec Internet Security Threat Report: Trends for July–December 2007 (Executive Summary) XIII. Symantec Corp. April 2008. p. 29
3. F-Secure Quarterly Security Wrap-up for the first quarter of 2008. F-Secure. March 31, 2008.



153

Intelligent Transport Systems

Automated congestion warning system

Abstract

The Orlando-Orange County Expressway Authority (Authority) has deployed an innovative congestion warning system to further improve safety on S.R. 408, the primary east-west expressway through downtown Orlando, Florida. This system incorporates a dynamic message sign (DMS) installed on the S.R. 408 mainline to provide automated congestion warnings for exit lanes to Interstate 4 (I-4). The full-color, full-matrix DMS allows the advisory messages to closely resemble messages that would be displayed on a static sign panel. The warning system is linked to microwave vehicle detection sensors that monitor traffic conditions and allow the sign to respond in real-time, without input from operators at the traffic management center. The automated congestion warning system was deployed in concert with other enhanced directional signing elements to provide a comprehensive system that increases driver awareness of real-time congestion as well as providing enhanced route guidance approaching the I-4 interchange. This system has been performing reliably since its commissioning in January 2012. The Authority and other stakeholders are pleased with the performance of the system and initial indications point to improved safety as a result.

David R. McCoy, PE

Sr. Engineer IV

National Tolls

Atkins North America

Charles R. Lattimer, PE, PMP

Group Manager

Intelligent Transportation Systems

Atkins North America

Introduction

The interchange of S.R. 408 (also known as the East-West Expressway) and Interstate 4 (I-4) forms the crossroads for motor vehicle transportation in downtown Orlando, FL. On average, over 113,000 vehicles pass through this interchange every day on the S.R. 408 mainline[1]. In 2009, an interim reconstruction of the S.R. 408/I-4 interchange was completed to improve the flow of traffic in the area. While new ramps brought partial relief to some traffic movements, additional ramp improvements are necessary to fully address traffic flow concerns during peak traffic periods. These remaining ramp improvements are dependent upon the ultimate reconstruction of I-4 through the downtown corridor, which is scheduled to commence in late 2014. In the interim, certain sections of S.R. 408, particularly the westbound segment approaching the I-4 interchange, are subject to spill-over congestion from the off-ramps to I-4. The westbound S.R. 408 mainline experiences slow moving traffic in the outside lanes approaching the exits and higher speed traffic in the inside thru-lanes. Due to limited sight distance approaching the interchange, this speed differential creates a potential hazard for motorists who are unaware of slowed or stopped traffic ahead. The area of concern is indicated by a red dashed line shown in **Figure 1**.

In 2010, the Orlando-Orange County Expressway Authority (Authority) began investigating ways to improve safety on S.R. 408 westbound approaching the I-4 interchange. The Authority decided to utilize a two-part approach to address the problem. The first part was to improve the static approach signage to provide clearer direction to motorists as they approach the interchange. The second part of the solution provides automated congestion warning messages to motorists when traffic conditions warrant. This paper will discuss the improved approach signage and the automated congestion warning system.

Clearer directions using enhanced approach signage

Localized traffic bottlenecks can commonly occur at locations with lane drops, weaving areas, or at expressway exit ramps[2]. On S.R. 408 approaching the I-4 exit ramps, the number of westbound travel lanes decreases from four lanes to two lanes in the span of approximately one mile. Specifically, the four-lane section consists of one exit lane to I-4 east, one exit lane to I-4 west, and two thru lanes. Because of these lane drops, some level of weaving traffic is inevitable, which can contribute to congestion during peak volume periods. One way the Authority can seek to improve conditions is to minimize the probability of weaving traffic by providing enhanced directional signage approaching the interchange. This will increase the probability that drivers unaccustomed to the area will position themselves in the correct lane as early as possible.

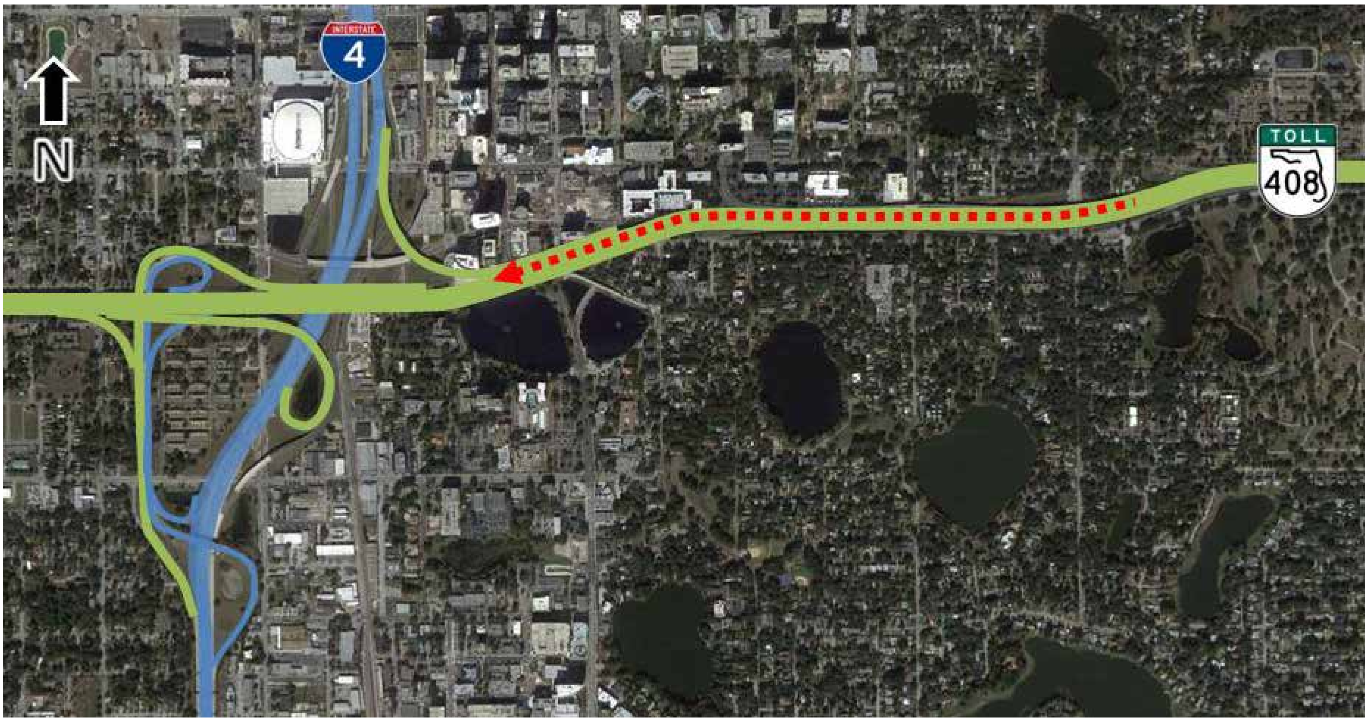


Figure 1. The S.R. 408 / I-4 interchange (red dashed line denotes the region of recurring congestion)

The Authority's enhanced signing approach involved two objectives: to clearly emphasize the exit lanes to I-4 and to clearly denote the "pull-through" lanes for traffic continuing on the S.R. 408 mainline. This was accomplished by means of enhanced static sign panels and horizontal signage elements. **Figure 2** and **Figure 3** show the before and after conditions of S.R. 408 Westbound. The details of the signage enhancements are described in the following two subsections.

Enhanced static sign panels

Two methods were used to enhance the message given by the static sign panels. The first was to install "pull-thru" signs for S.R. 408 westbound traffic. These signs provide lane confirmation for motorists who wish to continue on S.R. 408 westbound instead of exiting on I-4. The first pull-thru sign is installed approximately one mile from the first I-4 exit. The existing interchange sequence sign was replaced with a static pull-thru sign panel installed next to a dynamic message sign. The dynamic message sign provides automated congestion warning messages on an as-needed basis, which is described in a later section of the paper.

The first pull-thru sign uses the message "SR 408 WEST LEFT 2 LANES" to denote which lanes continue on S.R. 408. Additional pull-thru signs are installed at the exit gores to each I-4 ramp, with down-arrows over each thru lane instead of the text "LEFT 2 LANES." **Figure 4** contains graphical depictions of each type of pull-thru sign.

The second part of the approach was to increase the use of the down-arrow exit only notation for trapped exit lanes approaching the I-4 exits. Previous signage met the minimum requirements prescribed by the Federal Highway Administration's

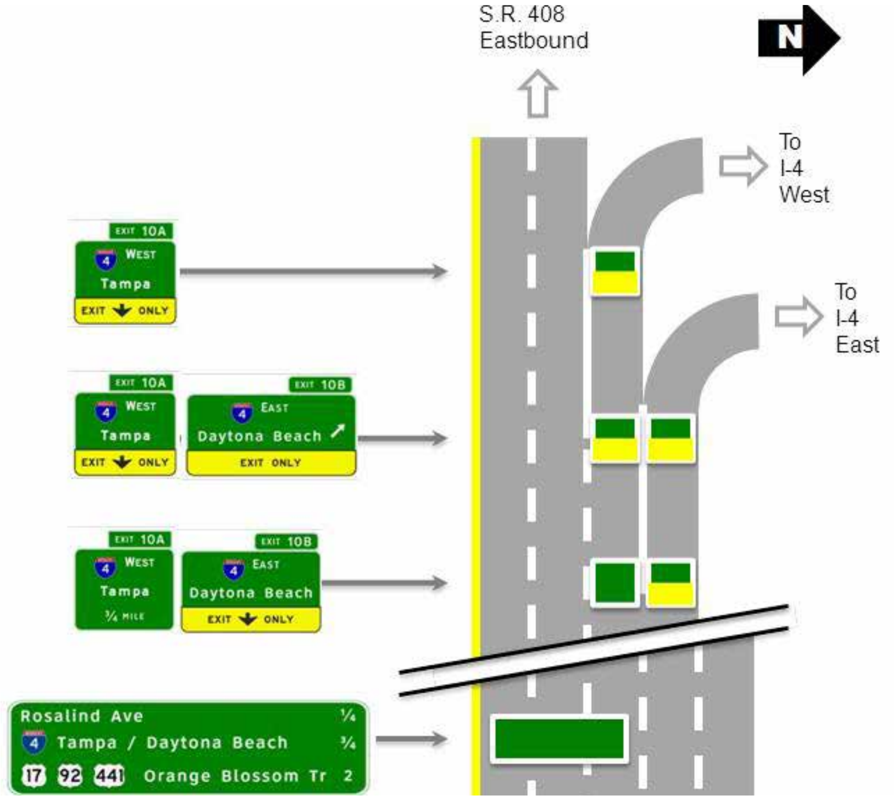


Figure 2. Original signage, prior to enhancement project

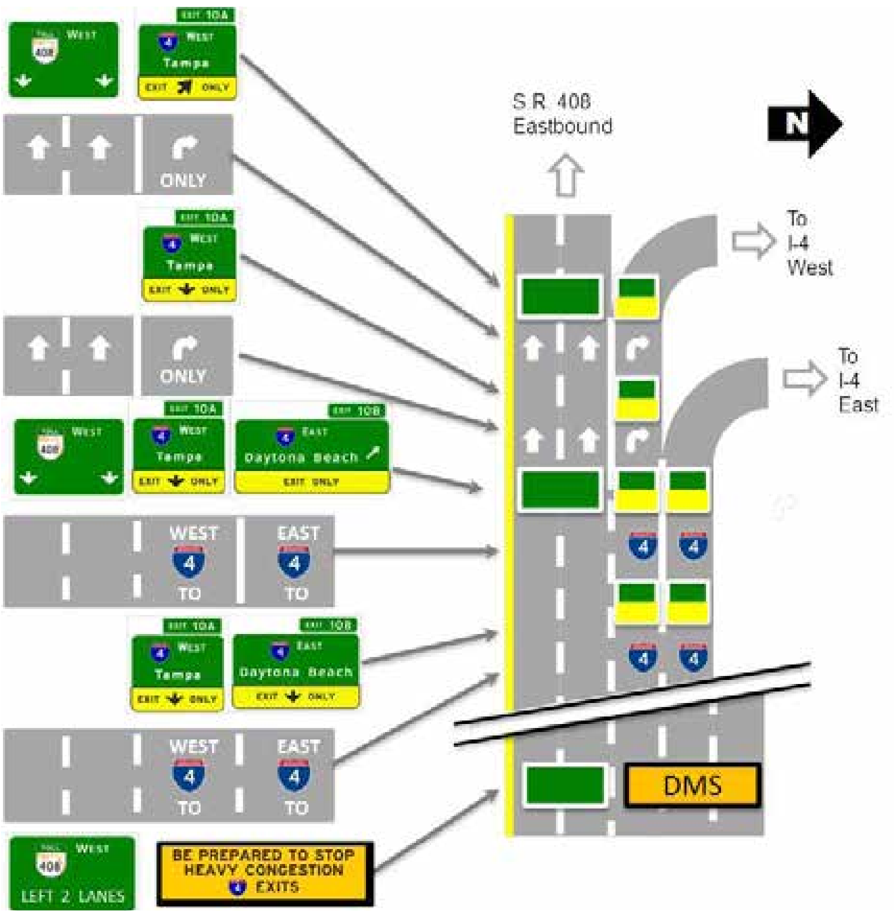


Figure 3. Signage after enhancement project

Manual on Uniform Traffic Control Devices (MUTCD), which establishes a consistent national standard for roadway signage in the United States. However, the Authority decided to make additional use of the down-arrow notation to further clarify the appropriate lane guidance message. An additional Exit Only sign was also installed at the exit gore for the ramp to I-4 westbound.

Horizontal signage

In addition to static sign panels, the Authority also elected to install horizontal signage elements in the area. These are thermoplastic elements that are adhered to the surface of the pavement, similar to reflective lane marking tape. Horizontal signage elements are not common to the Authority's expressways, so this project provided the opportunity to pilot test the elements in a special-use application. The boldest example of horizontal signage in the area are elements containing the text "TO I-4 WEST" and "TO I-4 EAST" which are affixed to the appropriate lane, as shown in Figure 5. In each case, I-4 is represented as a full color interstate road shield. Three sets of these elements are installed approaching the I-4 exits.

Once motorists pass the exit to I-4 east, the horizontal signage changes to straight thru arrows in the left two lanes and to a right-turn only arrow in the right-hand exit lane. This reinforces the message that the right lane becomes an exit-only lane and thru traffic must be in the two left lanes to continue. The arrows and text are made of white thermoplastic material.

Figure 4. Examples of S.R. 408 pull-thru signs

Figure 5. Horizontal Signage Elements

Automated congestion warning system

The second piece of the puzzle provides drivers with real-time alerts of congestion, raising awareness of a potential hazard beyond their line of sight. As the Authority explored this challenge, they sought a solution that provided the maximum amount of useful information to drivers at all times. For example, the installation of a static sign panel with beacons that read "Be Prepared to Stop When Flashing" was briefly considered. However, for the majority of the day the beacons would be turned off and the message on the sign would not be relevant. Likewise, a traditional amber DMS was also considered which would only illuminate during congested conditions. But, as before, this would mean the sign would be blank for much of the rest of the day. The Authority views DMS as significant investments of their customers' toll dollars. For this reason, each DMS needs to provide useful information to customers at all times leaving a DMS blank for any reason was not a viable option.

As the final solution to the problem, the Authority selected a high-resolution, full-matrix, full color DMS to provide congestion messages to drivers. Using this solution, the Authority displays an MUTCD-compliant interchange sequence message during non-congested conditions, which further improves guidance to drivers as they approach the I-4 interchange. During congested conditions, the default interchange sequence message is replaced by an appropriate congestion message. Incorporating a dynamic message sign as part of the solution provides flexibility to the Authority if they need to change or fine-tune messages in the future. Messages can be modified and uploaded to the sign whenever the need should arise.

Another requirement was for the congestion warning system to operate autonomously, without interaction from the staff at the Regional Traffic Management Center (RTMC) who manage the Authority's expressways. As such, the logic modules used to control the sign messages are installed locally in the sign enclosure. While exception messages can be posted remotely by RTMC personnel, the congestion warning system is designed to operate without communication between the sign location and a central control server.

To our knowledge, the design approach and system architecture described below is unique, at least in terms of deployments in the United States. **Figure 6** shows the locations of the sign and the two downstream sensors used to measure the speed of traffic. Both sensors are Wavetronix HD microwave vehicle detection sensors (MVDS), which are configured to observe the speed of traffic on a per-lane basis. While the sensors are configured to detect speeds in all lanes, only traffic speeds from the lanes that become the exit lanes to I-4 are used in the congestion detection algorithm.

Figure 7 shows the lane configuration for Speed Sensor 1. Notice that lanes 1 and 2 become the exit lanes to I-4, while the other lanes are thru lanes on S.R. 408. The sensor is installed on the upright of the full-span sign structure located at the exit gore to the ramp to I-4 Eastbound.

Figure 8 depicts the lane configuration at Speed Sensor 2. Since the sensors are only able to read ten lanes of traffic, the two-lane exit ramp to Rosalind Ave. was configured as a single lane. This is allowable in our application since the speeds in that lane do not serve as inputs to the congestion detection algorithm. Because of the exit ramp to Rosalind Ave., the exit lanes to I-4 are actually configured as lanes 2 and 3 at this location. The sensor

Figure 6. DMS and sensor locations for congestion alert system

Figure 7. Lane configuration of Speed Sensor 1

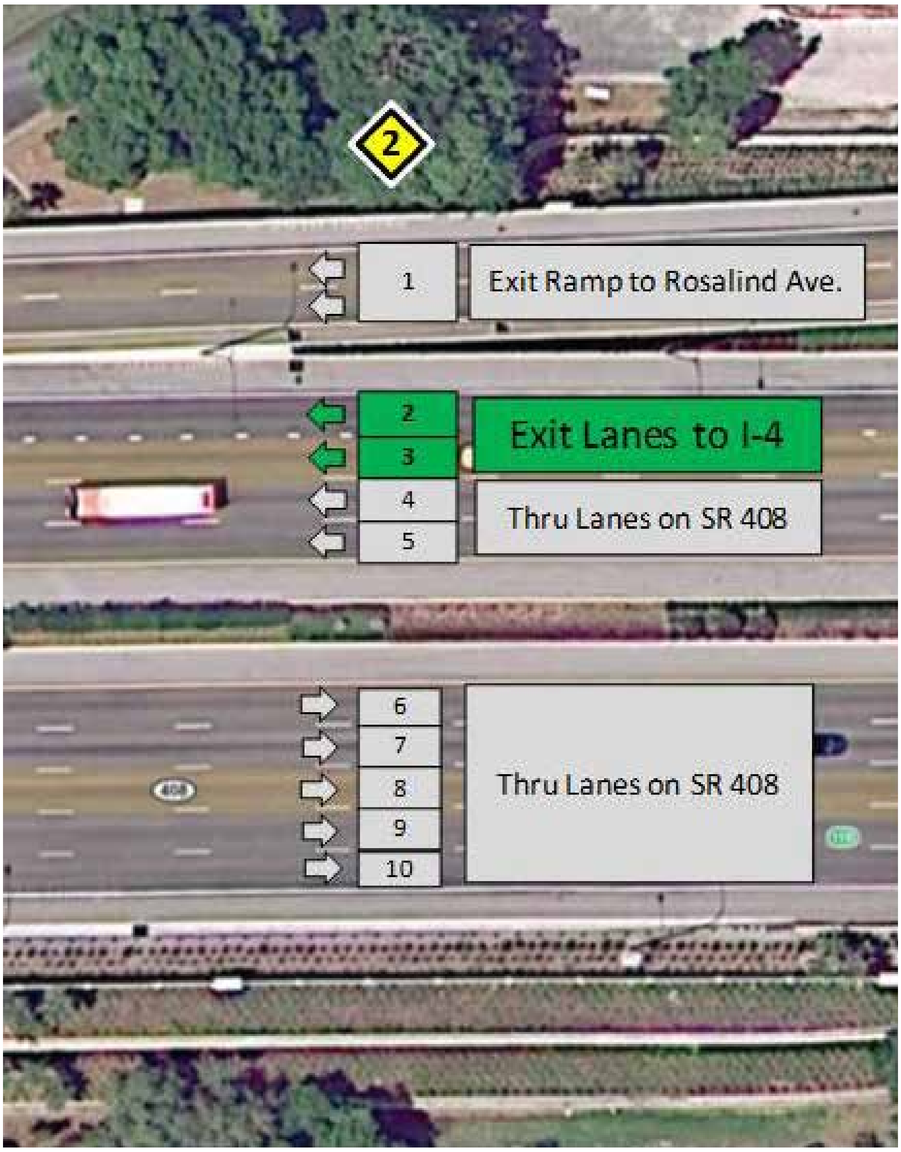


Figure 8. Lane configuration of Speed Sensor 2

is mounted on a dedicated pole installed on the shoulder of the expressway.

Each sensor connects to a logic module that calculates the average traffic speeds in the exit lanes to I-4 over a one-minute time interval. If the logic module detects that the average speed across the two lanes is below 45 MPH for two consecutive one-minute intervals, a contact closure is activated. The contact closure clears once the speed of traffic exceeds 45 MPH over a one-minute interval (for reference, the speed limit on this section of S.R. 408 is 55 MPH). All logic modules are contained in the DMS sign enclosure and are connected to the sensors using serial-to-fiber-optic media converters.

Three full-color DMS messages are stored locally in the DMS controller. These include one default message, which is displayed during non-congested conditions, and two separate congestion messages that are displayed based upon the severity of congestion detected downstream. By default, when both contact closures are open, the DMS displays an interchange sequence message for the upcoming I-4 ramps. This image is composed of white text on a green background and utilizes interstate highway shield graphics, similar to what would be seen on a static sign panel, as shown in **Figure 9**.

Figure 9. Sign message under normal freeflow traffic conditions

During the morning peak travel conditions, a queue will begin to build eastward, starting at the gore of the exit ramp to I-4. When the average speed in the exit lanes drops below 45 MPH for two consective one-minute intervals, the contact closure connected to Speed Sensor 1 activates. This contact closure is connected to a digital I/O board, which in turn is connected to the DMS controller. This input causes the DMS controller to post the first of two congestion messages, a soft "CAUTION CONGESTION AHEAD I-4 EXITS" message is displayed on the sign, as shown in **Figure 10**.

When the queue backs up further and reaches Sensor 2 (the sensor closest to the sign), the message changes to a hard "BE PREPARED TO STOP HEAVY CONGESTION I-4 EXITS" message, as show in **Figure 11**. A more forceful message was selected for this scenario since motorists who see the sign message will not have as much time to react to the message when encountering the back of a queue or shockwave. Once congestion clears, the message automatically returns to the default interchange sequence message, without intervention from operators at the traffic management center.

It should be noted if Speed Sensor 2 is triggered at any time, the hard "BE PREPARED TO STOP" congestion message is posted, even if traffic is clear in the Speed Sensor 1 location.

Experience and lessons learned

The automated congestion warning system entered active service in January 2012. In the time the system has been active, it has received compliments from Authority staff and other partner agencies for improving safety by drawing the driver's attention to the potential hazards ahead. This system has received praise from traffic management center staff and appears to have had a positive effect in reducing incidents.

There were sixteen (16) reported incidents in the AM peak direction over the period of January-March 2011 and July-September 2011. The automated congestion alert sign was activated on January 6, 2012. During the corresponding period of January-March 2012 and July-September 2012, only ten (10) accidents were reported. Due to the random nature of accidents, it is difficult to demonstrate a statistically significant reduction in accidents over the short time periods list above. However, the Authority will continue to monitor conditions to determine more conclusively the positive effects of the system.

This project is a visible and effective method of improving safety on the Authority's expressways. According to the morning shift lead operator at the District 5 Traffic Management Center (which operates the Authority's expressways), "there has been a noticeable decrease in the amount of incidents we have approaching those WB I-4 exits. I think the sign has been effective in letting traffic know to expect heavy delays in the area of I-4, making them more cautious when trying to take those ramps."

Figure 10. System operating with moderate congestion

Figure 11. System operating with heavy congestion

Conclusion

This project is an example of how an innovative combination of ITS technology and more traditional signage elements can be a cost-effective solution to certain on-going safety concerns, either as a permanent solution or as an effective measure until capacity improvements can be funded.

Acknowledgements

This paper draws on work accomplished for the Orlando-Orange County Expressway Authority and is published with the permission of the Authority.

References

1. Orlando-Orange County Expressway Authority. System's Traffic Data and Statistics Manual. Orlando : OOCEA, 2012
2. Federal Highway Administration. Recurring Traffic Bottlenecks: A Primer. Washington, DC : U.S. Department of Transportation, 2012.

Intelligent Transport Systems

The practical experiences of 3G UTMC communications in Slough

Synopsis

The paper gives an overview of 3G technology, reviews the deployment of the 3G based Urban Traffic Management and Control (UTMC) communications in Slough and reveals several lessons learnt, allowing other traffic practitioners to assess if using 3G is the right communication choice for their sites.

Keith Manston
Traffic Solutions
Siemens Mobility

Gareth Tilley
Senior Consultant
Highways & Transportation
Atkins

Nick Healy
Highways
Surrey County Council

Introduction

Set against a backdrop of challenging economic times, an aging communication system and the redevelopment of a town centre, the Slough 3G Project was born.

Before the advent of internet protocol (IP) based communications, most urban traffic control (UTC) implementations using split cycle offset optimisation technique (SCOOT) required timely delivery of data to and from street equipment to ensure efficient operation. Typically, this relied on 'second by second' data transfer, usually over expensive telephone circuits with dedicated connections between the outstations 'on-street' and the system in the control room.

The continued development of SCOOT and the deployment of SCOOT MC3 (Managing Congestion, Communications and Control) has enabled the core SCOOT algorithm to tolerate greater variation in communications technologies and even significant (in communications terms) data loss, which has further increased the communication options available.

Since the development of the UG405 UTMC Protocols1, traffic signal companies like Siemens have exploited IP based solutions over fibre and copper, using asymmetric and symmetric digital subscriber lines (ADSL and SDSL). However, wireless communications have been of particular interest to many UTC users because they offer significant overall cost benefits via the use of 3G.

In 2009, Siemens commenced field trials within Southampton and some other UK locations. The trials aimed to provide performance data for outstation transmission unit (OTU) equipment using 3G communications. The results and observations were recorded in and presented to the JCT Traffic Signal Symposium & Exhibition in 2012.

At that time, the initial trial results looked promising, but it was reported that further trials and a wider deployment would be necessary to determine if 3G really offered the potential to reliably be used for UTC communications.

Since the original trial, Slough Borough Council has embarked on the much wider deployment of 3G across approximately 50 sites and has gained extensive (and positive) experience of using this communication technology for SCOOT.

This paper reviews the 3G wireless solution development from a small trial to a full, system-wide introduction and deployment across Slough and its use in a SCOOT environment. Also, it reviews the background to the project, the lessons learnt and it offers practical advice for those considering using 3G in this way.

3G; the background

3G refers to "Third Generation Communications" and it is the latest evolution of mobile wireless technology to be extensively deployed. Although both 4G and even 5G solutions are being considered for future development, 3G represents the latest widely available solution and it will probably remain the leading technology for some years.

First generation mobile communications were provided by analogue solutions between user handsets and the network base stations and these offered no real data capacity. The first 'mass market' mobile solution with a digital data capacity arrived with the advent of GSM (Global System for Mobile Communications), in the early 1990s. Although data transfer is supported across GSM networks, the data rate is low (9.6Kbps to 14.4Kbps) and as calls are charged by time, rather than the amount of data used, "permanently on" connections can be expensive to implement. Nevertheless, the data transfer capacity of GSM networks is ideally suited for applications such as remote monitoring, where data rate requirements are generally low and connections are necessary only intermittently.

The introduction of GPRS (General Packet Radio Service) in 2001 brought the possibility of much higher data-rates and with charges based on data exchanged over the network, rather than how long the user was connected, and as such "permanently on" services became a realistic option. With IP communication support, GPRS networks also provided the means for users to browse the Web, opening up a whole new range of uses for the mobile phone. Further development of the GPRS technology has led to enhanced data rates for GSM Evolution (EDGE), which offer even higher data rates.

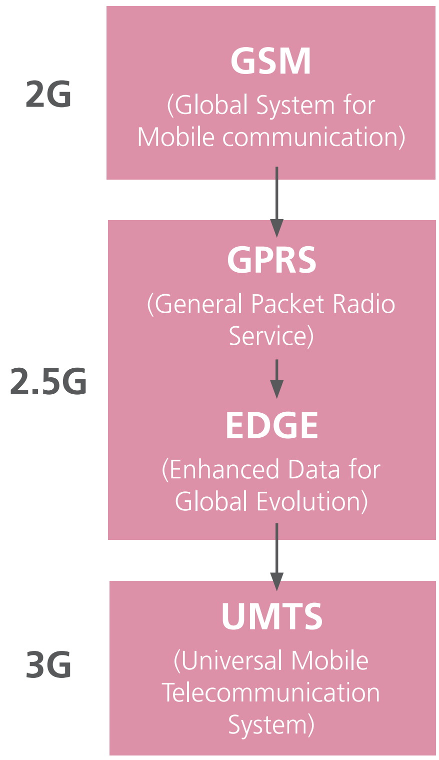


Figure 1. The evolution of 3G technology

Today, GPRS is often used for traffic applications, where relatively frequent communication with outstation equipment is needed (for example VMS signs), but the timely delivery of data is not critical. It generally doesn't matter if a sign setting message takes many seconds to reach its destination and if a particular message does not reach a sign; it can be resent with little or no impact on the travelling public. However, to be successfully deployed for UTC applications, it is necessary for any chosen communications solution to be able to offer timely delivery of data, ideally 24 hours a day, 365 days a year!

The introduction of SCOOT MC3 and UG405 has helped significantly to reduce the system reliance on very accurate 'second by second' communications. Small time delays and irregularities in data transmission times between the UTC software and OTUs can now be accommodated.

However, SCOOT still needs to receive and transmit data in a reasonably timely manner in order to maintain efficient traffic control. In MC3, data is 'time stamped' so that delays in data transmission can be accommodated whilst maintaining 'current' data in SCOOT. The system does not allow for frequent very long delays in data transmission, as these would result in poor traffic control.

For detector data transmitted from the OTU to the system, the aim is to maintain the delivery of second-by-second data for high performance traffic control, but modern IP communications systems can introduce uncertainty into the data transmission process.

This is allowed for by the data time-stamping within the UG405 protocol. Data may be delayed, out of order, repeated or missed. However, as long as a complete message is received (e.g. the four bits of detector occupancy from any given second), then the relationship of this data to on-street occupancy can be reconstructed and used by SCOOT. Control data transmitted to the street also includes time stamping, which allows messages to be sent a short time in advance of any actual stage changes to accommodate small delays in communication. In addition, the overall data usage necessary for control messages is significantly reduced because only state changes are transmitted. This is a key factor in assessing the overall cost of implementing any communications solution.

Figure 2. Original signage, prior to enhancement project

As a consequence, SCOOT MC3 will now operate successfully with communications that do not guarantee the time for transmission or order of data delivery and has been designed to operate efficiently with possible delays of up to four seconds.

Experiences in Slough

The Business Case

Slough Borough Council (SBC)'s original traffic control system was a Siemens Alpha solution which was installed in 1998, and was used to control junctions both under fixed time and SCOOT control. The authority updated to PC SCOOT in 2009, but with the communications system continuing to operate under Tele-command 12.

The decision by SBC to redevelop the town hall resulted in the requirement to relocate a major communications point for council services. This included the Tele 12 cabinets and associated BT termination points.

The need to relocate these critical connections into the St Martin's Place office also coincided with operational pressures to reduce revenue expenditure and future commitments. With a method of communicating to traffic signals that was effectively obsolete, a programme upgrade was opportune.

Although SBC would have ideally preferred a phased switch over to a new communications method, the date had been set to demolish the Town Hall, which resulted in a programme with a fixed end date.

In partnership with Atkins, SBC undertook a review of the current industry communication methods. This review investigated how to link traffic signals to UTC with SCOOT control. A cost comparison between moving the termination lines to a new council building and retaining the Tele12 operation (effectively doing nothing) was undertaken.

The three options considered were: do nothing, use a mesh wireless system, or implement the 3G communications. A brief overview of these options is outlined below.

Option 1 - do nothing:

This required the following works to be undertaken:

- Move the existing system, including the transfer of the BT leased lines and in-station transmission units from the old town hall to St Martin's Place
- Reconnect all BT lines within the controllers, as currently terminated outside in small cabinet
- Reconnection of the controller EPROMS to the OTUs (as a high number did not work correctly this would enable proper control and fault reporting to occur)
- Ensure program and funding for the next ten years to replace obsolete OTU equipment.

Option 2 - introduce a new mesh System:

To introduce the required hardware, new OTUs and mesh modems into all controllers and introduce a new server within St Martin's Place:

- Disconnect existing BT lines
- Install ADSL Back Haul line into the St Martin's Place office
- Base on a 20 year life expectancy.

Option 3 - 3G Communications:

To introduce the required hardware, new OTUs and 3G Modems into all controllers and install new server within St Martin's Place:-

- Disconnect BT lines
- Install ADSL back haul line into the St Martin's Place office
- Base on a ten year life expectancy.

The evaluation of the capital cost against the expected revenue savings showed that a capital spend in the region of £205,000 for a 3G solution would bring a saving to SBC of £36,320 pa under its current operation and result in a payback period of around five years. This was based on estimated annual revenue costs of £9,680 pa. This was one of the main reasons for the option being selected as the way forward.

However one of the risks of the 3G business case and the proposal was that the appropriate level of coverage by a 3G provider might not be available at each controller location. This issue was underlined by SBC's concurrent procurement exercise to obtain a new supplier for all its fixed and mobile communications. Until the provider was known, no formal confirmation could be given on whether the service provider could meet the 3G coverage requirements.

Also, the business case was dependent on the monthly data costs. The monthly data costs from the Southampton trials were taken as a basis for the business case. This equated to £150.00 pa for each site. In addition, the backhaul line cost £1,320pa.

Deployment

SBC decided on a two stage approach for the transfer across to 3G. This was so the transfer of critical sites could take place well in advance of the end date ensuring that if issues occurred, the loss of the non-critical sites not monitored by UTC (e.g. due to the Town Hall Tele 12 having to be completely switched off) would not affect the town's traffic flows.

The corporate communications contract was awarded to O2, which is also one of the major employers in the area. O2 engaged with a company called Wyless, who supplied the Subscriber Identity Module (SIM) cards and managed the network.

Siemens undertook a site survey in relation to the possible 3G network and works commenced to introduce routers, antennae and new Gemini units. The initial 25 sites linked to UTC and controlled under SCOOT were completed with a few teething problems and with some valuable lessons learnt, which will be discussed later.

The majority of the sites worked well and the availability performance levels (or "up time") was considered acceptable. However, a small number of sites suffered from a high level of intermittent faults. These faults were very random with variable lengths of downtime, ranging from seconds to tens of minutes or even longer. Some of the longer downtime could relate to the supplier communications network management that we had no control over. Essentially, the communication service availability would be acceptable for long periods of time and then drop out with no clear reason or pattern why this occurred. This resulted in extensive time spent to solve and review data by Wyless, Siemens and O2.

The issue of performance on this small number of sites resulted in the decision to compare an alternative provider's performance, so SBC also approached Mobius.

Mobius was asked to supply four SIMs to be deployed in four sites with connectivity issues. The SIMs resolved connectivity issues at two sites, which seemed to have a drop-off in performance that appeared to mimic the performance profile of mobile broadband SIMs. The reason for this is currently unknown.

Around this time of the project the Wyless SIM cards performance seem to improve and allowed for the project to move on to the next stage.

The second stage was to move the remaining 25 UTC sites to 3G. These sites were not as critical as those transferred during the first phase because they were monitored on UTC, with short spells of control under SCOOT, depending on traffic management requirements.

What went well, what were the problems?

Stakeholder engagement

The whole project and the programme to introduce 3G was very much a team activity, with a significant need for various organisations to work together. The team within SBC not only included the traffic management team, but also the IT and building management staff. This was because of the installation of new modems and equipment within the SBC communications room. They were engaged from the start, which helped significantly in the delivery of the early infrastructure of the project because they understood the background, were empowered and had ownership of the project.

Standardisation

The controller EPROMS at a number of sites required updating to allow for the new communication method. Also, the existing EPROM configurations at a number of sites had not been correctly wired into the existing OTUs. This issue was further compounded by a wide variation in the control and reply bit formats, requiring multiple format types within the database. This was caused by a lack of previous standardisation and a variety of past approaches to the set up of the UTC system.

Resilience

A further benefit that SBC has gained from the project through the transfer over to UTMC communications and the introduction of the Gemini units, has been the ability to remotely configure and integrate OTUs, increasing the flexibility to upgrade to MOVA control if required in the future, or act as a fallback method of control.

Focus on critical junctions

Critical junctions on the A4 needed to be transferred over to 3G. These transfers had to be planned in more detail than some of the other sites because the works needed to be undertaken in a timely manner. Sites were grouped as SCOOT regions and to retain SCOOT linking meant moving each region together, when possible.

Delays in the works would have had the same impact as if UTC was not operational, which was known to cause traffic delays that would have affected M4 junction 6.

Location of the 3G antenna

Unlike GSM, the general concept that the traffic signal pole nearest the controller will be suitable for mounting the antenna is not the case when introducing a 3G connection. A survey of ducting and suitability of pole locations was performed. The key constraining factor was the length of cable that could be installed between the controller and antenna. These works and the associated project management for the New Roads and Street Works Act (NRSWA), is clearly not necessary for non-3G alternatives.

Overall Learning

Following this project, some key points that are particularly linked to the use of 3G and highlight the possible pitfalls that other users can avoid following this project are:

- Work with suppliers to improve equipment performance (in this particular case; routers)
- Use specialist suppliers where necessary (M2M SIM cards)
- Antennae positioning is important, as is ensuring an adequate 3G coverage survey
- UTC communications failure reporting needs to be matched to 3G.

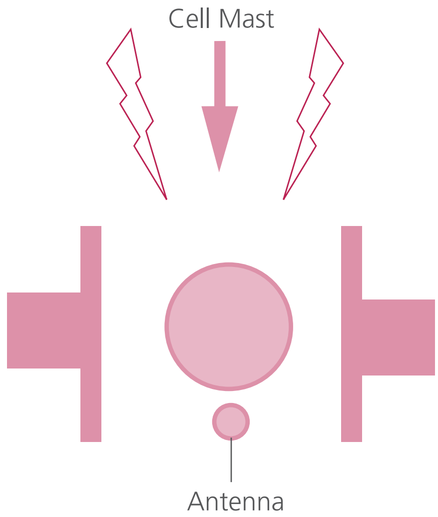


Figure 3.

In more detail, not all routers and 3G routers have the same performance levels when used to transfer data for UTC purposes. Mobius worked with Siemens to identify that there was a significant differential between the two types of wireless routers supplied to Siemens by two separate hardware suppliers. The adaptive module supplied router did not indicate any problems (and this is the router that Siemens have settled on as standard at the moment). However, the router supplied by Digi appeared to be less reliable on the Vodafone network (i.e. not reacquiring a connection after it failed, which demanded a manual reboot). Mobius and Digi identified a bug and a known software patch resolved the problems.

Brief comparisons between networks show there are fewer drop-offs using the Vodafone network compared with the O2 network. This also highlights the difference between using a SIM provided by the network and a machine-to-machine (M2M) SIM provided by a true M2M specialist.

The location and position of the antenna is critical. This was demonstrated at two of the initial batch of 25 sites. These sites were investigated by Mobius and were found to have good signal strength and the cells and network were not over loaded. The bespoke tools that Mobius had developed showed the SIM was available to send data, but the router was showing low signal strength. A site investigation found that the installation of the antenna was the most likely reason for poor connectivity. The pole and pole cap was shielding the antenna from the cell mast (see below).

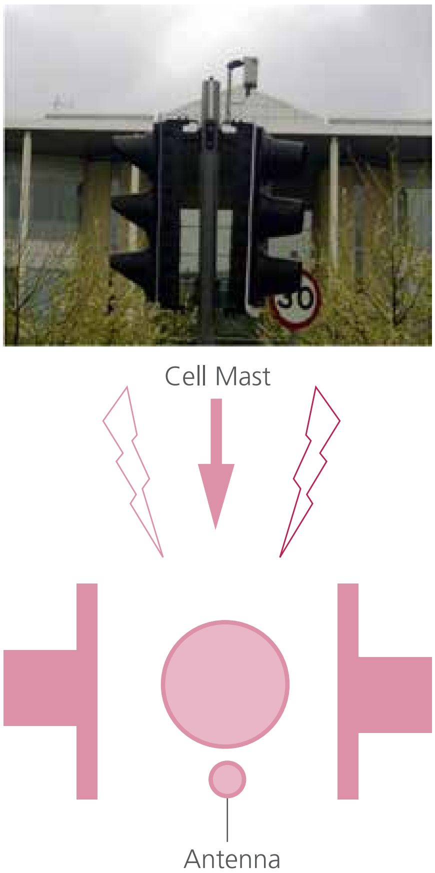


Figure 4.

Although each antenna had been installed on brackets, their location and height were making a considerable difference. This layout was altered with a modified mount for the antennae, which raised them higher than their associated pole caps, resulting in a significantly improved level of reception and performance of the sites. The lessons learned allowed a more detailed consideration when installing the second batch of sites, which reduced the time, effort and cost of tracking connection issues later on.

The reporting of 3G communication failures via the UTC system is very limited. The system does not have the tools to allow reporting of communications performance, or to give users advance warning of a possible failure of a site. Although this was only really an issue during the early stages of implementation, the random occurrence on the Slough system now is dealt with in a similar way to the previous communications methods and fault identification.

Conclusion

SBC experienced problems and challenges during the early stages of the project, whilst stakeholders were coming to grips with a new technology application. However, the council is now happy with the completion and the objectives it set out have been met. The project has demonstrated that 3G does now offer a viable alternative to other communications solutions, albeit with some caveats, particularly for sites of high strategic importance.

The benefits seen from the cost savings over a suitable period of time have allowed the council to look at a Phase 3 and to seek internal funding to review additional sites and options.

However, the project has also shown that 3G suppliers are not necessarily the same and standard 3G SIM cards do not always work, so procurement of slightly higher-cost M2M cards and modems may be needed to deal with certain "problematic" sites and locations.

3G is one of many new traffic signal-related technology options that designers, operators and suppliers of traffic control systems can deploy. However, careful consideration and initial trials to compare product performance should be undertaken to assist in the stakeholder's understanding and to choose the right technology. Traffic control systems can now communicate across a range of systems rather than just the one communication type, ensuring a high level of connectivity and ongoing reduction in operational costs. Gone are the days of one supplier for traffic signal communications.

Acknowledgement

An earlier version of this paper was previously presented at the JCT Symposium, Warwick University, September 2012. The authors wish to thank Andrew Holt, formerly of Atkins, for his valuable contributions to that earlier version.

References

1. TS004.004:2008, UTMC Objects Registry Annex D: Data Objects, published by UTMC Ltd, 2008

2. Manston, K., Wylie, M., "Advanced wireless communications for UTC", The 15th JCT Traffic Signal Symposium & Exhibition, October 2010.

Analysis of end zone cracks in prestressed concrete I-girders

Abstract

With the use of higher strength concrete, deeper girders, and significantly higher prestressing forces, longitudinal web cracks of precast prestressed concrete girders have become more prevalent. The purpose of this investigation is to identify the main causes of such cracks that occurred in the Cross Florida Barge Canal Bridge. First, brief descriptions of the bridge and cracking are given. Then, analytical models are presented. The analytical results show that the maximum vertical tensile force within a girder end zone due to the longitudinal prestressing strands is nearly 8% of the total prestressing force which is significantly higher than the force prescribed by the current AASHTO specifications to determine the quantity of vertical reinforcement required at the ends of pretensioned members. Finally, some recommendations for end zone detailing are discussed. The results presented can help bridge designers provide better bridge detailing and further mitigate cracking in prestressed spliced I-girders.

Dongzhou Huang
Ph. D., P.E.

