A Report to
Network Rail and the ORR
from
Asset Management
Consulting Limited

Version: 1.0
22nd May 2008

Independent Reporter Review of Statistical Sampling Used In CECASE

Controlled Copy No:	UN

APPROVAL				
Version	**Date**	**Compiled by**	**Reviewed by**	**Authorised by**
1.0	22nd May 2008	M A Sultan	C Johnson	R J Edwards
Asset Management Consulting Limited File Ref: 007.180				

DISTRIBUTION			
Name	**Controlled Doc. No.**	**From (version)**	**To (version)**
1 copy of document each to:			
Mervyn Carter	1	Draft C	C
Nick Stuart	2	Draft C	C
Andrew Newby	3	Draft C	C
Alan Dray	4	Draft C	C
Uncontrolled copies distributed as required			
C = Current version			

AMENDMENT HISTORY		
Version	**Sections**	**Amendment Details**
Draft A	All	Initial draft for comment
Draft B	All	Revised following meeting with Network Rail
Draft C	All	Revised following Network Rail's review
1.0	All	Revised following ORR's comments

Original held by Asset Management Consulting Limited at
221 St. John Street, Clerkenwell, London, EC1V 4LY
Tel +44 (0) 20 7688 2828 Fax +44 (0) 20 7688 2829

Executive Summary

Network Rail and the Office of Rail Regulation (ORR) have commissioned AMCL, through its role as Independent Reporter, to conduct a review of Network Rail's Civil Engineering Cost and Strategy Evaluation (CECASE) decision support tool which evolved from the 'Structures Annual Cost Profile' (SACP) work. The remit given to AMCL focussed on five key issues:

- The sufficiency of the sample size in being able to accurately model the major asset types and the various forms of construction;
- The sufficiency of the sample size in accurately modelling the distribution of structures across different route types;
- The degree of accuracy generated by the modelling and the likely tolerance of these;
- How sensitive the model is to sample size, asset selection and policy application; and
- Whether or not there is any evidence of Network Rail adopting an iterative learning process with the sampling.

AMCL has reviewed the CECASE modelling process which is made up of four key components, namely the prior model, the Markov Chain model (which considers intervention points and timing, etc), the posterior or Updated model and the Policy Planning Tool. Overall, AMCL found that CECASE is a robust decision support tool when modelling long lived Civils structures. Its strengths include the ability to learn from various sources of information, the ability to refine imperfect or lacking information using relatively small samples, and that it is a proven method of statistical analysis which has been used in the water and rail sectors. It is particularly effective for long range forecasting especially over the 30-100 year horizons.

One of the limitations of the CECASE model is that it is less robust over the short term with Network Rail acknowledging that it is not as accurate over the 1-5 year period. Sources of uncertainty include assumptions concerning the starting point of a structure in its lifecycle, likely interventions, rates of degradation, accuracy of the unit costs, policy assumptions, skill and judgement of the engineers and modellers. It is therefore important to understand the uncertainty associated with the CP4 work volumes and costs derived from CECASE.

Part of this uncertainty comes from the sample size used for the detailed studies and one way of assessing the sufficiency of sample sizes is to consider what would be reasonable for a model

of this kind based on standard sampling theory. AMCL has calculated that overall the uncertainty as a result of the sample size alone was in the order of ±8% for the whole portfolio at 95% confidence limits. The level of uncertainty at the asset type level ranges from ±13.2% to ±50% (the actual level of uncertainty is greater than this though as explained below). It should be noted that the asset types with the highest level of expenditure are at the lower end of this range, confirming that the detailed studies have been focused on the higher expenditure asset types.

Network Rail has calculated that the actual level of uncertainty for the Civils expenditure in CP4 is ±17% at the 95% confidence level for the whole portfolio, taking into account all aspects of uncertainty and this is within AMCL's expectations of the ±15-20% for a model of this type. When considering the asset type level, the level of uncertainty increases and ranges from ±33% to ±74% which is quite large and may warrant further sampling for the assets with the highest expenditure.

CECASE deals with additional sources of random and systematic[1] error and it should be recognised that even if a 100% sample was available there would still be uncertainty associated with engineers' judgement, likely interventions, timing and costs. Network Rail has estimated, for example, that for Metal Underbridges there would still be a level of uncertainty of ±26% associated with the cost of replacements and ±14% for repairs to decks if a 100% sample was available.