Sr Engineer IV

TS Transportation Design

Atkins North America

Introduction

Precast prestressed concrete girders are widely used in the United States and the world for bridge construction. Longitudinal web cracks at the ends of pretensioned concrete girders are commonly observed at the time of strand detensioning. With the use of higher strength concrete, deeper girders, and significantly higher prestress forces, these cracks are becoming more prevalent and, in some cases, larger. The recently constructed Florida Barge Canal Bridge (**Figure 1**) is a three-span continuous spliced I-girder bridge. Many cracks developed at the ends of the girders after the prestressing strands were released.

The objective of this investigation is to identify the main causes of the cracks and to propose a practical method for detailing the end zone of precast prestressed concrete girders. First, a brief description of the cracking is given. Then analytical models are presented. Finally, the analytical results and recommendations for end zone detailing are discussed.

Description of bridge

The Florida Barge Canal Bridge carries the US 19 (CR55 & US 98) Corridor over the Cross Florida Barge Canal, located in Citrus County, Florida. The bridge is a three-span continuous spliced I-girder bridge of 77.7-86.9-77.7 m (255-285-255 ft) (**Figure 1a**). The cross-section of each of the Northbound and Southbound bridges consists of four girders with a cast-in-place deck of 216 mm (8.5 in) (**Figure 1b**). Each of the girders consists of five segments: two end segments of 49.45 m (162.25 ft), two haunch segments of 36.58 m (120 ft), and one drop-in segment of 49.68 m (163 ft) (**Figure 2**). The segments are connected by two closure pours of two feet long and four post-tension tendons with 15- 15mm (0.6 in) ϕ strands each. The duct diameter is 95 mm (3.75 in) and the thickness of the girder web is 241 mm (9.5 in). The haunch girder depth varies from 3.96 m (13ft) at the interior support to 1.98 m (6.5 ft) at its ends. Both the end segment and the drop-in segment have a constant girder depth of 1.98 m (6.5 ft) with 44-15 mm (0.6 in) ϕ strands stressed to 195.5 kN (43.95 kips) each at the bottom flange and 4-15 mm (0.6 in) ϕ stressed to 44.5 kN (10 kips) each at the top flange.

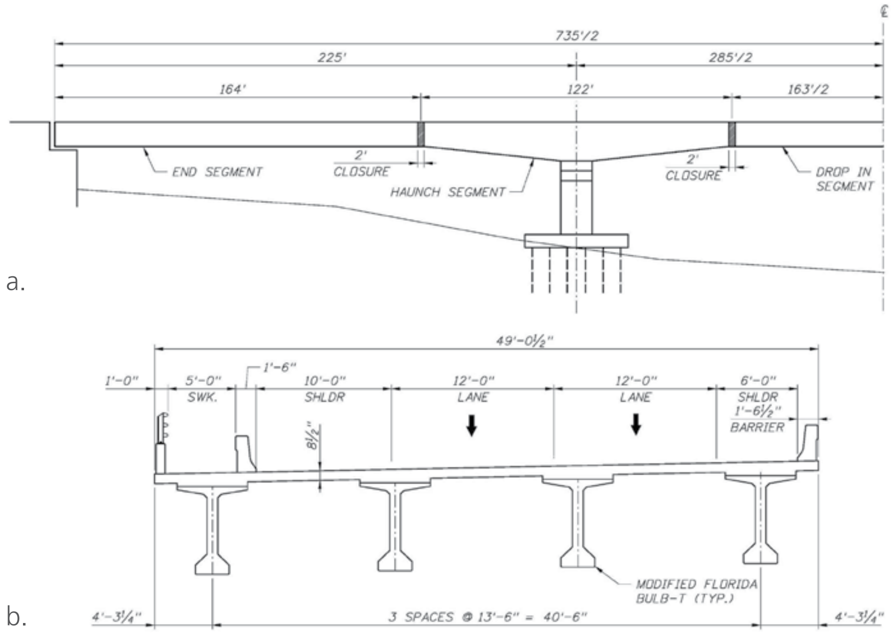


Figure 1. Cross Florida Barge Canal Bridge, (a) Elevation, (b) Cross Section (1'=305 mm, 1"=25.4 mm)

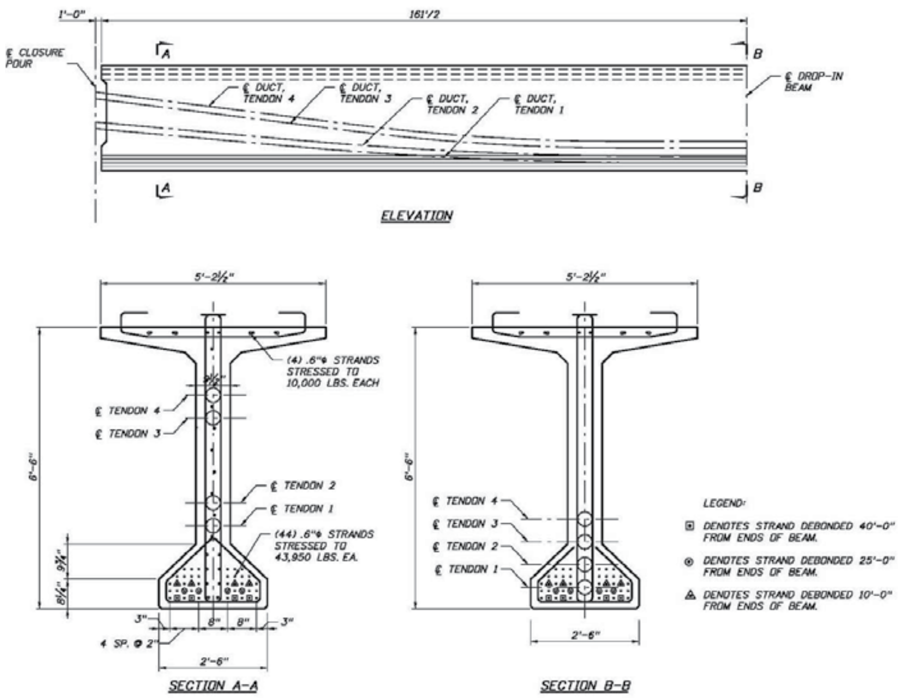


Figure 2. Elevation and cross section of the drop-in girder (1'=305 mm, 1"=25.4 mm)

The haunch segment has 24-15 mm (0.6 in) ϕ prestressing strands at the top flange and 10 strands in the web. The design strength of the concrete is 59 Mpa (8.5 ksi). The primary dimensions of the girders are shown in **Figures 2** and **3**.

Many cracks were discovered in the webs at the ends of girders after the prestressing strands were released. **Figure 4** presents crack patterns of two drop-in girders and **Figure 5** shows the crack patterns of two haunch girders. All other drop-in and haunch girders have similar crack patterns. Most of the cracks are along the duct locations and are nearly horizontal. The width of the cracks varies from 0.102 to 0.203 mm (0.004 to 0.008 in). The crack length ranges from 152 to 533 mm (6 to 21 in).

Analytical model

Bridge model

The spliced girders are modeled as a thin-wall structure and divided into a number of quadrilateral shell plate elements (refer to **Figure 6**) and analyzed using the finite element method. There are a total of 390 and 288 elements in the longitudinal direction of the drop-in and haunch girders respectively. Both drop-in and haunch girders were modeled with 34 transverse elements in the top flange and 6 in the bottom flange. The width and length of each shell-plate element are about 127 mm (5 in) except in the duct areas. The thickness of the elements in the duct areas is approximated as the web thickness minus the duct diameter. The aspect ratio of most elements is approximately unity. The total number of finite elements is 13,260 and 9,792 for drop-in and haunch girders respectively. In the analysis the girders are assumed to be homogenous and elastic.

Prestressing strands model

The strand pattern and debonding schedule of the prestressing strands for drop-in and haunch girders are illustrated in **Figures 3** and **4**. The prestressing strands are Grade 270, 15 mm (0.6 in) ϕ low-relaxation strands stressed to 195.5 kN (43.95 kips) each. The prestressing force is simulated as a number of concentrated forces and moments applied at the half-depth of the related elements[1]. According to the debonding schedule shown in **Figure 3** for the drop-in girder, the forces and moments are separated into five groups, which represent the prestressing forces of fully-bonded strands and the strands to be debonded at 3.05 m (10.0 ft), 7.62 m (25.0 ft), and 12.19 m (40.0 ft), respectively. The transfer length of the strands is assumed to be 60 strand diameters [2,3]. The prestressing force of the strands is assumed to vary linearly from zero at the point where bonding commences to a maximum at the transfer length and applied at the related nodes (see **Figures 8** and **9**). The prestressing force model for the drop-in girder is illustrated in **Figures 7** and **8**. From **Figure 7**, it can be seen that the prestressing forces of the fully-bonded strands are divided into eight concentrated forces (Pj1 to Pj8) and eight concentrated moments (Mj1 to Mj8) along the girder longitudinal direction at the bottom flange. Each of the concentrated forces and moments is further divided into several concentrated forces or moments along the transverse direction (**Figure 9**). The prestressing force model for the haunch girder can be developed similarly[1].

Analytical results

To evaluate the effect of the boundary conditions on the girder vertical tensile stress due to prestressing forces, both the drop-in girder and the haunch girder are assumed to be simply supported at each end, one fixed and the other a roller. **Figure 9** shows the vertical tensile stress distribution at girder ends due to the prestressing forces. **Figures 9a** and **9b** illustrate the vertical tensile stress distribution for the drop-in girder at the fixed and roller supports respectively.

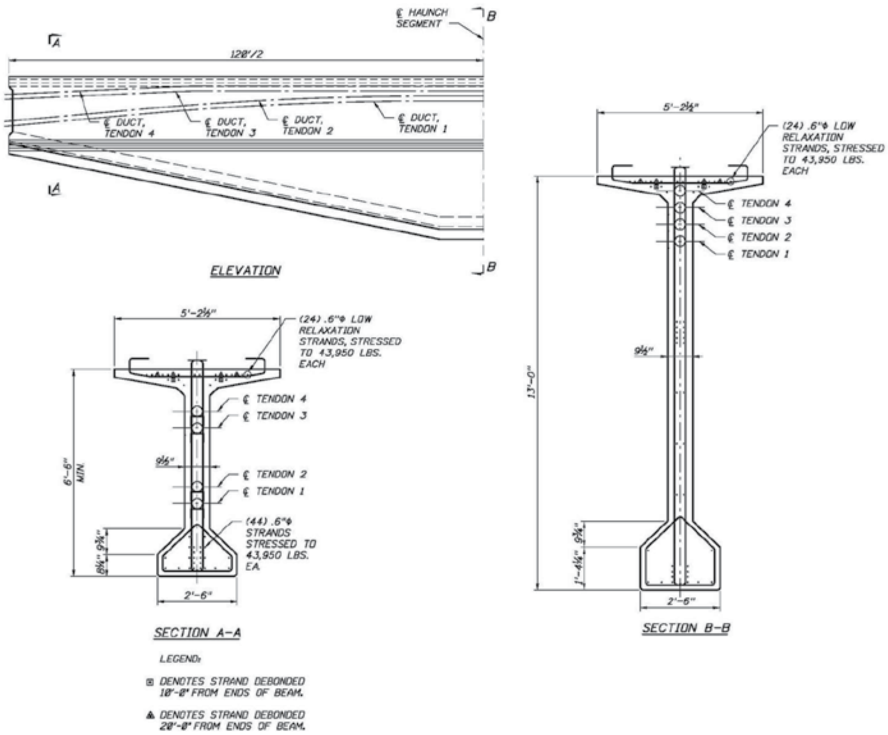


Figure 3. Elevation and cross section of the drop-in girder (1′=0.305 m, 1″=25.4 mm)

Figures 9c and **9d** present the vertical tensile stress distribution for the haunch girder at the fixed and roller supports. From **Figure 9**, we can see:

i. the prestressing force can create significantly high vertical tensile stress at the girder ends about at 1/3 to 2/5 girder depth[1] from the location of the prestressing strands, especially in the duct areas. The maximum tensile stress for the drop-in and haunch girder are about 8.97 Mpa (1.3 ksi) and 6.90 Mpa (1.0 ksi) respectively. These tensile stresses are large enough to cause cracking

ii. The tensile stress is distributed along the end of the web for a distance of about ¼ girder depth

iii. The end support conditions have little effect on the vertical tensile stress distribution. The maximum vertical tensile stress at the fixed ends is slightly larger than that at the roller supported ends.

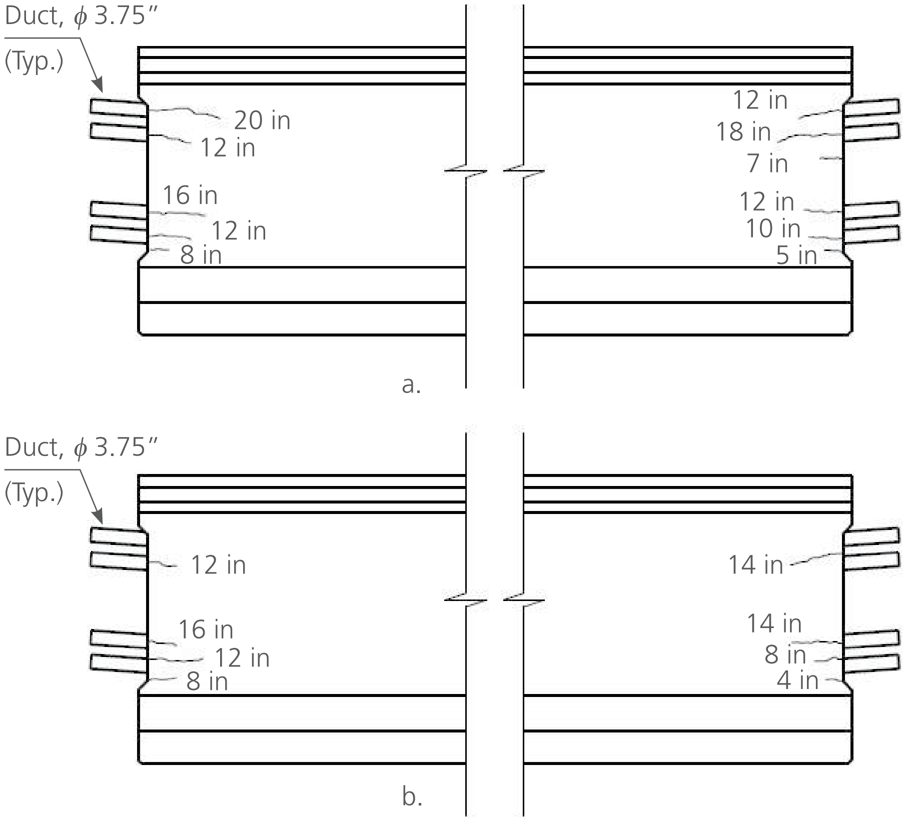


Figure 4. Crack pattern of drop-in girders, (a) Girder 2-2 NB, (b) Girder 2-3NB (1in =25.4 mm)

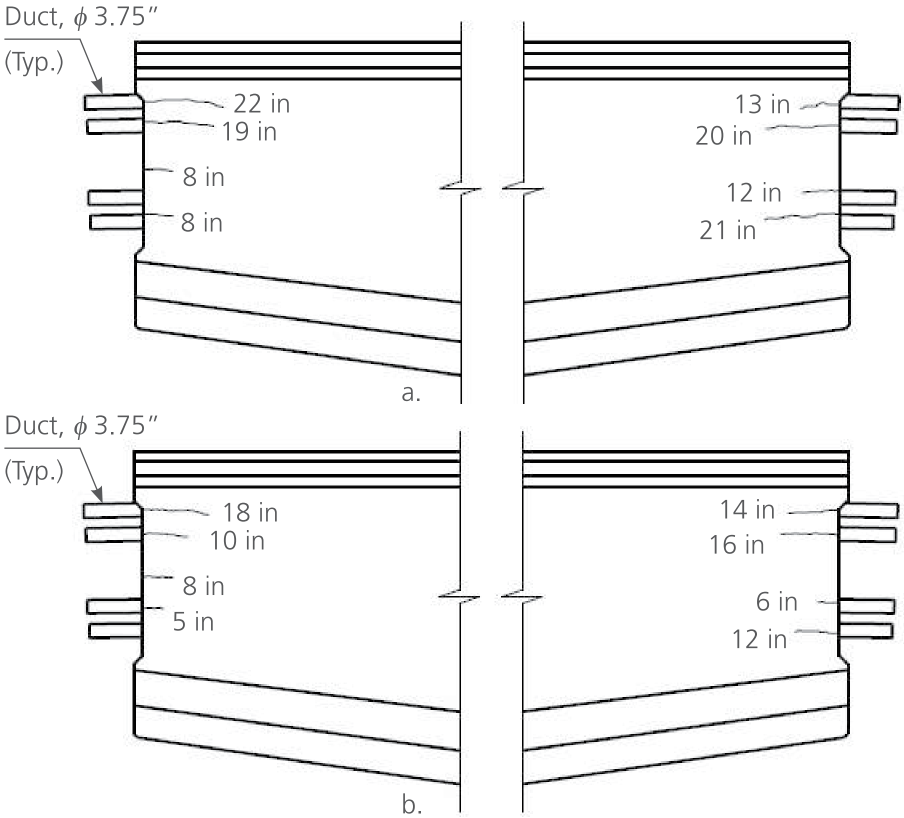


Figure 5. Crack pattern of haunch girders, (a) Girder P2-4NB, (b) Girder P3-1NB (1in =25.4 mm)

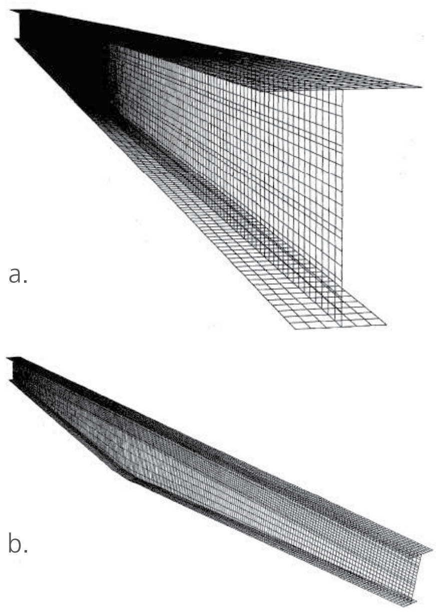


Figure 6. Bridge model, (a) drop in girder, (b) haunch girder

To simplify the quantitative analysis of the effect of the horizontal prestressing forces on vertical tensile stress, the web is assumed to be uniform without any ducts. The vertical tensile stress distribution at fixed and roller supports for the drop-in girder are shown in **Figure 10**. In **Figure 10**, the tensile stresses are caused by the bottom prestressing strands only. **Figure 11** shows the vertical tensile stress distributions along the web vertical direction for four different vertical sections and **Figure 13** shows the vertical tensile stress distributions along the girder longitudinal direction for 6 different horizontal sections. The section locations and numbering are shown in **Figure 11**. In **Figure 11**, V_i represents the vertical sections and H_i represents the horizontal sections. The total tensile force in each cross-section is equal to the tensile stress area multiplied the web thickness and is provided in **Table 1**. The total tensile forces shown in **Table 1** were calculated based on a 0.61 m (2 ft) length from the girder end (see **Figure 12**). The variation of the tensile stress between two nodes is assumed to be linear in determining the tensile stress area. To illustrate how the longitudinal prestressing force relates to the vertical tensile force, **Table 1** gives the ratios of the vertical tensile force to the total prestressing force. From this table, it can be seen that the maximum tensile force occurs at about two fifths of the girder depth from the bottom and is about 7.8 % of the total longitudinal prestressing force.

Using the same method as described above, we can obtain the vertical tensile stress distribution due to the top prestressing strands. The total tensile forces at several cross-sections and the ratios of these tensile stresses to the total prestressing force are shown in **Table 2**. From **Table 2**, it can be observed that the percentages are slightly higher than those caused by the bottom strands.

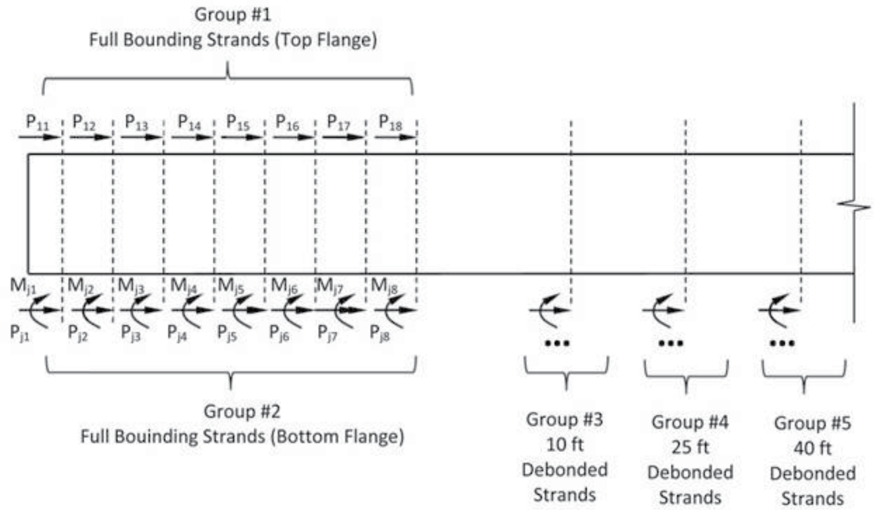


Figure 7. Prestressing force model in bridge longitudinal direction for drop-in girders (1ft =305 mm)

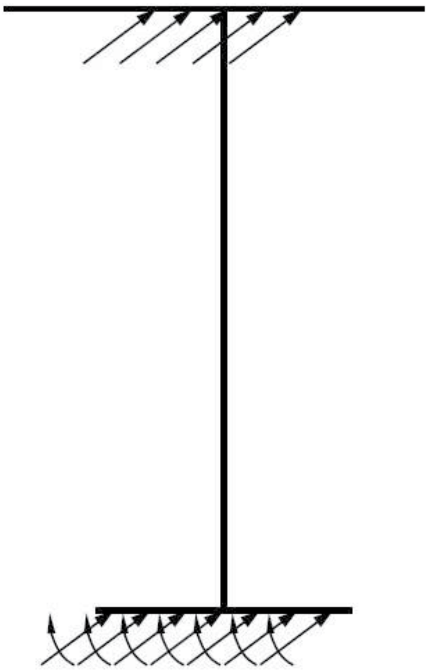
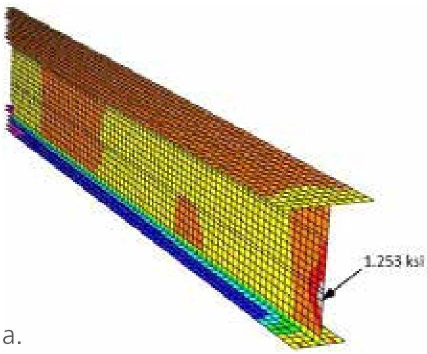
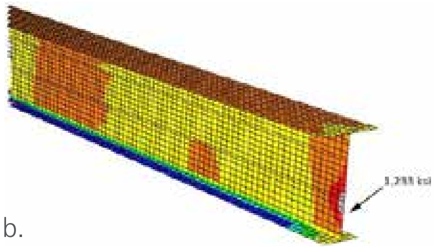


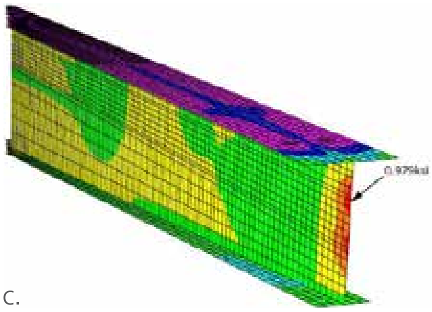
Figure 8. Prestressing force model in bridge transverse direction for drop-in girders



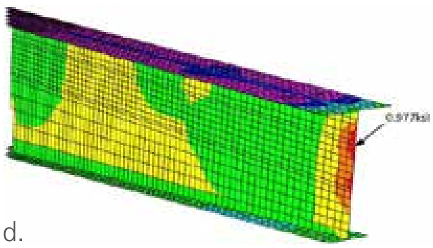
a.



b.



c.



d.

Figure 9. Vertical tensile stress distribution, (a) at fixed end of drop-in girder, (b) at roller support of drop-in girder, (c) at fixed support of haunch girder, (d) at roller support of haunch girder (1 ksi = 6.9 Mpa)

Item	Section	Total Prestressing Force, PTs(kips)	Total Vertical Tensile Force VT (kips)	Percentage VT/PTs x100%
Fixed End	H_1	1142.7	35.226	3.08%
	H_3	1142.7	88.226	7.72%
	H_5	1142.7	88.733	7.77%
	H_7	1142.7	81.546	7.14%
	H_9	1142.7	67.382	5.90%
	H_{11}	1142.7	50.346	4.41%
Free End	H_1	1142.7	35.266	3.09%
	H_3	1142.7	87.972	7.70%
	H_5	1142.7	88.446	7.74%
	H_7	1142.7	81.294	7.11%
	H_9	1142.7	67.626	5.92%
	H_{11}	1142.7	50.502	4.42%

Table 1. Vertical tensile force due to bottom strands (1 ksi = 6.9 Mpa)

Section	Total Prestressing Force, PTs (kips)	Total Vertical Tensile Force VT (kips)	Percentage VT/PTs x100%
H_5	40.0	1.521	3.80%
H_7	40.0	2.147	5.37%
H_9	40.0	2.615	6.54%
H_{11}	40.0	3.089	7.72%
H_{14}	40.0	3.165	7.91%
H_{16}	40.0	2.864	7.20%

Table 2. Vertical tensile stress at fixed end due to top strands (1 ksi = 6.9 Mpa)

Girder Depth (inch)	58 [Marshall and Mattock, 1962]	63 [Marshall, and Mattock, 1962]	70 [Huang and Shahawy, 2005]	78 [Current Research]
V_T/P_{Ts}	2.53%	3.44%	4.85%	7.8%

Table 3. Effect of girder depth

Table 3 presents the effect of girder depth on the percentage of the vertical tensile stress to the total longitudinal prestressing force, based on the results reported by Huang and Shahawy[1], Marshall and Mattock[4], and this research. It is interesting to see that the percentage increases with web depth.

Evaluation of AASHTO specifications

AASHTO Standard Specifications Article 9.22.1 states that in pretensioned beams, vertical stirrups acting at a unit stress of 137.93 Mpa (20 ksi) to resist at least 4% of total prestressing force shall be placed within the distance of one quarter of the girder depth of the end of the beam. AASHTO LRFD Article 5.10.10.1 is roughly equal to this provision and requires that the bursting resistance of pretensioned anchorage zones provided by the vertical reinforcement in the ends of pretensioned beams at the service limit state shall be taken as

$$Pr = f_s A_s$$

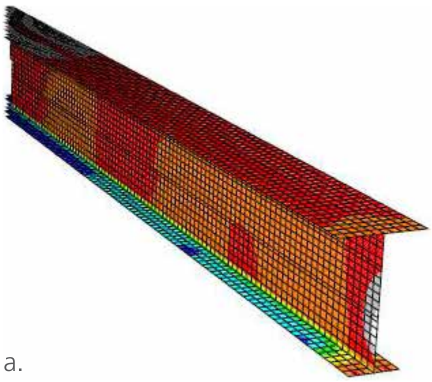
where f_s = stress in steel not exceeding 137.93 Mpa (20 ksi), As = total area of vertical reinforcement located within the distance h/4 from the end of the beam, h= overall depth of precast member. The resistance Pr shall not be less 4% of the prestressing force at transfer. Based on Equation 1, the required reinforcement is:

$A_s = P_r/f_s = (0.04*1182.7)$ kips/20 ksi= 1529 mm^2 (2.37 in^2)

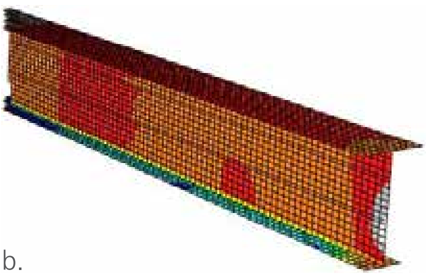
If we used the actual calculated maximum tensile force shown in **Tables 1** and **2**, then, the required reinforcement is:

$A_s = P_r/f_s = 90.25$ kips/20 ksi= 2912 mm^2 (4.513 in^2)

which is almost twice that which the current AASHTO Specifications requires.



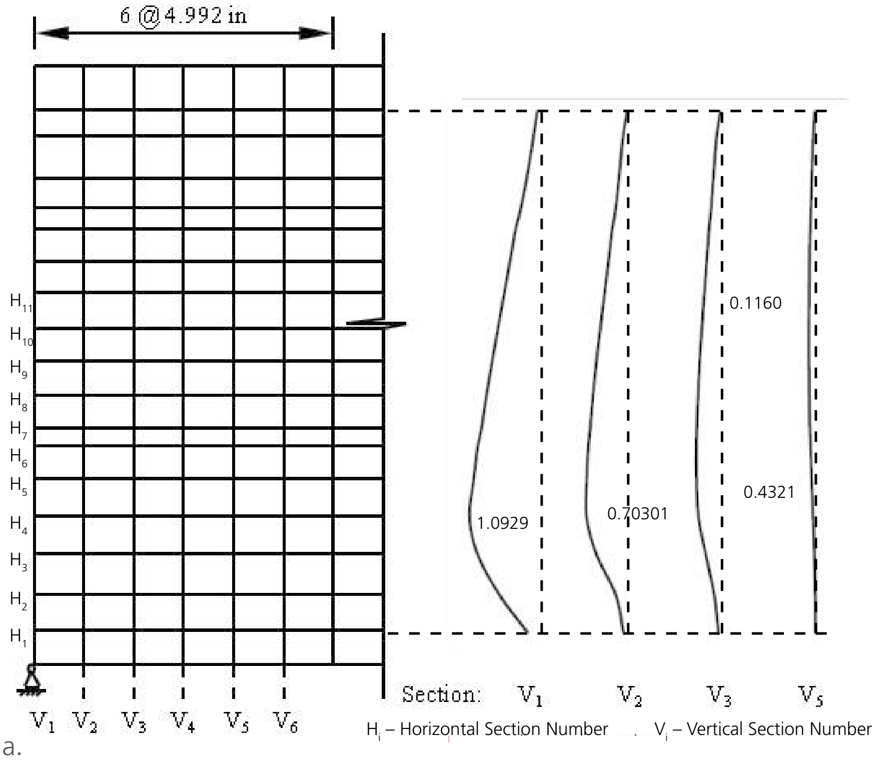
a.



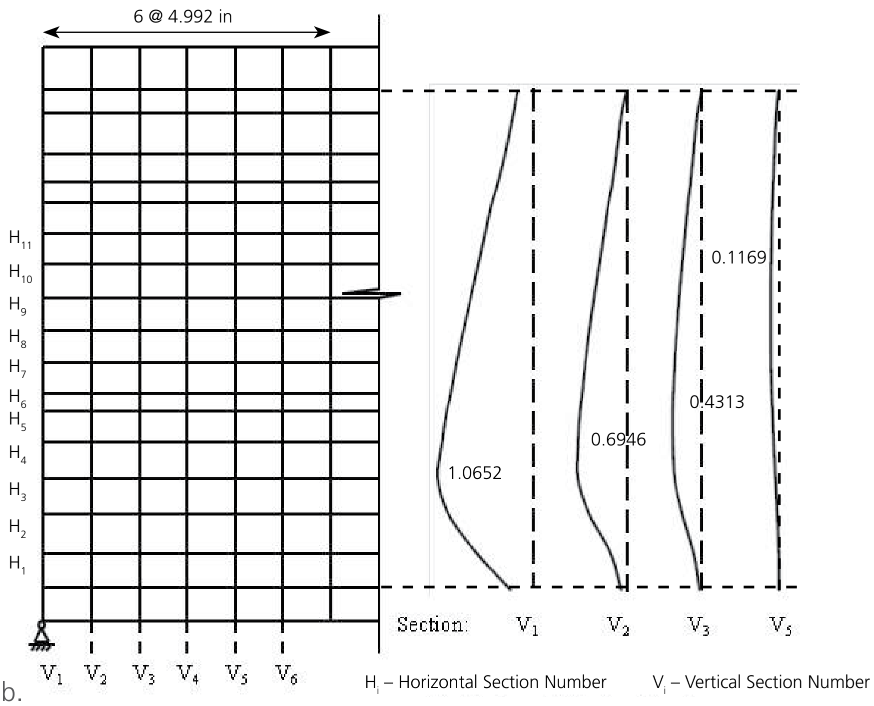
b.

Figure 10. Vertical tensile stress distribution with uniform web due to bottom strands, (a) at fixed support of drop-in girder, (b) at roller support of drop-in girder

There are 25-#5 two leg stirrups distributed at a spacing of 76 mm (3 in) within a length of 1.83 m (6 ft) from the girder end of both haunch and drop-in girders with four extra same size stirrups bounded to the first four stirrups in the end zone. The actual provided vertical reinforcement area within h/4 width is 2890 mm^2 (4.48 in^2) which is much larger than AASHTO requirement of 1529 mm^2 (2.37 in^2) and slightly smaller than the required reinforcement area of 2910 mm2 (4.51 in^2), which should not be a major concern in the cracking control of the beam if there are no ducts in the web. It appears that the source of the cracking is the elevated tensile stresses, especially at the ducts where the section is reduced. But



a.



b.

Figure 11. Typical Vertical Tensile Stress Distribution along Girder Vertical Direction, (a) At Fixed Support, (b) At Roller Support (unit = ksi, 1 in = 25.4 mm, 1ksi = 6.9 Mpa)

there are some other factors that may also contribute to the cracking: (1) the tensile stress is not uniformly distributed along the h/4 width; (2) stress concentration in the duct area which contributes about 15% higher vertical tensile stress than that without ducts; and (3) localized drying shrinkage and endogenous shrinkage which are resisted by hoop stresses in the duct with self consolidating concrete construction.

However, the principal cause of the horizontal cracks is the high vertical tensile stress due to the prestresssing forces which are significantly higher than current AASHTO Specification limits.

Conclusions and recommendations

The main factor that contributed to the spliced girder cracking is the high vertical tensile stress due to the horizontal prestressing force that developed at the end of the beam. The vertical tensile force due to the prestressing force typically increases as the girder depth increases. For the Florida Barge Canal Bridge, the ratio of the vertical tensile force to the total horizontal prestressing force is approximately 8% which is significantly higher than the ratio used by the current AASHTO Specifications to determine the quantity of vertical reinforcement required at the ends of pretensioned members. The post tensioning ducts result in about 15% increase in vertical tensile stress in comparison with that in a uniform web and increase the potential for cracking. To mitigate cracking at the girder ends, it is recommended to properly increase the vertical reinforcement and/or the width of the beam's web within a distance h/4, measured from the spliced end of the girder, especially for girders with a depth exceeding 1.98 m (6.5 ft). Place vertical stirrups acting at a unit stress of 137.93 Mpa (20 ksi) to resist at least 8% of total prestressing forces within a distance of one quarter of the girder depth from the end of the girder with a depth exceeding 1.98 m (6.5 ft).

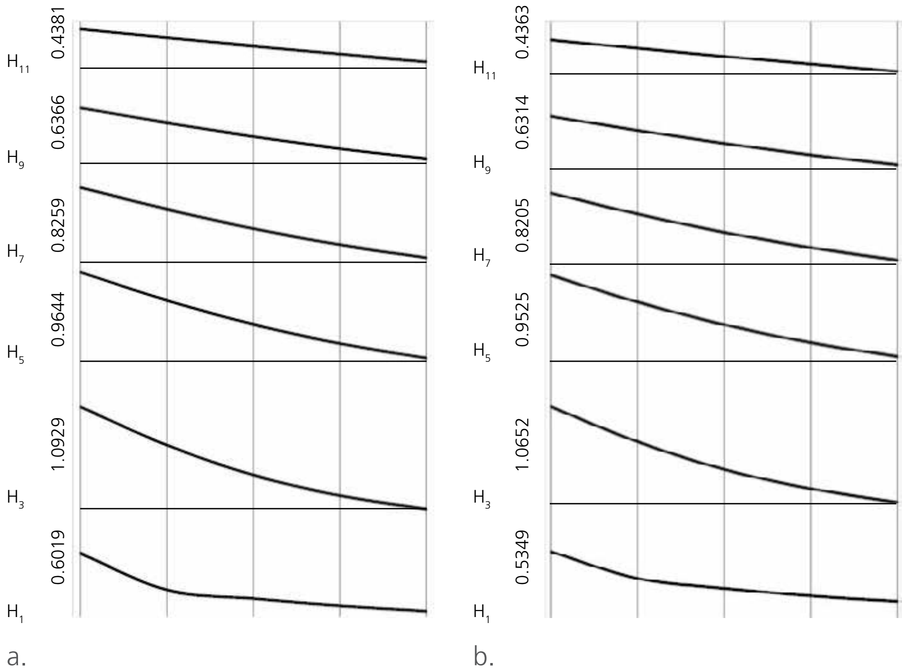


Figure 12. Typical vertical tensile stress distribution along girder longitudinal direction, (a) at fixed support, (b) at roller support (unit = ksi, 1 ksi = 6.9 mpa)

Acknowledgements

The authors would like to express their sincere appreciation to Mr. John Criss, Mr. John Previte, and Dr. Andy Richardson, for their valuable comments and support received during the research. The opinions and conclusions expressed in this paper are those of the author and are not necessarily those of the Florida Department of Transportation.

References

1. Huang, D.Z. and Shahawy, M., Analysis of Tensile Stresses in Transfer Zone of Prestressed Concrete U-Beams, Transportation Research Record: Journal of the Transportation Research Board, No. 1928. Transportation Research Board of the National Academies. Washington. D.C., pp. 134-141, 2005
2. AASHTO, Standard Specifications for Highway Bridges, 17th ed., Washington, D.C. (2002)
3. AASHTO, Load and Resistance Factor Design (LRFD) Bridge Design Specifications, 5th ed., Washington, D.C. (2010)
4. Marshall, W. T., and A. H., Mattock. Control of Horizontal Cracking in the Ends of Prestressed Concrete Girders, Journal of the Prestressed Concrete Institute, Vol. 7, 1962, pp. 56-74 (1962)
5. Gilbert, R.I., Shrinkage, Cracking and Deflection – the serviceability of Concrete Structures, Electronic Journal of Structural Engineering, EJSE International, Vol.1., No.1, pp 2-14 (2001).

156

Structures

Concrete sustainability and alternative concrete materials for highway structures

Abstract

This paper considers and discusses the different types of waste arising in the construction of roads and highway concrete structures with particular emphasis on using secondary materials in the construction of concrete structures. The paper discusses using cement replacement materials to reduce the CO_2 impact of concrete and the carbon footprint of reinforced concrete highway structures. Ground granulated blast furnace slag (GGBS) and high volume fly ash concretes containing 50-70% Pulverised fuel ash (PFA) or more than 65% GGBS are alternative sustainable, high performance concrete mixes with high workability, high ultimate strength, and high durability. This type of concrete has an excellent durability in terms of sulphate resistance and chloride mitigation for use in highway structures. The maximum 55% replacement of cement by PFA in concrete is allowed by the European standard; however, concrete with up to 70% PFA has been successfully used for structural applications.

Keywords: Highways structures, Cement, Concrete, Waste, Recycle, GGBS, PFA.

Homayoon Sadeghi Pouya

Materials Engineer

Highways & Transportation

Atkins

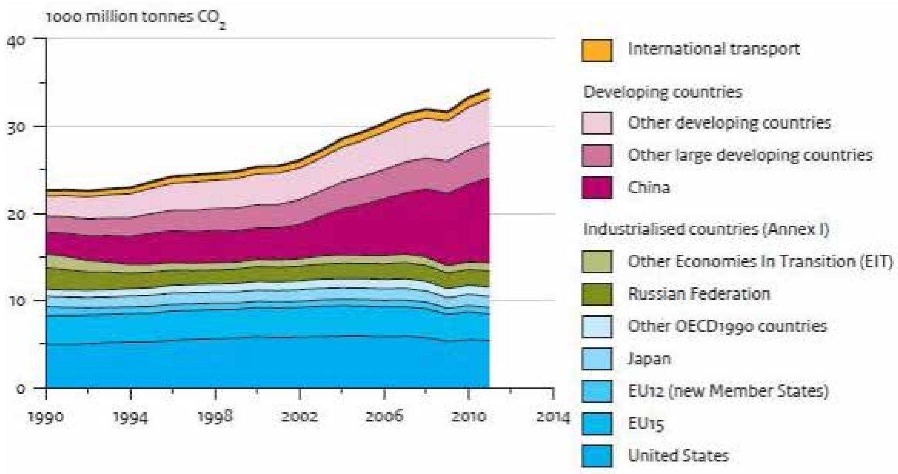


Figure 1. Alternative materials arisings, stockpiles and locations in the UK (million tonnes per annum)[5]

Introduction

Concrete, typically composed of aggregate, sand, water, and hydraulic binders (almost all based on Portland cement), is by far the most widely employed construction material worldwide in terms of volume, and as such, has a huge impact on the environment and also on sustainable development. Produced using readily available raw materials, being easy to use and possessing good strength and durability, concrete is indispensable for meeting modern society's needs for infrastructure, industry and housing. However, significant environmental problems result from the manufacture of Portland cement. Worldwide, the manufacture of Portland cement accounts for 6-7% of the total carbon dioxide (CO_2) produced by humans. This amount of greenhouse gas is equivalent to 330 million cars each driving 12,500 miles per year. In the UK the production of one tonne of Portland cement produces about 0.85 tonnes of greenhouse gases as reported by Cambureau[1].

This paper will briefly discuss the issues of sustainability and the carbon footprint of cement and concrete production, with emphasis on the available options for reducing the use of Portland cement in construction of roads and highway structures.

Portland cement and blended cements

Portland cement (PC) clinker is manufactured by heating a mixture largely consisting of limestone (calcium carbonate) and clay, generally in a rotary kiln. The clinker is then blended with other materials such as gypsum, ground granulated blast furnace slag (GGBS) and pulverised fuel ash (PFA) to produce cement. Approximately 2 million tonnes of GGBS and 0.5 million tonnes of PFA are used in blended cements in the UK. The cement manufacturing process leads to the generation of CO_2 through:

a. The de-carbonation of limestone

b. Energy used in heating the kiln

c. Energy used in grinding the clinker.

Total CO_2 emissions depend on cement composition and the nature of raw materials used, and the efficiency of the manufacturing process. In modern, best practice works, 0.83 tonnes of CO_2 are produced per tonne of Portland cement (about 0.54 tonnes are produced from the de-carbonation of limestone). For less efficient kilns emissions can exceed one tonne of CO_2 per tonne of cement produced.

Cement manufacture generates approximately 10 million tonnes of CO_2 in the UK per annum. Of this, about 59% arises from de-carbonation of limestone and 41% from fuel combustion. Additionally about 34,000 tonnes of NO_x and 18 million tonnes of SO_x are released. Solid waste materials are also produced, including about 63,000 tonnes of cement kiln dust and 31,000 tonnes of other waste. This is currently land-filled although cement kiln dust is increasingly being used in the remediation of contaminated land. **Figure 2** shows environmental impacts of cement manufacture.

In the UK, about 12.7 million tonnes of Portland cement are consumed per annum mainly in the form of ready mix concrete[4].

Use of industrial waste in concrete highway structures

Highways structures

In the context of this paper, highways structures are considered to include bridges, culverts, retaining walls, gantries and those include concrete elements such as decks, beams, slabs, piers, abutments, walls, foundations and piles.

Industrial waste and alternative materials

Table 1 lists the sources, tonnages and arisings of the main alternative materials for construction use in the UK. Recycled aggregates (RA), recycled concrete aggregate (RCA) and foundry sand are well known and suitable for use in certain structural concretes in relation to standards (BS8500 and BS EN12620). WRAP's Aggregain[5] also covers detailed properties and applications of secondary aggregates derived from various waste resources.

Other materials which may have greater potential for use as aggregates are china clay by-products

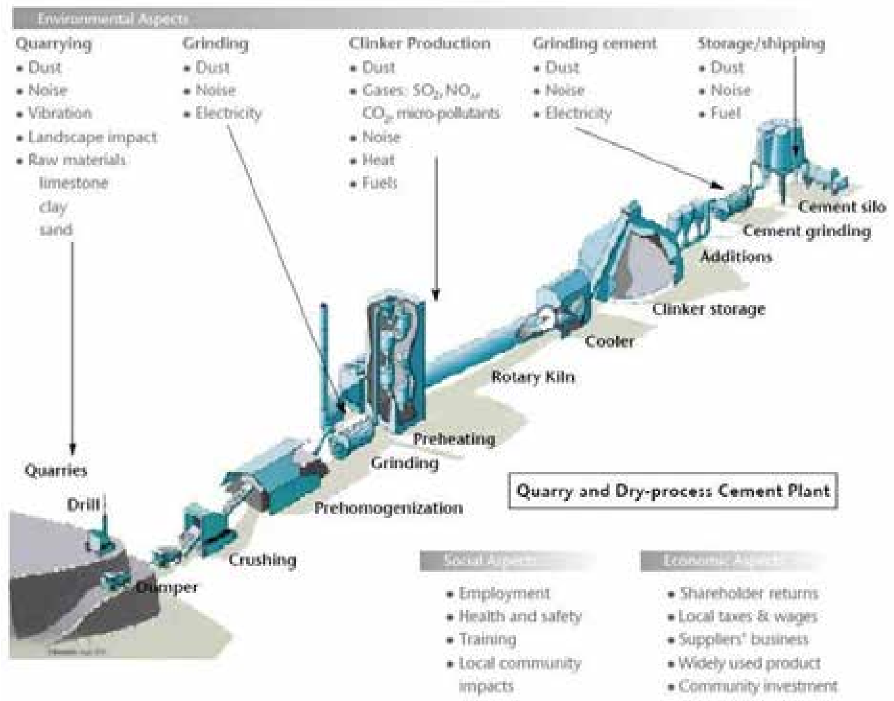


Figure 2. Environmental impacts of quarry and dry-process cement plant[4]

Materials	Arising Mt/a	%	Stockpile Mt	Location
Blast furnace and steel slag	4.0	3.2	NA	N and E England, S Wales
China clay by-products	22.6	14.7	600.0	Cornwall, Devon
Coal and other mining waste	7.5	4.9	15.0	Coalfields, mineral workings
Construction and demolition wastes	94.0	61.0	0.0	Principally urban areas
Crushed brick	5.0	3.2	NA	Demolition sites
Foundry sands	0.9	0.6	0.0	West Midlands and Yorkshire
Glass cullet	2.2	1.4	0.0	Municipal waste
Municipal solid waste incinerated ash and bottom ash	1.35	0.9	0.0	Generally urban areas
Power station fly ash	4.9	3.2	55.0	Coal-fired power stations
Scrap tyres	0.4	0.3	NA	NA
Slate waste	6.3	4.1	466.0	Wales, Scotland, Pennines
Spent oil shale	0.0	0.0	100.0	West Lothian
Spent railway track ballast	1.3	0.8	NA	NA
Waste plastic	2.8	1.8	0.0	NA
Total	154.0	100.0	1236.0	

Table 1. Alternative materials arisings, stockpiles and locations in the UK (million tonnes per annum)[5]

(23 Mt per annum) and municipal solid waste bottom ash, which is processed into Incinerated bottom ash (IBA) aggregate (1.4 Mt per annum).

Use of industrial waste as aggregates

Construction and demolition waste

Construction and demolition waste (CDW) covers a wide range of material including concrete, masonry and asphalt road materials that arise from the demolition of buildings, airfield runways, roads etc. CDW can be reused as substitutes for cement or natural aggregates. There are three main categories of construction and demolition waste: 1) clean crushed concrete, 2) clean crushed brick and 3) crushed demolition debris - including waste from buildings (e.g. rubber, plastic, gypsum etc.)

CDW makes up the most significant component of total alternative material arising as shown in **Table 1**, but could be variable in composition, properties and impurities. The reliability and quality control of the material is therefore an essential requirement.

In 2005, waste protocol projects were defined by the Environment Agency and the Waste and Resources Action Programme (WRAP) to develop clear guidance on how to recover materials in key waste streams, including construction and demolition wastes. The Quality Protocol seeks to promote the environmentally sound use of inert aggregates in various applications such as construction of highway structures.

In spite of the implementation of quality control systems and the improved consistency of the material, in road construction most CDW is still used in general bulk fill, capping and sub-base applications.

While crushed concrete, bricks and reclaimed asphalt from clean sources can easily be used to replace natural aggregate in various applications, including those of structural concrete such as bridges, the presence of contamination in the forms of wood, plaster, soil etc. results in limitation on their wider high-value utilisation[8].

Municipal solid waste incinerator bottom ash

Municipal solid waste incinerator bottom ash (MSWIBA) is the by product produced during the combustion of municipal solid waste in incinerator plant facilities. Fly ash (which is different from power station fly ash) is also collected from the incineration process but with potential high concentrations of toxic materials impeding its use in the construction industry.

The incineration residues mainly contain clinker, glass, ceramics, metal and unburnt organic matters. In the UK, current processing of the material involves only mechanical treatment without chemical processing or washing. This treatment includes extracting metal, screening, removal of unburnt organic matters and natural drying[9]. Storage of MSWIBA for up to three months under controlled conditions is recommended to allow swelling and oxidation ageing to occur, to improve the chemical integrity and structural durability of the ash[10]. Current markets for MSWIBA include aggregate replacement in cement and bitumen bound materials, granular sub-base, masonry blocks and bulk fill applications. However, there are some considerations of chloride and sulphate contents when used in concrete. MSWIBA was also classified as a material that would require measures to protect against leaching[10-12]. The material therefore has more potential for use in cementitious and bituminous bound applications, with reduced leaching potential.

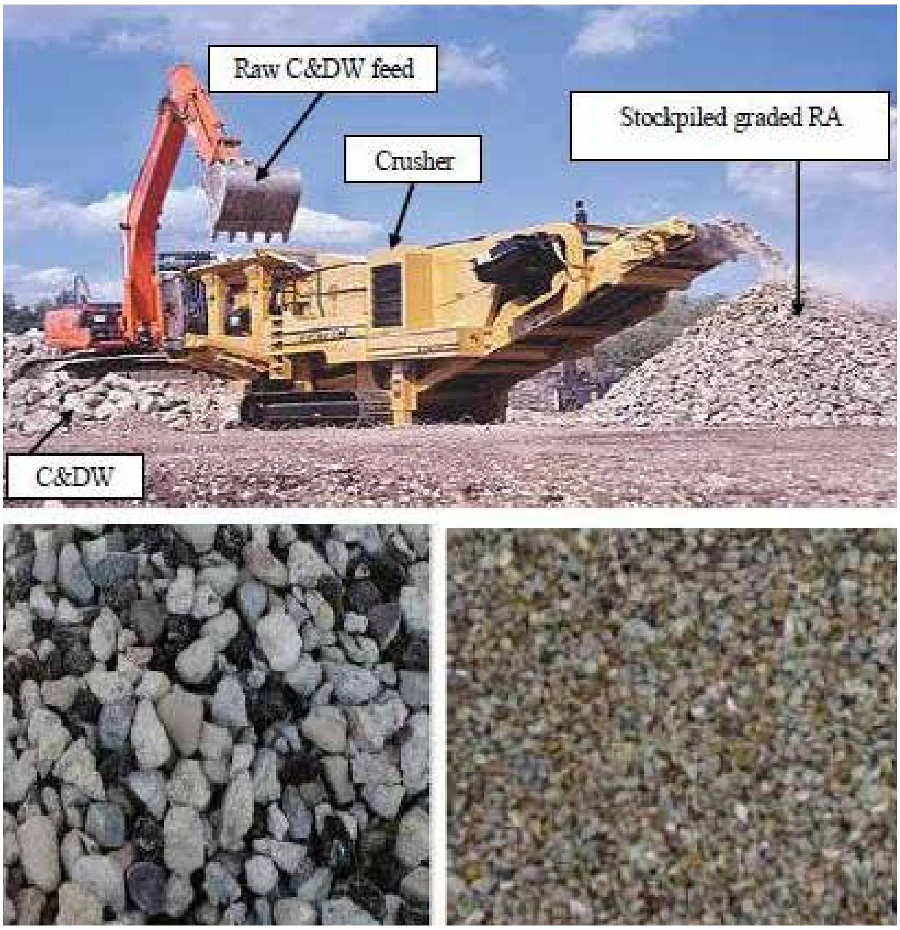


Figure 3. Turning C&DW to recycled aggregates of different sizes

Glass

Glass is another alternative material with a potential use in high value applications. Potential sources of glass include cullet (crushed waste glass), bottle bank glass and other waste flat glass from demolition and replacement window industries. Recycled glass is currently used in various applications including recycled back to glass, aggregate in asphalt and concrete, sand for bedding paving blocks and general replacements of sand in fill applications[5, 7].

Glass cullet is produced from crushing waste glass, collected in municipal and industrial waste streams, to a uniform size. Glass cullet of mixed colours has been used in highway construction as an aggregate substitute in asphalt mixes and also as backfill to structures in the USA and Europe.

There are however some concerns about using the materials as an aggregate in concrete due to the susceptibility to alkali-silica reaction. Work by BRE[13] and the University of Sheffield[14] has shown that glass cullet is highly reactive in terms of ASR but that expansion can be suppressed through combinations including cement replacements (such as PFA, GGBA and metakaolin) and/or appropriate size fractions. Sand-sized glass (1-2 mm) has been shown to cause little or no expansion and thus has potential to be used in concrete at lower risk of expansion than coarser fractions.

It is technically possible to use glass cullet as an aggregate in concrete. However, to date, most specifiers have been unwilling to take the risk, probably due to the relatively low volumes available (around 2 million tonnes per annum), relative to the demand in lower risk uses - such as bedding sand. Currently glass is not permitted in pavement quality concrete in the Design Manual for Roads and Bridges.

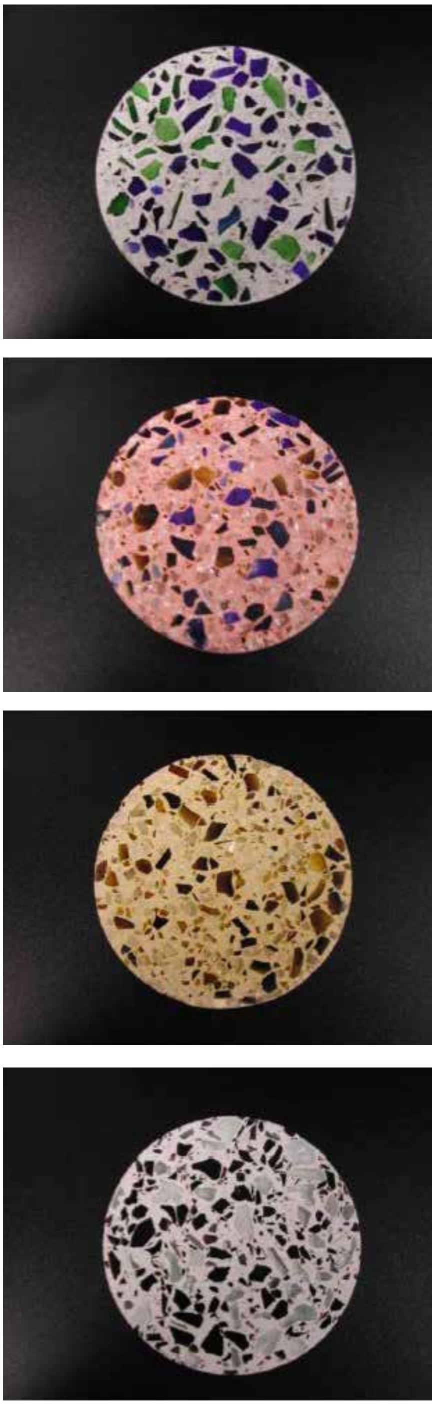


Figure 4. Polished coloured glass concrete as paving block

Rubber

In the UK some 435,000 tonnes of used tyres are produced each year and require some form of treatment prior to disposal. EU Council Directive 1999/31/EC now prevents the disposal of whole tyres and shredded tyres to landfill, so there is a strong drive to find uses for tyre-derived rubber.

Tyre rubber can be recycled in a variety of forms, ranging from bales of tyres, whole tyres through to crumb rubber (< 5 mm) or powdered rubber. Chips and crumb are most appropriate, in terms of size, for aggregate use. The processing of tyres is extremely energy intensive.

Tyre rubber products have a range of potential and existing uses in highway construction including embankment construction, drainage, backfill and in retaining walls. Chipped tyres have been used to provide free draining backfill to retaining structures such as bridge abutments.

A number of researches have been conducted on the use of tyre-derived rubber in concretes. It has been shown that if the amount of rubber in the concrete is limited (to around 10% replacement of coarse aggregate by volume), a normal strength concrete could be produced with enhanced properties such as improved flexibility and reduced weight[17, 18].

Plastics

The amount of plastic waste currently generated in the UK is about 2.8 million tonnes per annum. The two main types of plastic are thermoplastics, which soften when heated and harden again when cooled), and thermosetting plastics, (which harden by curing and cannot be re-moulded. Thermoplastics are by far the most common types, forming approximately 80% of the plastic waste stream, and are also the most easily recyclable. Sources of waste plastics that are of sufficiently low value to be used in construction are generally mixed plastics, derived from municipal solid waste or the automotive industry. There are some track records for the use of pelletised or shredded plastic waste in concrete and cold mix asphalt pavements. Lightweight aggregate has also been produced. This can be used in the manufacture of low-density concrete. It also has potential for use as backfill to structures or drains.

A study conducted by Tarmac Ltd. and the Transport Research Laboratory (TRL) assessed the use of waste plastics from the automotive industry to replace the sand fraction in concrete. The inclusion of plastic led to a reduction in density and also in compressive strength. The compressive strength was 40 MPa at 28 days, with plastic replacing 15% of the sand, compared with 49 MPa for the control mix without plastic[18].

Plastic aggregates are well suited to products such as kerb stones or temporary crash barriers, where weight can be important. The structural properties and durability of products containing plastic aggregates in terms of shrinkage, creep etc. should be assessed prior to use in wider applications[19, 20].

Standards and specifications for use of industrial waste as aggregates in concrete

The current specifications and standards for concrete already permit a range of secondary materials - residues from industrial processes including mining to be incorporated.

- BS EN 12620 (aggregates for concrete)[21] specifies the properties of aggregates for concrete and does not distinguish between primary, secondary or recycled aggregates. Inert filler aggregates added to concrete are mentioned (material generally passing a 0.063 mm sieve), a role which may suit some industrial by-products

- BS8500-2 (Concrete)[22] specifies the use of aggregates meeting the requirement of BS EN 12620. It specifically mentions the circumstances for use of recycled aggregates (RA and RCA) and air-cooled blast furnace slag aggregates
- Series 1700 (structural concrete) of the Manual of Contract Documents for Highway Works Volume 1 (Specification for Highway Works)[23] allows for the use of aggregates conforming to the standards listed in BS 8500-2 (i.e. BS EN 12620) but specifically excludes the use of RA and RCA, unless allowed in contract-specific information
- Design Manual for Roads and Bridges (DMRB Volume 7, Section 1, Part 2, HD35/04, Pavement design and maintenance, Preamble, Conservation and the use of secondary and recycled materials) gives a range of materials that are permitted in pavement quality concrete (Table 2)[24]. This provides a useful reference point for aggregates that could potentially be used in concrete structures. The specification currently has no specific provision for the use of different materials. These include glass, plastic waste and tyre rubber.

Use of industrial waste and by-products as cement replacements and alternative binders

Cement replacements include natural materials such as certain volcanic rocks and industrial by-products such as GGBS, PFA and silica fume. Such materials contribute to the long-term properties of the concrete, reducing the amount of Portland cement required. Ignoring the additional benefits of using these materials, at their simplest level, they act as cement extenders, thus reducing the amount of greenhouse gases associated with each tonne of cement product (between 7 and 13%) and the natural resources extracted per tonne. The UK has an extensive history of the use of GGBS and PFA in cement products[25, 26]. However, the benefits of this approach are limited by the availability of cement replacements.

Application	Pavement Quality Concrete (SHW Series 1000)
Recycled aggregate	Not permitted
Recycled concrete	Specific or general provision
Foundry sand	Specific or general provision
Recycled asphalt	Not permitted
Blast-furnace slag	Specific or general provision
Steel slag	Not permitted
Burnt colliery spoil	Not permitted
Unburnt colliery spoil	Not permitted
Spent oil shale	Not permitted
Pulverised fuel ash	Specific or general provision
Furnace bottom ash	Not permitted
China clay sand/stent	Specific or general provision
Incinerator bottom ash (IBA)	Specific or general provision
Recycled glass	Not permitted
Phosphoric slag	Specific or general provision
Slate aggregate	Specific or general provision

Table 2. Permitted uses of secondary aggregates in the Specification for Highway Works [23]

The benefit in terms of CO_2 reduction gained from the use of cement replacements can be roughly estimated as the CO_2 associated with the cement not used. In the case of GGBS, large replacement rates of up to 70% mean that the CO_2 reduction can be considerable.

It should be remembered that GGBS supply is limited to approximately 2.5 million tonnes in the UK and that GGBS contents in concrete are subject to limits based on the short-term and long-term properties of the concrete.

Type	Addition	Portland cement replacement, %
CEM I		0 – 5
IIA	Silica fume	6 – 10
	Fly ash	6 – 20
IIB-V	Fly ash	21 – 35
IVB-V		36 – 55
IIB-S	GGBS	21 – 35
IIIA		36 – 65
IIIB		66 – 80

Table 3. Cements and combinations available with various cement replacement materials

High volume cement replacement materials in concrete

A new approach to the production of blended or high-volume mineral additive (HVMA) cements helps to improve its environmental credentials. HVMA cement technology is based on the inter-grinding of Portland cement clinker, gypsum, mineral additives and special admixtures. Alternatively, the mineral additive can be added to the concrete at the mixer. This approach can lead to improved durability and reduce costs through the use of inexpensive natural mineral additives or industrial by-products such as fly ash or ground granulated blast furnace slag. This improvement leads to a reduction of energy consumption per unit of the cement produced. Higher strength, better durability, reduction in CO_2 emissions at the clinker production stage, and decrease of landfill area occupied by industrial by-products, all provide ecological advantages for HVMA cements. The use of various pozzolanic and hydraulic materials to replace a high volume of Portland cement in concrete is described in the following sections.

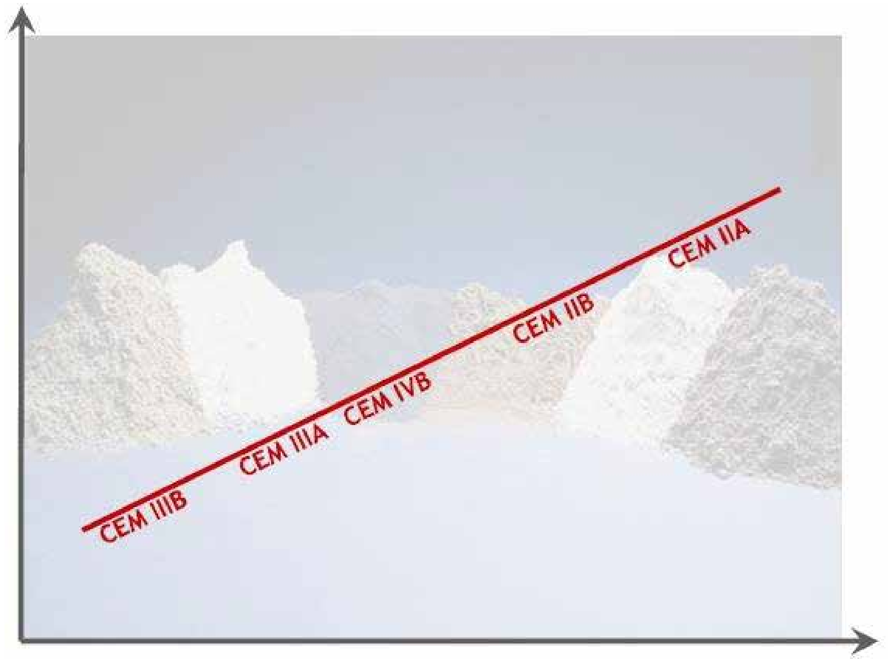


Figure 5. Embodied carbon in cement containing various cement replacement materials

High volume PFA concrete

Fly ash, a principal by-product of coal-fired power plants, is well accepted as a pozzolanic material that may be used either as a component of blended Portland cements or as a mineral admixture in concrete. Fly ash has been widely used in industrial products; however, a large amount fly ash is still being sent to landfill[27]. In commercial practice, the dosage of fly ash in concrete, was limited to 15%-20% by mass of the total cementitious material. Usually, this amount has a beneficial effect on the workability, chloride diffusion and cost economy of concrete but it may not be enough to sufficiently improve the durability to sulphate attack, alkali-silica expansion, and thermal cracking. For this purpose, larger amounts of fly ash, of the order of 25%-35% are now being used.

Although 25%-35% fly ash by mass of the cementitious material is considerably higher than 15%-20%, this is not high enough to classify the mixtures as High volume fly ash concrete (HVFAC) according to the definition proposed by Malhotra and Mehta[28]. From theoretical considerations and practical experience Malhotra et al. determined that, with 50% or more cement replacement by fly ash, it is possible to produce sustainable, high performance concrete mixtures that show high workability, high ultimate strength and high durability.

The HVFAC is one specific type of fly ash concrete with higher fly ash contents, lower water-to-cementitious materials ratio and lower cement contents. This is to take full advantage of the increased workability and durability provided by fly ash and the low w/c, and to produce a more environmentally friendly concrete by reducing its cement content. The main difference between the HVFAC and the usual fly ash concrete is that in the former concrete, the amount of Portland cement is minimized through proper mixture proportioning using large amounts of fly ash and judicious selection of materials and chemical admixtures while maintaining, and often improving its performance as compared to conventional concrete. To obtain the superior performance of this type of concrete, the w/c of the HVFAC should be kept well below 0.40 and, preferably of the order of 0.35 or less. To produce a workable concrete at such low w/c ratio, the use of superplasticiser is generally essential.

Fresh properties of HVFAC

It is generally observed that a partial substitution of Portland cement by fly ash in a mortar or concrete mixture reduces that water requirement for obtaining a given consistency. Experimental studies by Malhotra[28, 29] have shown that with HVFA concrete mixtures, depending on the quality of fly ash and the amount of cement replaced, up to 20% reduction in water requirements can be achieved. This means that good fly ash has a plasticising effect when used in high-volume. Cabrera et al [30] also reported that carbon content of the fly ash affects the concrete properties to different degrees. The high carbon content may result in increased or reduced flowability of mixtures depending on type and shape of carbon particles and the process by which coal is burnt. Therefore, a high volume fly ash (>35%) can be used in concrete, providing that the type of carbon is identified and fly ash with the right combination of carbon content and particle shape is selected.

Properties of hardened concrete

i) Drying shrinkage

Perhaps the greatest disadvantage associated with the use of neat Portland-cement concrete with high cement content is cracking due to drying shrinkage. The drying shrinkage of concrete is directly influenced by the amount of the cement paste present. It therefore increases with an increase in the cement paste-to-aggregate ratio in the concrete mixture, and also increases with the water content of the paste. The water-reducing property of fly ash can be advantageously used for achieving a considerable reduction in the drying shrinkage of concrete mixtures (by requiring less water for a given workability). This is due to the resulting significant reduction in the cement paste-to aggregate volume ratio (up to 30%) in the HVFA concrete compared to ordinary concrete[31].

ii) Permeability and durability

In general, the resistance of a reinforced-concrete structure to corrosion, alkali-aggregate expansion, sulphate and other forms of chemical attack depends on the permeability of the concrete. The permeability is greatly influenced by the amount of water, type and amount of supplementary cementing materials, curing, and cracking resistance of concrete. HVFA concrete mixtures, when properly cured, are able to provide low permeability and durable concrete[31]. However, it should be noted that the performance of concrete containing high volume mineral admixtures may not be covered by existing codes of practice.

With respect to chloride diffusion, HVFA concrete with and without water reducing admixture significantly reduces chloride ion penetration and the corrosion of reinforcing steel. Percentage replacement of Portland cement by fly ash from 35% to 65% is effective in decreasing chloride permeability, especially in the concrete cover zone. The risk of reinforcement corrosion in HVFA concrete with percentage cement replacement from 50% to 65% was shown to reduce significantly compared to control concrete (with the same cover to reinforcement. Also it was reported that the effectiveness of the corrosion risk reduction was independent of compressive strength. It should be noted that chloride penetration and corrosion risk reduction also depend on quality of concrete ion terms of compaction and cohesiveness[31].

Electrochemical data using polarization resistance techniques on cement paste subjected to chloride indicated that, even with limited initial curing of seven days, the corrosion rate of steel in concrete containing 50 per cent fly ash was similar to that of plain cement at high water-to-binder ratio (>0.6) and were lower than that of plain cement at low water-to-binder ratio (0.45). Laboratory studies also indicated that steel passivation characteristics improved with age of hydration and that there was no negative effect caused by lowering in pH due to the pozzolanic reaction[32].

iii) Compressive strength

The effect of fly ash content on compressive strength of concrete specimens with the same w/c ratios was investigated by Guangcheng et al[33]. It can be seen in **Figure 1** that although the compressive strength of mortars reduces with increasing of the amount of fly ash above 30%, the strength of mixes containing 60% fly ash was still in the range of 40 to 50 MPa and greater than the control OPC mixes. As shown in **Figure 2** the unit strength factor (ratio of strength of HVFAC to OPC concrete) for mixes incorporating fly ash was slightly greater than 1.0 indicating that structurally acceptable mixes can be produced with a high volume replacement of cement by fly ash. According to Cyr et al[34] fly ash particles have physical filler effects and pozzolanic effects in cementitious matrix.

Naik et al[35] also reported that concrete made for highway applications with large amounts of ASTM (American Society for Testing and Materials) Class C and Class F fly ash performed well over the long-term.

With respect to above data, although several studies indicated that structurally acceptable concrete with excellent durability can be produced using high volume fly ash, the properties of hardened concrete and long term performance of these mixes depend on cement content, w/b ratio, curing, storage condition and age of specimens which may vary in various experiments carried out by several researchers. Therefore, care should be taken where these types of concrete are specified for structural applications.

Concrete construction practice

Experience obtained in field trials with high-volume fly ash concretes showed that they can be mixed, transported, placed, and finished using conventional concreting equipment and techniques.

Due to the high volume of fines and a low water content, fresh concrete mixtures of the HVFA system are generally very cohesive and show little or no bleeding and segregation. They show excellent pumpability and workability at slumps as low as 75 mm. The material moves well to fill space without much effort and behaves almost like a self-consolidating concrete. Consequently, the surface finish is usually smooth and without honeycombing and blowholes.

Due to the lower Portland cement content, HVFA concrete mixtures may take one to two hours longer to set. Accelerating admixtures can be used providing that their compatibility with the actual concrete mixture has been adequately tested. Usually HVFA concrete mixtures do not suffer excessive slump loss in a short period. Jobsite re-tempering of ready-mixed HVFA concrete is permissible to restore severe slump loss with a small amount of superplasticiser or water, provided the w/c ratio does not exceed the specified limit[36].

It should be noted that low water/cement and non-bleeding concrete mixtures are vulnerable to plastic shrinkage cracking, as well as autogenous shrinkage cracking. Therefore, concrete surfaces must be protected from any water loss by operating a water-fogger around the structure during the placement, or by covering the surface with a heavy plastic sheet immediately after the placement of concrete. A minimum of 7 days of moist-curing is necessary to achieve the optimum strength and durability characteristics that are possible from the use of HVFA concrete. With foundations, piers, columns and beams, leaving the form work in place for at least a week is acceptable in lieu of moist-curing[37].

Site experiences

One of the first applications of HVFA was an unreinforced concrete pavement constructed in the US in the 1970s[38]. In Canada, beginning with a massive concrete foundation built in 1987, reinforced columns, beams and floor slabs of an office complex were cast using HVFA concrete in 1988[39]. Mehta and Langley[40] reported the construction experience with a large (36 by 17 by 1.2 m) HVFA concrete foundation that has remained crack free and performed well in the long term. Similar experience with large foundation slabs, cast-in-place drilled piers, and caissons was reported from recently built structures in Houston and Chicago[40]. The HVFAC also has been used in a number of structures in Eastern Canada since the late 1980s. It was found that those field applications that properly designed and cured, HVFAC demonstrated

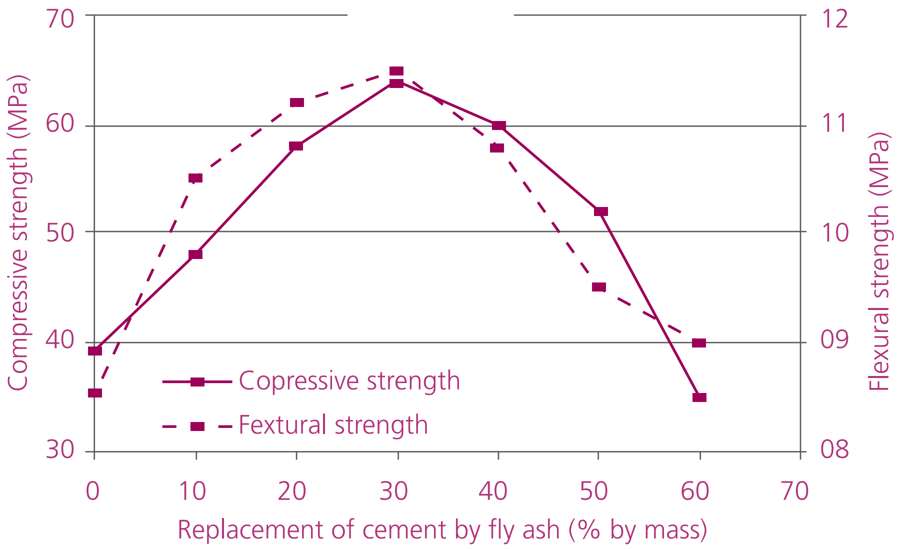


Figure 6. Effect of fly ash content on compressive strength of mortar mixes (Water/Binder: 0.3)[32]

Figure 7. Effect of fly ash content on unit strength factor of mortar mixes (Water/Binder: 0.3)[32]

excellent performance, both in mechanical as well as in durability aspects. HVFAC with medium to high strength using 55-60% fly ash was used to construct the walls, columns and slabs of Brentwood Skyline station in Canada. Cube strengths of 50 MPa and over 100 MPa have been recorded in laboratory specimens after 28 days and one year moist curing[41]. Mehta[42] and Manmohan[43] have also documented the successful construction experience with another HVFA concrete project involving reinforced belt foundations, shear walls, and collector beams for the seismic upgrade of a building at the University of California campus at Berkeley.

A recent Danish study[44] reported that a low carbon foot print concrete ("green" concrete) incorporating about 40% fly ash was used for construction of a road bridge, demonstrating the most promising "green" concrete composition at full scale. The concrete used was based on blends of fly ash and Portland cement based on low-energy mineralized clinker. Durability of the concretes was tested with respect to chloride ingress, carbonation, freeze–thaw resistance and alkali–silica reaction. No significant differences between reference and "green" concretes were found. The embodied CO_2 emission of the "green" concrete was evaluated and a total reduction of about 38% compared to conventional concrete was reported.

Standards

European standard EN197-1:2000 allows maximum percentages of fly ash of 35% in CEM II, 55% in CEM IV (not marketed in Belgium), and 50% in CEM V. Apart from use in these composite cements, fly ash can be added - possibly in higher quantities as a separate component to concrete mixtures; however, the effect on structural properties and durability has to be taken into account. The Canadian CANMET Institute first coined the term "high-volume fly ash concrete (HVFA)" for structural concrete with high volumes (>50% of binder) of fly ash[45].

BS 8500-1:2006 also allows the use of high volume fly ash (36% to 55%) cement known as CEM IVB-V in structural concrete in all design concrete classes. This type of cement is specified by the BS particularly for those exposure conditions with high concentrations of sulphates and chloride where a high resistance to aggressive environment is required.

Blast furnace slag and steel-making slags

Slags are non-metallic by-products from metal manufacturing. The slag occurs as a molten liquid melt and is a complex solution of silicates and oxides that solidifies upon rapid cooling by immersion in water[46, 47]. Ground granulated blast furnace slag (GGBS) is a by-product of the iron industry. Its cementitious properties have been known for some time. Since the 1950s, use of GGBS as a separate cementitious material has become widespread in many different countries. GGBS is not only used as a partial Portland cement replacement, but also (in its air cooled form), as an aggregate.

The optimum cement replacement level of GGBS is often quoted to be about 50% and sometimes as high as 70 and 80% (by weight of cement). Like fly ash, GGBS also improves many mechanical and durability properties of concrete and generates less heat of hydration than the same amount of Portland cement. For example, recently the 2.7 m thick foundation slab for a water treatment plant in New York City was constructed using 70% slag and 30% Portland cement. One of the major design objectives was to minimize temperature differentials due to heat of hydration without the installation of a potentially costly internal cooling system and thereby satisfy the rather stringent specifications regarding the elimination of cracks. In many situations so-called ternary systems, that is, blends of Portland cement and silica fume, GGBS, has become popular. In Europe the practice of pre-blending cements and various pozzolans is widespread. The cost of slag is generally of the same order as that of Portland cement. Primarily because of its known beneficial properties, customers are willing to pay as much for the slag as for the cement it replaces. It should be remembered that GGBS supply is limited to roughly 2.5 million tonnes in the UK and that GGBS contents in concrete are subject to limits based on the short-term and long-term properties of the concrete[48].

Although the steel industry probably generates the largest amount of slag, several other metallurgical slags are produced today that are still being mostly stockpiled, landfilled, or "downcycled" into low-value applications such as road base. Such disposal methods carry their environmental costs, especially since these materials often contain toxic metals that may leach out and contaminate the ground water. Basic oxygen steel (BOS) slag from steel manufacture has an established track record as coarse aggregate in asphalt. Processing of the material generates a large amount of fine dust, about 40-50% of which is not utilised. The main barrier to the use of BOS in concrete is the tendency of the material to expand due to the presence of free oxides of magnesium and calcium. In most applications, the slag needs to be weathered before use to minimise the dimensional instability of the material[49]. Recent studies have shown that such slag can be used beneficially in concrete applications. Concrete industry offers ideal conditions for the beneficial use of such slags and ashes because the harmful metals can be immobilized and safely incorporated into the hydration products of cement[50].

Environmental impact of HVMA cements

Blended cements incorporating the supplementary cementitious materials described above can partly replace the Portland cement clinker in concretes. Clearly, blended cements meet the challenges of modern society by increasing bulk production and conserving energy[52,53,54,55]. In the case of CO_2 emissions, which have been "frozen" by the EU and US at the level of 2000, the share of conventional Portland cement in the market must be reduced drastically by the year 2015 (**Figure 8**)[53].

At the same time, the extensive updating of existing facilities for the manufacture of clinker consumes a great deal of capital investment and yields only a slow return[56]. Therefore, the expansion of an existing cement plant requires a proportionally high rate of investment. The major part of this investment is associated with the installation of heavy equipment and construction. However, in the case of HVMA cement, new investments are required only to upgrade the grinding unit. This may increase production capacity by 40–50% with consequent increase in profit, but without any additional increase of clinker output. Since HVMA cement uses about 30–50% less clinker than Portland cement, it will create less ecological damage. It should be also noted that increasing the amount of cement replacement materials in a particular concrete mix may not provide a net environmental benefit in comparison with using concrete mixes containing lower amounts of these materials spread more widely.

The environmental and cost benefits accrued by using HVMA concrete in place of normal concrete for each project can be estimated. For instance at the compressive strength achieved by HVFA concrete containing 65% PFA (equal to a normal concrete with strength of 35 to 40 MPa and cement content 400 kg/m^3), the Portland cement content is reduced by 140 kg/m^3, resulting in total savings of about 140 tonnes of Portland cement for each 1000 m^3 of concrete which is used for a typical construction project. This is equivalent to about 140 tonnes of carbon emissions, along with the energy savings associated with Portland cement production. Similarly, with an average fly ash content of 260 kg/m^3, a typical construction project uses 260 tonnes of fly ash for each 1000 m^3 of concrete that would otherwise have been sent to landfill[56].

Glass waste as a binder

Glass ground to cement fineness has been shown to act as a pozzolanic cementitious binder in concrete with a strength activity comparable to that of PFA[57, 58] with up to 30% replacement of Portland cement by glass. The main limitation to use is the limited supply due to the high cost and energy input required to grind glass to cement fineness. Ground glass at certain grading can also suppress ASR expansion[58].

Sulphate waste as a binder

Sulphate industrial by-products include the following:

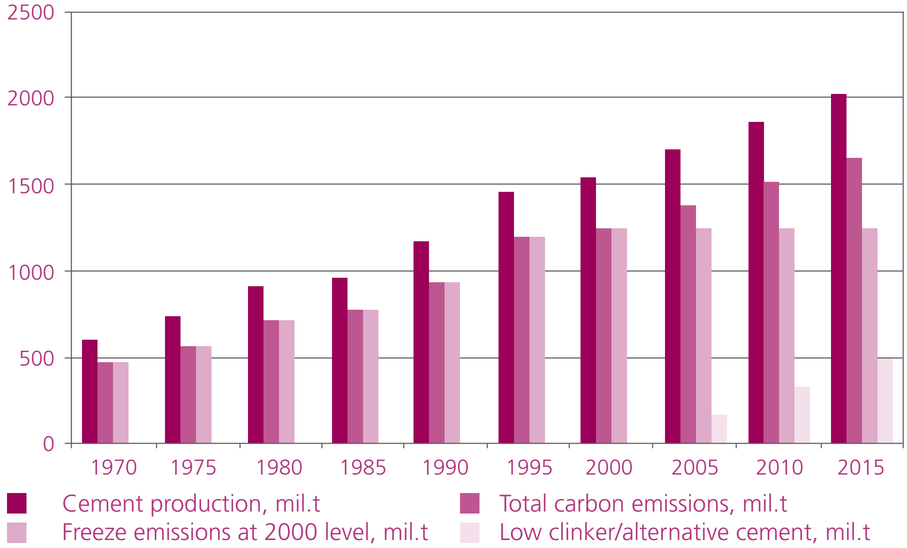


Figure 8. Prediction of emission and market share of low-clinker[53]

a. Flue gas desulphurisation (FGD) gypsum (gypsum from flue gas desulphurisation associated with coal fired power generation)

b. Fluorogypsum (from fertiliser manufacture)

c. Titanogypsum (gypsum from titanium dioxide manufacture)

d. Red gypsum (contaminated gypsum from titanium dioxide manufacture)

e. Waste gypsum derived from construction and demolition activities (plasterboard) - a significant processing industry is developing around this

f. Miscellaneous (e.g. gypsum neutralisation of waste battery acids).

FGD, titanogypsum and (to an extent) waste gypsum from construction and demolition find ready markets in the plasterboard industry[59]. The potential for use of sulphate or gypsum wastes in conventional concrete structures is limited due to the possibility of deleterious sulphate interactions and rapid setting reactions with Portland cement. Nevertheless, there are opportunities for gypsum or sulphate residues in concrete combination with low carbon calcium sulphoaluminate (CSA) cement binder. This cement is, however, not widely available.

Studies have shown that calcium sulphate-GGBS-Portland cement binders can develop significant strength. However, where calcium sulphate is the dominant ingredient, such formulations are only suitable for dry environments (due to expansion in the presence of moisture)[60, 61]; waste gypsum has, however, been applied successfully in stabilised road-base construction[62].

Conclusions

Concrete is the most widely consumed substance in the world with the production of 12 billion tonnes per annum, thus the impact of concrete on sustainable construction is crucial. The production of cement as the main component of concrete generates approximately 10 million tonnes of CO_2 in the UK per annum. There are a number of approaches to reduce the energy consumption and CO_2 emission in cement manufacture including the use of blended cements, use of alternative binder materials in concrete and the use of recycled materials. These approaches have been reviewed and the following conclusions can be drawn:

- Lack of knowledge of material performance characteristics and their behaviour in composites, together with the lack of end-performance specifications constitute the main technical barriers to the use of new materials and applications
- High Volume Mineral Additive (HVMA) cements contain less cement clinker than existing cements and so help to meet the challenges of a sustainable society; HVMA cements also contribute to the reduction of CO_2 and other emissions at source. Due to its better eco-credentials, the market share of HVMA cements can increase in the future provided there is good availability of raw materials. Also because of its strength and durability, the production and application of high performance blended cement has the potential to reduce the environmental degradation associated with construction activities
- High volume fly ash concretes containing 50-70% PFA are sustainable, high performance concrete mixes that show high workability, high ultimate strength, and high durability. The higher fly ash contents and lower water-to-cementitious materials ratio give this type of concrete an excellent durability in terms of sulphate resistance and mitigation of chloride diffusion for use in highway structures. The maximum 55% replacement of cement by PFA in concrete is allowed by the European standard.