Regarding sensitivity of the model to the sample size, Network Rail has indicated that increasing the sample size by 10% would only reduce the original variance estimate marginally which is supported by the analysis undertaken by AMCL using standard sampling theory. Network Rail has stated that, since the learning from the detailed studies was significant, additional studies would have a limited effect on the overall variances. AMCL expects that the sampling would need to be doubled or tripled to make a significant improvement but such increased sampling would only affect the random error component of the estimates. The sources of systematic error would also need to be considered further.

[1] **Random error**: difference between an estimated or measured value and the true value that is caused by random, and inherently unpredictable fluctuations in the measurement or the system being studied.

Systematic error: the difference between an estimated or measured value and the true value that is caused by non-random fluctuations from known or unknown sources.

In terms of the sensitivity of the model to asset selection, the prior model is based on a sample of 16,000 structures which is derived from SCMI studies and a set of assumptions have been applied to extrapolate the attributes to the total asset base. The detailed studies are then used to update the prior model which consisted of 740 studies in total. Network Rail has provided the basis for selecting the detailed studies which were stated to be representative of the typical structures found on the network.

It is understood that funding for CP3 as set out in the ORR's ACR2003 document[2] was calculated on a different basis to the current PR08 policies. For instance policy C 08 as it stands now, assumes a higher level of preventative maintenance as does policy B 08 which has led to significant increases in forecast expenditure in CP4. A comparable baseline should have been modelled in order to understand the reasons for the increases over and above the CP3 expenditure. It is fundamental that the Territory engineers are clear on the differences between ACR2003 policies and the current PR08 polices in order that the detailed studies are carried out on a consistent basis.

It is understood that Network Rail did consider modelling the Policy C03 approach generally adopted in CP3. However, after consultation with Mouchel, they rejected this option because they considered that the high number of intervention options made this approach difficult to model effectively. AMCL is of the opinion that this option could be modelled and calibrated against the activities in CP2 and CP3. Given that Policy E (which was modelled in CECASE) could have been used for the minor works elements as a starting point.

Actual works undertaken in the Territories can now be used to inform the modelling going forward which should help to reduce the overall uncertainty. The knowledge gained from the detailed studies together with actual ongoing interventions should also help to refine the assumptions used in the Markov model. This information together with ongoing SCMI assessments can then be used to refine the assumptions and estimates used to produce the overall forecasts. This will also give the ORR confidence that the actual work undertaken in the Territories is broadly in line with that predicted by CECASE.

[2] www.rail-reg.gov.uk/upload/pdf/184.pdf

In summary, it is AMCL's opinion that the CECASE modelling process is a robust approach and the level of uncertainty as a result of both the sample size of detailed studies, and of the wider uncertainty associated with likely interventions, timing and costs, is within the levels expected for models of this type.

The levels of uncertainty are within the range expected for this type of modelling, but if there is a requirement to reduce this uncertainty, additional sampling should be focused on the highest cost assets of metal and masonry underbridges, earthworks (soft) and metal overbridges where the level of uncertainty ranges from ±33% up to ±63%.

In addition, there is an opportunity to increase the confidence in the CP4 expenditure forecasts by further examining the difference and justification for policies B and C (i.e. ACR2003 versus PR08) and how the uncertainty associated with these policies and justifications are modelled within CECASE. It would also be beneficial to have modelled a baseline expenditure for CP4 (in line with the CP3 assumptions) in order to understand the reasons for and the impact of the new PR08 policies on a value for money basis.

Table of Contents

1 Introduction

Network Rail and the Office of Rail Regulation (ORR) have commissioned AMCL, through its role as Independent Reporter, to conduct a review of Network Rail's Civil Engineering Cost and Strategy Evaluation (CECASE) decision support tool which evolved from the 'Structures Annual Cost Profile' (SACP) work. This tool is used to support and determine Network Rail's civil structures funding requirements for Control Period 4 (CP4).

CECASE generates a network-wide budget for maintenance and renewals by using selected sample studies to calculate expected volumes of interventions over time and the application of different policies to the various asset types. It relies heavily on statistical methods, including Markov chain and Bayesian techniques to model various investment outcomes (in terms of level of intervention and frequency). This in turn is used to derive the work volumes and costs for the population of 40,000 bridges (plus other structures noted below) based on a small selected sample (580 was mentioned in the brief but this appears to be closer to 740 based on the latest information provided).