References

1. www.cembureau.be
2. IEA (2011b). CO2 from fuel combustion. 2011 Edition. International Energy Agency, Paris
3. The UK Cement Industry, Benefit and Cost Analysis, Forum for the Future, Sept 2005 (available from BCA website)
4. BCA website www.bca.co.uk
5. www.aggregain.org.uk
6. Ellis, S.J., P3: Construction materials, sustainable construction in practice (SCiP). TRL Limited 200
7. www.wrap.org.uk
8. Rilem Report 22, Sustainable raw materials construction and demolition waste, Rilem publications SARL, 2000
9. York, D.M., MSWI plants: Sustainable quarries. Proceedings of the International Symposium on "Use of incinerator ash", Scotland, University of Dundee, 2000, pp.259-265
10. CIRIA C513 (1999). The reclaimed and recycled construction materials handbook. Construction Industry Research and Information Association Report. London: CIRIA
11. British Standards, BS 6543 (1985). Guide to use of industrial by-products and waste materials in building and civil engineering
12. CIRIA Report 167, Use of industrial by-products in road construction – water quality effects, construction Industry Research and Information Association Report, 1997
13. Moulinier, F., Lane, S and Dunster, A (2006). The use of glass as aggregate in concrete: final report (project report GLA0046), WRAP 2006, 21 pp http://www.wrap.org.uk/downloads/conglasscrete_final_March06_WRAP_AD3_2_.cefcf0c4.3215.pdf
14. Byars, E. A., Morales, B and Zhu, Y (2004) Conglasscrete 2- final report, WRAP 2004, 34 pp http://www.wrap.org.uk/downloads/ConGlassCrete2FinalBodyi.e6295b60.504.pdf
15. El-Dieb, A.S., Concrete using rubber tyre particles as aggregate. Proceedings. Recycling and reuse of used tyres (Eds Dhir et al), Thomas Telford, 2001. 251-259
16. Pierce, C.E., Scrap tire modified concrete: past, present and future. Proceedings Used/post consumer tyres (Eds Limachaya et al), Thomas Telford, 2004. 1-16
17. Kew, H.Y., The use of recycled rubber tyres in concrete. Proceedings Used/post consumer tyres (Eds Limachaya et al), Thomas Telford, 2004, 135-142
18. Hassan, K.E., Sanders, P.J., & Nicholls, J.C., April 2004. Development of asphalt and concrete products incorporating alternative aggregates, 51 pp

19. Kashi, M.G., Synthetic lightweight aggregate from fly ash and waste plastics. Beneficial use of Recycled Materials in Transportation Applications, Washington DC 2001. (http://www.rmrc.unh.edu/Resources/PandD/useful.asp), 2001

20. Koide H., Tomon M. and Sasaki T. (2002). Investigation of the use of waste plastic as an aggregate for light weight concrete, Sustainable Concrete Construction September 2002. Thomas Telford, pp. 177-185

21. BS EN12620, 2002. Aggregates for concrete

22. BS EN 8500-2 (2002) Concrete —Complementary British Standard to BS EN206-1 — Part 2: Specification for constituent materials and concrete

23. Series 1700 (structural concrete) of the Manual of Contract Documents for Highway Works Volume 1 (Specification for Highway Works)

24. Volume 7 of the design manual for roads and bridges

25. British Standard Institution, BS 3892-1:1997, "Specification of pulverized-fuel ash for use with Portland cement", BSI, London, 1997

26. British Standards Institution, BS6699: 1992, "Ground granulated blast furnace slag for use with Portland cement", BSI, London, 1992

27. J.G. Cabrera, C.D. Atis¸, Design and properties of high volume fly ash high-performance concrete, Proceedings of Gramado Conference, Brasil, SP-186, American Concrete Institute Pubs., USA, 1999, pp. 21– 37

28. Malhotra, V.M., and P.K. Mehta. High-Performance, High-Volume Fly Ash Concrete. Supplementary Cementing Materials for Sustainable Development, Inc., Ottawa, Canada, 2002, 101 pp

29. Malhotra, V.M., "High-Performance, High-Volume Fly Ash Concrete." Concrete International 24(7), 2002, pp. 30-34

30. Cabrera, J.G., Atis, C.D., Design and properties of high volume fly ash high-performance concrete, Proceedings of Gramado Conference, Brasil, SP-186, American Concrete Institute Pubs., USA, 1999, pp. 21– 37

31. Malhotra, V.M., and P.K. Mehta. High-Performance, High-Volume Fly Ash Concrete, Supplementary Cementing Materials for Sustainable Development, Inc., Ottawa, Canada, 2002, 101 pp

32. V. Sirvivatnanon, H. T. Cao, and P. Nelson, Mechanical and Durability Properties of High Volume Fly Ash Concrete, ACI Special Publication, Vol. 145, 1994, pp. 967-984

33. Guangcheng, L, Yunhua, L and Youjun, X., "study on fly ash-saturation in cementitious materials", Sustainkable Construction Materials and Technologies – Chun,Claisse, Naik and Ganjian (eds), Covemnrty, UK, 2007, ISBN 978-0-415-44689-1

34. Cyr, M., Lawrence, P., Ringot, E., Efficiency of mineral admixtures in mortars: Quantification of the physical and chemical effects of fine admixtures in relation with compressive strength, Cement and Concrete Research 36 (2006) 264 – 277

35. Naik, T.R., Ramme, B.W., Kraus, R.N. and Siddique, R., Long term performance of high volume fly ash concrete, ACI Materials Journal, 100 (2), 2003, pp.150-155

36. Mehta, P.K. "Durability: Critical Issues for the Future." Concrete International 19(7), 1997, pp. 69-76

37. Mehta, P.K. "Concrete Technology for Sustainable Development." Concrete International 21(11), 1999, pp. 47-52

38. Naik, T.R., B.W. Ramme, R.N. Kraus, and R. Siddique. "Long Term Performance of High-Volume Fly Ash Concrete Pavements." ACI Materials Journal 100(2), 2003, pp. 150-155

39. Langley, W.S., and G.H. Leaman. "Practical Uses for High-Volume Fly Ash Concrete." AC, SP-178. American Concrete Institute, 1998, pp. 545-574

40. Mehta, P.K., and W.S. Langley. "Monolith Foundation: Built to Last a 1000 Years. Concrete International 22(7), 2000, pp. 27-30

41. Marsh, Bryan, High-volume fly ash concrete, Concrete, April 2003, pp 54-55

42. Mehta, P.K., "Use of Superplactisizers in High-Volume Fly Ash Concrete: U.S. Case Histories." Proceedings of the Nelu Spiratos Symposium. Committee for the Organization of CANMET/ACI Conferences, 2003, pp. 89-105

43. Manmohan, D., and P.K. Mehta. "Heavily Reinforced Shear Walls and Mass Foundations Built With Green Concrete." Concrete International 24(8), 2003, pp. 64-70

44. J.S. Damtoft, M. Glavind, C. Munch-Petersen, Sustainable development and climate change initiatives, Cem. Conc. Res. vol. 38, issue 2, 2008, pp. 115-127

45. Konstantin Sobolev, Research Article The Scientific World JOURNAL (2003) 3, 308-318 ISSN 1537-744X; DOI 10.1100/tsw.2003.23 Sustainable Development of the Cement Industry and Blended Cements to Meet Ecological Challenges

46. http://www.greenresourcecenter.org/, access Feb 2009

47. ACI Committee 233, Ground Granulated Blast-Furnace Slag as a Cementitious Constituent in Concrete, American Concrete Institute Report ACI 233R-95, Farmington Hills, MI; 1995

48. Siddique, R., Waste Materials and By-Products in Concrete, Springer, Berlin; 2008

49. Reid, J.M. and Chandler, J.W.E., Recycling in transport infrastructure, Crowthorne, TRL, 2001

50. Behnoud, A., 2005, Effects of High Temperatures on High-Strength Concretes Incorporating Copper Slag Aggregates, Proc, 7th International Symposium on High-Performance Concrete, Washington, D.C; 2005

51. Mehta, P.K., Concrete Technology for Sustainable Development – Overview of Essential Elements, Concrete Technology for a Sustainable Development in the 21st Century, Gjorv, O.E. and Sakai K, eds., London, E & FN Spon; 2000. p. 83-94

52. Sobolev, K.G. and Soboleva, S.V. (1996) High Performance Indigenous Cement. XXIVth World Housing Congress, Ankara, Turkey. pp. 189–202

53. Aitcin, P.C., (2000) Cements of yesterday and today concrete of tomorrow. Cement Concrete Res. 30, 1349–1359

54. Roumain, J.-C. and Sarkar, S.L. (2000) Cements of the 21st Century. Cement and Concrete Technology in the 2000-s: II International Symposium, Istanbul, Turkey. pp. 32–37

55. Swamy, R.N. (2000) Holistic Design of Concrete Technology the Only Route to Durability and Sustainability in Construction. Cement and Concrete Technology in the 2000-s: II International Symposium, Istanbul, Turkey. pp. 58–71

56. Sobolev, K. and Arikan, M. (2002) Production Process of High Volume Mineral Additive ECO-Cement: Part II — Feasibility Analysis. Cement and Concrete World, March–April. pp. 56–65.

57. Dunster, A., Conglasscrete (2007): certification of glass pozzolan, 20 pp. http://www.wrap.org.uk/downloads/ConGlassCrete2AppendixDi.108e37cd.446.pdf

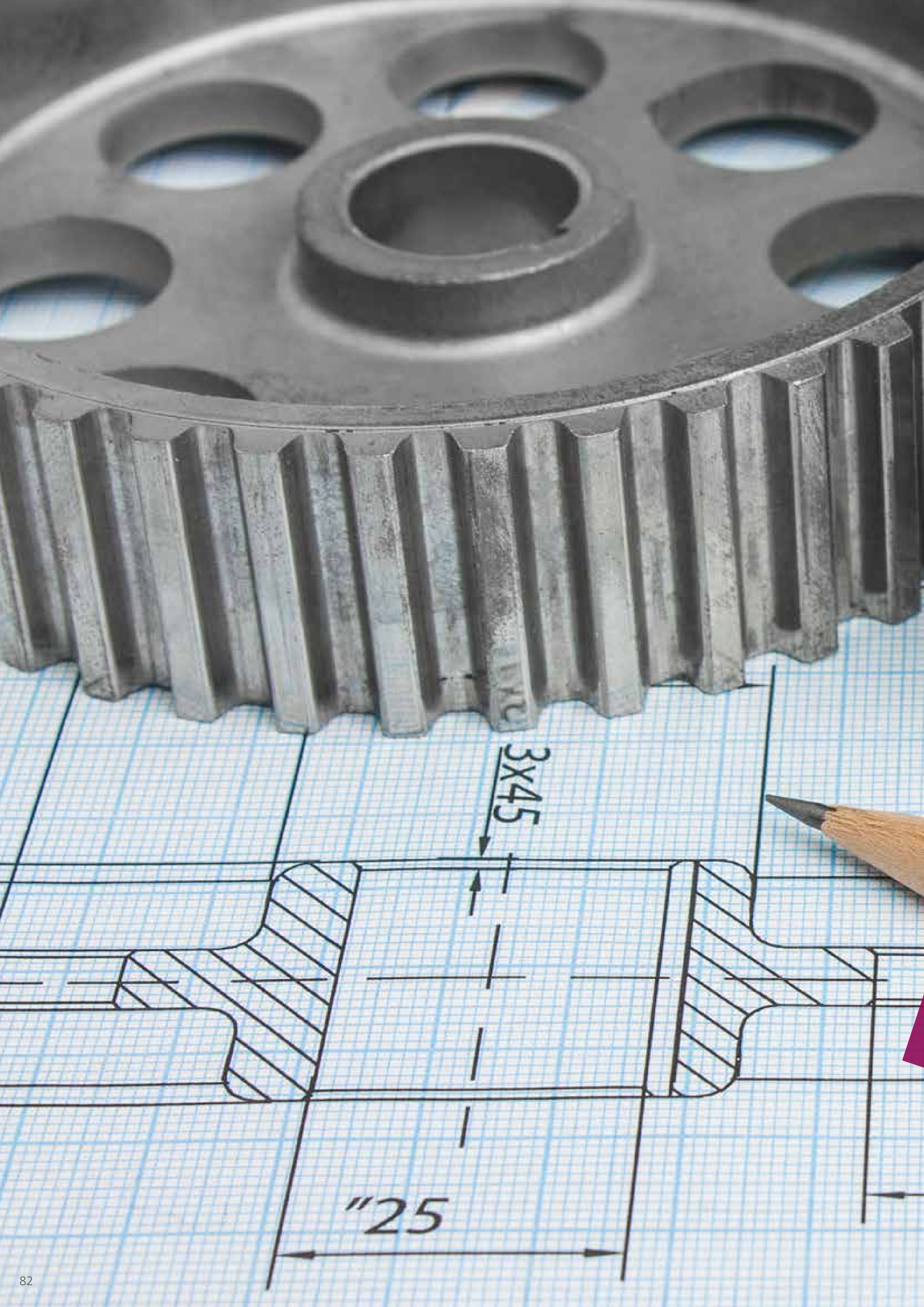
58. Dyer, T. D. and Dhir, R., Use of glass cullet as a cement component in concrete. Proceedings International symposium- recycling and reuse of glass cullet, 2001, pp. 157-166

59. Minerals planning factsheet- gypsum http://www.bgs.ac.uk/mineralsuk/downloads/mpfgypsum.pdf

60. O'Rourke, B., Development of calcium sulfate-ggbs-Portland cement binders Construction and Building Materials, 23 (2009), 340-346

61. Gartner, E., Industrially interesting approaches to "low CO2" cements. Cement and Concrete Research 2004, 34 (9) 1489-98

62. Ganjian, E., Claisse, P., Sadeghi Pouya, H., Recycled gypsum in a novel cementitious material, 27 pp. WRAP http://www.wrap.org.uk/downloads/Case_study_Recycled_gypsum_in_a_novel_ cementitious_material.fbec2059.4563.pdf.


3X45
"25

157

Structures

Management of the M4 Elevated Section Substructures

Abstract

The M4 Elevated Section in West London is a 1.9 km concrete viaduct structure providing a major arterial route in to London. An intervention model bringing together structural assessment and forecast deterioration, corrosion modelling and cracking has been developed to prioritise structural rehabilitation. To evaluate residual strength an initial assessment of crossheads was undertaken, which identified a deficiency in tensile capacity at the ends of the crosshead cantilevers. Further assessment has been undertaken including three dimensional strut and tie, non-linear finite element analysis and plastic analysis to justify continued trafficking of the structure, confirm public safety and to determine the need for strengthening.

Additionally, extensive monitoring of the crossheads has been implemented including mapping of all cracks and remote crack monitoring to safeguard the substructures and identify early signs of structural distress. The long term maintenance strategy brings together strengthening and cathodic protection concrete preservation methods with removal and repair of concrete delamination.

This paper discusses the development of the prioritisation process including deterioration modelling, together with the extended structural analysis to formulate a strengthening programme for these substructures.

Chris Hendy

Head of Bridge Design and Technology

Highways & Transportation

Atkins

Chris Brock

Group Engineer (Bridges)

Highways & Transportation

Atkins

Introduction

Overview of the elevated structure

The M4 motorway in West London is supported on an elevated reinforced concrete viaduct 1.9 km in length. The elevated section carries the M4 two lane motorway above and parallel to the A4 dual carriageway, forming a major arterial route into central London. The structure incorporates 103 concrete crossheads supporting the mainline and a further 23 crossheads supporting the Junction 2 sliproads. The elevated structure comprises simply supported decks, typically with 16.2 m span lengths, formed of prestressed concrete beams with an insitu concrete deck. The substructures are typically single pier reinforced concrete with cantilever crossheads supporting the decks via elastomeric bearings. The crossheads incorporate nibs that extend out from the lower section of the crosshead cantilevers, forming the bearing shelves. The central section of the crosshead extends up to the top of deck level and directly supports the carriageway between spans. Asphaltic plug deck joints are provided between decks above the upstand.

Strategic management principles

The M4 elevated section falls within the 30 year concession contact awarded to Connect Plus as Managing Agent. The contractual requirements relating to the M4 elevated section require the Managing Agent to hand back the structure in the same condition as at the start of the contract and to ensure the total cost of works in the 15 years following the concession is no greater than the final 15 years of the contract. To achieve these requirements it is essential to understand the current condition of the crossheads, rate of deterioration and external factors such as drainage which are driving the deterioration.

Figure 1. General view of the M4 Elevated Substructures

It was therefore necessary to develop a strategy for the management of the M4 elevated substructures over a 45 year period. The objective of the strategy was to maximise the operation of the assets and minimise total spend by evaluating the optimum points at which interventions should be undertaken. Availability of the highway was a key factor because the contract includes financial penalty for periods of unavailability. The best theoretical strategy for the M4 elevated was initially found to be to schedule maintenance works interventions only where they became necessary to maintain structural capacity. However, as discussed in the section on sub-standard structure management strategy, it soon became clear all crossheads first needed strengthening to overcome an existing deficiency and this needed to be implemented ahead of operation of the long-term maintenance strategy. Additionally, other considerations affected longterm maintenance prioritisation including ease of access to different crossheads and control of other risks to public safety such as that posed by spalling concrete.

To inform the management strategy, the time to intervention based on the point when the structural capacity was no longer adequate, was calculated. This process brought together corrosion modelling to estimate the degree of reinforcement corrosion and strength loss with overall structural assessment. From chloride sampling taken from the M4 elevated crossheads, the corrosion modelling evaluated the time duration for the level of chlorides to reach 0.3% of cement content at the depth of reinforcement. This threshold level was considered to be the point where active corrosion of the reinforcement was instigated and an assumed rate of corrosion was applied to the affected bars (BRE 2009). The resulting strength reduction was brought together with the structural assessment to forecast a maximum time to intervention.

To develop the long-term maintenance strategy it was necessary to put in place a system to prioritise intervention. A scoring system was therefore developed considering primary deterioration indicators on which each substructure could be evaluated. This considered in particular:

- structural capacity compared to demand
- deterioration rate (based on chloride levels and halfcell survey results)
- concrete delamination extents
- crack extents

- existence or otherwise of cathodic protection
- traffic management requirements, access issues and disruption to the travelling public.

Each of these characteristics was weighted depending on the level to which they dictate the intervention prioritisation of the substructures. A threshold level was determined representing a level of deterioration where intervention is necessary and a time to intervention for each crosshead and indicative works programme, produced.

Current condition

The crosshead substructures are suffering from extensive corrosion and concrete delamination caused by salt water ingress (from winter maintenance) through deck joints, evidenced by extensive water staining. Large areas of delaminated cover concrete have been removed from the structure to mitigate the risk of loose concrete falling on the road below. Site testing has confirmed the chloride concentration is significantly in excess of the 0.3% threshold associated with high risk of active corrosion occurring (up to a maximum of ~8%). The presence of reinforcement corrosion is further supported by halfcell potentials taken from the structure exceeding 350mV (CET Safehouse 2012), indicating a greater than 90% probability of corrosion occurring.

Cracking has also occurred on many of the crosshead ends concentrated at the re-entrant corners. This location acts as a halving joint and is subject to high stress concentration, therefore some degree of cracking is expected. However concern was raised as to the crack extents. In some locations the cracking has propagated between the re-entrant corners to 'join up'. This behaviour is representative of a tensile failure of the concrete and subsequent transfer of loading on to the available vertical reinforcement which is of concern

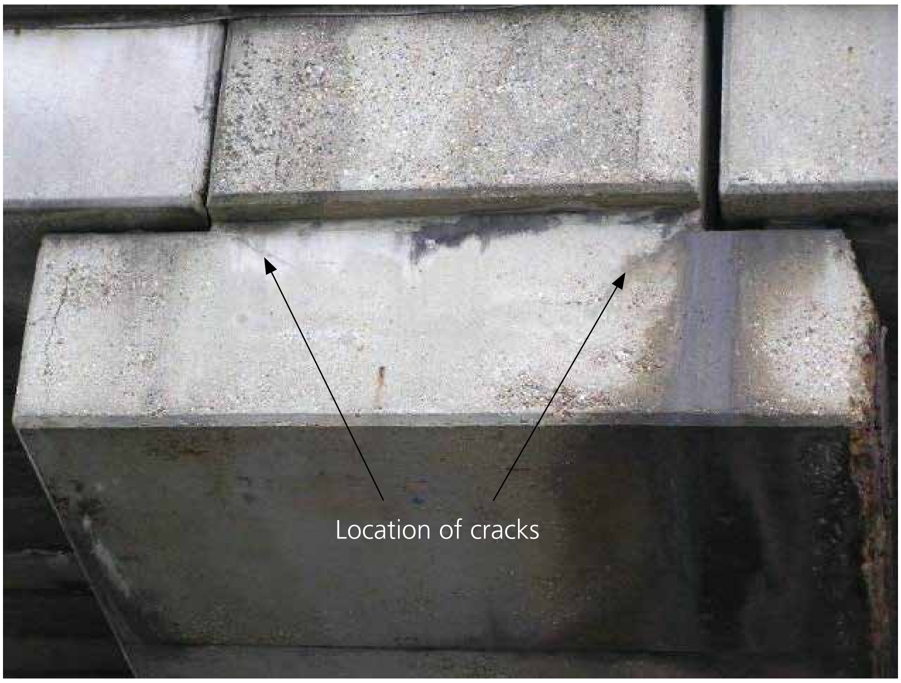


Figure 2. Location of cracks at crosshead cantilever ends

as discussed in the Initial Structural Assessment.

Due to the shape of the crossheads, it is impossible to inspect the condition of the concrete faces adjacent to the deck ends. Principally, testing has focused on the accessible vertical faces and crosshead soffits, both of which are typically in poor condition. It is considered highly likely that conditions at the inaccessible faces are suffering at least the same level of corrosion as the visible faces.

Deck joints

Water penetration through the deck expansion joints has been reported since the mid 1980s. A programme of expansion joint replacements to all crossheads has recently been completed, which involved removing the life expired asphaltic plug joints (APJ) and replacing these with high modulus asphaltic plug joints with improved whole life cost performance. The original design had a discrete APJ above each expansion joint. The replacement joints extend over both expansion joints across the full width of the crosshead.

Provision for subsurface drainage in the joints was improved but subsequent surveys have shown significant amounts of water continuing to penetrate the deck joints, probably beneath the surfacing and under the joint, resulting in ongoing contamination of the crossheads. An additional problem is that the APJ extends only between kerblines; water can thus penetrate through the verges and central reserve.

Control of concrete spalling

Significant areas of concrete delamination are present on the elevated substructures. The potential for delaminated concrete to fall on to the A4 dual carriageway directly beneath the M4 elevated section presents a hazard to road users.

To mitigate the risk from falling concrete, a netting system has been installed around the crosshead cantilevers. This netting is designed to catch the full cover concrete without significant deflection or damage to the netting occurring. Additionally a regular regime of delamination surveys has been put in place to identify areas of deteriorated concrete before they become delaminated from the substructures.

These delamination surveys are carried out on a three month basis during the summer months and on a more frequent six week basis during the winter months where the risk of delamination is greater due to freeze-thaw action. Further mitigation is provided through regular walk inspections examining the netting for any captured concrete, which would indicate acceleration in deterioration.

Figure 3. Concrete delamination on crosshead cantilever faces

Loss of concrete section has also exposed large areas of reinforcement as shown in **Figure 3**, exposing reinforcement to accelerated atmospheric corrosion.

Sub-standard structure management strategy

Initial Structural Assessment

An initial assessment of the crosshead structures was undertaken based on record drawings to determine live load capacity and its sensitivity to reinforcement corrosion. The assessment demonstrated that the majority of each crosshead was able to carry 40 tonne assessment loading, but the light vertical reinforcement in the end 3m of each cantilever was overstressed. The location of the bearing shelves below the top of the crosshead, where the main flexural reinforcement is provided, necessitates load transfer up to this reinforcement. The vertical reinforcement provided at the ends of the cantilevers was found to be inadequate to cater for this load path. This deficiency was further exacerbated by the high risk of active corrosion occurring to the existing reinforcement which will further weaken the structure. Additionally, codified minimum reinforcement requirements were not met, indicating a sudden brittle failure was possible should the concrete tensile strength be lost. This made the presence of cracks as shown in **Figure 2** a concern.

Management Strategy Formulation

Following the conclusions of the initial structural assessment a rigorous management regime was implemented in accordance with BD 79/06 (Highways Agency 2006). This regime included a thorough examination of the likelihood and implications of structural failure, together with instigating a detailed inspection regime focusing on the ends of the crossheads where a potential failure mechanism would manifest first. The management regime focused on a risk based approach to protecting public safety, putting in place defined trigger levels in the event any evidence of structural deterioration was identified.

The following actions were undertaken to ascertain the optimum management strategy for the M4 elevated substructures:

- Extended analysis considering three dimensional strut and tie analysis, non linear concrete modelling and considerations of reinforcement-strain hardening
- Assessment of ductility of potential failure mechanisms and hence the amount of warning these would give
- Detailed inspection of the substructures at cantilever ends to identify evidence of initiation of a failure mechanism
- Analysis of the impacts and disruption to the travelling public in the event of emergency closure, including the elevated accident risk associated with traffic diversions
- Design and procurement of contingency measures to be mobilised immediately in the event evidence of accelerated structural deterioration was identified.

Refined Structural Analysis

Refined analysis was undertaken to optimise the predicted load carrying capacity of the crossheads and to inform of the potential

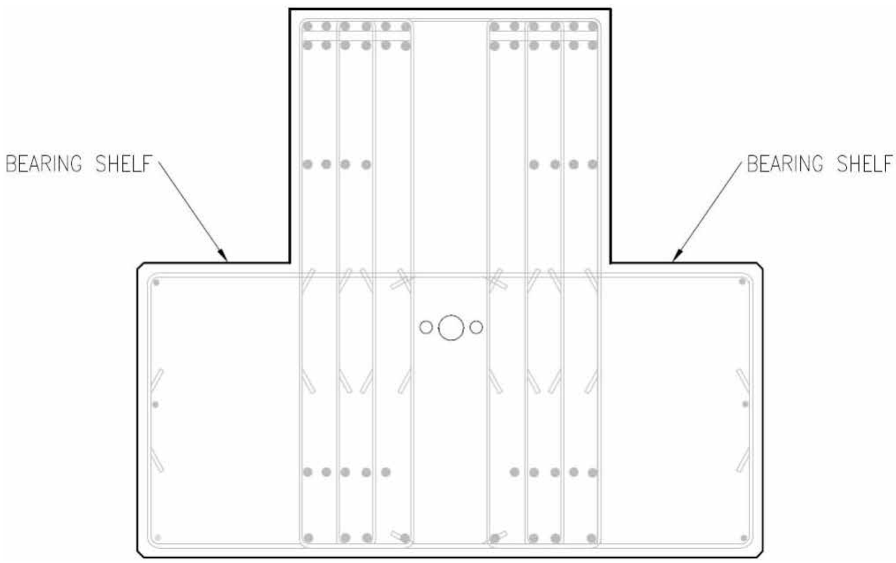


Figure 4. Existing reinforcement within the crosshead cantilevers

failure behaviour. Extended analysis considering both strut and tie modelling and non-linear concrete modelling was undertaken.

The strut and tie analysis considered optimisation between two load paths available at the end of the crosshead. The first load path represented the transfer of load vertically from the bearing locations to the flexural reinforcement at the top of the crosshead section (**Figure 5**).

The second strut and tie load path considered took advantage of concrete compression struts in the lower portion of the crosshead to transfer the bearing loads to the bent up bars closer to the pier. This load path reduces the load required to be carried on the vertical links but requires a tensile tie to be available directly below the bearing shelves. Record drawings indicate there is only one 25mm diameter bar located at the top of the nibs which could provide this resistance. However cover meter surveys has shown there is no end anchorage to these bars and it is considered highly likely significant corrosion has generally occurred (which is observed in locations where the cover concrete has delaminated).

The load capacity of the crosshead cantilever ends was optimised by superimposing the resistances provided by both load paths. However the contribution of the second load path is limited by the contribution of the longitudinal tension bar within the nibs, therefore it was not possible to generate significant benefits over and above the predictions of the initial analysis.

In tandem a non-linear concrete model of the crossheads was developed to investigate potential strength benefits and to provide greater certainty as to how a failure mechanism would ultimately develop. The model incorporated all reinforcement and considered a rigid connection to the concrete, therefore not taking in to account bond characteristics. The analysis was performed in Abaqus using its 'concrete damaged plasticity' model for concrete behaviour. The analysis considered tensile cracking, tension softening and compressive crushing of the concrete.

This analysis showed even for relatively small tensile concrete characteristic strength (~1N/mm^2) the concrete was able to transfer the bearing loads and the concrete at the end of the cantilever remained uncracked. Failure was transferred to the cantilever root at much elevated load. This was not considered realistic because the model did not properly take account of the potential brittle fracture of the concrete arising from high stress concentrations at the crack tip, nor the presence of pre-existing shrinkage cracks or voids.

To reduce the influence of concrete tensile strength, the analysis was re-run for lower concrete tensile strengths which then gave very

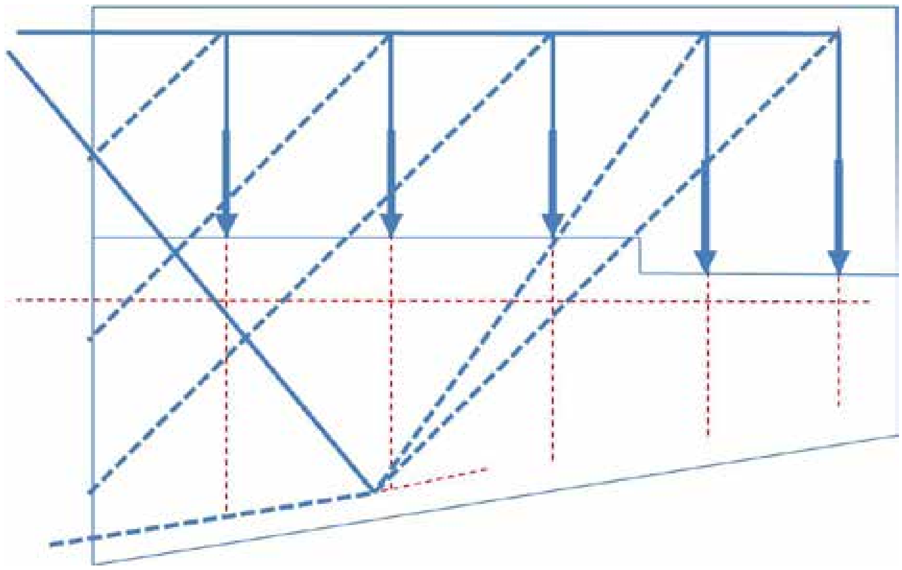


Figure 5. Strut and Tie Analysis Load Path 1

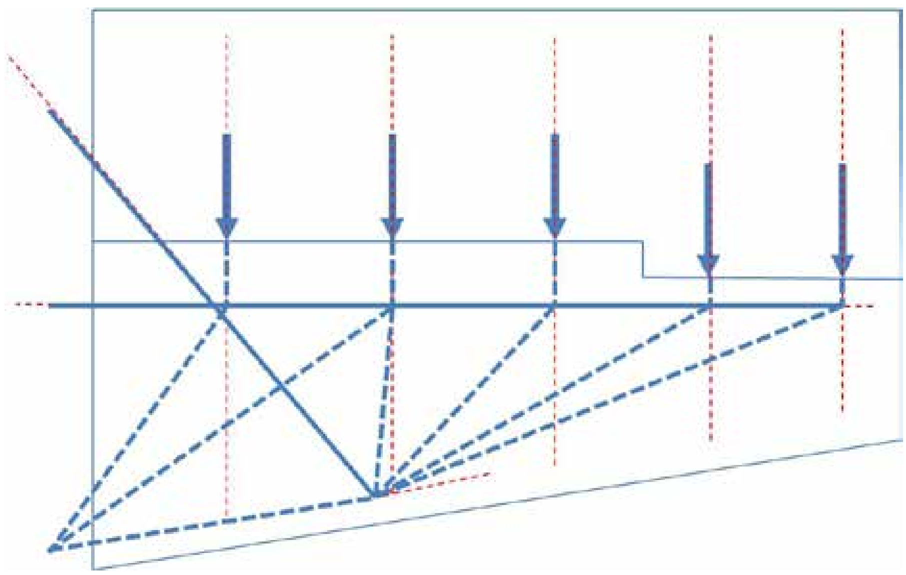


Figure 6. Strut and Tie Analysis Load Path 2

similar results to the strut and tie approach, as would be expected since it ignores concrete tensile strength. For low concrete strengths cracks propagate at the crosshead ends under lower loading conditions, resulting in redistribution of force to the vertical bars in the crosshead ends and ultimately ductile failure at a load lower than that required. For higher concrete strengths, failure occurs above the required load but is governed by concrete tensile failure and is hence brittle. This resistance model could not be relied upon and hence adequate strength of the crossheads could still not be demonstrated.

Consideration of Overall Failure Mechanism and Strain Hardening

The refined structural analysis still focused on the crosshead in isolation and not the potential for any redistribution of load to elsewhere in the structure upon localised failure. To understand the full implications of a failure the combined superstructure and substructure was considered as a single structural system.

A ductile failure, ignoring concrete tensile strength, occurring anywhere in the final 3.0m of a crosshead cantilever was considered. This part of the crosshead supports the verge, 600mm hard strip and a proportion of the inside traffic lane (the M4 motorway having no hard shoulder in this location). In the event the cantilever end displaces and support is lost to the outer deck bearings, the deck itself can span transversely due to the presence of end diaphragms and concrete infill between the beams. In reality, the ductility of the vertical steel in the crossheads means that the two mechanisms can be superposed, thus adding strength in excess of the predictions stated in ‘Refined Structural Analysis’.

Reinforcement test data also concluded the ultimate steel strength is greatly in excess of the yield strength by 45% as a

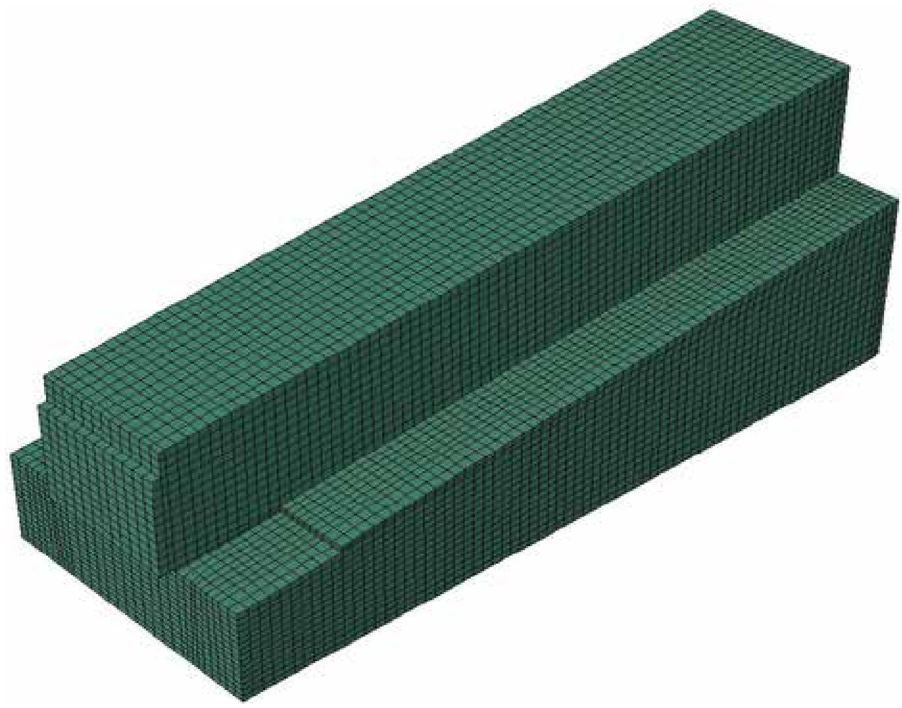


Figure 7. Non-linear Finite Element Crosshead Modelling

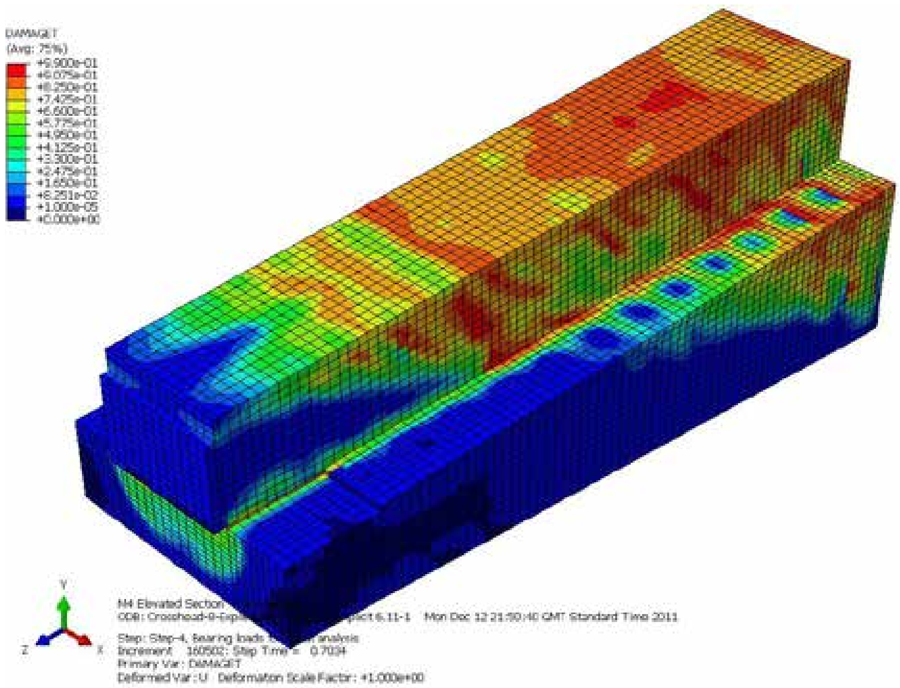


Figure 8. Non-linear Finite Element Crosshead Analysis

minimum. When strain hardening was additionally considered it was possible to demonstrate that the full traffic loading could be carried at the ultimate limit state, albeit with significant plastic deformation and damage occurring to the crosshead cantilever and adjacent decks leaving the structure unserviceable and very expensive to repair.

Whilst this gave confidence overall collapse would not occur, a residual concern that lumps of concrete might spall under the high strains and fall on to the A4 below was considered. This risk was, however, mitigated by the presence of the protective netting described in the section on control of concrete spalling.

In light of this analysis and mitigation, the overall risks to public safety from continued operation in the short to medium term were considered to be low if suitable monitoring was also put in place as discussed in the section on structural monitoring. However, it was concluded strengthening was required to maintain adequate reliability longterm. This is discussed in the section on strengthening and protection.

Structural monitoring

Whilst the structural analysis provided reassurance that any failure associated with the substructures would not be catastrophic, it predicted the potential for significant structural damage to occur and ultimately complete closures of the M4 and A4 arterial routes, generating traffic disruption on a national scale. Thus a detailed regime of structural monitoring was put in place to identify early signs of structural deterioration and distress and to allow intervention before serious damage occurs.

Monitoring has been implemented on a risk based approach with focus being on the piers showing the greatest cracking. A series of trigger levels was defined based on risk level. DEMEC studs have been installed at all re-entrant crack locations and a 3 monthly programme of remeasurement has been implemented. Where cracking extends the full distance between re-entrant corners the affected crossheads are classed as high risk and real time monitoring devices have been installed over the cracks. These devices take regular measurements of displacements parallel and perpendicular to the cracks and the data can be remotely accessed through a web based portal. The monitoring data is constantly interrogated to identify any signs of progressive deterioration through steady increasing crack length and width.

Typically the monitoring shows a very high correlation with temperature affects, with both daily and seasonal changes observed. Various methods have been used to remove these temperature affects with varying degrees of success. They do not prevent long term trends being identified, which in the case of some crossheads do show progressive crack increase.

Trigger levels for crack width (based on localised reinforcement yielding) have been set. If they are exceeded, temporary interim strengthening measures have been designed and fabricated and are available to be installed rapidly. In accordance with the requirements of BD 79, strengthening of all the deficient regions is programmed within three years.

Strengthening and protection

To provide adequate reliability over the residual life of the structures, localised strengthening is required to them all. Due to the large number of structures to be strengthened (approximately 200 cantilevers) it was essential to produce a costeffective solution which would additionally have minimum impact on the travelling public during installation. Consequently a trial strengthening scheme was first undertaken to one crosshead, number 82. The strengthening method additionally was developed to minimise any deterioration of the recently replaced expansion joints (thereby not inadvertently leading to greater water ingress of reinforcement corrosion).

The strengthening comprises installing, by drilling and fixing, new reinforcement both transversely and vertically to supplement load path 1, shown in **Figure 5**, at the ends of the cantilevers. The requirement not to damage the expansion joints or close the M4 during the works meant the vertical bars could not be anchored on the top of the crosshead. Therefore the vertical bars are resin anchored as high as feasible in the concrete section and additional longitudinal reinforcement installed where the vertical bars are fully anchored. This combination of additional reinforcement provides a truss system to transfer the bearing loads at the ends of the cantilevers to the sections of greater reinforcement. To minimise disruption to the traffic on the A4 below, all works were undertaken during nighttime.

As the crosshead concrete is contaminated with high levels of chlorides the maintenance strategy needed to take future corrosion of reinforcement in to consideration. To address chloride induced corrosion stainless steel reinforcement has been specified which has enhanced protection for high chloride environments. Additionally resin grout has been specified both to take advantage of its greater bond strengths and its chemical isolating properties. Fully grouting the additional reinforcement will minimise the risk of chloride attack and the stainless steel will further mitigate the extent of future deterioration.

A key issue in the strengthening installation was the drilling of holes for new reinforcement insertion without damaging the existing reinforcement. This led to the need both for a prescriptive drilling protocol, requiring core holes to be abandoned and relocated in some situations, together with the addition of temporary external longitudinal prestress applied by prestressing bars anchored to a steel frame at each end of the crosshead. This prestressing force supplemented load path 2 shown in **Figure 6** (by increasing the resistance of the horizontal tie member) and thus provided mitigation against damaging the sparse existing vertical reinforcement during drilling. The layout of core holes was carefully designed to minimise the potential for hitting reinforcement and 3D modelling proved helpful in this regard as discussed in the section on corrosion protection.

Corrosion protection

A key consideration in the maintenance strategy was the link between the assessed capacity of the structure and the forecast deterioration due to chlorides. As such, a choice had to be made whether to carry out significant concrete repairs, potentially several times in the life of each structure, or to install cathodic protection and hopefully arrest deterioration. Installation of cathodic protection was evaluated as the optimum cost solution, arresting structural deterioration from corrosion and also localised spalling of concrete.

An impressed current protection system has been designed to deliver protection to the reinforcement locations identified as at risk of corrosion and was trialled first at crosshead 82. This comprises discrete anodes installed below the bearing shelves to provide protection to the vertical and transverse bars. This location is particularly at risk of corrosion resulting from water ingress through the cracks propagating from the re-entrant corners. Additionally discrete anodes have been designed to be installed in longitudinal cores located at the top of the crossheads to deliver protection to the main flexural reinforcement. These cores are made from both ends of the crosshead and discrete anodes secured with cementitious grout. Finally a cathodic protection ribbon system is added to the accessible faces of the crossheads. It is installed in chases cut in to the concrete, removing the requirement for a spray applied cementitious overlay and maximising future durability. While

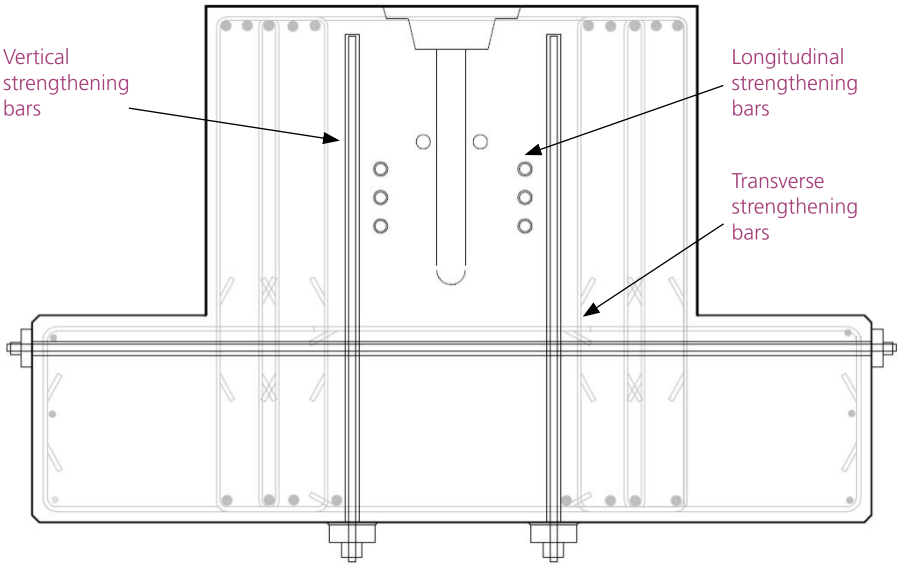


Figure 9. Strengthening reinforcement to be installed within the crosshead cantilevers

the stainless steel strengthening bars do not explicitly need protection against chloride induced corrosion, they are made electrically continuous with the other reinforcement and connected to the negative side of the cathodic protection system to prevent stray current induced corrosion.

Cathodic protection is combined with removal and repair of delaminating concrete. Areas of delaminated or spalled concrete are first removed and concrete repair material 'tied' into the existing reinforcement cage. Combining repair and cathodic protection removes the need to remove chloride impregnated concrete, minimising the breakout and repair extents. The anticipated design life of concrete repairs alone (without cathodic protection) is considered to be approximately 5 – 10 years, at which time further repair works would be necessary.

Building information modelling

Building Information Modelling (BIM) was used during the interim planning and strengthening design phases and continues to be used for ongoing asset management of the crossheads.

The geometry of all the crossheads and the surrounding carriageway below the viaducts was laser scanned and a 3D model created from the point cloud data. This provided valuable measurements of the asbuilt structure to crosscheck against the record drawings. It also allowed emergency temporary propping systems (for use in the event that monitoring trigger levels were exceeded) to be visualised and developed such that their founding locations on the A4 carriageway below had the minimum impact on traffic flows and safety.

Figure 10. Laser scan image of the M4 Elevated Section

At the detailed design phase for the strengthening, asbuilt reinforcement was incorporated into the 3D models. This model was then used to route virtual core holes for the new reinforcement and cathodic protection anodes through the crosshead in locations to minimise clashes with existing reinforcement.

The BIM model also contains numerical and visual attached data relevant to the ongoing asset management of the viaducts. Each crosshead is split into smaller uniquely referenced surfaces to which condition data is attached including:

- chloride levels
- half cell readings
- strength utilisation
- estimated reinforcement section loss
- crack widths
- presence of spalling and exposed reinforcement.

The data can be viewed numerically or by colour scale in Navisworks (see **Figure 12**) and an embedded deterioration model in the source data allows predictions of this condition data to be viewed with time.

The crosshead as a whole has further attached data which can similarly be viewed including:

- BCI condition score
- maintenance prioritisation score
- proposed maintenance plan and intervention date
- inspection reports and defect diagrams
- as-built information.

Aspirationally, the model will be developed to include full details in 3D of all existing and new reinforcement together with details of the cathodic protection systems. Additionally all defects will be added directly to the model so they can be viewed live rather than through attachments.

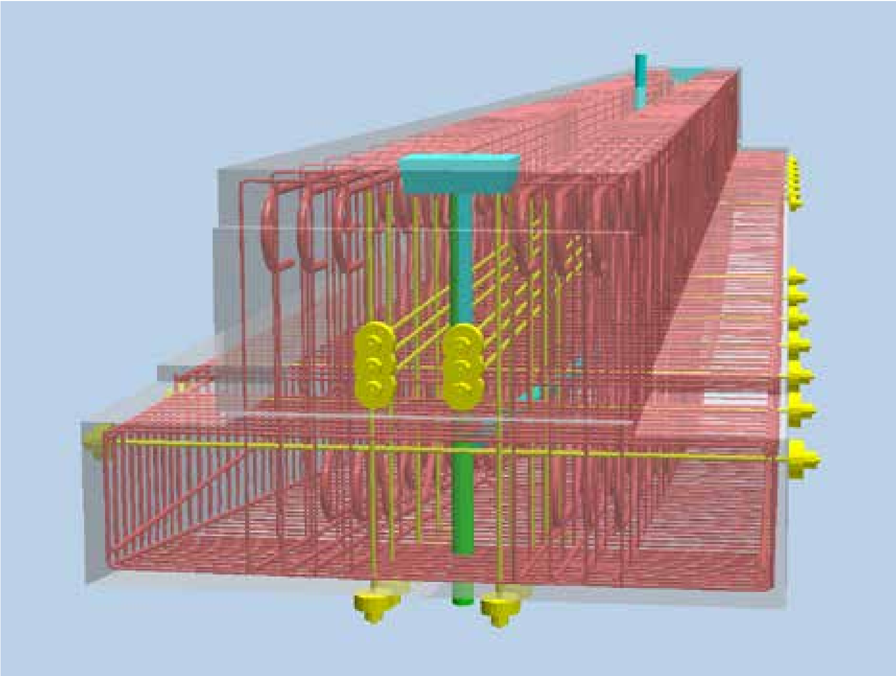


Figure 11. Image of crosshead showing existing and strengthening reinforcement

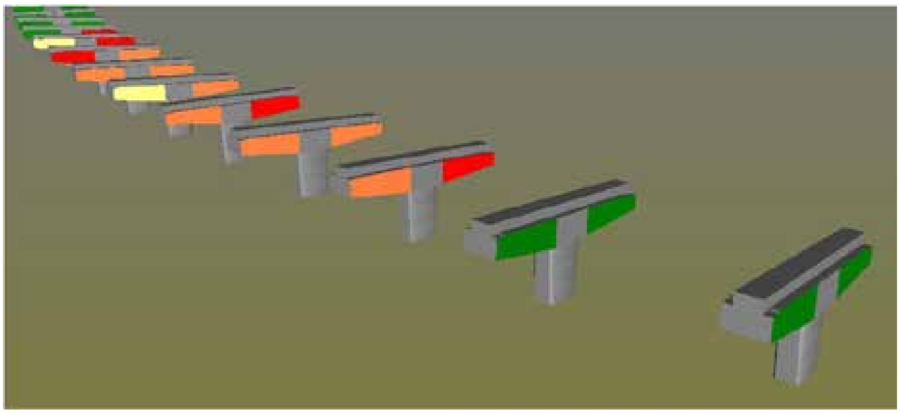


Figure 12. BIM environment showing variation in chloride levels

Conclusion

The M4 elevated section provides a major arterial route in to London and is of strategic national importance. However the substructures especially have been suffering significant deterioration over a number of years. The substructures are experiencing large areas of concrete delamination with exposed reinforcement. Additionally water ingress through the deck joints has resulted in very high levels of chloride impregnation with an associated high risk of reinforcement corrosion actively occurring.

The original scope of the maintenance strategy was to predict a future time to intervention, bringing together structural capacity and corrosion modelling. This process identified a significant existing strength deficiency at the ends of the cantilever crossheads. A management strategy in accordance with BD 79 has been instigated, which has considered further analysis including strut and tie and nonlinear concrete modelling to understand the extents of this deficiency and the anticipated resultant failure mechanism. Plastic analysis incorporating the postyield behaviour of the reinforcement and transverse contribution of the decks was used to justify continued trafficking of the structure together with rigorous monitoring to identify structural deterioration early. However, strengthening is required to maintain longterm performance.

A comprehensive prioritisation process has been developed, including deterioration modelling, to formulate a programme both for the short term strengthening and longer term intervention requirements for these substructures. Concrete rehabilitation methods have been evaluated and a programme of cathodic protection combined with concrete repairs has been developed. It is proposed this intervention will be undertaken over the 30 year concession period and prioritised on a risk basis.

A BIM model was developed to assist emergency propping and strengthening design and to provide a visual representation of the maintenance strategy, allowing the user to see how the substructures deteriorate over time and when intervention is required. Additionally the BIM model provides an interactive and intuitive record system.

References

1. ACI 222R-01. 2001. Protection of Metals in Concrete Against Corrosion, American Concrete Institute, Farmington Hills, MI, USA
2. BRE Digest 444-2, 2009. Corrosion of steel in concrete
3. CET. 2012. M4 Elevated Crossheads Visual Monitoring Re-sults – October 2012
4. Clear K.C. 1976. Time-to-Corrosion of Reinforcing Steel in Concrete Slabs, Federal Highway Administration, PB 258 446, Vol. 3
5. Concrete Society TR54. 2000. Diagnosis of deterioration in concrete structures – identification of defects, evaluation and development of remedial actions
6. Concrete Society TR 61. 2004. Enhancing reinforced concrete durability
7. Highways Agency. 2006. BD79/06: Design Manual for Roads and Bridges Volume 3: Section 4: Part 18. The Management of Sub-Standard Structures
8. Hillerborg, A., M. Modeer, & P. E. Petersson. 1976. "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," Cement and Concrete Research, vol. 6, pp. 773–782
9. Lee, J., & G. L. Fenves. 1998. "Plastic-Damage Model for Cy-clic Loading of Concrete Structures," Journal of Engineering Mechanics, vol. 124, no.8, pp. 892–900
10. Lubliner, J., J. Oliver, S. Oller, & E. Oñat. 1989. "A Plastic-Damage Model for Concrete," International Journal of Sol-ids and Structures, vol. 25, pp. 299–329
11. Nokken, M., Boddy, A., Hooton, R.D., & Thomas, M.D. 2006. Time dependent diffusion in concrete—three laboratory studies, Cement and Concrete Research, Vol. 36, pp. 200-207
12. Yong Ann, K.. & Song, H. 2007. Chloride threshold level for corrosion of steel in concrete, Corrosion Science vol. 49, pp. 4113–4133.

158

Structures

Management of the Severn Bridge suspension bridge

Abstract

The M48 Severn Bridge is a 988 m span suspension bridge carrying the M48 motorway across the River Severn in the UK. The bridge is substandard to current UK assessment codes but is permitted to continue in operation on the basis that application of the Highways Agency document, BD79, which allows a risk-based management approach to bridge operation, indicates the risk presented is acceptable given the mitigation and monitoring in place; this includes Weigh in Motion monitoring of traffic, acoustic monitoring of the main cables and a permanent dehumidification system to protect the main cables from corrosion.

A key aspect of the work involved in verifying that the bridge can remain in operation is undertaking con-tinuous structural assessment, taking into account the results of previous cable intrusive investigations and data from the various continuous monitoring systems. This involves developing the predictive future deterioration model for the cable; deterioration has now largely stabilised with respect to corrosion with the dehumidification system in place but wire breaks may continue due to overload events and fatigue. Weigh in Motion data for the traffic crossing the bridge is used to continuously adjust the Bridge Specific Assessment Live Load model used for the bridge assessment. This is used to justify the continued usage of the bridge where the traffic loading remains below a critical value and to design interventions where this critical traffic loading is exceeded, either for extreme events or due to a steady increase in traffic volume with time.

This paper discusses the determination of the main cable strength allowing for its deterioration to date. It also discusses determination of the bridge specific assessment live loading from Weigh in Motion data.

Chris Hendy

Head of Bridge Design and Technology

Highways & Transportation

Atkins

Chris Mundell

Senior Bridge Engineer

Highways & Transportation

Atkins

David Bishop

Bridge Engineer

Highways & Transportation

Atkins

Introduction

The Severn Bridge is a 988 m suspension bridge which, since opening in 1966, has carried over 300,000,000 vehicles. Due to increases in traffic load and deterioration, the bridge has received several strengthening measures since construction, including replacement of the inclined hangers, fatigue strengthening of the deck and bearing replacements.

In 2006, a 40 year inspection of the cables was carried out to determine the condition of the cables. It was found that a number of the wires were broken or corroded, sparking concern over the bridge's overall capacity. As a consequence, Heavy Goods Vehicles (HGVs) were restricted to the outside lanes, with weight restrictions in place elsewhere. Additional measures to protect and monitor the cables were installed, including a corrosion inhibitor within a dehumidification system, and acoustic monitoring to detect further wire breaks.

In conjunction with the preventative measures in place, the M48 Severn Bridge is subject to continuous structural assessment, in particular for the main cables. This paper summarises the strength evaluation performed by Atkins and the methods used to monitor and calculate bridge specific live loads derived by the authors.

Bridge specific assessment live loading

Purpose of BSALL

In the UK, the Highways Agency Design Manual for Roads and Bridges (DMRB) includes extensive guidance for the assessment loading criteria for bridges spanning up to 50 m. For the assessment of longspan bridges such as the M48 Severn Crossing, the use of such assessment loading becomes excessive, and not representative of real traffic loading conditions. Instead, a bridge-specific assessment live load (BSALL) can be determined, giving an assessment loading derived - in this case - from a probabilistic assessment of the actual traffic flows. With this approach comes the necessity to continuously monitor the traffic, as any changes in traffic trends will also affect the assessment loading.

Traffic Monitoring

There are two methods that are generally used to provide the necessary traffic data to calculate a BSALL. Traffic flows can be simulated using probabilistic traffic flows and vehicle weights (e.g. using Monte Carlo simulations), or they can be measured on site using live monitoring from Weigh-in-Motion (WiM) stations - as is the case on Severn Bridge. Historically, the latter method has not been viable due to the computing power necessary to deal with such large volumes of traffic data; however modern processors have overcome such issues.

The M48 Severn Crossing has four WiM sensors. These are located at the west side of the Wye Bridge under each lane. The lanes are numbered from 1 to 4, with 1 and 2 being Eastbound traffic, and 3 and 4 being the Westbound traffic. Lanes 1 and 4 are the offside lanes, with 2 and 3 the nearside lanes. The data acquisition equipment is recalibrated on a regular basis and captures the following information:

- Lane of traffic;
- Date and time;
- Vehicle type;
- Vehicle length;
- Number of axles and axle loadings;
- Axle spacings.

WiM Trigger Levels

The WiM data must be routinely monitored to ensure that no changes in traffic trending are occurring. To achieve this, rather than continuously rederiving the BSALL, three trigger criteria are considered for traffic flow in each direction. These are the number of HGVs as a percentage of the total traffic, the standard deviation of the HGV weights and the mean of the HGV weights. The trigger levels are considered for four scenarios during any given day:

i. Early morning, when traffic volumes are low, but HGV percentages are respectively high;

ii. Morning rush hour, when traffic volumes peak, but relative percentages of HGV traffic are lower than scenario i;

iii. Evening rush hour, when traffic volumes again peak, but relative percentages of HGV traffic are lower than scenario i;

iv. A five hour period between scenarios ii and iii.

It has been determined that scenario iv) is the critical case, when the traffic volumes and percentages of HGV traffic combine to give the most onerous condition. It has also been found that these triggers are very sensitive to the sampling period, and large variations can occur on a daily basis. This highlights the necessity for caution when extrapolating trends and BSALL values from small samples of data, which may be subject to similar fluctuations.

The WiM data is monitored on a monthly basis, with averages calculated for the critical scenario iv) time period. The exceedance of any one trigger criterion (% HGV, mean or standard deviation) is not indicative of an increase in the BSALL. Even the exceedence of two or more triggers in any given monitoring period does not give certainty of an increased BSALL; this acts only as a prompt that the BSALL should be checked and compared against previously calculated levels. Generally when trigger levels have been exceeded to date there has been either no or minimal increase in derived BSALL.

Derivation of BSALL

A BSALL is a probabilistic traffic load based on a given set of traffic data. Live loading due to traffic flows will generally conform to a Gaussian distribution. However, to determine BSALLs for the purpose of assessment, these 'normal' traffic conditions must be extrapolated to provide safe but realistic load effects. In the UK, the characteristic highway traffic load model is defined as the load effect that has a 5% probability of being exceeded within a 120 year period.

The BSALL methodology discussed in this section has been carried out independently for each lane of traffic. Initially, the WiM data is examined for a given time period. In this instance, five consecutive weekdays are considered in a month. If the BSALL is to be used for assessing annual trends, this will be the same month as for any previous assessments to reduce the impact of seasonal variations.

The traffic from each lane is first "squashed" into queues formed of the vehicles crossing the bridge in each hour (the simulations presented in this paper have assumed that the traffic remains in the lanes as recorded by the WiM sensors). Therefore for each 24 hour period of data assessed, 24 No. hourly traffic queues are formed for each lane, giving 96 separate queues in total. These queues are formed with the clear gap between each vehicle of 5 m, measured as the distance between the rearmost axle of a vehicle and the front axle of the following vehicle.

Each lane is considered independently of the others, with the hourly traffic queues used to determine the worst load effects for that train of vehicles. The worst case midspan bending moment, uniformly distributed load (UDL) and nearsupport shear forces are calculated for each load train. The assessed bridge length is taken to be 1298m, equating to the M48 Severn Bridge main span and one side span (988m + 310m).

The maximum load effects for each hour are tabulated and formed into a cumulative frequency distribution. These probabilities are then normalised to provide the Gumbel probabilities using the following formula:

$$p_{\text{Gumbel}} = -\text{Ln}\{-\text{Ln}(\text{cumulative frequency})\}$$

The data, once normalised, can be presented as shown in **Figure 1**. This graph shows the normalised Gumbel probabilities from five days of WiM traffic queues. The Gumbel probabilities are given on the X-axis, with the associated load effects given on the Y-axis. The load effects for these data can also be shown instead as a ratio of either BD 21/01 or BD 37/01 values as required.

Figure 1 also shows two trend lines; the whole-data trend line is generated by considering the entire dataset, whilst the red dashed trend line considers only the tail of the data where more linear behaviour is observed. If the entire dataset is considered, the extrapolation to the extreme 120 year event could be greatly over estimated, as would be the case in **Figure 1**. By tailoring the trend line of each graph to suit the linear section of the data, the 120 year probability can be more accurately determined. In instances where the dataset is completely linear this tailoring may not be required.

In all cases, increasing the volume of data considered will improve the accuracy of the BSALLs derived, as any linear trending will be made more apparent. The methodology presented in this paper uses five consecutive days of traffic, with 24 traffic queues per day, giving 120 data points for each lane. It has been found that in most cases this provides sufficient clarity to clearly determine a linear trend line.

Once the maximum load effects for each traffic queue have been obtained, the nominalised value of the BSALL probabilities as derived from the hourly maximums as exemplified in **Figure 1** is obtained as follows:

The 5% probability of exceedence in 120 years equates (for small probabilities) to an hourly probability of exceedence of:

$$p120 = \frac{5\%}{(\text{total hours in 120 years}) \times f}$$

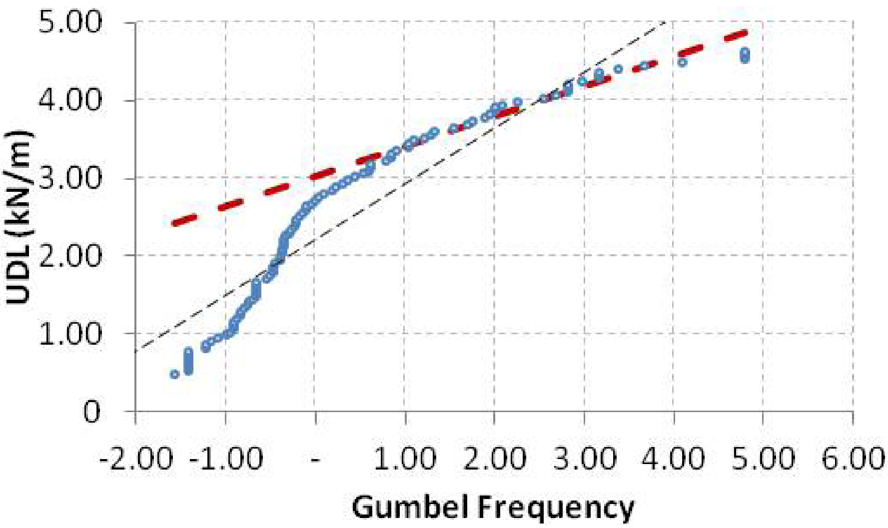


Figure 1. Typical Lane 1 Gumbel Distribution

The probability of a queue forming, f, may be derived as a function of the hourly traffic, as discussed in the TRL report CR016 on long span bridge loading (1986), where:

$$f = \frac{2 \times \text{hourly flow}}{1200}$$

Therefore, the probability of a queue forming will vary depending on the time of day and the traffic flow. The average peak hourly traffic for the WiM data obtained to date is 1984 vehicles per hour, giving a frequency f of 3.3. However, a frequency of 5% has been conservatively assumed, thus:

$$p120 = \frac{0.05}{(24 \times 365 \times 120) \times 0.05} = 9.513 \times 10^{-7}$$

(for any given hour).

This allows the 120 year Gumbel probability to be determined from the normalised distribution as:

$$p_{\text{Gumbel}} = \text{-Ln}\{\text{-Ln}(1 - p120)\}$$

In the case where $p120$ is 9.513 × 10^{-7}, the pGumbel is 13.86. For the example dataset shown in **Figure 1**, for a Gumbel frequency of 13.86 and using a manually determined trendline, the characteristic BSALL = 8.43 kNm^{-1}. As per BD 50/92, the nominal value is characteristic value / 1.2 and therefore becomes 7.03 kNm^{-1} for Lane 1 in this instance. Should the queue frequency of 3.3 be used, this value would reduce to 6.90 kNm^{-1}; a reduction of only 2%. If the derived load was used with Eurocodes, the characteristic value would be used directly to form combinations.

Calculated BSALL Values

The methodology presented in the section on Derivation of BSALL was carried out for eight months between June 2012 and January 2013 to monitor seasonal variations. However, the BSALL would only normally be calculated annually, with the WiM triggers used to identify any abnormal occurrences.

Table 1 presents the nominal BSALL UDLs for each lane, with the associated BD21/01 and BD37/01 coexistent values shown for comparison. The BSALL loading for each lane was conservatively taken to be coexistent; a short study of the traffic showed this to be reasonable for Severn.

It should be noted that for the M48 Severn Crossing, there are often incidents or routine maintenance activities that require the closure of one or more lane on the bridge. In such instances, the traffic flows will not create more onerous loading conditions (as there will be fewer loaded lanes), but they may skew the calculated BSALLs for the individual lanes. For this reason, any days including lane closures or Traffic Management are removed from the dataset prior to calculation of the BSALL.

Cable strength evalution

Inspection

Inspection of the main cables of the Severn Bridge was carried out in 2006 and subsequently in 2010, following which a dehumidification system was commissioned in 2008 (Cocksedge et al, 2010). The main cables have been subject to a yearly evaluation and this section of the paper describes the most recent evaluation performed.

The inspections were carried out in accordance with the NCHRP Report 534, Guidelines for Inspection and Strength Evaluation of Suspension Bridge Parallel-Wire Cables, Mayrbaurl & Camo (2004). The inspection was carried out by removing the wrapping wire to expose the 8322 No., 4.8 mm diameter, high tensile steel wires. Wedges were inserted between the parallel wires at number of intersections to enable a detailed visual inspection and the removal of sample wires for testing.

Initially tensile tests were carried out to determine yield strength (0.25% offset method), tensile strength, elongation in 254 mm gauge length, reduction of area and modulus of elasticity.

The detailed visual inspection records the number of broken wires and categories each wire into corrosion stages. This inspection compliments the testing regime and enables wire properties derived through testing to be extrapolated to the full cable cross section.

	Nominal Uniform Distributed Loads				
	Lane 1	Lane 2	Lane 3	Lane 4	Total
	(kNm-1)	(kNm-1)	(kNm-1)	(kNm-1)	(kNm-1)
June 2012	7.12	1.85	3.17	10.81	22.9
July 2012	7.88	2.37	2.11	11.18	23.5
Aug 2012	8.89	3.09	2.08	10.17	24.2
Sept 2012	8.51	2.67	2.12	7.88	21.2
Oct 2012	11.56	2.74	2.28	12.33	28.9
Nov 2012	10.56	2.24	2.28	10.75	25.8
Dec 2012	9.12	3.36	3.12	8.03	23.6
Jan 2013	7.68	2.75	2.93	8.98	22.3
Average	8.92	2.63	2.51	10.06	24.1
BD 21/01	17.58	8.79	7.03	17.58	50.95
BD 37/01	17.58	10.55	10.55	11.78	50.46

Table 1. Nominal BSALL UDLs

Strength Evaluation

Initially the cable strength was evaluated using the brittle wire model described by Mayrbaurl & Camo (2004) to determine nominal cable strength capacity. However, for assessment purposes, a subsequent re-evaluation was carried out to determine the cable strength at a 5% probability of exceedence. This enabled the authors to apply nationally determined partial factors of safety to the derived capacity and calculated cable force, thus allowing an Ultimate Limit State assessment to be performed rather than a factor of safety calculation.

The following sections of this paper describe the evaluation methodology, provide comments on the evaluation results, sensitivity of the results and provide recommendation for future internal inspections performed.

NCHRP Report 534 Strength Evaluation

The strength is calculated at a specific inspected location between two cable hangers which is called the evaluated panel. The strength of the cable in a specific panel is dependent not only on the deteriorated condition of the wires in that panel but also on the deteriorated condition of the wires in adjacent panels. A broken wire does not become inactive over the entire cable length. It redevelops its force as the distance from the break increases, caused by friction due to wrapping wire (ignored in calculation) and clamping force at the cable bands (hanger locations). The strength of a cable at the evaluated panel is the sum of the strengths of wires in three categories:

i. all wires in the evaluated panel minus already broken wires in that panel and nearby panels;

ii. wires that are already broken in nearby panels. These wires partially redevelop a proportion of their strength for the evaluation panel;

iii. wires in the nearby panels that are assumed to become broken i.e. that are predicted to break by the brittle-wire model. These wires partially redevelop a proportion of their strength for the evaluation panel.

Evaluating Inspection Data

Each wedge allows visual inspection of a number of wires within the main cable, but only a small percentage of internal wires can be visually inspected within each wedge. Therefore it is assumed that a wire observed at the surface represents all wires at the same radius in that half-sector between adjacent wedges. For illustration, **Figure 2** below shows the location of the observed broken wires along the wedge lines within one typical panel for Severn Bridge. This principle of assigning properties to unobserved wires is used for estimating the corrosion stage of each wire and the number of broken wires.

Grouping of Wires

Each wire is assigned a stage of corrosion as described by Mayrbaurl & Camo (2004) as follows:

Stage 1: spots of zinc oxidation on the wires;

Stage 2: zinc oxidation on the entire wire surface;

Stage 3: spots of brown rust covering up to 30% of the surface of a 75mm to 150mm length of wire; and

Stage 4: brown rust covering more than 30% of the surface of a 75mm to 150mm length of wire.

The stage of corrosion assigned to each observed wire is based on the maximum corrosion observed throughout a panel length. All wires in a specific stage of corrosion are assumed to have the same mechanical properties and distribution of these properties wherever they occur.

Following tensile testing, samples wires are grouped as follows:

i. Corrosion stage 1 and 2 wires are combined into a single group, called Group 2;

ii. Corrosion stage 3 wires that are not cracked are Group 3;

iii. Corrosion stage 4 wires that are not cracked are Group 4;

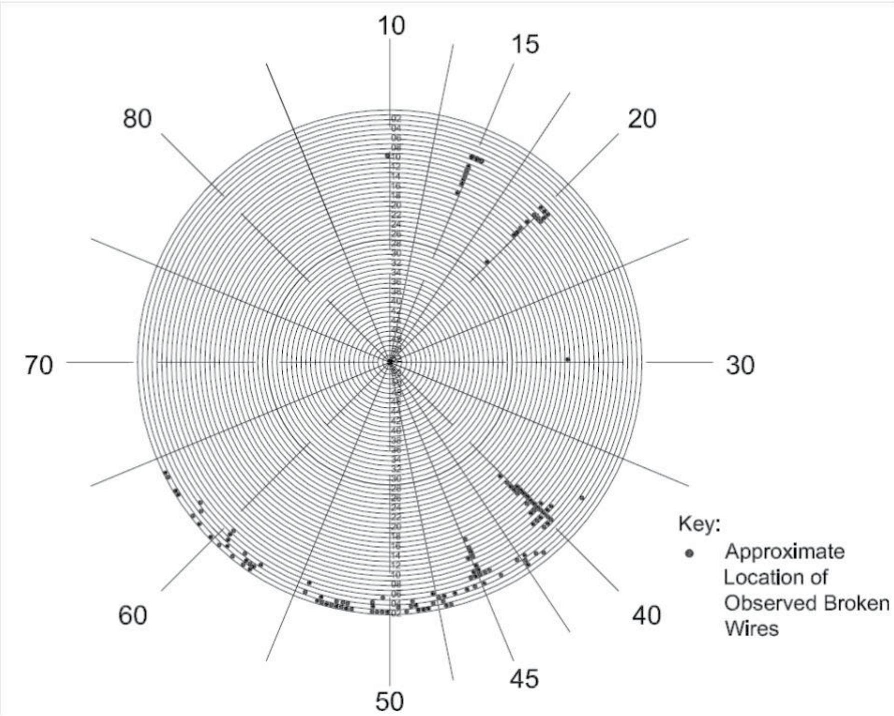


Figure 2. Location of observed broken wires

Structures

iv. Cracked wires are Group 5, irrespective of corrosion stage; they are identified by a brittle fracture type failure in tensile testing;

v. Broken wires in the effective development length are treated as a separate group, irrespective of corrosion stage.

The tensile characteristics of each group are determined. The percentage of cracked wires is used to determine the number of discrete cracked wires in the effective development length. A discrete cracked wire is a wire that is cracked in panel i but is not cracked in all the panels nearer than i to the evaluated panel. The number of discrete cracked wires and the panel location of the crack is determined using the survivor function and the corrosion stage of each of the adjacent panels.

The number of discrete cracked wires in the effective development length is subtracted from the appropriate group and assigned to a separate group for cracked wires.

It was determined that for Corrosion Stage 3 sample wires 14.3% and 25.9% of Corrosion Stage 4 contain cracks. This is the assumed percentage of cracked wires within one evaluated panel length. The survivor function is used to calculate the percentage of discrete cracked wires (percentage of cracked wires in the entire development length) for each wire group. For the Severn Bridge this results in 72% of Stage 3 wires and 89% of Stage 4 wires being treated as discrete cracked wires in Group 5. The number of Stage 3 and Stage 4 wires visually observed during the inspection therefore had a significant impact on the overall strength of the main cables.

Broken Wires and Re-development length

As discussed in the NCHRP Report 534 Strength Evaluation, wires broken in adjacent panels are assumed to partially redevelop their strength. The amount of redevelopment depends on the redevelopment coefficient.

The force in a wire that is redeveloped by clamping friction at a cable band was estimated from site measurements of the gap between the ends of broken or cut wires. This essentially measures the amount a wire slips under service load. Note that although the inspected panel is unwrapped the adjacent panels are still wrapped, and therefore the measured gap may include the effects of the adjacent panel wrapping in resisting slip. It is assumed in the evaluation panel however that the redevelopment force is applied only at cable bands and no intermediate friction occurs due to the wire wrapping. This is prudent as the potential friction from wrapping wire in the evaluation panel is lost during intrusive inspection.

The cable bands orientation and connection detail for the M48 Severn Bridge is unique due to the inclined nature of the hangers (**Figure 3**). Live load applied to hangers will cause a reduction in the clamping force. This reduction in clamping force was taken into consideration during the calculation of the redevelopment coefficient and redevelopment length by assessing the maximum reduction in clamping force applied to the cables.

For many suspension bridges the cable band clamping force increases as the slope of the cable increases, which results in a reduced redevelopment length near the towers. The clamping force for the M48 Severn Bridge however remains constant because of the unique arrangement of the cable bands and the horizontal load applied from the inclined hangers. Each cable band arrangement is uniform throughout the bridge with the number of cable band bolts remaining constant. This uniform clamping force resulted in a uniform effective development length for all panels.

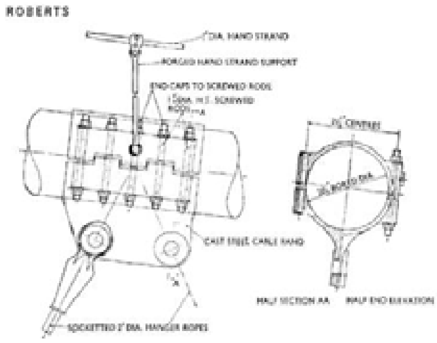


Figure 3. Cable Band and Hanger Detail Roberts (1968)

Brittle-wire Model

The Brittle-Wire Model described by Mayrbaurl & Camo (2004) was used during the assessment to evaluate the strength of the cables. It assumes that all of the wires in each group follow the same stress-strain diagram, so that the stress in all of the intact wires is the same at any specific value of strain. An individual wire will share in carrying the tension in the cable until the stress in that wire exceeds its minimum tensile strength, whereupon all its strength is lost at the break location.

Determining the cable strength requires increasing the cable stress in steps and calculating the number of wires that fail at each increment as they reach their tensile strength. The number of newly failed wires is subtracted from the number of previously intact wires to determine the number of unbroken wires. The cable force is calculated as the area of unbroken wires multiplied by the wire stress at that increment. When wires break, the stress in adjacent wires must increase to compensate. At some level of stress any further increase in stress will cause too many wires to fail and the overall strength will not increase. The stress at this optimal level determines the maximum force attained in the cable. The tensile stress which causes the highest cable capacity varies for each evaluation panel depending on the number of wires in each group for the given panel.

Discrete cracked wires which fail in adjacent panels using the Brittle-Wire model are partially redeveloped by intermittent cable bands in the same way as wires which were already found to be broken.

Evaluation of Tensile Test Data

The results of the tensile testing for each wire group on Severn Bridge was used to determine the compound cumulative distribution curves shown in **Figure 4**. The tensile data was assumed to follow the Weibull distribution for all wire groups. The parameters of the Weibull distribution are not implicit, but are determined through iterative trial and error as recommended. To ensure that the data did approximately follow the Weibull distribution, the parameters were initially estimated using the empirical survivor function. The tensile strengths were assigned to a bin boundary, each bin was assumed to be a stress range of 10 N/mm^2 and the percentage of samples within each bin was determined. The log of the empirical survivor function was then plotted vs. ln x, where x is, in this case the bin boundaries, the plot should be approximately linear if the data are consistent with a Weibull distribution. **Figure 5** shows an approximate linear relationship for each group which confirmed the appropriateness of adopting the Weibull distribution.

A sensitivity check undertaken showed that the cable strength increased by only approximately 1% if a normal distribution was adopted so the additional effort associated with using a Weibull distribution was questionable.

The minimum probable strength of a single wire which comprises groups 2, 3 and 4 were determined assuming a normal distributed onesided confidence level of 97.5%. The minimum probable strength of group 5 cracked wires, however, was not determined at a minimum confidence level. It is assumed in accordance with Mayrbaurl & Camo (2004) that the minimum tensile strength of cracked wires between hangers for a given sample was recorded. This is because insufficient cracked specimen data points are available for each sample wire. This assumption could be optimistic as a relatively small length of each wire is subjected to tensile testing. For each sample wire 10 tensile test specimens are carried out for specimen lengths of 254mm. This equates to 2.54 m out a possible 18.1m wire length being tested. It is recommended that for longer sample wires are tested in future to reduce the potential error in deriving the minimum tensile strength of cracked wires.

Embrittlement of Corroded Wires and Effect of Reduced ductility

Embrittlement phenomena (hydrogen embrittlement, stress corrosion cracking and corrosion fatigue) are typically associated with localised attack and not with general uniform corrosion. These failure mechanisms are not as easily described by their surface effects, and are usually difficult to detect in a visual inspection and become apparent only after wire failure has occurred.

Embrittlement is predominantly associated with being caused during the manufacturing process,

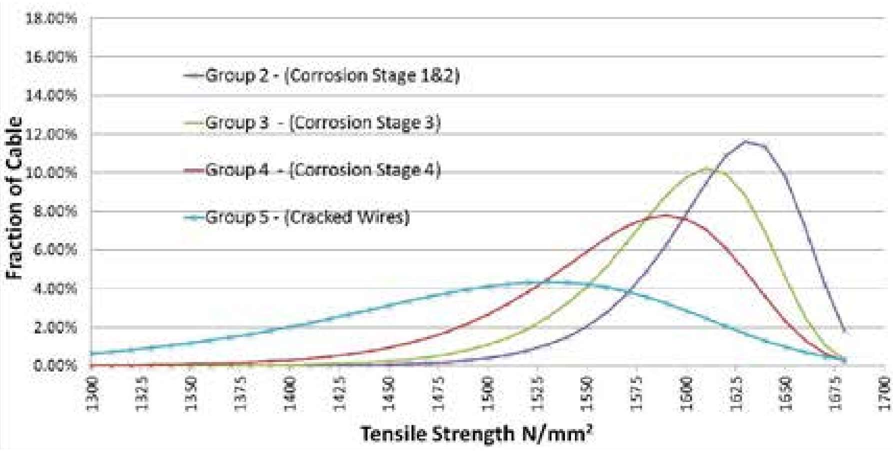


Figure 4. Tensile Strength (Weibull Compound Cumulative Distribution Curve)

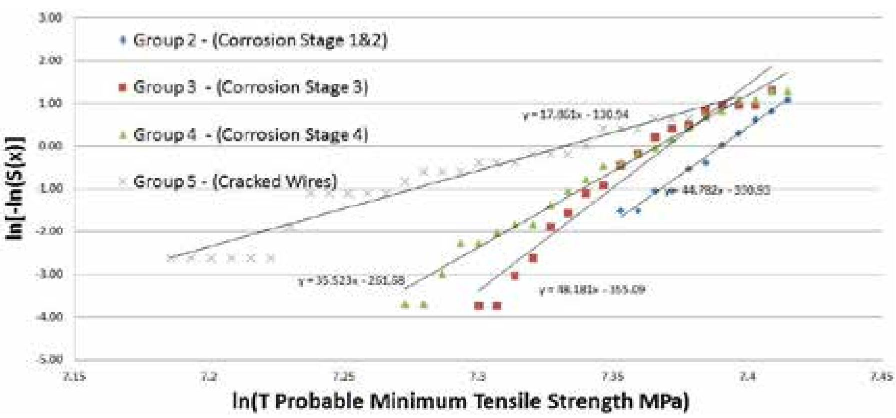


Figure 5. Weilbull Plot of Survivor Function (by Wire Group)

Structures

however, it was hypothesized that an additional mechanism of introduction of hydrogen, which could cause wires to become embrittled, is galvanic corrosion. The dehumidification system has halted any further corrosion.

Reduced ductility of corroded bridge cable wires have been recorded by Nakamura & Suzumura (2011). An evaluation was therefore carried out to assess the effects of any reduced ductility. **Figure 6** shows the average stress-strain diagram recorded for all wire groups together. It was noted however that for Group 5 cracked wires the ultimate strain is less than the average ultimate strain recorded below. To assess the distribution of ultimate strain for each wire group the authors calculated the cumulative distribution curve at the 97.5% confidence level. The method adopted was similar to the method used to calculate the minimum tensile strength of each wire between cable bands assuming the normal distribution. **Figure 7** shows the results of this assessment and the distribution of ultimate strains for the Severn Bridge.

It can be seen from **Figure 7** that the ductility of wires which contain cracks is significantly reduced from those that do not contain cracks. This is analogous to results observed on previous studies carried out for suspension bridge wires.

Although the brittle-wire model takes different strengths into consideration for different wire groups it does not take into account varying stress-strain behaviour of each wire. This behaviour may be more different than the assumption of curtailing the plateau at the peak stress since each curve has a different shape and yield point.

If the convergence stress calculated using the brittle-wire model was high, for example 1500N/mm², some wires will not have reached the strain needed to mobilise the required level of stress.

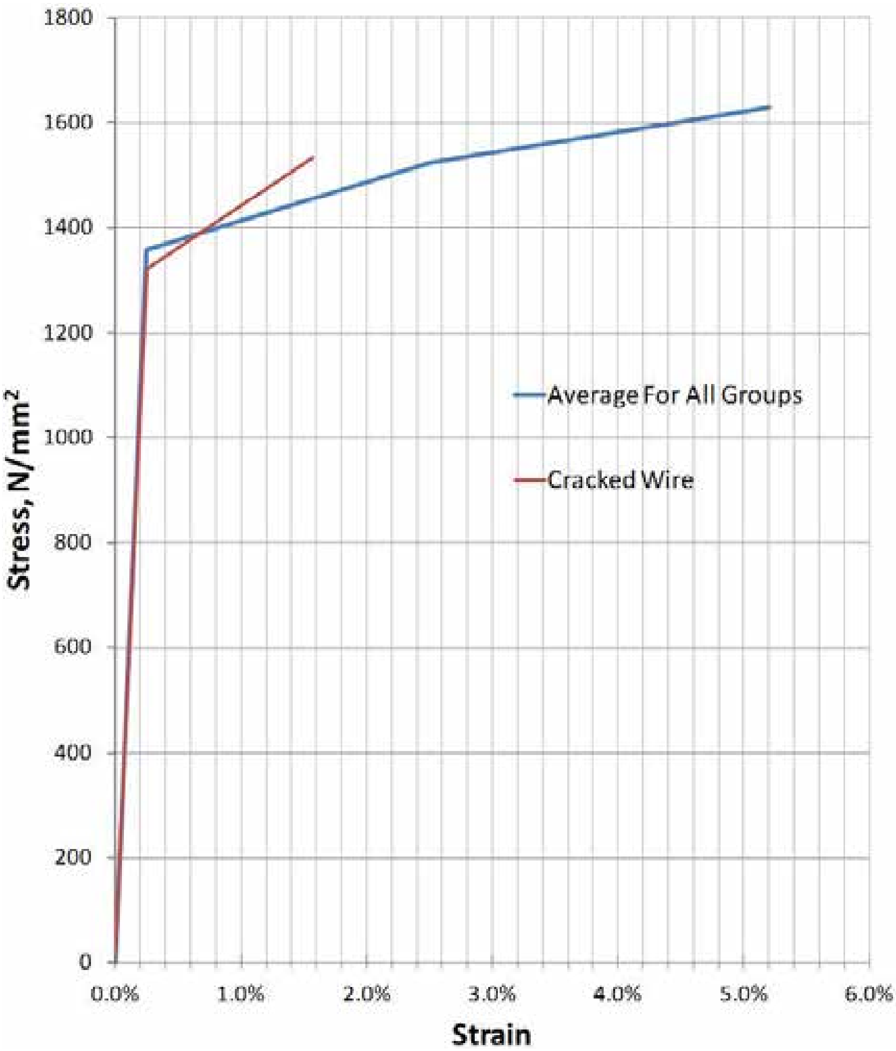


Figure 6. Stress Strain Diagram for All Wire Groups and cracked wires

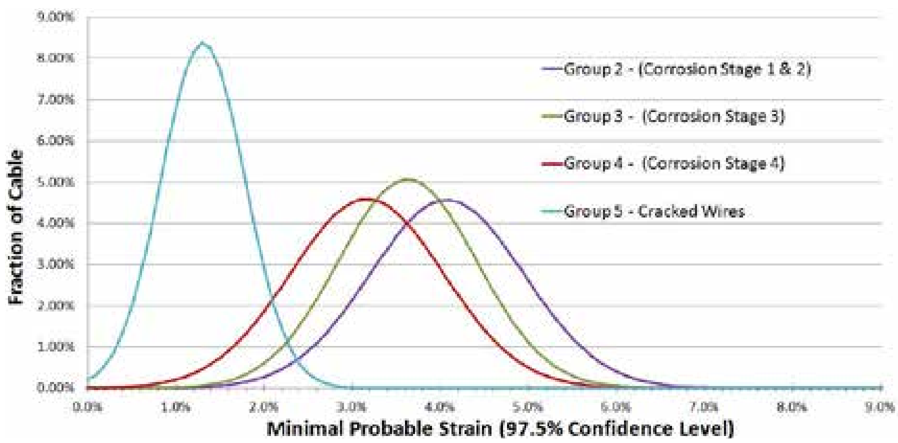


Figure 7. Ultimate Strain (Weibull Compound Cumulative Distribution Curve)

Figure 6 also shows the stress-strain diagram for a single cracked wire recorded during the internal inspection. This wire highlights the variations in stress strain behaviour.

The stress-strain curves are variable so using an average is not accurate particularly for very brittle failures. For the M48 Severn Bridge, however, convergence stress for the critical panel is reached at 1380N/mm^2. It can be seen from above that the fraction of cracked wires at this strain would be less than that determined using the brittle-wire model. Therefore the effect the varying stress-strain behaviour had on the assessment was low and verified the use of the brittle-wire model in this case.

It is recommended that this check is carried out in addition to the analysis described in the NCHRP to confirm the assumptions of the brittle-wire model.

Ultimate limit state assessment

Further to the above evaluation, a method was adopted to calculate the cable strength at the 95% confidence level to allow an ultimate limit state assessment to be undertaken following Eurocode principles but using the bridge specific assessment live loading discussed in the section on 2 'Bridge specific assessment live loading' and bridge specific partial material and load factors determined for the Severn Bridge.

The cable strength method adopted evaluated the likely error for each of the following input data:

i. Error in calculating the minimum probable tensile strength of each wire, taking into account the limited number of tests carried out on each wire length;

ii. Error in calculating the cable strength due to the error in estimation of the no. of broken wires. This method assumes that that each sector is a specific sample set with a fraction of wires observed and broken wires recorded. The normal distribution is used to determine the number of broken wires for the entire panel population based on the observed sample set;

iii. Error in calculating the cable strength due to the approximation of the fraction of cracked wires. This simply assumes a distribution based on the number of samples obtained. Since a wire is either cracked or not cracked, the binomial distribution was used to estimate the possible error in the number of cracked wires;

iv. Error in calculating the cable strength due to errors in assignment of corrosion stage;

v. Error caused by extrapolation of the cable strength over the entire length of the cable. See section below for further details.

To calculate the error caused by extrapolation of the cable strength over the entire cable length Mayrbaurl & Camo (2004) suggests that an exponential distribution is used. This type of distribution, however, does not fit in with the data returned for cable strength at each panel location. A relatively uniform reduction in cable strength has been calculated as the cable approaches the centre of the bridge. This relatively uniform change in cable strength has allowed the authors to allocate a number of cable zones for which the cable strength has been grouped.

The zones were decided upon following review of the characteristics of derived strength, location and internal environmental conditions. It has been found during this evaluation that the normal distribution fits the strength of the cables within each zone, however, it is recognised that very few data points are available and therefore the students t-leg distribution was adopted.

The error calculated between the mean nominal strength and the 95% confidence level of the datasets was determined. The error in strength calculated was then applied to the mean strength of each zone calculated from the 95% confidence level of each panel. This derives the ULS characteristic strength for all panels which were not investigated.

Although the 95% confidence level is deemed sufficient to carry out an ULS assessment it is noted however that there is limited information available on tensile strength distributions of corroded wires. Therefore, a larger confidence level of 97.5% was adopted to calculate the strengths.

Combining the derivation of the BSALL loads and strength evaluation allowed an assessment to be performed which showed adequate reliability for continued operation of the M48 Severn Bridge, provided there is no further deterioration of the main cables or significant increases in load.

Conclusions

This paper has briefly described the methods and techniques adopted to determine the bridge specific assessment live loading (BSALL) and deteriorated strength of the main cables to enable an Ultimate Limit State assessment to be carried out for the Severn Bridge. This has shown that the bridge has adequate reliability provided that traffic loading does not increase and the main cable does not deteriorate further.

Through the calculation of the BSALL over a number of months it has been shown that the bridge's assessment live loading based on the WiM data remains unchanged from previous assessments and is considerably lower than required by UK national loading standards. The BSALL can now be calculated on an annual basis and used to check that the existing assessment assumptions remain valid.