Assets Covered by CECASE analysis	Quantity
Bridges Earth Structures (embankments & cuttings) Culverts Retaining Walls Coastal defences	40,000 No. (58,000 spans) 13,000 miles 23,000 No. 17,000 No. (1500 miles) 120 miles
Assets not covered by CECASE analysis	**Quantity**
Tunnels Major structures	700 No. (total length 200 miles) 27 (short list) 300 (long list)

Table 1. Network Rail's Civils Assets

The ORR is interested in the modelling approach that has been adopted by Network Rail with particular regard to sample selection and sizes. The remit given to AMCL focussed on five key issues:

- The sufficiency of the sample size in being able to accurately model the major asset types (underbridges, footbridges, etc.) and the various forms of construction (single and multi span, construction materials iron, steel, brick/masonry, concrete);
- The sufficiency of the sample size in accurately modelling the distribution of structures across different route types;
- The degree of accuracy generated by the modelling and the likely tolerance of the figures;
- How sensitive the model is to sample size, asset selection and policy application; and
- Whether or not there is any evidence of Network Rail adopting an iterative learning process with the sampling i.e. do actual interventions that are used to revise the STAMP (Structures Asset Management Plan) model influence and improve the factual accuracy of the decision tree?

2 Methodology

In order to address the sufficiency of the sampling size and accuracy of the CECASE outputs AMCL produced a list of some 51 questions which was used as a framework to address the ORR's concerns mentioned in the introduction of this report.

The framework examined the key elements of the model and in particular the following areas:

- How the model is constructed and how it accounts for different types of structure, construction method, materials etc;
- The completeness and quality of the base data and any corrections that are applied to missing data, weighting etc;
- The inputs to the modelling process (e.g. policy, sampling methodology and sizes, etc);
- How representative the samples are and the extent of any bias in the sampling and how corrections are applied;
- The sensitivity of the model to input changes;
- Elicitation of probabilities and variances methodology (reliability of expert opinion, etc.); how these are applied and understanding how effective the learning process is;
- Understanding the limits of accuracy of the outputs from the modelling process;
- Examine how well the outputs correlate to actual programmes of work including minor works, substantial repairs and renewals; and
- Understand the feedback mechanisms that are in place to refine the base assumptions and used for instance in the STAMP model.

A copy of the questions and the responses from Network Rail has been provided in **Appendix A** of this report.

3 Findings

3.1 CECASE Process

Diagram 1 provides an overview of the CECASE modelling process. There are a number of papers and presentations available which describe CECASE in much greater detail.