It is hoped that future deterioration of the main cables will be arrested by the dehumidification system and the bridge may continue to be in operation long into the future. Subsequent internal inspections are scheduled to ensure accurate and current data is available for future evaluations.

References

1. Cocksedge C, Hudson T, Urbans B, Baron S (2010) Proceedings of the ICE - Bridge Engineering, Volume 163, Issue 4, 01 December 2010, M48 Severn Bridge – Main Cable inspection and rehabilitation
2. Mayrbaurl & Camo (2004) National Cooperative Highway Research Program (NCHRP), Report 534, Guidelines for Inspection and Strength Evaluation of Suspension Bridge Parallel Wire Cables, Transportation Research Board of the National Academies
3. Nakamura & Suzumura (2011) Corrosion of Bridge Cables and the Protection Methods, Journal of IABSE 0042, 20-23, 2011
4. Roberts (1969) Freeman Fox and Partners, Severn Bridge - Design and Contract arrangements, Paper 7138
5. TRL Contractor Report 16 (1986). Interim Design Standard: Long Span Bridges.

159

Structures

Optimized paneling solution - design rationalization and optimised paneling for architectural freeform surfaces

Abstract

Programmatic freedom and modeling software tools have led to a spectrum of geometrically challenging freeform surfaces. The problem lies in defining these freeform design surfaces in terms of constructible components. Different custom tessellation algorithms have thus been developed in response to this problem. These tessellations produce a large number of different panel sizes and there isn't any standard solution for rationalization of such surfaces. A paneling solution plays an important role in this rationalization process.

This paper will investigate ways in which a freeform surface can be rationalized to produce an optimised paneling solution. The research develops a generative algorithm combining dynamic relaxation and a particle spring optimization with paneling layout principles.

The aim of the paper is to minimize the number of panel variations that occur in freeform surface. Finally this leads to achieve a trade-off between the rationalized geometry and its original counterpart.

Keywords: Rationalization, Paneling, Geometric Optimization, Freeform Surfaces, Particle Spring System, Dynamic Relaxation

Abir Acharjee, BArch, MSc.

Bartlett School of Graduate Studies

University College London

London

Introduction

With the emergence of large scale architectural freeform surfaces, the main challenge is to proceed from a geometrically complex design to a feasible and affordable way of production. This leads to the process of "rationalization" and achieving a "paneling solution". This process deals with the approximation of a design surface by a set of different sizes of panels that can closely approximate the design surface and can be fabricated at reasonable cost meeting the architect's perception. The main challenge in paneling these freeform surfaces lies in the complex interplay of various objectives related to geometric, aesthetic, structural and fabrication constraints that need to be considered simultaneously[1].

A smooth and continuous flow of the panel edge lines add to the rhythm, aesthetics and continuity of the structure. Quad meshes have lower node complexity. The triangular panels produce better surface approximation and continuity. Curved panels produce superior inter-panel continuity but the cost of mould fabrication often dominates the panel cost. Planar panels are easiest to produce and cost effective. The cost of construction not only depends on the number of panels and the complexity of the paneling layout but also on the frequency of reuse of different sizes of panel, referred to as "panel types". The aim of this research is to investigate the issue of buildability of a freeform surface to achieve a cost effective paneling solution which has minimum number of panel types that closely approximates the design surface. The research question is, therefore, what is the degree to which the reduction in the number of panel types affects the cost against the degree to which it deviates from the original surface?

Background

Dynamic relaxation and particle spring systems are used in many cases for form finding. Dynamic relaxation is a numerical method which is often used in structural formfinding to find minimum surface for fabric structures of cable-nets. The aim is to find a geometry where all forces are in equilibrium. One of the early examples of the use of dynamic relaxation in architectural design was the Great Court roof of the British Museum[2]. The particle system has received a lot of attention from the early pioneers of digital architecture and is used as a tool for form-finding using digital simulation of various architectural designs. Particle spring systems are used in the development of a three dimensional design and analysis tool which allows the user to find structural forms in real time[3]. A particle spring approach to geometric constraints solving was presented by Thierry[4].

The parametric design approach has been used since the early 21st century for advanced surface rationalization. Using planar quad panels for covering general freeform surfaces with new ways of supporting beam layout was proposed for the computation of multi-layer structures[5]. This was extended to the covering of freeform surfaces by single-curved panels arranged along surface strips[6]. The concept of symmetrization was proposed to enhance object symmetry by controlled deformation of underlying meshing structure[7,8]. The idea of optimizing for repeated elements by altering the vertex positions of a given mesh was explored in the context of quad meshes[9]. A mathematical approach using discrete equivalence classes has been used for triangulated surface such that each polygon falls into a set of discrete equivalence classes[10]. This assumes a fixed topology and uses the k-means clustering of triangles. A related problem of panel mould reuse using different classes of panel geometries was proposed by using a novel 6-dimensional metric space to allow fast computation of approximate inter-panel distances[1]. This does not try to use a small number of congruent shapes but addresses a related problem of what type of surfaces to use for minimizing construction costs.

Research method

Overview of the algorithm

The methodology outlined in this section consists of six steps that need to be addressed in the generation of an optimized paneling solution for a freeform design surface. The first phase of the algorithm defines the design surface by a mathematical construct and deals with tessellations which subdivides the surface into a series of triangles that forms the basic mesh. In the second step dynamic relaxation is used to get a better distribution of nodes on the surface[11]. The third step utilizes a particle spring system. The particles are released from the surface to get the specified panel edge lengths for the respective edges falling under specified ranges. In the fourth step, panels are cast and laid down on this released surface to achieve a "Panel Binning Solution". In the fifth step panels in each of the panel type are studied in detail and "Mother Panels" for all panel types are declared. In the sixth step, the panels in a panel type are replaced by the mother panel of that panel type.

Description of NURBS surface

The basic surface geometry is defined by "NURBS", which is an industry standard tool for the representation and design of geometry[12]. NURBS stands for Non-Uniform Rational B-Splines. It offers a common mathematical form for both analytical and freeform shapes. The main components of a NURBS surface are the "control points": its associated "weights", "knots" and "degree". Various surfaces can be generated by moving their control points and changing the density of tessellation. The control points have an associated polynomial equation named as the "basis" function. A rational B-Spline is defined as the ratio of the two basis function in "u" and "v" which are the two directions of the parametric space of the UV coordinate system[13]. Two polynomial equations are defined i.e. basis-U (Ni,p) and basis-V (Nj,q), where the shapes of the basis functions are determined by the knots vectors xi, and defined by the following formula for the u-direction and alike for the v-direction[11,12].

Equation 1

$$N_{i,1}(u) = 1 \text{ if } x_i <= u < x_{i+1}$$
$$= 0 \text{ otherwise}$$

Equation 2

$$N_{i,p}(u) = (u - x_i)\, N_{i,p-1}(u) \,/\, x_{i+p-1} - x \,\mathrm{I} + (x_{i+p} - u)\, N_{i+1,p-1}(u) \,/\, x_{i+p} - x_{i+1}$$

Subsequently the final calculation of the NURBS curve is determined by a parametric equation which calculates the points on the curve for u and v respectively.

Equation 3

$$P(u) = \sum_{i=1}^{m} N_{i,p}(u)\, P_i \quad \text{and} \quad P(v) = \sum_{j=1}^{m} N_{j,p}(v)\, P_j$$

m is the number of control points vertically and n is the number of control points horizontally. From (**Equation 1**) and (**Equation 2**), $N_{i,p}$ (u) and $N_{i,q}$ (v) are the B-spline basis functions with degree p and q; P_i and P_j are the array of m x n control points. From (**Equation 3**) the resultant P(u) and P(v) define the points on the surface for a specific u,v location. The code uses a double loop that calculates the NURBS equation for all the control points and returns a 3D vector containing the XYZ position of the points on the surface.

Description of dynamic relaxation

The aim of the relaxation process is to find a geometry where all forces are in equilibrium and to have a better distribution of nodes throughout the surface. The panel edge lengths are used as weights in the NURBS equation and determine the direction that gets the majority in the optimization. The relaxation process only affects the position of the nodes in the parametric space; therefore the nodes are free to move around on the surface through manipulation of their respective "u" and "v" coordinates.

Description of the particle spring system

The main aim for the inclusion of a particle spring system is to fix the initial panel edge dimensions to a fixed number of lengths before the panels are formed. This reduces the variations in panel edge lengths for the overall topology. Four variations of lengths (6, 8, 10, 12 variations) will be tested. The system consists of a series of particles which act as the nodal points for the original surface and a set of springs which connect the nodes via the specified tessellation pattern. The preliminary positions of the nodes are derived from the mentioned NURBS algorithm. At each iteration, the movement of the nodes is established depending on the ratio of the actual length to ideal spring length[14]. The spring lengths are compared against a series of ideal lengths, which are calculated prior to optimization. Each of the three sides of the panels are analysed individually. The springs are released from the surface one at a time and if their lengths are within a defined range then they are resized based on the ideal length of that range (**Figure 1**).

Description of the paneling layout

"Panel types" refer to the different sizes of panels on the surface. Two panels with the same panel types must have their respective edge lengths within specified "tolerances". "Kink Angle" is the angle between these two panels (**Figure 2**).

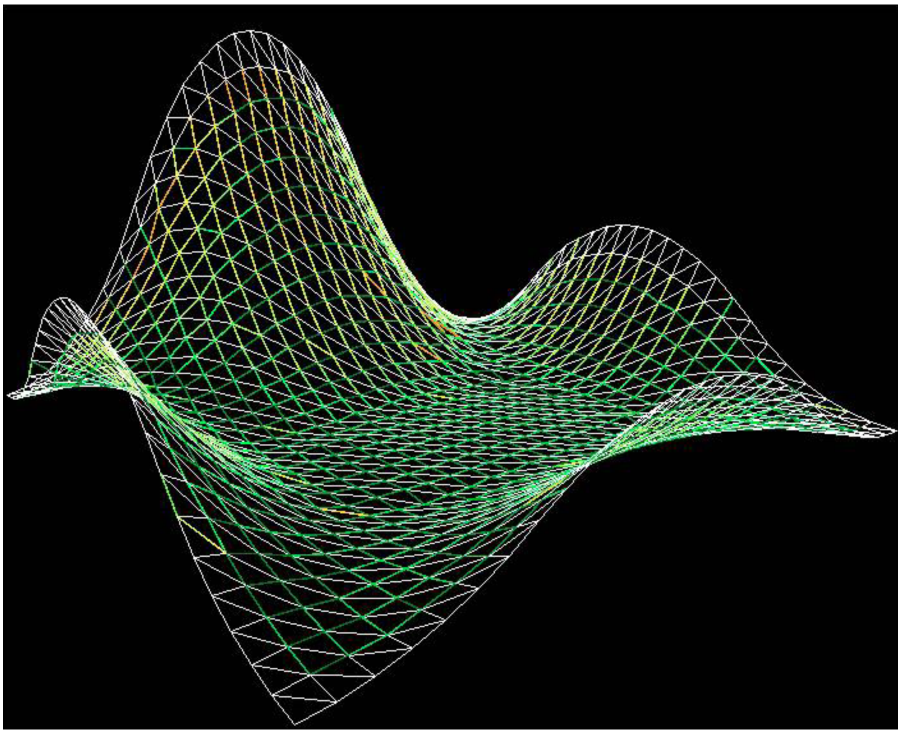


Figure 1. Surface after releasing with set number of lengths

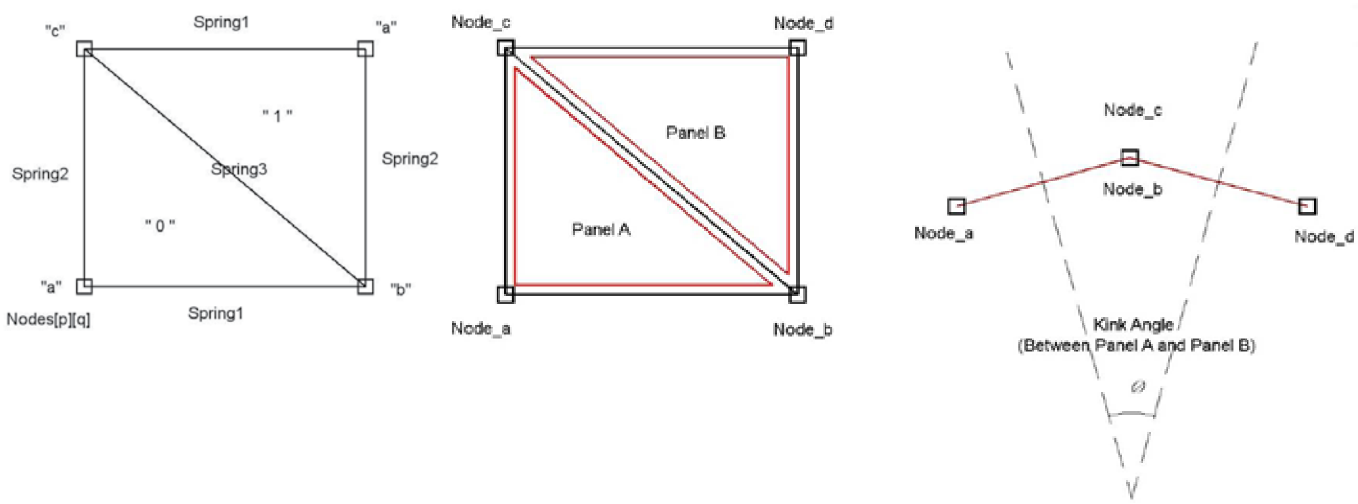


Figure 2. A node with two panels and the kink angle between them

Description of panel binning solution

The "panel binning solution" is the process of extraction of the number of panel types. Two panels are selected and each of their edges are analysed. The respective shortest sides, medium sides and longest sides for these two panels are compared. If the differences in lengths are within the specified tolerance they are assigned the same panel type or if they do not match with the existing panel types a new panel type is declared. This is repeated with all the panels (**Figure 3**).

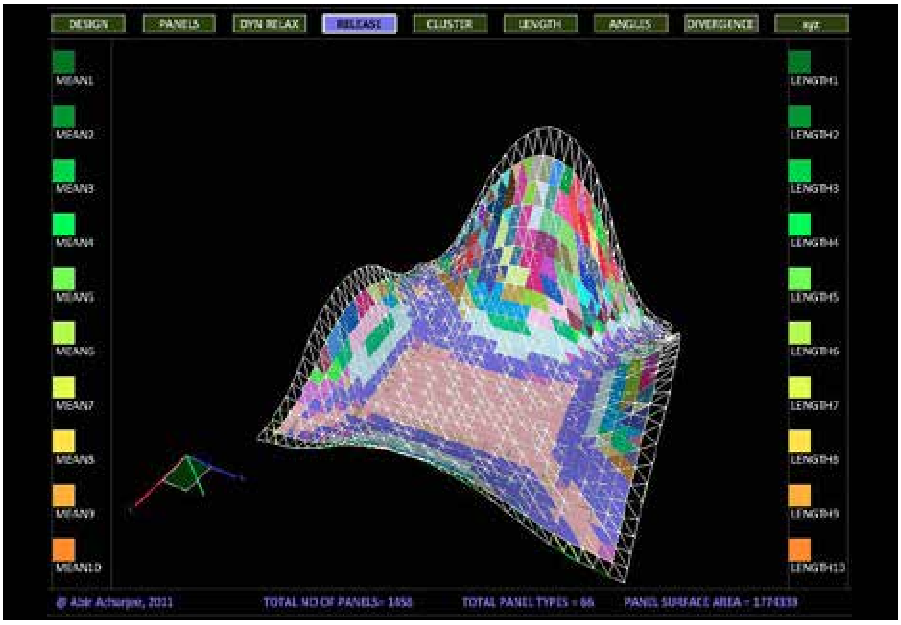


Figure 3. Panel types extraction

Formation of mother panel and description of panel mapping

Panel edge lengths and areas of each panel in a panel type are analysed and they are sorted with respect to their areas in that group. The minimum panel for each group is declared as a "mother panel" of that group (**Figure 4**). The panels in a particular group are replaced by the mother panel of that group. The mother panel dimensions can be used for fabrication purposes. A variable offset for every panel is calculated based on the panel's area, its panel type and the mother panel for this panel type. The tolerance used in the panel edge lengths contribute towards the divergences.

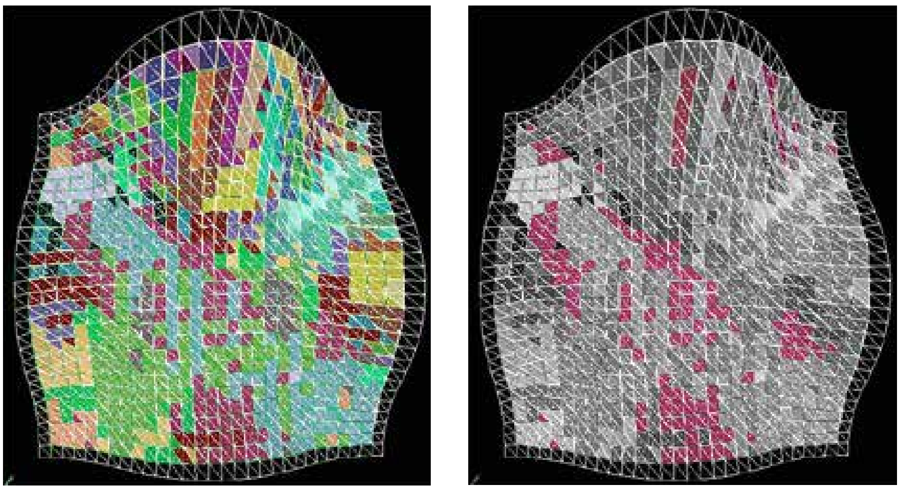


Figure 4. Showing all panel types (left), showing one panel type in red (right)

Experiments

Parameters for the experiments

The experiments aim to pick up from the constructional phase of a design project and assume a defined topology which needs to be resolved for fabrication. Thus a single surface is defined as opposed to testing multiple surfaces. The system unit is defined as "mm" (millimetre) with a maximum tolerance of 0.1 mm. Seven categories of node samples, which are controlled by the density of triangular grid, are taken in total and the standard deviation of the panel dimensions for each node category help to determine the cost-effective density of the tessellation for initial experimentation.

The first set of experiments deal with dynamic relaxation to analyse the extent of its ability to reduce the variations in the panel edge lengths. After releasing the nodes using particle spring optimization the reduction in the number of panel types is analysed. The deviation of each node from the original surface is tested to understand the overall deviation. The difference in the kink angle between two consecutive panels of the original surface and that of the optimised surface estimates the angular deviation. The change in structural efficiency is analysed by the change in structural stress. All these factors are tested for all the four different length variations and add up collectively to the success of the algorithm.

Analysis of dynamic relaxation

Dynamic relaxation settles down to a stable state after 300 iterations. Initial analysis of the spring lengths shows that out of a total of 2352 panel edge lengths there are 2071 different panel edge lengths. The performance of the algorithm is tested and it shows reduction in the range of panel edge length by about 27% and reduction in their standard deviation by about 12%. The maximum and minimum values are brought down by approximately 28% and 23% respectively (**Figure 5**).

Analysis of the node movements after relaxation shows that the nodes move differently to attain the equilibrium lengths (**Figure 6**). Analysing the difference in the kink angle between the original surface and the dynamically relaxed surface, helps to understand the effect of relaxation on the surface smoothness. A small change in the kink angle is observed (**Figure 7**).

Seven ranges of average kink angles are mapped to panel colours of the surface. The change in kink angle is mostly seen in the area near one of the saddle points (**Figure 8**).

Optimization of panel edge lengths and release function

Having reduced the level of variations and total length range through dynamic relaxation, the following investigations are initialised from the relaxed position of the nodes, with all the four variations. The ideal rest lengths are achieved through an analysis of all the panel edge

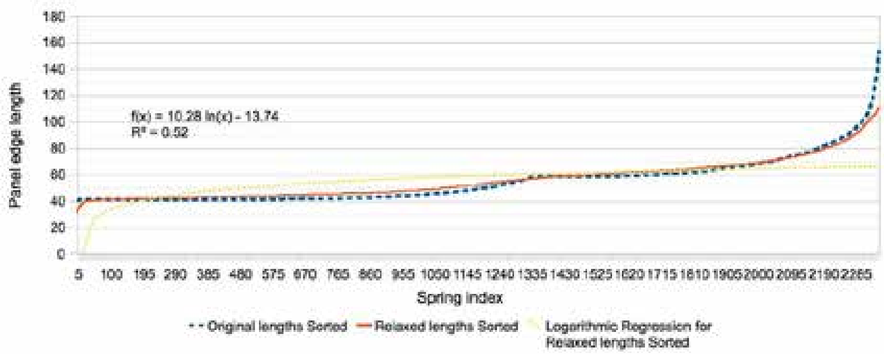


Figure 5. Reduction in panel edge length variation

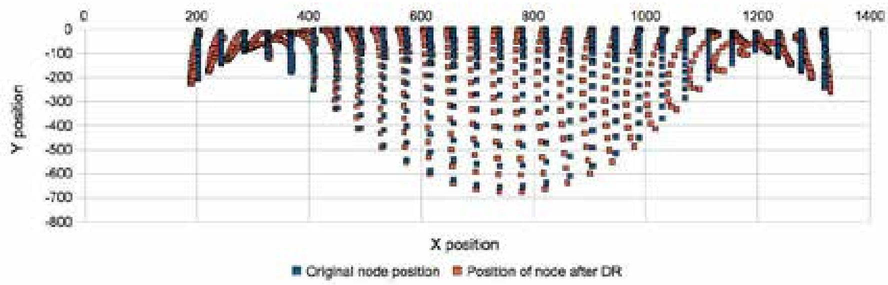


Figure 6. Movement of the nodes after dynamic relaxation

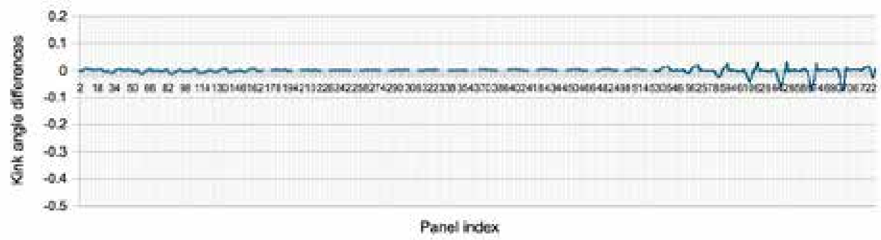


Figure 7. Difference in kink angle between the original Surface and the dynamically relaxed surface

lengths and dividing the actual range of spring lengths by the number of variations required. For each average value two ranges are assigned which serve as a guide for the springs (**Figure 9**).

Spring 1 and Spring 2, which make the main quadrilateral of two panels, are released. They were constrained to various numbers of average lengths. For testing the accuracy of the optimization technique, the final rest lengths of the springs are analysed. Most of the springs are successful in achieving the averages (**Figure 10,11**).

Analysis of panel types

The purpose of the panel binning solution phase is to group similar panels with certain tolerances. The springs are released with 6, 8, 10 and 12 variations in lengths. Then the panels are cast and the panel binning solution is applied (**Figure 12-14**).

When dealing with a tolerance of less than 0.1 mm the 6 and 8 variations produce similar numbers of panel types. Also there is no significant reduction in the number of panel types between 12 and 10 variations. For a tolerance of less than 0.25 mm there is a steady drop in the number of panel types for all the variations. With the increase in tolerance, the rate of decrease in panel types is reduced (**Figure 14**).

Use of dynamic relaxation before the release function slightly increases the kink angle in the area of transition of high to low curvature (**Figure 15**).

It is interesting to note that, as the number of variations is decreased from 12 to 10 more kink angle variation is seen in the area near the saddle point. With 8 or 6 variations, kink angle gets distributed over the surface. Eight variations give a moderately reduced kink angle as compared to six variations (**Figure 16**).

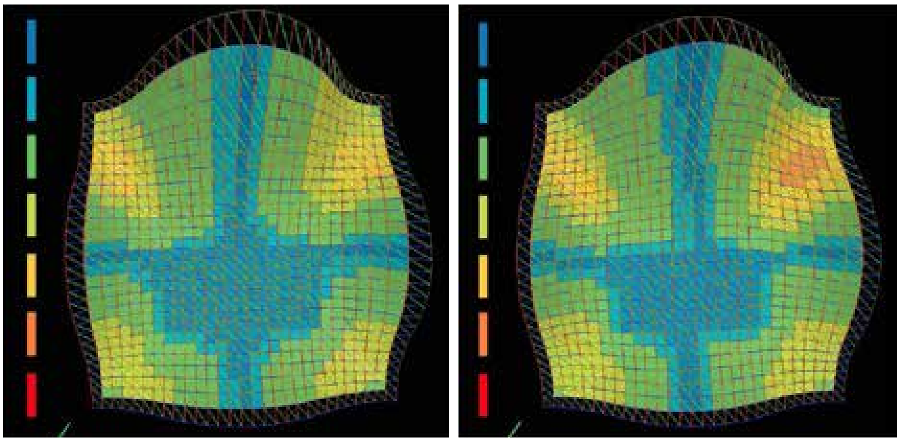


Figure 8. Mapped kink angle colours for original (left) and relaxed surface (right)

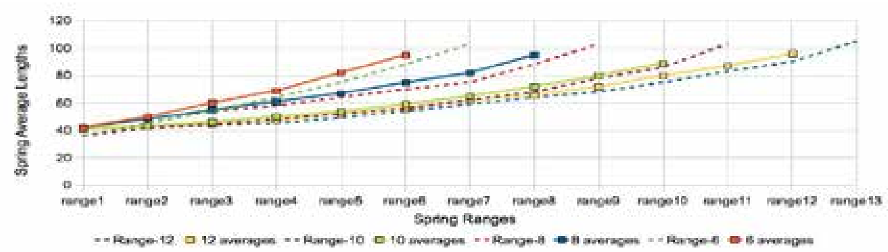


Figure 9. Average lengths for different variations in spring lengths

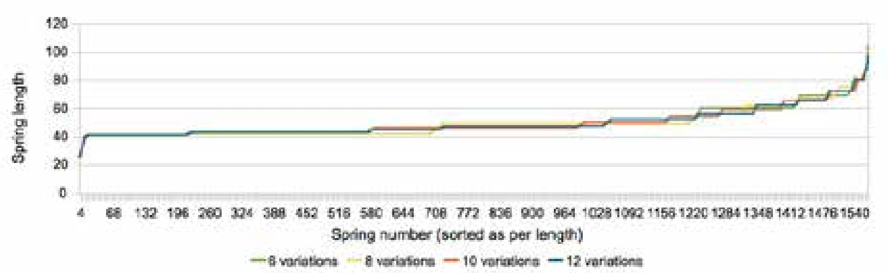


Figure 10. Final rest lengths for all the four variations

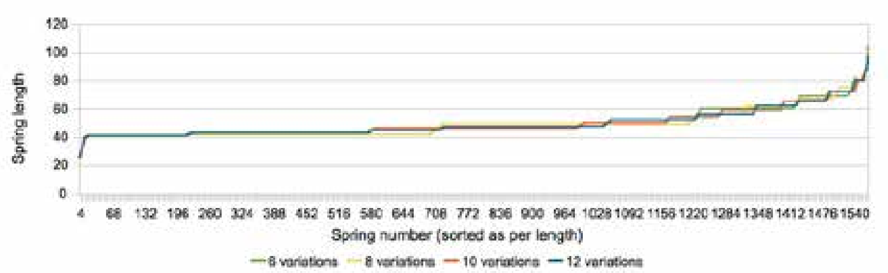


Figure 11. Final rest lengths of spring 1 for 10 variations

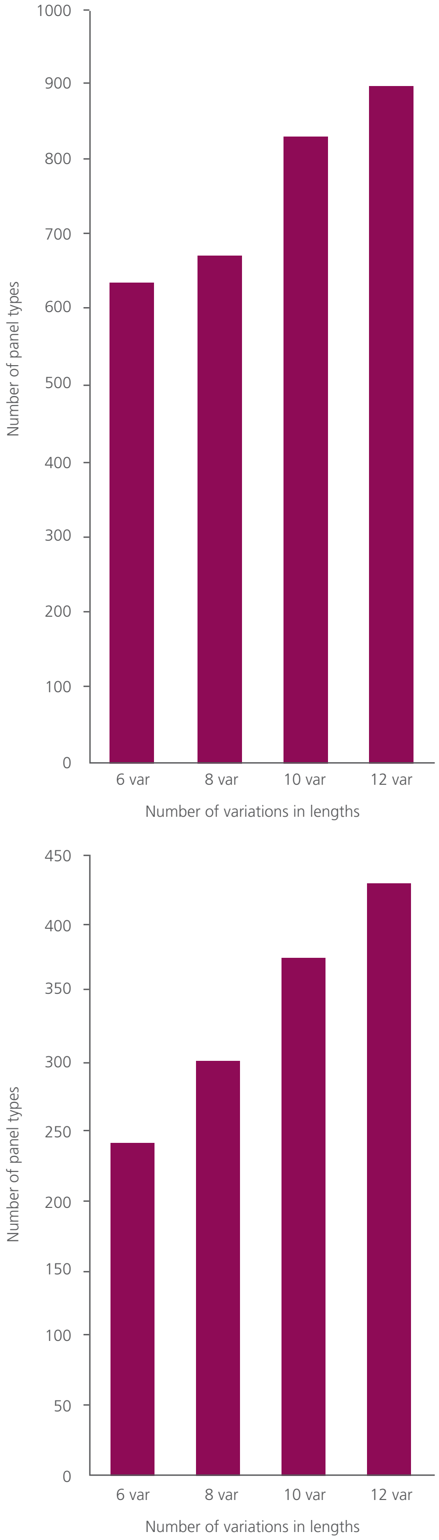


Figure 12. (Top and bottom) Reduction in the number of panel types with different variations in lengths

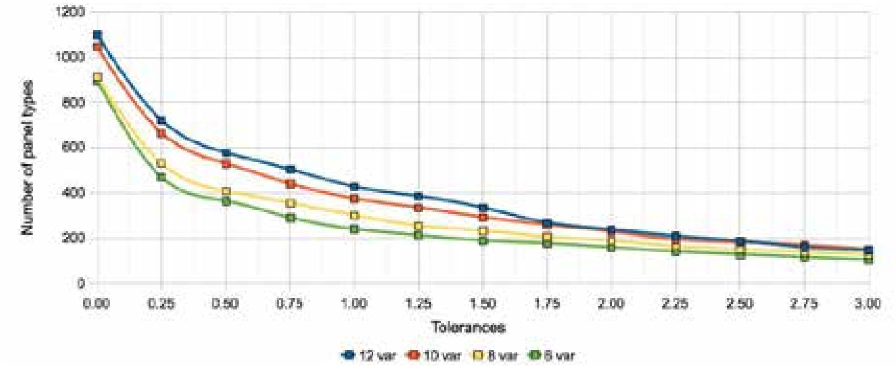
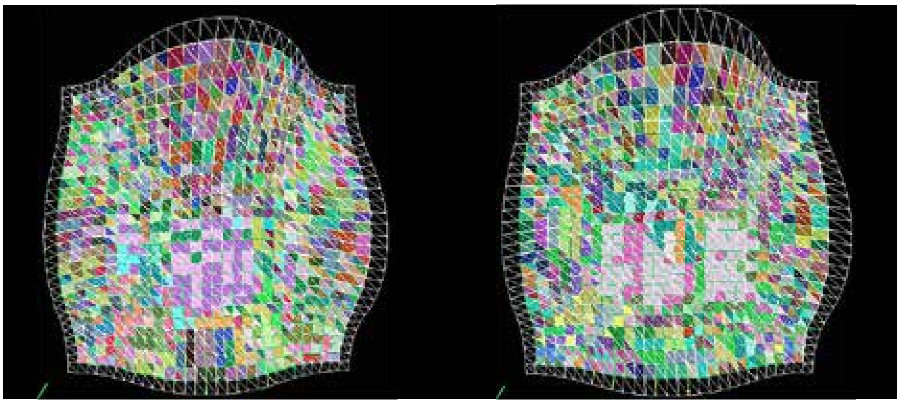
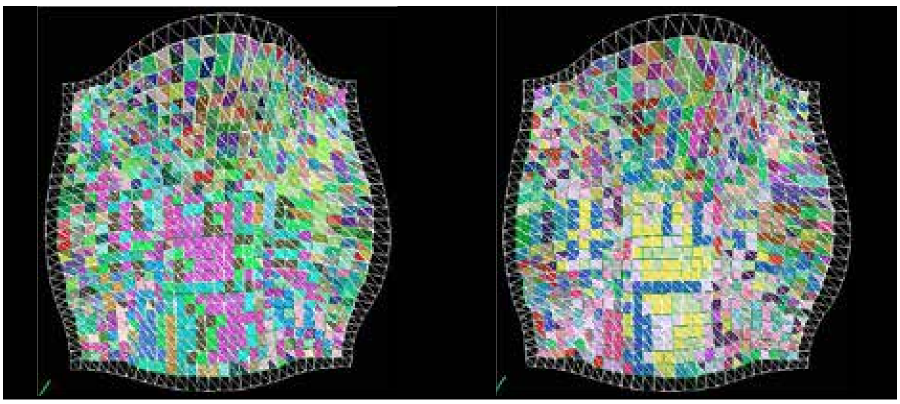


Figure 13. Reduction in the number of panel types with panel tolerances



(12 variations – 427 panel types) 10 variations – 341 panel types)



(8 variations – 330 panel types) (6 variations – 239 panel types)

Figure 14. Reduction in the number of panel types with different variations in lengths (fixed tolerance 1.0)

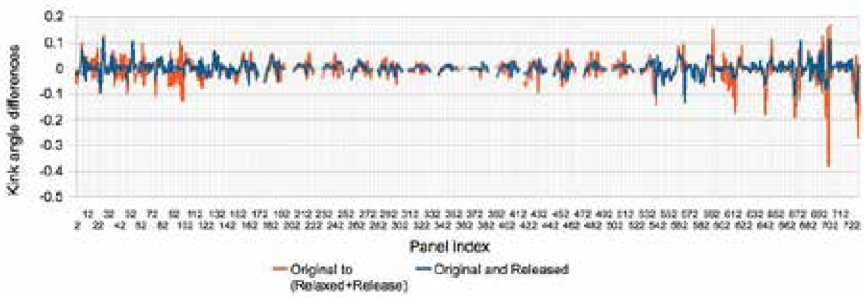


Figure 15. Difference in kink angle after releasing the springs (with and without relaxation in the first stage)

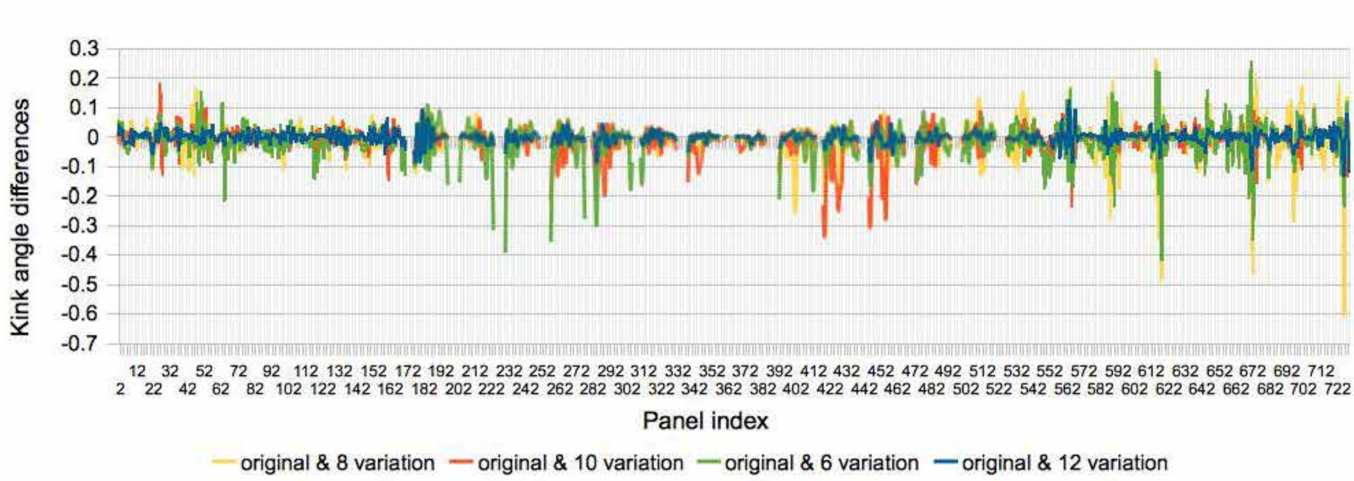


Figure 16. Variations of kink angles with different variations of spring lengths

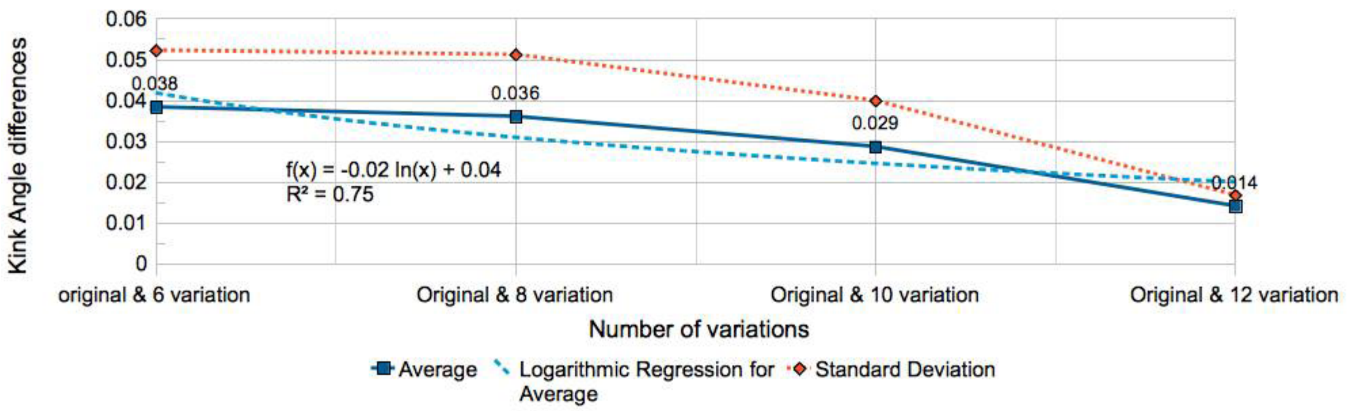


Figure 17. Variations of kink angle differences with variations of spring lengths

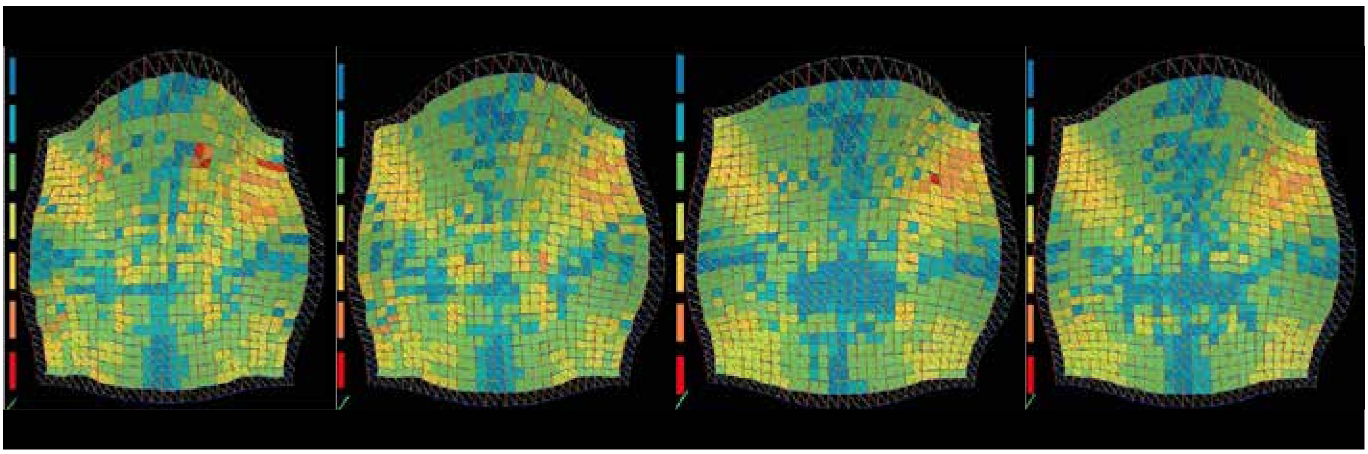


Figure 18. Average kink angle for 6 (left) to 12 (right) variations

Seven ranges of average kink angles ranging from low (blue) to high (red) are mapped on to the panel colour of the surface (**Figure 18**).

In order to understand the deviation from the original surface, the actual XYZ deviations of the node from the original position are studied (**Figure 19, 20**).

Twelve variations have a consistent and smaller deviation throughout the surface. Eight variations have more deviation in the whole spectrum; it is less than six variations at the start but is higher in other areas. There is a significant drop in kink angle from 8 to 10 variations (**Figure 21**).

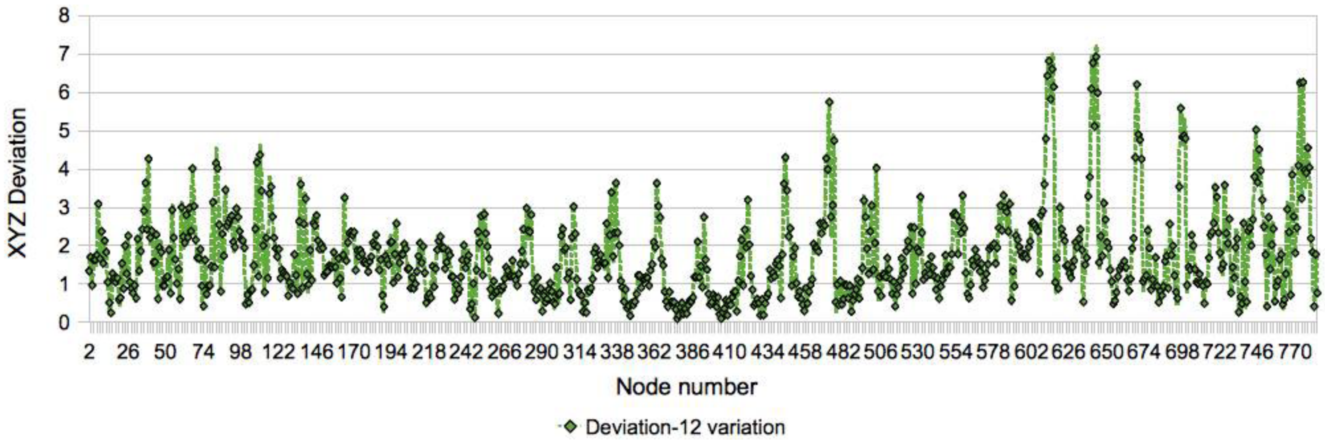


Figure 19. Deviation for 12 variations in spring lengths

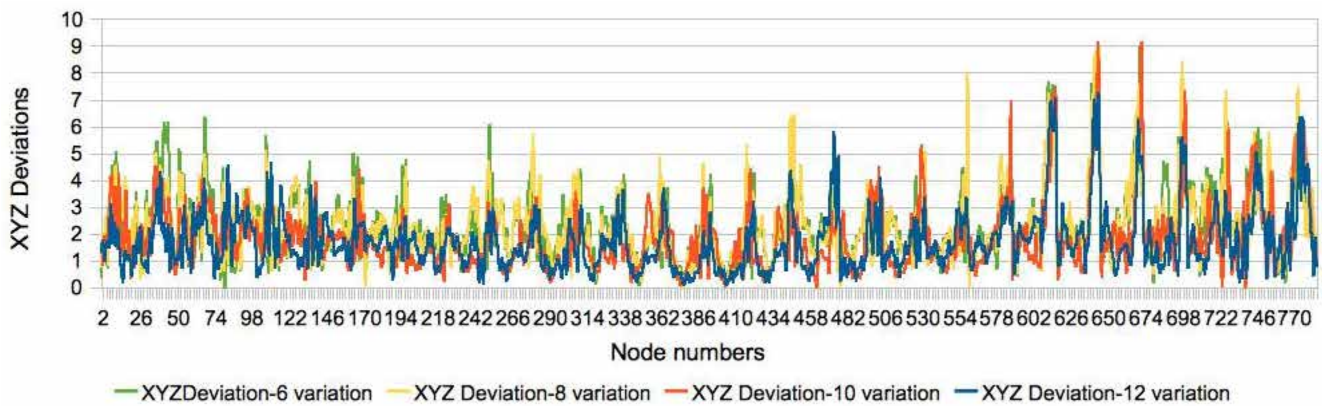


Figure 20. Deviation for all variations in spring lengths

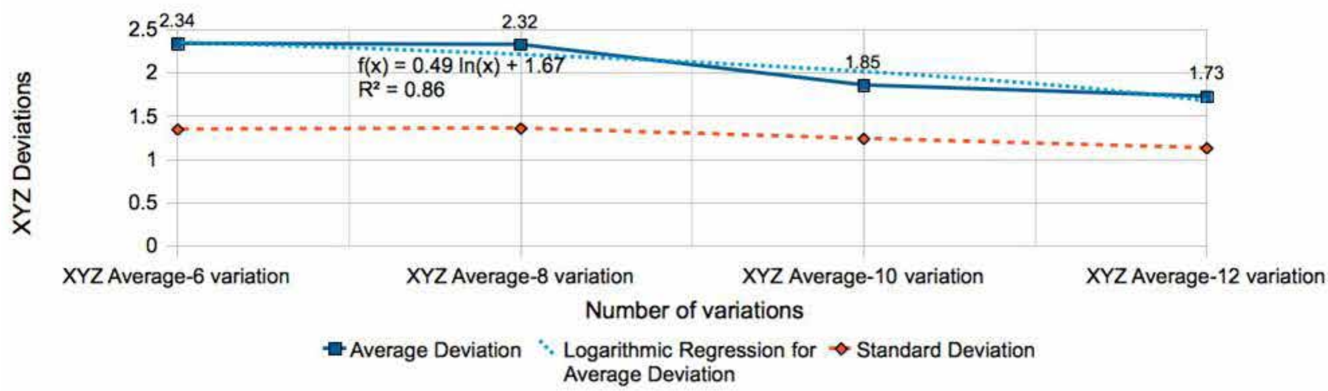


Figure 21. Deviation for all variations in spring lengths

Structural analysis

The following experiments aim to find the variation in structural efficiency with the decrease in the number of variations in the structure. Structural analysis is performed with pinned edge supports at the edges and fixed connections. Only the self weight of the structure is considered. If the loads and sections are the same, the measure of the structural efficiency of the geometry is done approximately by the stresses. The maximum and minimum stress developed in the members due to axial forces and moment help to get a preliminary understanding of the overall change in the structural forces. It is seen that as the number of length variations decreases in the members there is a tendency for the stress in the members to increase (**Figure 22**). This is only an indication of how the stresses are developed.