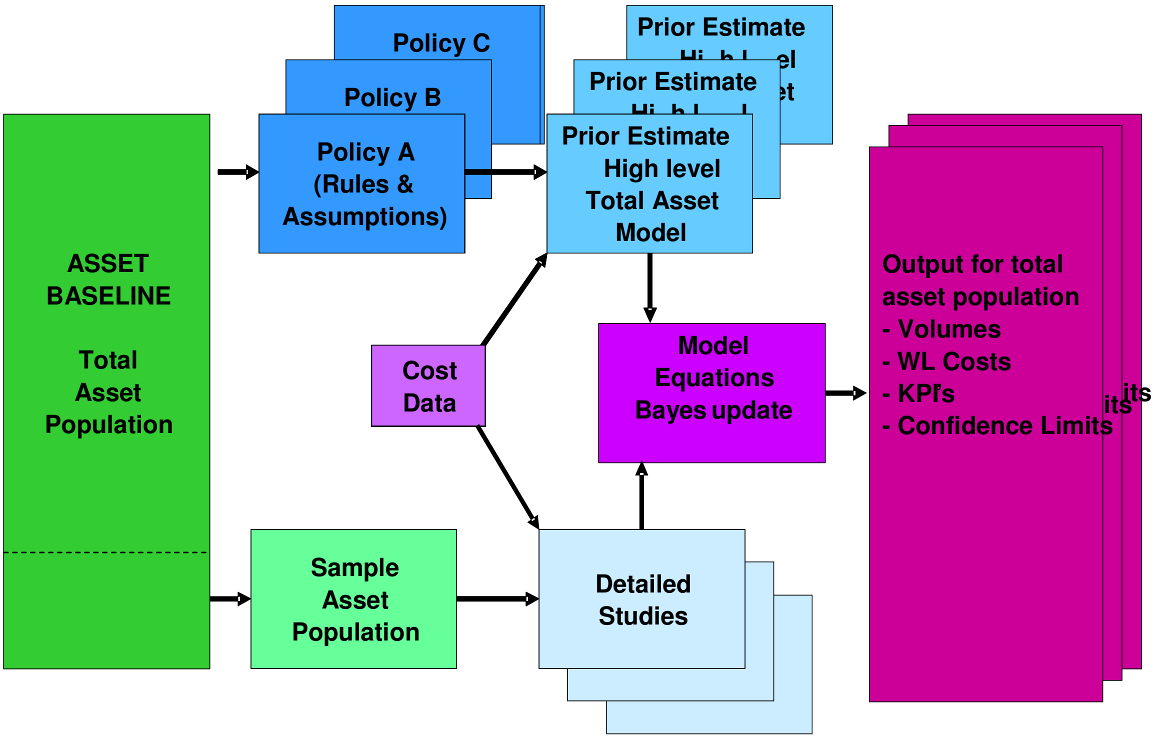


Diagram 1. CECASE Process – Source ORR Workshop March 4th 2008

As mentioned previously, Network Rail owns and manages a substantial portfolio of civil engineering assets which includes 40,000 bridges, 13,000 miles of earthworks, and 120 miles of sea defences. CECASE was developed to predict the strategic expenditure requirements over the next 30 years or so, and in fact the CECASE model is able to predict expenditure over the next 100 years.

There are four main components of the CECASE tool which are:

- Markov Chain (provides the degradation and intervention cycles, etc)
- Prior Model (uses 16,000 samples from the SCMI studies as a starting point)
- Updated Model (740 detailed samples are used to refine the prior model)
- Policy Planning Tool (a deterministic model used for scenario modelling etc)

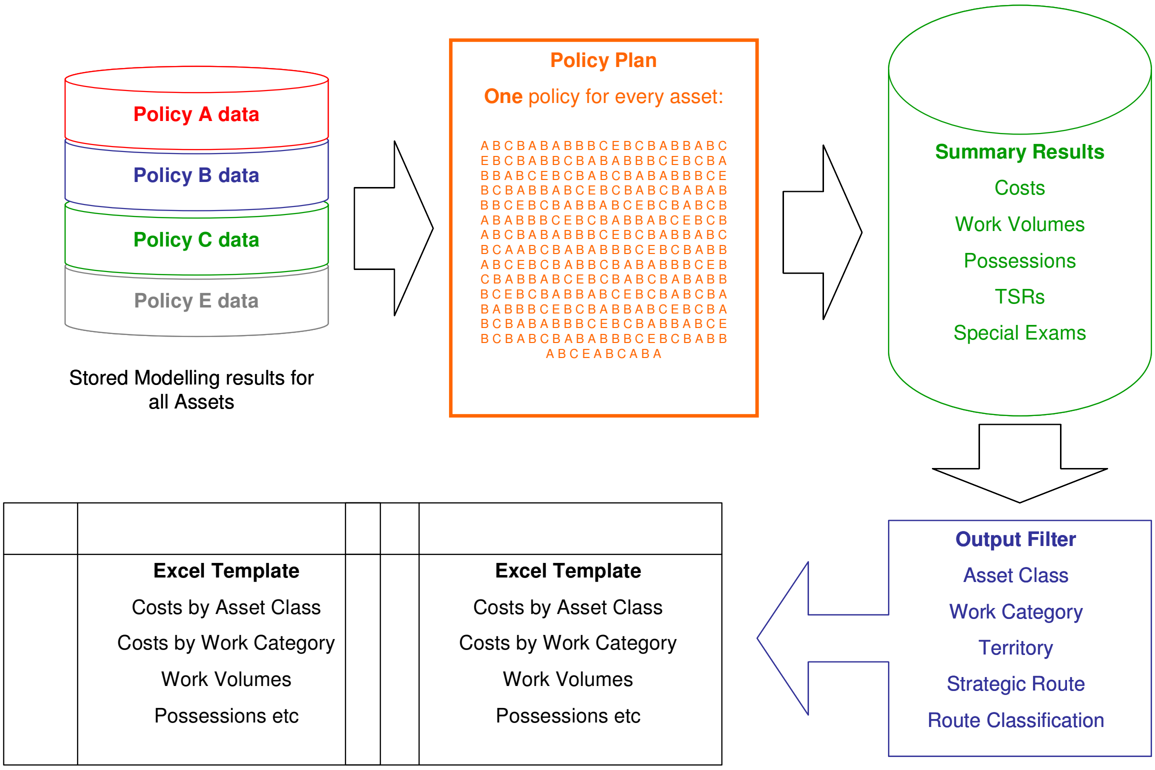


Diagram 2. CECASE Policy Planning Tool – Source Mouchel

It should be noted that the Policy Planning Tool operates outside of the statistical modelling and is used to undertake what-if analysis and scenario modelling etc. This is essentially a deterministic model which can be used to change the assigned policies more efficiently than using the statistical model.

The results from the statistical process is stored in the policy planning tool (PPT) database and the sensitivity of the overall results to changes in policy is produced by applying different policies to different assets within the PPT. The CECASE Model allows Network Rail to plan its maintenance work on civil engineering assets in order to achieve its long-term objectives and provide substantiation for the programme of works as part of the periodic submissions for funding to the Office of Rail Regulation.

It is recognised that a civils asset can be managed in a number of ways and that the exact intervention point can be difficult to predict since rates of degradation can vary greatly from structure to structure. Factors affecting the degradation include: primarily the environment, loading, materials, standards, etc.

Over the years there have been a number of different approaches taken in maintaining the structures to varying degrees. Network Rail has now developed a more systematic approach to managing its civils assets by adopting a set of policies that can be applied to an asset over its lifetime. The policies can be applied by route, reflecting additional traffic or importance in terms of abatements, etc. The key policies that Network Rail has developed are as follows:

- **Policy A Life Extension** – used for assets of historical value or on assets where the disruption caused by replacement would be unacceptable. The aim is to maintain the structure as long as possible;
- **Policy B Least Whole Life Cost** – the basis of this policy to minimise the whole life cost so that it poses minimal amount of risk to the railway;
- **Policy C Least WLC with some restrictions** – this is similar to B but gradually allows the assets to deteriorate until it is sensible to replace the asset.
- **Policy D Short term budget constraint –** manage the asset short term through unlimited speed/load restrictions which causes a higher level of disruption to the network (note: this policy was not modelled in CECASE);
- **Policy E Least cost for non operational structures –** this is generally applied to closed lines to meet Health and Safety requirements.

It is understood that Network Rail's ACR2003 policies were not modelled using CECASE. Network Rail considered that the previous policies were largely reactive and therefore not suitable for use in strategic whole life cost modelling as they were considered sub-optimal in the long term. Policy B08 is a step up from Policy C08 and assumes a greater amount of preventative maintenance. An important issue is to understand the extent to which these policies are actually applied at the Territory level because the detailed studies should reflect the actual decisions adopted on the ground.

The Territory Engineers work banks are constrained by the available funding. If Policy B08 is adopted then Engineers are able to maximise the possession time to include additional preventative works which can be more cost effective in some circumstances. The Policy Planning Tool is able to model the impact of a mixed approach using Policy B08 and C08.

Diagram 3 below is a representation of the intervention points related to Policies A, B, and C that Network Rail has modelled in 2008.