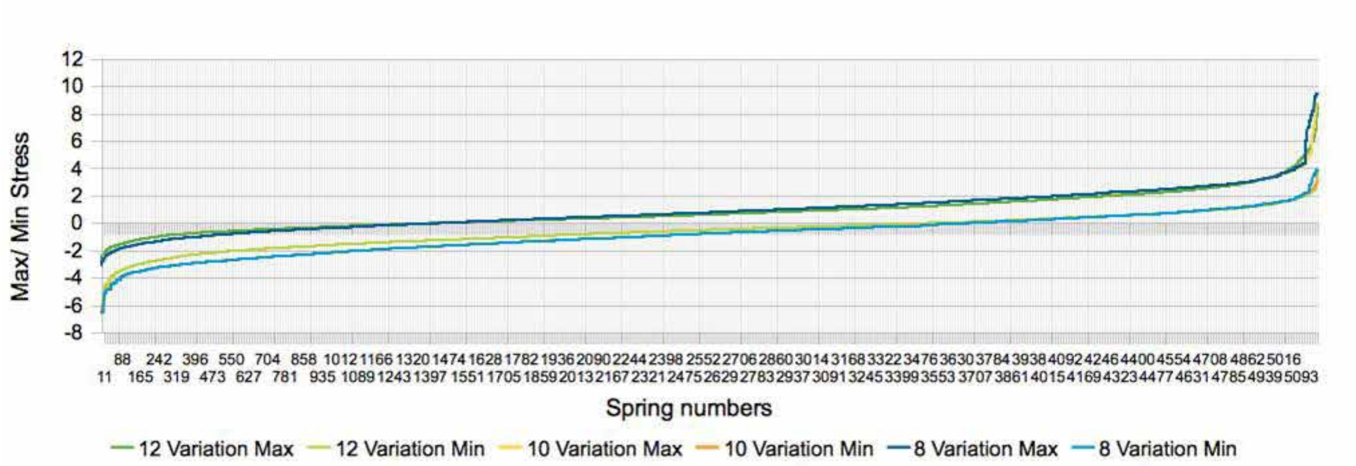


Figure 22. Max and min stresses for release surface with 12, 10 and 8 variations

Experiments with "example surface"

An example surface similar to the Great Court roof of the British Museum is taken to perform a comparative study. With a similar setup of experiments like the "original surface" all related experiments are performed. These experiments help towards the success of the algorithm and also to understand the additional parameters related to the original form of the surface geometry, which affect the paneling solution[15]. The surface is released with the set lengths for the panel edges and panel binning solution is applied (**Figure 23-26**).

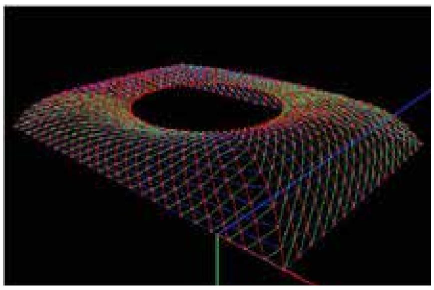


Figure 23. Original surface

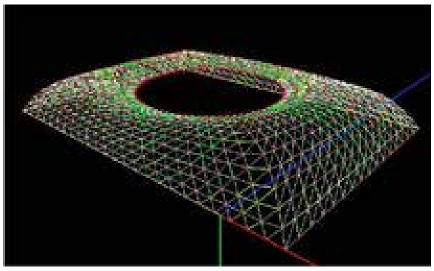


Figure 24. Released with 6 variations

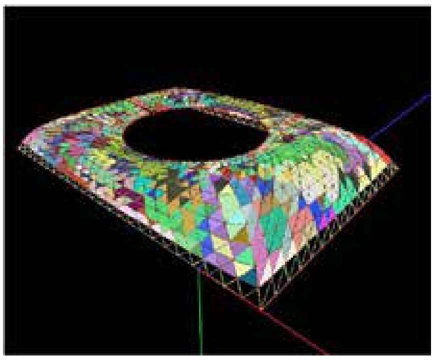


Figure 25. Original paneling layout

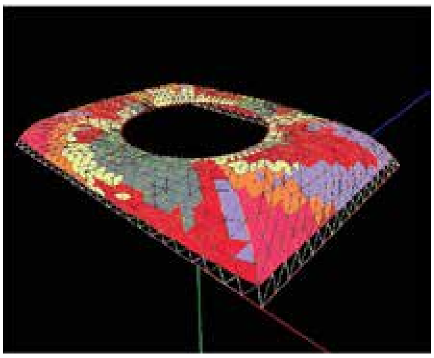


Figure 26. Optimised paneling solution

Discussion

With regard to the aim of the paper, the results show that out of the initial 1280 panel types on the original surface the applied algorithm reduces this to 374 panel types, which is a reduction of 70.78% with acceptable average deviation of 1.85 and an average kink angle of 0.029 for a tolerance of 1.0. This can be further reduced with different length variation as per the acceptable deviations, surface kink and tolerances. Analysis of dynamic relaxation results show that the use of relaxation helps to reduce the number of panel types when dealing with smaller tolerances resulting in a minor increase in kink angle. Dynamic relaxation mostly contributes to the paneling aesthetics by distributing the nodes on the surface and reducing the range and standard deviation of the spring lengths. Further experiments with the example surface reflect that the original surface geometry, its curvature conditions and symmetry influence the optimum paneling solution. The initial curve network and the original organisation of the nodes have a unique influence on the release of the surface. The variations in edge length and the panel geometries influence the effect of the binning solution. Curvature conditions, like the presence of saddle points affects the kink angle change. Interestingly a similar overall rate of decrease in the number of panel types is seen in the example surface as seen with the original surface, with the increase of tolerance.

Feeding in actual design constraints, a project specific optimum solution can be achieved. Further details of the panel mapping for the computation of inter-panel distance and the intersection of panel with centreline need to be analysed. The variation of panel types has been reduced but the resultant node and connections between the panel frameworks still remain differentiated and are open for further experimentation. Angle constraints and spring particle geometric solver can also be implemented to constrain the kink angles in areas of high visibility. Such a structural design has a lot of complexity and further investigation with loads, predefined member sizes, movable nodes and related structural parameters is required for detail analysis. The study focused on a specific topic of modularity for triangular panels but it can also be extended to different geometric shapes of panels. This can lead to the creation of a tool that could embed the geometric behaviour, manufacturing constraints and paneling logic into a single system.

Conclusion

This research sets out to analyse some panelization issues concerned with the construction of a freeform surface. A method to deal with the number of variations of panel sizes in such surfaces was proposed by using a generative algorithm combining dynamic relaxation and particle spring optimization. The experiments are conducted in three main stages which include the effectiveness of the dynamic relaxation, exploration of panel edge length variations and finally reduction in the number of panel types. The variations in lengths are tested against the node deviations, kink angles, structural efficiency and design tolerances to derive the optimum panel types for specific projects.

A trade off to reduce the extra panels against the deviation from the design surface can ideally be possible in an actual project scenario. This is a multidimensional issue related to the complex interplay of various objectives related to design, geometric, aesthetic, structural, fabrication and cost constraints that need to be considered in a similar scale with respect to the actual project. The scale on which the two issues of cost and deviation are plotted to find the optimal point of panel complexity would change from one project to another and a unique project optimum would be achieved for each project. This tool can be used to trade off the cost of extra panels against the deviation from the design surface. It allows the designer to achieve a paneling solution not only as an intuitive design rationalization tool but also as a method of post rationalization of an optimised geometry for achieving an optimum paneling solution.

Acknowledgements

I am grateful to Sean Hanna for his insightful feedback and advices. I would like to express my gratitude to Atkins. I would like to sincerely thank Martha Tsigkari and Prarthana Jagannath for technical advice. I would also like to thank Peter Debney for advice with Oasys GSA for structural analysis.

This paper was previously presented at the 14th Generative Art Conference in December 2011 in Rome.

References

1. Eigensatz, et al., 2010. Panelling Architectural Freeform Surfaces, ACM SIGGRAPH
2. Williams, Chris J. K., 2001. The analytic and numerical definition of the geometry of the British Museum Great Court Roof, University of Bath, UK
3. Kilian, A. Ochsendorf, J., 2005. Particle Spring Systems for structural form finding, Journal of the international association for shell and spatial structures, 148, pp. 77-84
4. Thierry, S.E.B., 2011. A Particle-Spring approach to geometric constraints solving, University of Strasbourg-CNRS,LSIIT, "Computer Image and Geometry" team France
5. Liu, et al., 2006. Geometric modeling with conical meshes and developable surfaces, ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH
6. Pottmann, et al., 2010. Advances in architectural geometry, Vienna: Springer-Verlag
7. Mitra, N. Guibas, L. Pauly, M., 2007. Approximate Symmetry Detection and Symmetrization, ACM SIGGRAPH.Golovinskiy, A. Podolak, J. Funkhouser, T.,2009. Symmetry-Aware Mesh Processing, IMA Conference on the Mathematics of Surfaces.
8. Fu, C. Lai, C. He, Y. Cohen-or, D., 2010. K – Set Tilable Surfaces, SIGGRAPH, Technical Paper, ACM Transactions on Graphics, Article No 44, Vol 29

9. Singh, M. Schaefer, S., 2010. Triangle surfaces with discrete Equivalence Classes, Proceedings of ACM SIGGRAPH, pp. 46-47
10. Darvishani, S., 2009. Honeycomb Topologies: Design rationalization of a freeform space frame structure, MSc Thesis, Bartlett School of Graduate Studies, UCL
11. Rogers, F., 2000. An introduction to NURBS: with Historical Perspective. London: Kaufmann
12. Kolarevic, B.,2005. Architecture in the digital age: design and manufacturing, New Edition. London: Taylor and Francis Ltd
13. Kanellos, A., 2007. Topological Self Organisation: Using a Particle- Spring System Simulation to generate structural space filling lattices, MSc Thesis, Bartlett School of Graduate Studies, UCL
14. Boer, S. Oosterhuis, K., 2002. Architectural Parametric Design and Mass Customization, ONL Architecture, Rotterdam, Netherlands.

Structures

160

Sustainability

The Blackwater Valley Road: using green infrastructure for ecological mitigation

Abstract

The Blackwater Valley Road comprises 17km of dual carriageway on the Hampshire/Surrey borders. The road was planned and built long before the term ‘green Infrastructure’ became widely used, but it provides an excellent practical example of the approach. The success of ecological mitigation was first assessed in 2004 and reviewed again in 2011.

The cost of ecological mitigation is usually just a small part of the budget for a road scheme. Nevertheless, a cost is involved. Despite this, it is very rare for schemes to be reviewed to see whether the mitigation provided actually gave the desired results and to allow others to learn from the experience. One exception was the Blackwater Valley Road (A331) on the Hampshire/Surrey borders. In 2004 the success of ecological mitigation on this road scheme was assessed (Atkins et al. 2004). In 2011, 15 years after the road was completed, the authors review the results of monitoring studies. This paper is presented to allow others to share the experience, and also to ask whether the recommendations originally made in 2004 are now being put into practice more widely.

Suzanne Glencross MCIEEM

Senior Ecologist

Water & Environment

Atkins

Claire Wansbury FCIEEM CEcol CMLI CEnv

Associate Director

Water & Environment

Atkins

Steve Bailey MCIEEM

Blackwater Valley Countryside Partnership

The Blackwater Valley Road comprises 17km of high speed dual carriageway linking the A30 with the A31 and M3 (**Figure 1**). It was conceived in the 1960s and then built in four stages between 1985 and 1996 by Hampshire and Surrey County Councils. Each of the four stages was the subject of a separate Environmental Impact Assessment (EIA); at the time these schemes were some of the earliest road schemes to be subject to EIA in the UK. The scheme was planned and built long before the term 'green Infrastructure' became widely used (CIRIA 2011), yet the Blackwater Valley Road provides an excellent example of the approach.

The Blackwater Valley Countryside Project (BVCP) was formed in the 1960s to tackle the neglect and pollution in the Blackwater Valley and now works to increase the valley's importance for biodiversity and as a recreational resource for local people. The BVCP was involved in the scheme throughout its construction and remains responsible for management of the green infrastructure that was retained and created as part of the scheme.

The Blackwater Valley is a wedge of open space separating major urban areas on the Surrey/Hampshire/Berkshire borders. The landscape is dominated by a chain of lakes formed by sand and gravel extraction. Rapid urban expansion led to the degradation of the local landscape and resulted in traffic congestion in the urban areas along the valley. The scheme to build the Blackwater Valley Road was constrained by significant engineering challenges including the narrowness of the valley, the presence of lakes and rivers and by the need to construct an aqueduct for the Basingstoke Canal Site of Special Scientific Interest (SSSI).

The EIAs for the four schemes recognised that the road would have significant effects on ecology as a result of:

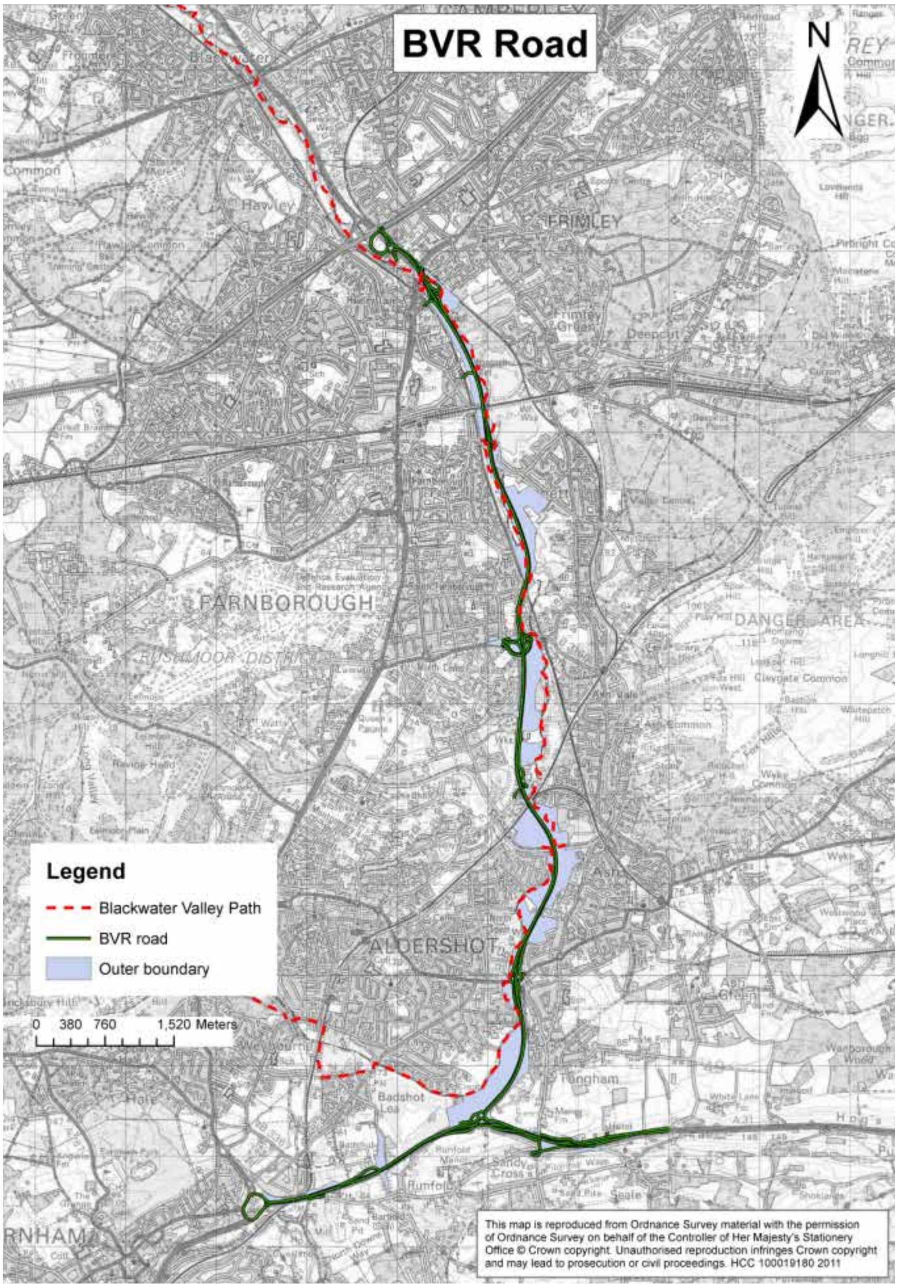


Figure 1. The Blackwater Valley Road (map provided by Hampshire County Council)

- habitat loss – a net loss of 38ha;
- habitat fragmentation reducing the effectiveness of the valley as a wildlife corridor;
- noise disturbance;
- potential pollution of the River Blackwater from surface water run-off;
- effects on legally protected species due to habitat loss and severance; and
- loss of recreational facilities.

In order to mitigate significant effects, the two County Councils purchased substantial areas of land alongside the scheme. This enabled a comprehensive package of green infrastructure measures to be designed and constructed. The overall package of ecological measures included:

- avoidance of existing sensitive areas wherever possible;

- temporary fencing to prevent damage to adjacent areas during construction;
- design of river diversions to improve riparian habitat;
- habitat creation resulting in an increase in water bodies and woodland of 90ha;
- habitat management, such as tree removal from grassland and swamp areas;
- translocation of heathland, aquatic and marginal vegetation and individual rare plants;
- natural regeneration of chalk grassland communities (**Figure 2**);
- capture and translocation of reptiles, amphibians and fish;
- design of drainage ponds to provide wildlife habitat;
- construction of a tunnel for roosting bats and erection of bat and bird boxes;
- measures to protect the water quality of the river; and
- provision of a public footpath, doubling the area of open access land and improving the quality of an angling lake.

Figure 2. A verge that was not sown with a seed mix and was allowed to regenerate naturally supports chalk grassland plant species. Photo by Tony Anderson, BVCP

The results of monitoring exercises undertaken in 2004 and 2011 indicated that the habitat creation schemes were largely successful, although some of the new habitats, such as woodland, will still take many years to be of equal quality to those lost. Wildfowl populations have largely benefited from the borrow pits, which provided new water bodies, and woodland bird populations use the extensive new tree belts. Translocation of aquatic plant species was successful, whereas few of the translocated grassland plants survived. Populations of legally protected species have been retained. The habitat changes brought about by the road scheme also benefited many species not targeted by the mitigation work, such as the wildfowl that have benefited from the new water bodies created from the borrow pits. However, some species have been adversely affected by the changes both directly and as a result of the changes to the overall balance of habitats within the valley; for example, attempts to translocate the common spotted orchid were unsuccessful.

Lessons for other schemes

The BVR was one of the earliest schemes subject to EIA in the UK. The methodology of EIA has advanced rapidly, and it was hoped that some of the lessons learnt in this study would help with the continuing evolution of EIA and ecological mitigation. Major issues and ideas identified in 2004 as a result of post-construction monitoring are set out below, with comments on whether they appear to have been taken on board more widely based on the hindsight available in 2011.

No.	2004 recommendation	2011 update
1	Provide sufficient area for the essential environmental works. In this study as much as four times the area of land required for the road was required for the mitigation measures.	There is now a recognition that plans for mitigation should be prepared early, so that land genuinely required for mitigation can be included in any Compulsory Purchase Order. Nevertheless, the area affected by habitat creation and management for the BVR remains unusually large, although this reflects the sensitivity of the location.
2	Prior to preparing a landscape plan assess what important habitats and species will be affected and in what quantity. The plan can then establish a correct balance of habitats to meet the needs of the target species.	EIA procedure ensures that baseline conditions are identified. Policy measures such as national and local biodiversity action plans help ecologists identify the important habitats and species with clear justification.
3	Avoid sensitive areas rather than relying on translocation. Translocation often fails and should be regarded as a valid method of salvage when loss is unavoidable, but essentially a means of enriching a newly created area and not a way of saving existing habitat.	The value of translocation as a salvage operation is increasingly recognised, and some improvements have been made in methods. The message that it is a method of last resort if in situ protection is not viable remains important.
4	Ensure works are correctly timed. Habitat creation work in advance of the habitat destruction during road construction allows time for establishment of habitat, and increases chances of success of translocated material.	Careful programme planning can ensure that time is allowed for mitigation measures. Commitments made to such programming can be set out in planning conditions or equivalents. Nevertheless, early consultation with ecologists is essential for project managers to take this into account.
5	Habitat creation and protection should concentrate on large scale and permanent features. Ephemeral features and small areas require intensive management to maintain in the long term.	Large scale permanent features are generally a better use of resources. However, there can be benefits in providing some ephemeral features, such as 'brownfield' bare ground, without entailing excessive costs if they are recognised as ephemeral and secondary vegetation is allowed to colonise naturally. As brownfield sites have been lost to development, even temporary provision of features that can be colonised by mobile species such as some specialist bare ground invertebrates contributes to their habitat (Box & Stanhope, 2004).
6	Address habitat fragmentation. If underpasses are not possible consideration can be given to installing 'green' bridges, specifically for wildlife; a reduction in road kills will also benefit road safety.	Understanding of successful design for wildlife underpasses, either purely for wildlife or combined with human access routes, has improved. There have also been a few examples in the UK of green bridges. However, the expense of green bridges means that they are rarely built, particularly if they cannot be incorporated into bridges required for people. For example, a green bridge was built over the Lamberhurst by-pass on the A21 in Kent, but this structure also provides the main access to the National Trust's Scotney Castle estate.
7	Do not over landscape. Current landscaping practice appears to be measured by the number of trees planted. This approach leads to tree planting on inappropriate areas, blocked footpaths and planting too dense to allow natural woodland flora to develop.	Landscaping understanding has improved, although tree planting schemes can still appear overly influenced by the desire to supress other plants at all costs and if dense planting is not followed by intense thinning, as would happen in planting for timber production, problems can arise. Establishment of scrub and woodland habitats through natural regeneration alone can be successful but is dependent on the distance and suitability of a suitable seed source and intensity of deer grazing. The authors feel that landscape schemes can also seem to sometimes focus on tree planting at the expense of good quality grassland creation.

No.	2004 recommendation	2011 update
8	Use native provenance vegetation. The long period of advance planning required for roads allows plenty of time to source and propagate all plant and seed requirements from local sources.	Use of native and, where justified, local provenance plants has increased. The problems that can arise if non-native stock is used are better understood. However, it can still be difficult to get funding for advance seed collection and propagation due to uncertainty of project approval and timescales. In addition, new questions will appear in the near future about the appropriateness of using plant stock more adapted to a drier and warmer climate than native plants. Such stock is already in use for some forestry schemes, and careful consideration needs to be given to the issues in ecological mitigation schemes. The authors' personal feelings are that such an approach underestimates the adaptability present in plants from the UK itself and this complex area needs more research. In 2013 the potential implications of 'native' stock being collected in the UK but grown overseas for cost savings are also being highlighted as the scale of such operations is viewed in the light of risk of spreading plant diseases.
9	Establish good working relationships between highway engineers and ecologists at an early stage. Communication between the different professions can be difficult but is essential to meet common goals. A dedicated Landscape Clerk of Works, involvement of local expertise, and a working group, are all ways of tackling the problem.	Use of an Ecological Clerk of Works is increasingly common on large schemes affecting sensitive habitats and species. Communication between disciplines at the design stage remains critical, and project managers should take control of this process and ensure that the specialists are talking to each other.
10	Monitoring should be put in place from the very beginning. Be prepared to alter designs and management to reflect monitoring results.	Some degree of monitoring is a common requirement prescribed in Environmental Statements. However, the feedback loop that allows this to influence designs and management and share results with others needs improvement. A particular challenge is provided where works are subject to licence due to effects on European Protected Species, as changes to the design would often require the submission and approval of a new method statement under the licence. This requirement could increase resistance to change.
11	Permitted development ancillary to construction needs to be strictly controlled with restoration conditions imposed. Site compounds, access roads, storage and processing areas etc can be highly damaging to the environment yet can be outside normal planning restrictions.	This recommendation is a reminder to ecologists, project managers and planning authorities to investigate requirements for ancillary development at an early stage.
12	Ensure continuing management so that beneficial impacts of mitigation measures are not lost. Mitigation measures should remain effective for the life time of the road. Funding to support mitigation measures should be an integral part of the long term highway maintenance budgets.	Funding in perpetuity cannot really be guaranteed. However, a commitment to a long term management plan helps ensure that this is taken into account in the highways maintenance programme.

Three key factors were identified as being instrumental in the success of the mitigation for this scheme:

- Ecologists worked closely with the highway engineers during design and construction of mitigation;
- The Blackwater Valley Countryside Partnership works closely with local authorities, private landowners and local community groups such as the Tongham Woodland Improvement Group (**Figure 3**) to manage the green infrastructure;
- Maintaining a flexible approach to management, based on monitoring of habitats and species, helped to direct and reshape mitigation measures, while continuing to focus on the original overall aims. For example, trees were removed from a number of plantations to allow naturally regenerating grassland flora to flourish.

The review of the Blackwater Valley Road recommendations shows that some aspects that were quite novel at the time have now become common practice, such as the use of an Ecological Clerk of Works. However, some lessons still need more consistent application. These include the need to mitigate impacts of ancillary permitted development and a tendency among some designers to plant more trees than necessary. Ecologists and land use planners are now experimenting with concepts of 'biodiversity offsetting'. Within such schemes, habitat translocation can add value to newly created habitat. However, this must not diminish the message that translocation is a method of last resort if in situ protection is not viable.

In summary, long before the phrase 'green infrastructure' came into common use, the Blackwater Valley Road scheme retained, created and managed 117 ha of land to provide multi-purpose benefits for people and wildlife. This green infrastructure was placed in local authority ownership and is being sympathetically managed to ensure that it provides green space for local people, habitat for wildlife and mitigation against the impacts of the road. At a time when budget constraints are likely to have an increasing influence on road and other major infrastructure schemes, it is essential that mitigation is as effective as possible. Green infrastructure with its multi-purpose approach to realising benefits is an excellent way of ensuring this. However, the exercise of monitoring and of reporting the results of mitigation is also important, to build a stronger evidence base to help demonstrate which techniques work and allow others to learn from experience.

Figure 3. Members of TWIG – Tongham Woodland Improvement Group – in their 'wood henge' seating area. Photo by Tony Anderson, BVCP

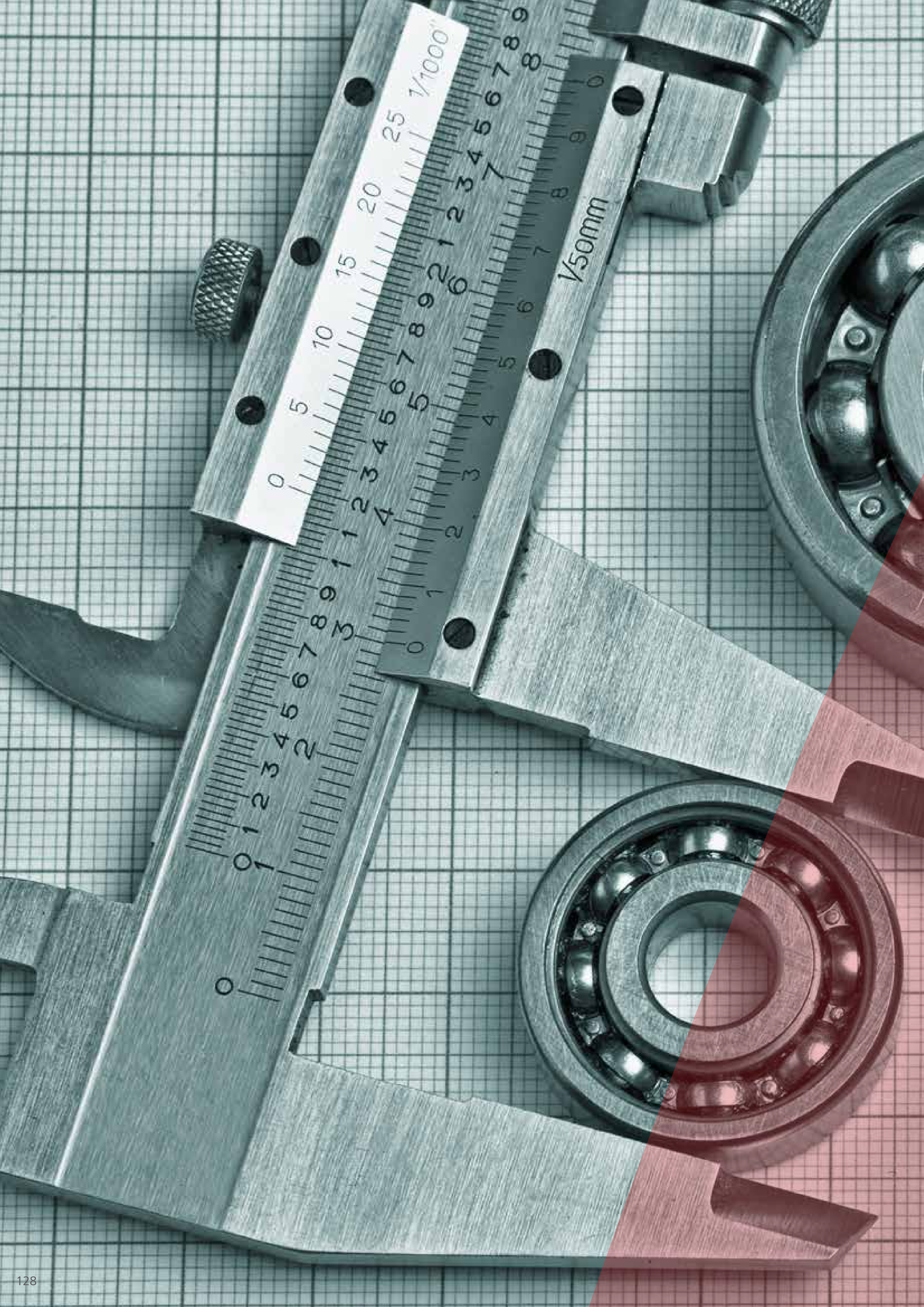
Acknowledgements

The original 2004 study was supported by Hampshire and Surrey County Councils and the Blackwater Valley Countryside Project. Sue Dent of Hampshire County Council and Jules Price of Atkins kindly provided comments on the draft article.

This paper was previously published in the Bulletin of the Chartered Institute of Ecology and Environmental Management, Issue 80, June 2013.

References

1. Atkins and Blackwater Valley Countryside Partnership (2004). The Blackwater Valley Road – Shaping the Landscape. Surrey County Counciland Hampshire County Council
2. Box, J. and Stanhope, K. (2004). Post-industrial and brownfield habitats. In Habitats Guide – volume 1, ed. H. Wilson, CIWEM
3. CIRIA (Dale, K., Thomson, C., Kelly, J., Hay, D., & MacDougall, K. 2011) Delivering biodiversity benefits through green infrastructure. CIRIA.



161

Technical Assurance

High-resolution technical assurance for projects with geographically distributed elements

Abstract

Atkins' success in winning the C222 (Country South preliminary civils) contract for the HS2 project brought a new level of complexity to the design and verification of railway infrastructure. With over three hundred separate elements of design, each inheriting requirements from the project specification, it was seen as necessary to develop a new solution to manage the requirements and their links to the design elements in order to provide an adequate level of technical assurance.

A database tool with diagramming capability was used to create a traceability structure that allowed us to display the traceability between the client's original requirements and our refined requirements, created to be specific to our scope of work. We used a layered model of the system architecture to apply requirements to the correct sets of individual design elements. From this connectivity we generated spreadsheet checklists for designers to perform verification, then imported the resulting data and processed it into a detailed report for submission to the client.

This solution is an advance in the capability of the Rail Solutions assurance team, and represents an opportunity to consolidate hitherto separated verification exercises for disciplines such as RAM, safety, human factors, EMC and environmental impact, as well as integrating the tasks of requirements management, interface and stakeholder management, verification and assurance.

Joe Silmon

Senior Systems Engineer

Rail Solutions

Atkins

Introduction

The HS2 project and Atkins

The UK government has made it their policy to construct a high-speed railway between London, Heathrow Airport, Birmingham, Leeds and Manchester, with interoperable services continuing along the West Coast Main Line to Liverpool and Glasgow[1].

Phase 1 of the project connects central London with the West Midlands. The post-consultation route, developed by Arup in the previous design stage, starts at a new extension to Euston Station and exits London to the west, before turning north towards the Chilterns and, ultimately, the West Midlands.

Phase 1 is currently in the preliminary design phase, and Atkins has won two contracts within this phase. Contract C222 comprises the design of earthworks, rail alignment and structures on the stretch of line between the edge of London and the edge of Warwickshire, known as "Country South". Atkins is delivering C222 using designers from across the business, based at sites including Birmingham, Croydon, Epsom and Euston.

Items such as signalling, trackform, tunnel evacuation and electrification systems are designed route-wide by other contractors.

Atkins' Water & Environment business has also won Contract C253 to perform an environmental impact assessment for the Country North section of the route, but that is not within the scope of this paper.

C222 traverses the picturesque Chiltern hills of Buckinghamshire, making it particularly sensitive to the needs of local residents, wildlife and the environment. Under these circumstances, a robust portfolio of technical assurance information is essential.

Requirement ID	REQ000-00000
Source document Requirement text:	DOC000 Example Requirements Spec Spatial provision shall be made for the assembly of replacement switches adjacent to crossovers.
Rationale	In-situ assembly allows switch replacements to be carried out in a shorter timeframe thus reducing disruption to the railway.
Compliance status **Verification argument** **(free text justification of compliance status)**	Compliant The general arrangement drawings show that extra space has been provided at crossovers 1A/B and 2A/B to either side of the alignment. The design is described in section 'The difference between direct and rich traceability'.
Verification evidence **(controlled document/drawing numbers or EDMS links)**	GA drawings: Sheet 1: C222-ATK-....... Sheet 2: C222-ATK-.... Design statement: C222-ATK-...

Table 1. Verification of a single element against a single requirement

Rail Solutions' technical assurance approach

For the purposes of this paper, we define "technical assurance" as the provision of evidence to the client that their requirements have been satisfied.

In past projects, such as the Thameslink and London Underground station at Farringdon, Crossrail North East Stations (Stratford-Shenfield) and Crossrail Victoria Dock Portal, technical assurance has been provided through the use of a "verification matrix" generated as a report from DOORS (a widely-used requirements management database tool).

Table 1 shows an example of the traditional approach to verification of a single requirement. Note that this would normally be displayed transposed (i.e. with a separate row in the table for each requirement and the data fields as columns). The verification data fields are highlighted in bold text.

Customer requirements are usually fixed (from the point of view of the customer) by the time Atkins starts work on a project. It is usual for the client to hold their requirements centrally and for Rail Solutions to provide verification information that traces directly to these originating requirements, without performing detailed analysis or refinement. Because most railway project requirements (at the level of Atkins' involvement) are constraints or citations of standards, rather than functional requirements, this approach serves adequately for small and geographically concentrated projects.

Where requirements contain ambiguities, conflicts, omissions or apply more widely than the project scope, an interpretation is usually proposed as part of the verification argument and either agreed in advance with the client or submitted along with the verification information.

Generally speaking, most of the projects that have been assured in this way are located in small geographical areas, commonly stations, and therefore a requirement takes a verification argument that adequately describes the whole design in one data field. Table- or matrix-based verification is perfectly adequate in such situations, although interpretation of the requirement is usually necessary. The limitations of this approach have become clear on the Crossrail North East Stations project, where individual evidence citations for up to thirteen separate stations needed to be inserted in one data field in the database.

A new level of technical assurance for a new type of project

A high-speed railway design is a very different kind of task from the design of a station or other works in a small geographical area. System-wide requirements apply, in reality, to each element that is distributed along the line of route. A traditional verification matrix would aim to provide sufficient evidence for that requirement's fulfilment using only one text field in a DOORS database or spreadsheet. However, for HS2, that field must then cover assurance of each of the 330 or more separate elements of the design. **Figure 1** illustrates how one requirement may need to be verified in multiple parts of the design. This theme is developed further in the section on requirements management methodology.

Anticipating a high level of scrutiny of the technical assurance information, it was clear that our traditional approach would not be sufficient even to provide ourselves with adequate assurance internally that we had fulfilled the requirements, let alone fulfilling the needs of a Government body and a public consultation/hybrid bill process.

To address this anticipated need for high-resolution assurance, an approach was developed that would ultimately produce a requirements listing for each design element, so that evidence could be developed and cited on an element-by-element basis, ensuring that for any point on the route, full, accurate and relevant evidence could be called up from the verification report without the need for interpretation or further querying.

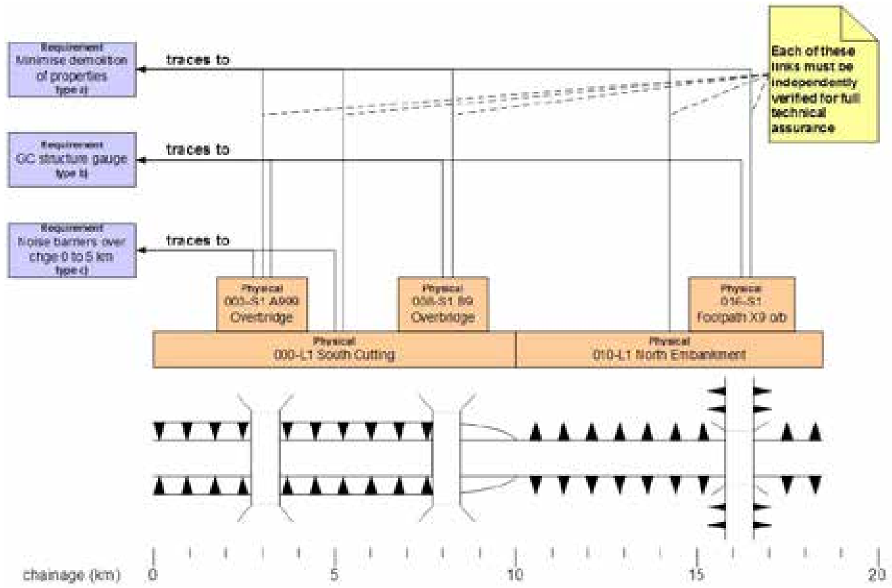


Figure 1. Applicability of requirements shown in a geographical context

Outline of this paper

This paper describes the nature of the requirements management task for the project and the structure of the railway system under design in section labelled 'Characterisation of the HS2 design and requirements set'. The principles of rich traceability are outlined in section on rich traceability, and their application in the project methodology is described in the section on requirements management methodology. A summary of the tools and automation scripts used is presented in the tools and automation section, followed by the interim outcomes of the work and the intentions for future development in a section on outcomes.

Characterisation of the HS2 design and requirements set

The Country South route

The Country South section of HS2, the part that Atkins is designing, runs for some 90 km between the edge of London and the edge of Warwickshire.

The design consists of over 300 separate elements of design (with around thirty different types of element), each assured via two routes: a design element statement document, and the verification report against the project requirements.

Certain types of element are repeated many times at distributed locations along the route, meaning that each may be subject to unique sets of local constraints (such as ecological restrictions). **Table 2** shows the top five element types by frequency (as at Configuration Control Point 4, 23 November 2012).

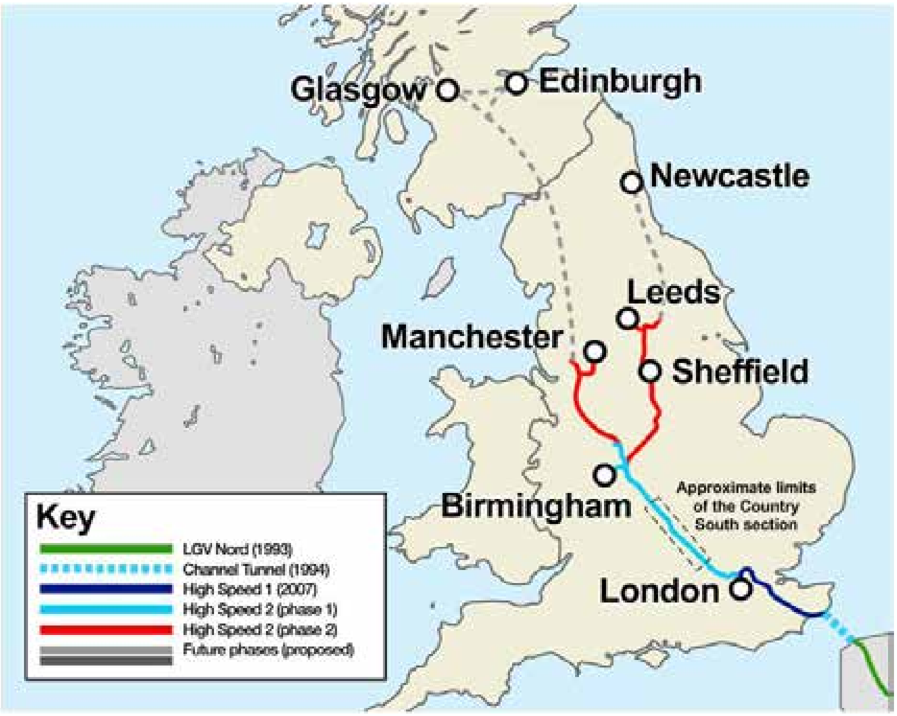


Figure 2. Proposed HS2 routes with the area of the Atkins scope labelled[6]

Element type	Frequency
Public right of way	72
Embankment	48
Overbridge	45
Watercourses and culverts	35
Cutting	29

Table 2. Frequencies of the top 5 most common types of design element in the Country South area of HS2

The challenge for the technical assurance effort on this contract is to ensure that for each one of these elements, adequate evidence is in place to demonstrate fulfilment of the client's requirements. Given the anticipated high level of public scrutiny of the design, it was felt necessary to ensure that each element has its own separate evidence against each requirement.