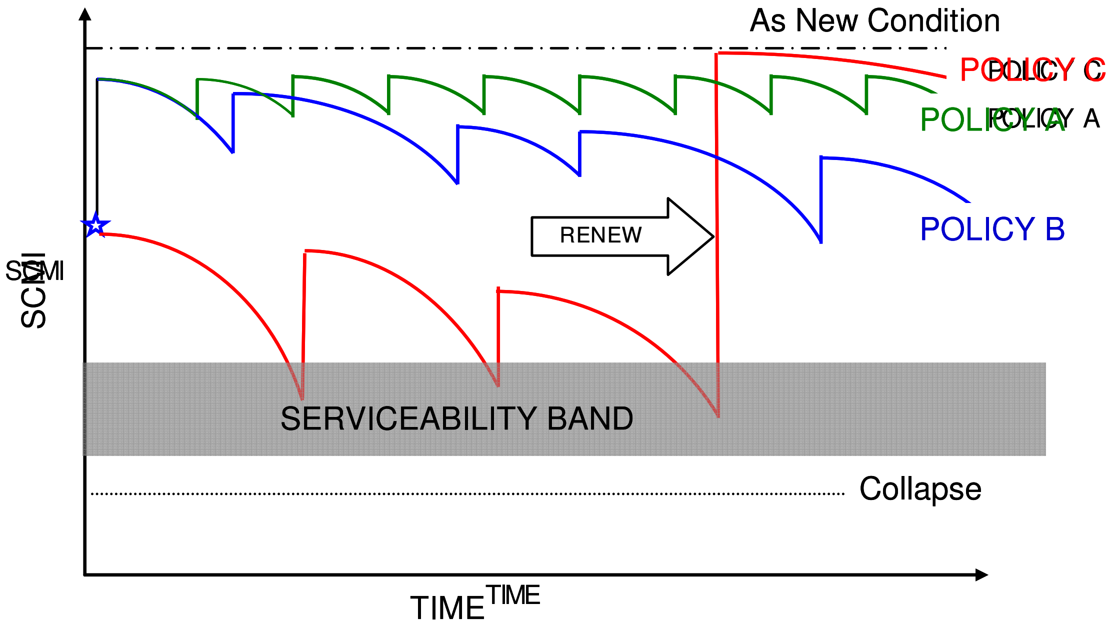


Diagram 3. Application of Policies – Source ORR Workshop March 4th 2008

A baseline estimate has been produced by using the 19,300 structures information from the SCMI reports. Of this some 13,900 structures were underbridges. It is understood that some 16,000 structures were finally modelled in CECASE as some of the structures could not be matched for modelling purposes. Where an asset had multiple examination dates the most recent was used. A set of assumptions was used to extrapolate the available information and produce a complete prior data set.

Once the policies have been applied, a *prior* model is derived as a starting point which can be thought of as a theory or a hypothesis. In essence, for structures, the objective is to test the probability of an intervention taking place such as a repair, strengthening, waterproofing, etc. As an example Diagram 4 shows the volumes of interventions for underbridges in CP3 compared to CP4. The important trend to note is that the volume of repairs has increased significantly in CP4.

It is necessary to test the *prior* model with real life data (detailed studies) in order to refine the model and reduce the uncertainty of the theoretical or prior model. Once the volumes of work have been generated then the unit costs can be applied to calculate the required expenditure over a period of time and this is the basis of Network Rail's submission for CP4.