This challenge is balanced by the opportunities presented by a green field site, where the technical choices are not very heavily constrained by existing solutions, for example spatial and alignment constraints, and interoperability with legacy systems dating back to the mid-1800s.

HS2 requirements set

Project requirements specification

The HS2 project requirements specification is held in a central database and in order to perform requirements engineering tasks, the Country South team took an export by means of a spreadsheet and imported it to the system architecture model.

Deliverable approach statements

During the initial stages of the project, discipline teams from all contractors met to discuss their design approaches and develop informal standards called deliverable approach statements (DASs) to guide the interpretation of the project requirements. These will form part of the second layer of refining requirements but have not yet been instructed for use and therefore must be regarded as preferences. Applicable national and international standards are addressed by the statements, so for the most part, compliance with the requirements in the DASs implies compliance with relevant standards, making it possible to satisfy multiple standards through a single set of requirements.

Rich traceability

The term "rich traceability" refers to the principle that a set of originating requirements should be progressively refined during a project's lifecycle, with explanation of the reasoning at each stage[2].

For the purposes of this paper, we define an originating requirement as one that is fixed by some external source; this is usually the client but may also be a standards body or external stakeholder. A derived requirement is one that the design team develops in order to specify the design more directly and thereby fulfil the originating requirement.

The difference between direct and rich traceability

Originating requirements are supported by derived requirements that reflect more accurately the shape of the design, given developments during the design process, new information, or the results of formal requirements analysis. The relationships between requirements provide **direct** traceability between a derived requirement and an originating requirement.

Each stage of refinement is accompanied by a "satisfaction argument" that states why the derived requirement is **necessary** to support the originating requirement, and how the derived requirement is **sufficient** to satisfy the originating requirement. Rich traceability is the information contained in the relational path from a derived requirement, via the satisfaction argument, to the originating requirement.

The set of requirements that is refined to the highest degree of detail is called the set of "leaf-level" requirements, analogous to the leaves on a tree being the outer, visible layer of foliage. In contrast to a functional decomposition, where each layer of the decomposition is relevant in order to get a full picture of the system's behaviour, the leaf-level requirements are a complete representation of the system requirements, without reference to the requirements from which they are derived.

The rich traceability approach has a number of significant advantages when compared to viewing the requirements as a flat set:

- Designers need only deal with the leaf-level requirements, which are either derived or validated by a requirements engineer, thereby guaranteeing that they will be well-formed
- Each layer of derived requirements can be reviewed with the client ahead of design submission, providing assurance progressively and allowing for feedback
- The details of the design can be robustly traced back to the client's requirements without significant effort
- The accompanying rigorous requirements analysis effort reduces project risks by uncovering deficiencies, conflicts and ambiguities in the requirements
- Multiple originating requirements can be demonstrated to be satisfied by a single requirement at leaf level, using the satisfaction argument; this allows multiple aspects such as safety, EMC and environmental impact to be incorporated in the requirements set without multiplying the effort needed to close out all requirements.

Rich traceability has not traditionally been applied on Rail's infrastructure projects (although it has been used by Rail Solutions on London Underground projects such as the Victoria Line Upgrade, Sub-Surface upgrade and the Deep Tube project). However, due to its complexity and criticality, the HS2 project both needs and provides the opportunity to employ this methodology.

The "leaf-level" requirements set

Developing a rich traceability structure for requirements will result in a set of requirements that is detailed and relevant to the design. The designers can then be provided with the most relevant layer of requirements to work into the design, and do not need to address how those requirements trace to the Client's requirements, because that structure has been interpreted, established and approved beforehand.

Figure 3 illustrates a simple requirements hierarchy, after analysis has taken place. Note that to establish the leaf-level set, it is only necessary to use the **direct** traceability relationships. Rich traceability tells us **why** the direct traceability exists.

Traceability may decompose requirements from the system to the sub-system or component level, but there are other possible reasons for a requirement to be refined, for example:

- As the resolution of a conflict between requirements
- As a detailed refinement of an originating requirement that was ambiguous or uncertain
- As a correction of a compound originating requirement.

In **Figure 3**, originating requirements 1 and 3 have been refined into derived requirements 4, 5 and 6. Originating requirement 2 has not been refined, as no need for refinement has been identified. The leaf-level requirements set therefore consists of derived requirements 4, 5 and 6, and originating requirement 2.

There is no reason for the leaf-level requirements to be at a homogeneous level. Requirements are only refined as far as necessary to ensure they are well-formed. Where originating requirements are sufficiently specific and well-formed without refinement, they form part of the leaf-level set without further interpretation.

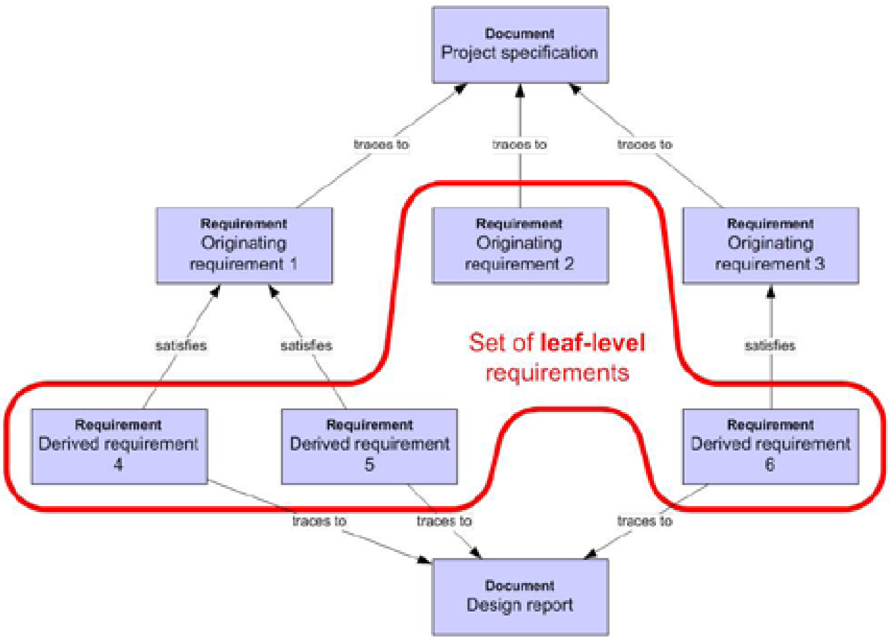


Figure 3. A requirements traceability diagram, illustrating direct traceability only and the leaf-level requirements

Technical Assurance

The derived requirements are documented in design reports in order to report back to the client on the reasoning behind the derivation.

Requirements management methodology

Processes for managing requirements

The approach for managing requirements was recorded in our requirements management plan for the project[3]. It is illustrated in flowchart form in **Figure 4**, and consists of several stages that will be described in the following subsections.

Note that several of the stages of the process are either automated or semi-automated, using scripting in the Enterprise Architect database. Further details of how this works can be found in the section on tools and automation.

Capturing requirements

Originating requirements were supplied by HS2 in varying types of format. A procedure was established to capture requirements in a consistent style as singular items in a spreadsheet template.

Once requirements were itemised in Excel spreadsheets, a script in Enterprise Architect read through the rows in the spreadsheet and automatically created all the requirement elements, populating custom data fields (known as "tagged values") to store the following attributes of the requirement:

- Requirement text
- Requirement status (proposed, accepted, superseded, out of scope)
- Requirement rationale.

A numbering convention was established, with three decimal

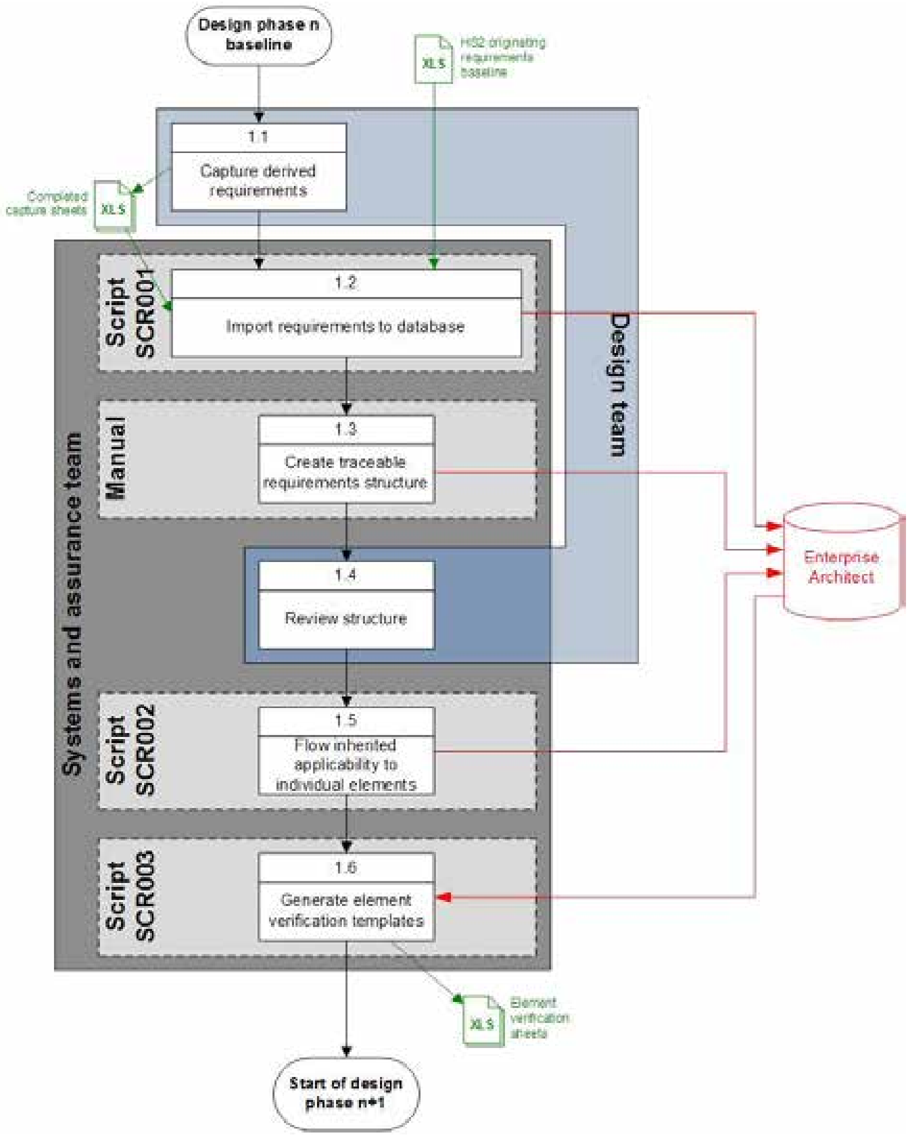


Figure 4. Procedure for capture, allocation and communication of requirements

digits representing a code for the identity of the source document (making it possible to identify the source document of a requirement from its identification number) and five digits as a serial number for the requirement, unique within the source document requirements set. An example would be "REQ001-00001" where the document code is 001 and the serial number is 00001.

Design engineers were instructed to carry out a requirements capture procedure and populate a standard spreadsheet with any requirements that may arise during meetings. This spreadsheet template could then be directly operated on by the script in Enterprise Architect.

Create traceability structure

Each set of requirements that was captured was initially stored in the database as a flat and untraced set of elements. These were then linked through relationships to the rest of the requirements set and to the elements or element types to which they apply.

This was carried out by the requirements engineer in the system architecture model, using a special viewpoint that extends the TRAK MV-03 Requirements viewpoint[4] through the addition of rich traceability elements as described in section 4.2.

TRAK is an open-source architecture framework defining a grammar of elements, relationships and viewpoints for describing enterprise architecture.

In order to apply a requirement only to the right set of design elements, requirements were classified as belonging to one of the following three sets:

a. Requirements that apply to all elements in the scope ("system-wide" requirements)
e.g. "The design shall minimise the need to demolish residential property"

b. Requirements that apply to all elements that are of a particular type or set of types ("type-wide" requirements)
e.g. "The structure gauge shall be GC" - applies to tunnels and overbridges

c. Requirements that apply to one or more individual elements ("location-specific" requirements)
e.g. "Noise mitigation measures shall be implemented between chainages x and y" - applies to all elements in that chainage range.

Figure 5 illustrates the initial traces that are created manually in the architecture model for each of the three types of requirement. Traces are only created from **leaf-level** requirements to design elements, because the leaf-level requirements are sufficient (as recorded in the satisfaction argument) to specify the design fully. To determine which of the client's requirements are ultimately responsible for a particular leaf-level requirement, the traceability hierarchy must be examined.

Allocating requirements to individual elements

The system model is used not only for requirements analysis but also to develop and maintain a model of the system architecture.

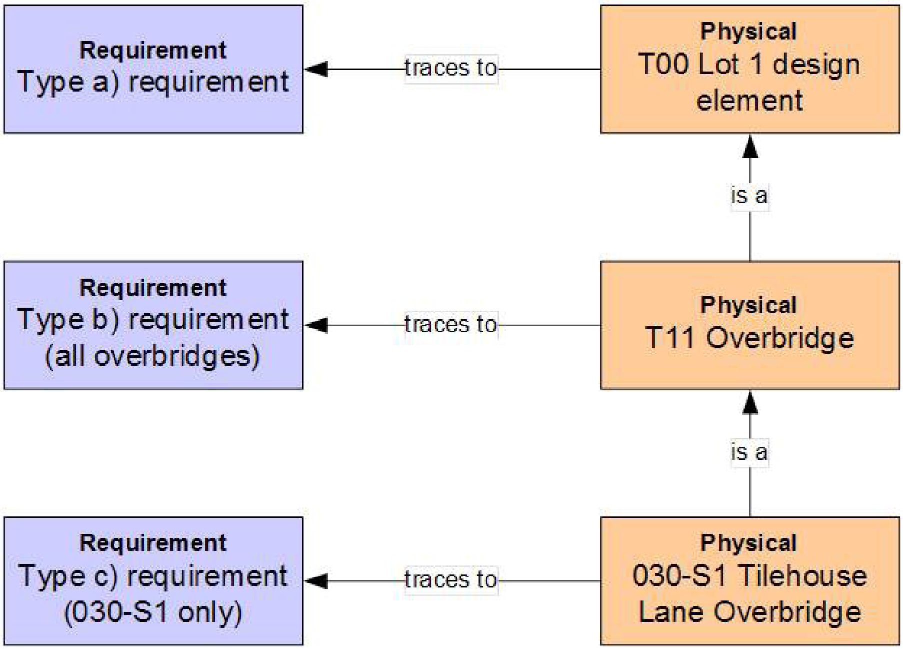


Figure 5. Manually-created traces for requirements of type a), b) and c)

A taxonomy of elements, using the "is a" relationship, was developed to represent the types of element in the system, making it possible to allocate requirements of types a, b and c (see **Figure 5**) to the appropriate class for their applicability to be inherited by the correct individual elements.

This is illustrated in **Figure 6** and **Figure 7**, which show how system-level and subsystem-level requirements were allocated to individual elements by inferred inheritance through the taxonomy. Class c) requirements are traced directly to the applicable elements during the process described in the section on creating traceability structure.

An alternative method for creating the individual elements would have been to use instances, in Unified Modelling Language (UML), of the original classes. However, an instance can only ever be of one class, whereas in the model we have developed, an individual element may need to inherit the properties of more than one element type. They therefore need to be classes in their own right with the possibility to have multiple generalisations.

Inferred traces in the model are created automatically by a script that examines the generalisation ("is a") relationships between the individual elements and the generic elements in the system architecture model. The script then implements an inheritance of requirements from the base classes to which each individual element belongs.

Following successful execution of this script, the model contains all the links between leaf-level requirements and each individual element of design.

Creating element requirements specifications

Each individual element now has a tailored requirements set, showing only what is applicable to that particular element. A script automatically identifies all the requirements that apply to each element, and creates a listing as an Excel spreadsheet, including blank fields for verification data to be entered by designers. These spreadsheet files are subsequently reprocessed to import the verification data back to the database for use in generating the verification report.

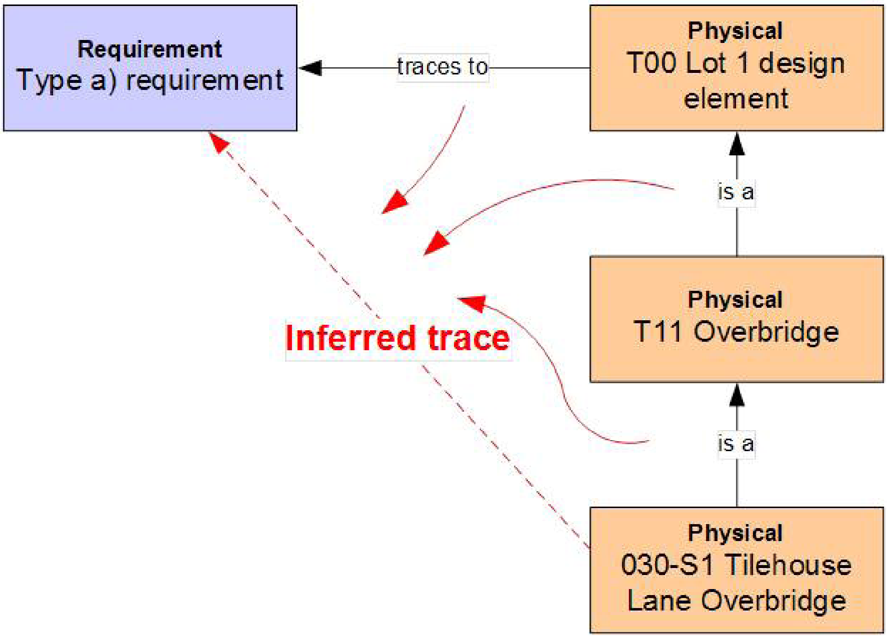


Figure 6. Inference of requirement applicability for type a) requirements

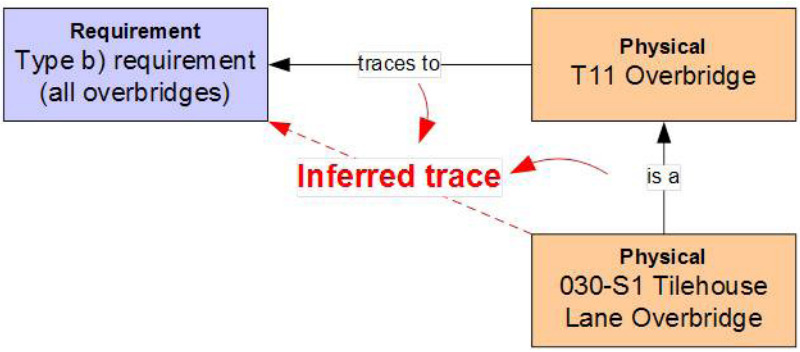


Figure 7. Inference of requirement applicability for type b) requirements

Rich traceability metamodel

In order to enter traceability information consistently in the database, a set of rules was created to govern the data entry. The database tool is diagram-based, allowing the traceability structure to be entered through individual views on the database. A new metamodel, compatible with the TRAK metamodel (as an optional extension), has been developed in this project in order to illustrate rich requirements traceability. This metamodel is illustrated in **Figure 8.**

Existing TRAK elements have been linked with new relationships; the UML elements "Issue" and "Change" have been used to store the satisfaction argument ("Issue" and "Change" were chosen arbitrarily for ease of scripting; there is no significance to the name of these elements, the key in terms of rich traceability is that the Issue element holds the argument of necessity, and the Change element holds the argument of sufficiency).

Table 3 lists and describes the relationships used in the requirements management metamodel.

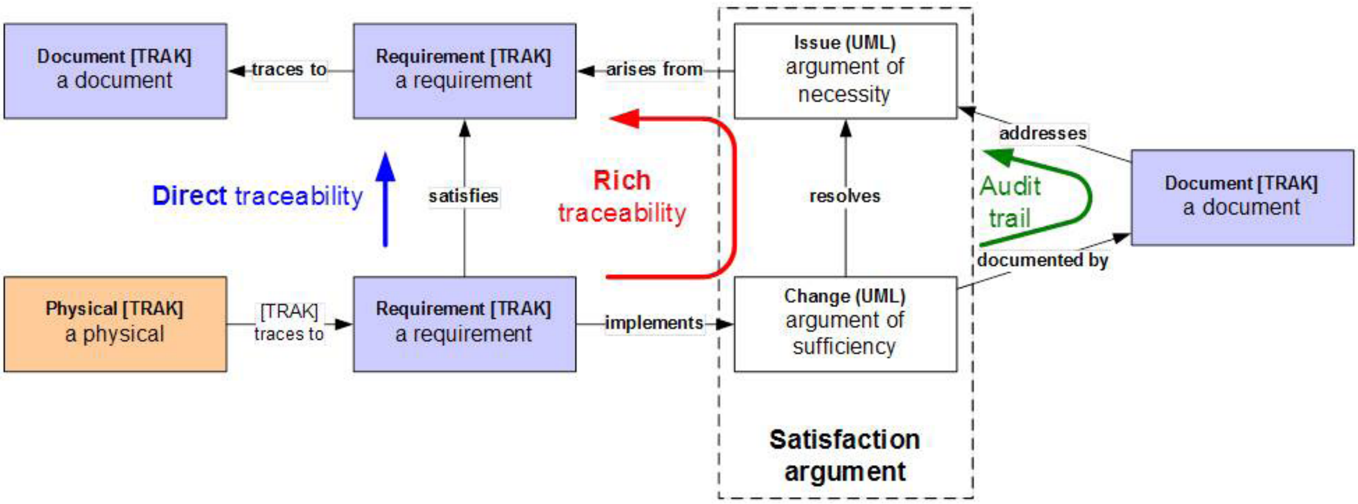


Figure 8. The metamodel for rich traceability, showing its use of standard TRAK elements

Subject	Relationship	Object	Purpose
Physical	traces to	Requirement	Existing TRAK relationship allowing us to show that a Physical (or other resource) has requirements that apply to it.
Requirement	traces to	Document	New relationship that provides traceability from a single Requirement back to the document from which it originates. Necessary to show why a requirement applies. Examples of suitable Documents would be specifications, employer's instructions, meeting minutes or interface control documents.
Requirement	satisfies	Requirement	New relationship illustrating direct traceability. Used to illustrate the development and refinement of requirements in a simple hierarchy.
Issue	arises from	Requirement	New relationship illustrating part of rich traceability. The Issue element contains the argument of necessity as to why the Requirement must be refined.
Change	resolves	Issue	New relationship illustrating part of rich traceability. The Change element contains the argument of sufficiency as to how the refining Requirements address the Issue.
Requirement	implements	Change	New relationship illustrating part of rich traceability. The Requirement element is one of the new Requirements created, as proposed in the Change, to be sufficient to address the necessity expressed in the Issue traced from the Change.
Change	documented by	Document	New optional relationship for audit trail/information purposes. The Document is any documentary evidence of the process followed to develop the new requirements e.g. meeting minutes where the initial requirement was discussed, requirements validation reports.
Document	addresses	Issue	New optional relationship for audit trail/information purposes. The Document records the discussion arising from the identification of the Issue.

Table 3. Elements and relationships in the HS2 requirements management metamodel

The rest of the system architecture model in the database uses the TRAK architecture framework[5], which provides a structured vocabulary for use in modelling systems.

Tools and automation

Understanding the requirements for tools

At the outset of the project, two database tools were available for use by the team. These are DOORS, by IBM, and Enterprise Architect (EA), by Sparx Systems. After careful consideration, Enterprise Architect was chosen as the platform for a complete systems architecture and requirements management solution.

The reason for this choice is that DOORS lacks some of the capabilities required by this project. **Table 4** gives details of the comparison made between the two tools at the time of planning for the project's execution.

A brief description of Enterprise Architect

Sparx Enterprise Architect (EA) is a UML modelling tool intended for software design. It is a lightweight and cheap tool that can run with standalone files or use a database repository.

The user interface is primarily diagram-based, with a model browser tree to one side of the display. All elements of the model are stored in database tables, ensuring that an element's manifestations in several views are consistent. A typical screenshot is shown in **Figure 9**.

EA was chosen by London Underground as their preferred platform for developing an implementation of their architecture framework TRAK. This architecture framework is ideal for modelling the architecture of the system under consideration, and is provided as an open-source plugin to EA.

TRAK includes provision to model the composition of elements and assemblies, as well as bi-directional interfaces and interactions between system elements. Thus, a complete interface model can be created and linked to the requirements that govern each interaction. As a tool becomes more about connectivity as well as storing volumes of data, graphical interfaces become more critical.

Enterprise Architect allows for reporting and automated operation using a Javascript engine that is capable of calling up ActiveX constructs. This allows the engine within EA to call functionality outside

Requirement	DOORS evaluation	Sparx EA evaluation
The tool shall be capable of exchanging data between requirements elements and Excel spreadsheets [Rationale: in order to exchange information with design teams who do not work in this tool]	Function included as standard	Function not included as standard but achievable using Javascript and ActiveX functionality [a team member with some Javascript experience was available]
The tool shall be capable of graphically representing relationships between elements, including requirements traceability [Rationale: to enable graphical explanation/justification for rich traceability, enabling more efficient assurance]	Very limited graphical functionality	UML diagrams formatted as TRAK viewpoints (or their extensions) provide this capability [a team member with extensive TRAK experience was available]
The tool shall be capable of generating reports and statistics on requirements verification [Rationale: to semi-automate the production of assurance deliverables]	Possible through DXL scripting [no DXL-trained team members were available]	Possible through Javascript and ActiveX functionality
The tool shall be capable of automatically generating inferred requirements traces, using the relationships in the system architecture model [Rationale: to semi-automate the production of requirements checklists for each design element]	Limited (uni-directional) traceability relationships are possible System architecture cannot be modelled to a sufficient level in DOORS	System architecture can be modelled through TRAK Solution Perspective viewpoints SVp-01 and SVp-02, with a supporting taxonomy view Inference and generation of traces possible through Javascript and ActiveX functionality

Table 4. Summary of the evaluation of DOORS and Sparx Enterprise Architect against the requirements for tool capabilities on he HS2 C222 project

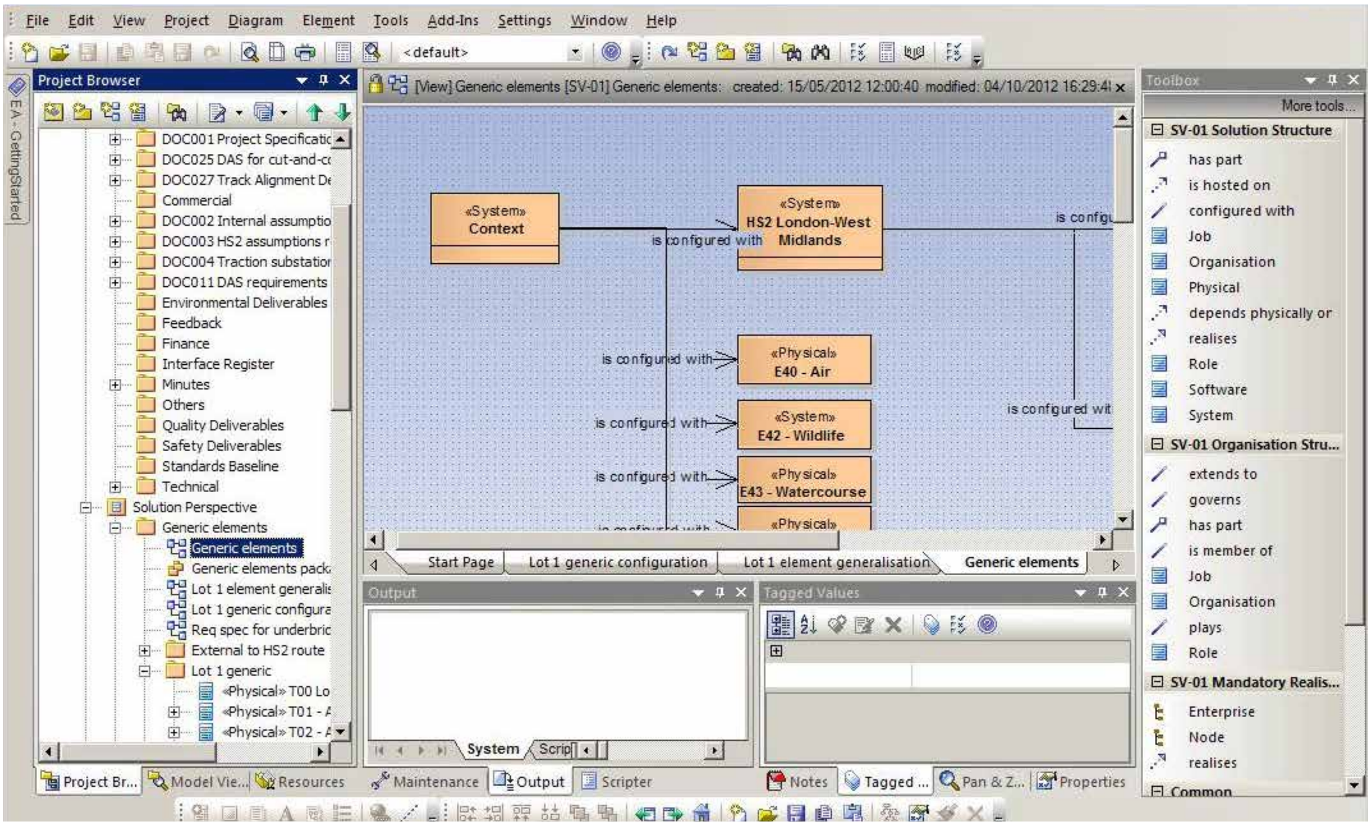


Figure 9. typical Enterprise Architect screenshot

the program, such as opening Microsoft Excel, creating and populating new files, or reading in data from chosen directories on the project shared network drive.

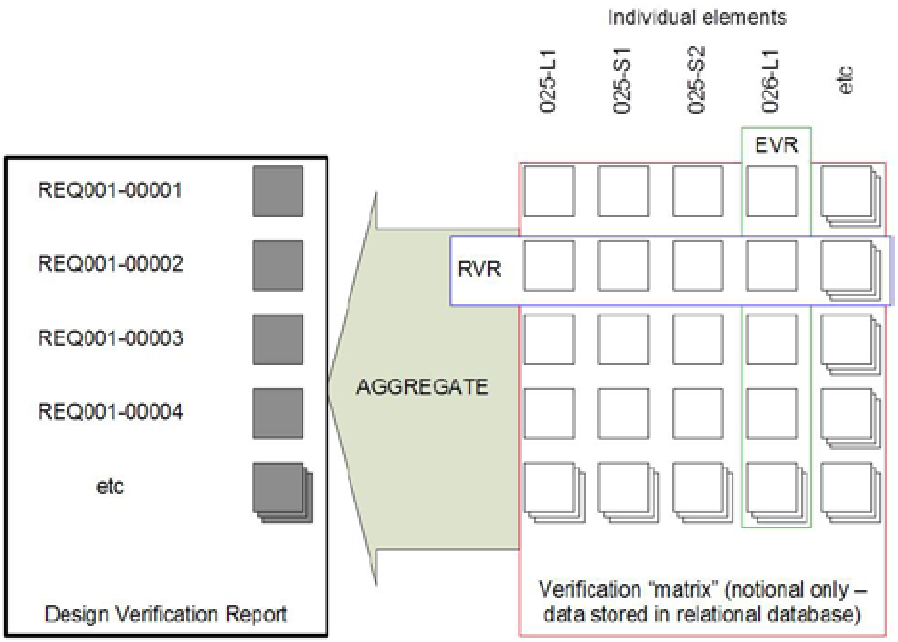


Figure 10. Structure of verification reports

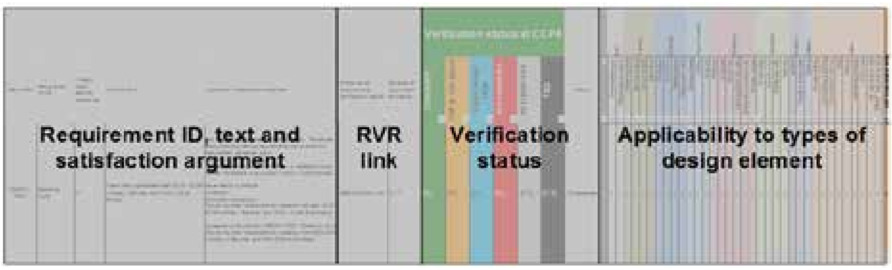


Figure 11. Screenshot of the aggregate Design Verification Report

Outcomes

Reporting of requirements verification

By creating custom reporting scripts to run through the system architecture model, it was possible to compile compliance statistics for requirements compliance at the first design submission milestone.

A traditional compliance matrix would have as many rows as requirements, and as many columns as design elements; in this case it would have been 125 x 340, a size that would make it impractical to navigate. The automatically-generated design verification report instead listed, for each originating requirement, the numbers and percentages of applicable design elements compliant, non-compliant and so on.

Individual reports were also created for each element (answering the concern "does this element comply with all its requirements?") and for each requirement (answering the concern "is this requirement fulfilled throughout the design?"). These are known respectively as element verification reports (EVR) and requirement verification reports (RVR).

Future development

The new solution developed during the C222 project is in its first iteration of use. There is a small number of bugs in the automation scripts which are outstanding to be addressed before the next verification exercise, but the solution is, overall, capable of producing accurate, element-by-element verification statistics against the client's requirements.

Although the Sparx EA tool is extremely cheap, flexible and user-friendly in terms of its graphical interface, it suffers some limitations, such as the restriction of UML constructs and the underlying object model.

In 2014, the team will be evaluating new toolset options, including fully-featured model-based systems engineering tools, which would have many of the automation features as standard that we were required to create ourselves using scripts.

Conclusions

Over the first nine months of the project, we have developed a novel requirements analysis and allocation solution, using the flexibility of a UML modelling tool coupled with the principles of rich traceability and system architecture modelling.

We have demonstrated the use of this solution on one of the most sensitive and demanding infrastructure projects Atkins has worked on. Whilst work is ongoing to continually improve the solution, the principle of element-by-element requirements allocation and verification has been shown to be possible using only a small number of staff and automation to carry out the bulk of the linking and report generation.

This is a step towards a stronger requirements and systems engineering capability, providing a means to capture information at a more detailed level to provide strong assurance cases on a project with a distributed set of design elements.

Acknowledgment

The author would like to express his sincere appreciation to David Nicholson (Principal Engineer, Rail Communication Systems) and to Mike Brownsword (Systems Engineering Consultant, Rail Solutions) for their valuable guidance and constructive reviews of this paper.

References

1. "Making HS2 Happen"; HM Department for Transport, in Transport Times Conference (speech by Rt Hon Justine Greening, Secretary of State for Transport). Pub. HM DfT 26th January 2012, URL https://www.gov.uk/government/speeches/making-hs2-happen
2. "Rich traceability"; Jeremy Dick, in IEEE Software, p.2005, 2002. DOI 10.1.1.114.6487, URL http://www.soi.city.ac.uk/~zisman/WSTPapers/Dick.pdf [retrieved 30/01/2013]
3. "C222 Requirements Management Plan"; Joe Silmon & Alun Jones, HS2 document reference C222-ATK-SA-PLN-020-000005 rev. P02. [not published]
4. "TRAK Viewpoints"; HM Department for Transport, published online at http://trakviewpoints.sourceforge.net [accessed 30/01/2013]
5. "TRAK Metamodel"; HM Department for Transport, published online at http://trakmetamodel.sourceforge.net [accessed 15/02/2013]
6. Graphics modified from original "UK High Speed rail map" by Wikimedia Commons user "cnbrb", distributed under CCSA3.0 licence.

Appendix: TRAK in brief

TRAK is an open-source enterprise architecture framework developed initially for rail applications by London Underground. It is loosely based on the DoDAF/MoDAF pedigree of architecture frameworks, but aims to be smaller, simpler and more pragmatic.

TRAK uses perspectives to subdivide the set of possible viewpoints, according to the level of view required.

The enterprise perspective deals with the structure, goals and capabilities of an enterprise, including transitions from one temporal phase to another.

The concept perspective models the problem space in terms of entities (nodes), needs and conceptual activities (similar to use cases).

The solution perspective models the structure and interactions of a particular solution, including elements for human resources.

The management perspective contains elements and viewpoints for model management, including requirements. The management perspective has been extended by the work in this paper.

The procurement perspective allows for the modelling of project phases and delivery as related to system and capability transitions.

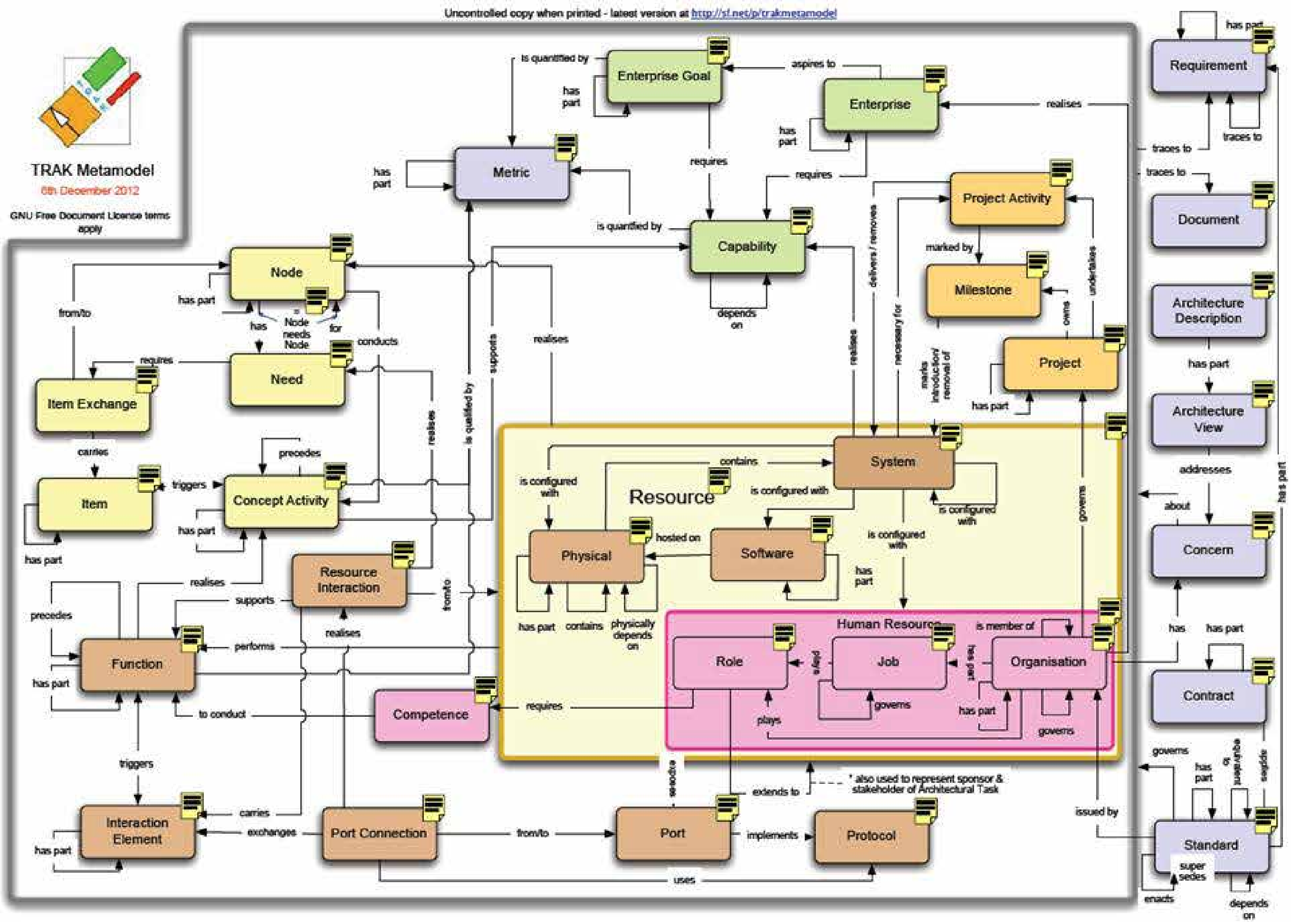
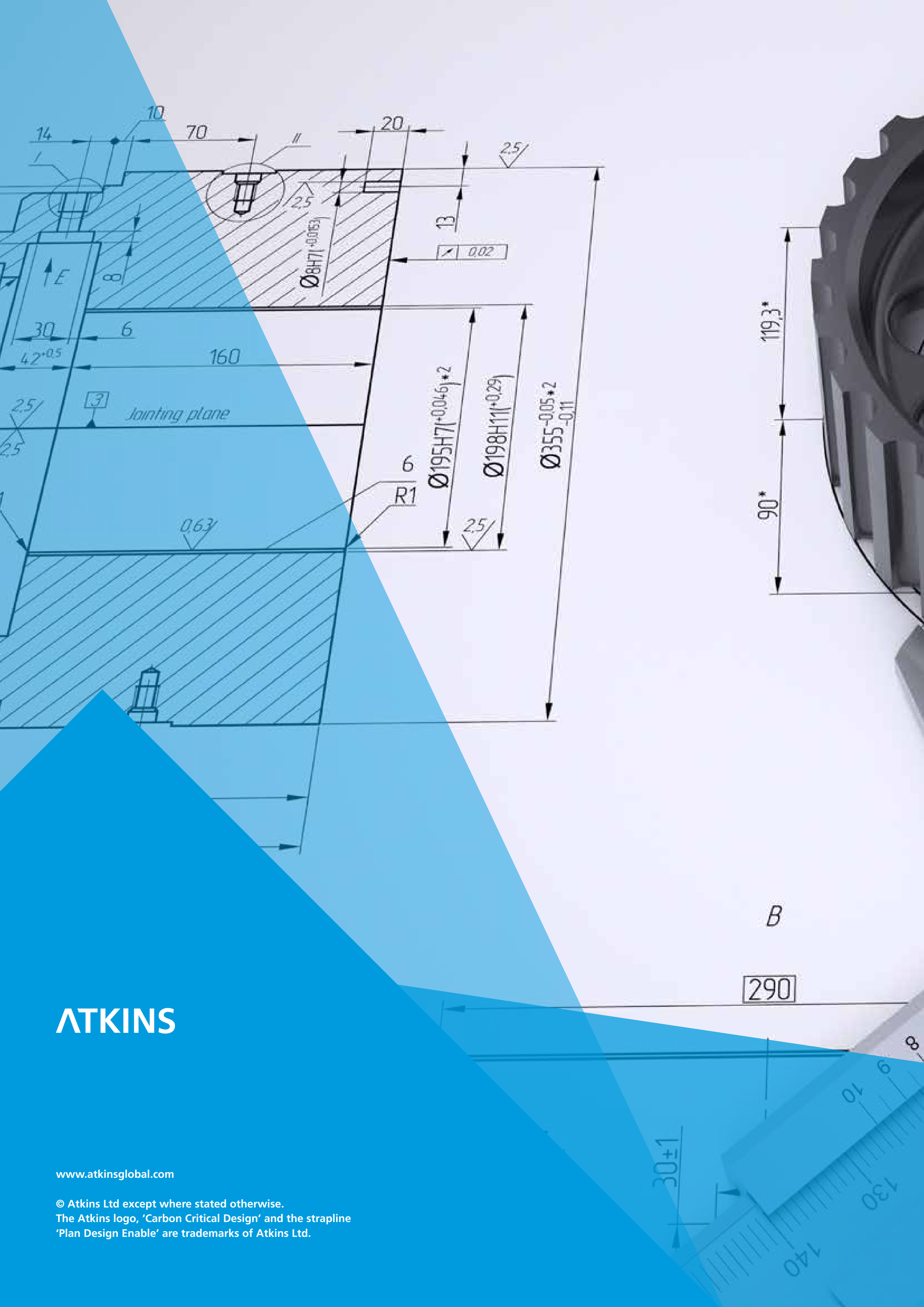


Figure 12. TRAK metamodel in full (Reproduced as permitted by GNU Free Documentation Licence - these terms apply to any future reproduction)



ATKINS

www.atkinsglobal.com