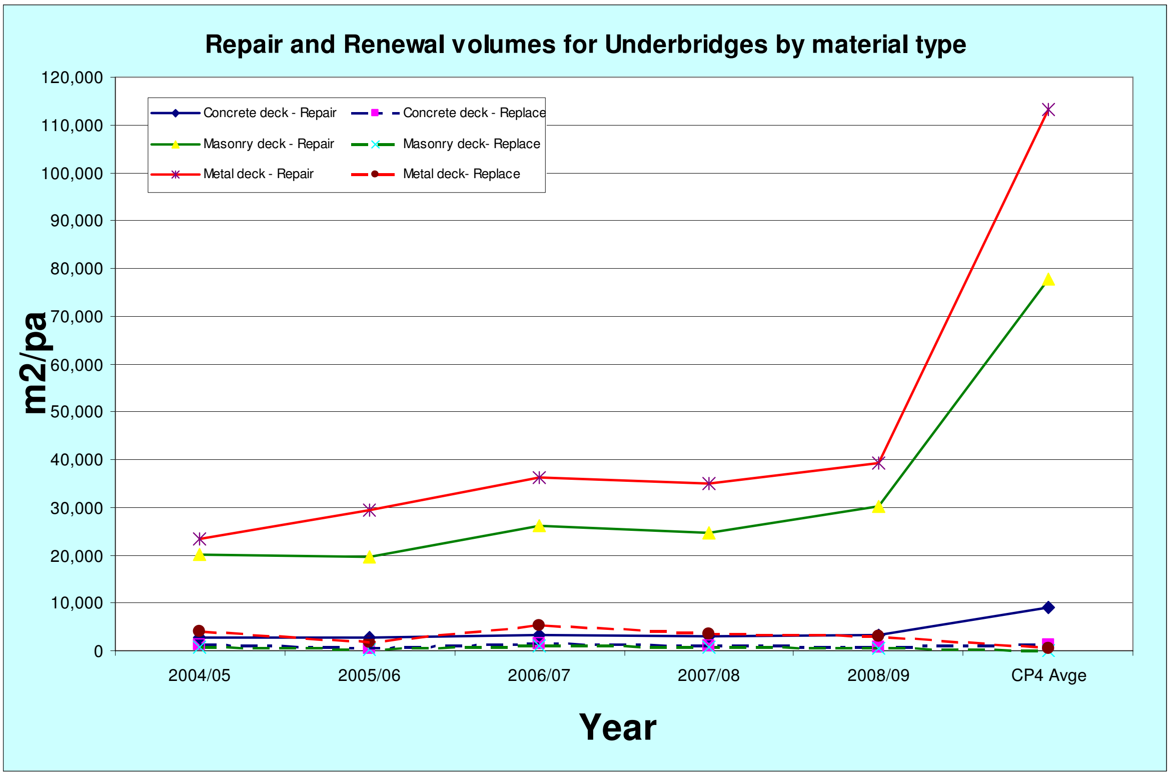


Diagram 4. Types of intervention modelled in CECASE – Source Network Rail

Once the ABLE model is run it generates a *posterior* model which aims to improve the *prior* forecast estimates. Our understanding is that these estimates are then later used in a Policy Planning Tool which is simply a deterministic model that is able to consider the implications of applying different policies to different structures or routes, etc.

It should be recognised that the Policy Planning Tool models the effect of different policies outside of the statistical modelling. Network Rail has stated that it is valid to use the Policy Planning Tool since the manner in which the structures are maintained are broadly the same across the routes and there is only an issue of timing and level of intervention across the routes that may differ.

3.2 CP4 Costs

Table 2 is an extract from Network Rail's Strategic Business Plan for CP4. The levels of expenditure and hence risk should inform the sampling process in order to obtain the best return considering the cost of the sampling.

CIVILS SBP SUMMARY			
Asset Group	**CP4 £m (2006/07 prices)**	**%**	**Modelling**
Underbridges	**831**	42.0%	CECASE
Earthworks	**344**	17.4%	CECASE
Overbridges	**325**	16.4%	CECASE
Major structures	**132**	6.7%	Major Structures Model
Tunnels	**131**	6.6%	Tunnels Model
Culvert clearance (Other)	**53.4**	2.7%	Culvert Model
Culverts	**49**	2.5%	CECASE
Footbridges	**43**	2.2%	CECASE
Retaining walls	**21**	1.0%	CECASE
Coastal / estuarial	**21**	1.0%	CECASE
Closed & mothballed (Other)	**16.1**	0.8%	CECASE
Ancient Mines (Other)	**12.7**	0.6%	Ancient Mine Model
UBT (Other)	**0.8**	0.0%	CECASE
OBT (Other)	**0.2**	0.0%	CECASE
Total	**1,979**	100%	

Table 2. Control Period 4 Civils Expenditure (Source Network Rail SBP)

On the basis of the expenditure breakdown, it is clear that underbridges, earthworks and overbridges account for the majority of the expenditure and therefore, they are the most important assets which are modelled in CECASE. Hence, sampling these assets should yield the greatest benefits in terms of reducing the overall uncertainty. The detailed studies undertaken by Network Rail appear to have focused on these three asset groups.

3.3 Analysis of Variance

Table 3 overleaf shows the total population of assets, sample sizes (740 in total) along with the calculated confidence intervals at the 95% confidence level. AMCL's calculations shows that the expected variance based on sampling theory together with required sample sizes needed to reduce the uncertainty to (say) ±10% at the 95% confidence limits.

The sampling for metal and masonry underbridges shows that Network Rail has targeted the important asset groups as expected with expected variances ranging from ±13.2% for masonry underbridges to ±50% for rock cuttings. The sum of the variances (based on sampling theory) would give an overall variance of ± 8% at the 95% confidence limit. CECASE contends with greater systematic and random errors which means that the overall variance will be much greater.

Based on Network Rail's calculations the overall variance is in the order of ±17% at the 95% confidence limit. This is within the ±15-20% that AMCL would have expected for a model of this kind. To some extent this reflects the difficulty of modelling long lived Civils assets, the limitations of the modelling process given uncertainties associated with intervention, timing, costs and so on.

Having said this, it is evident that the variances associated with some of the key asset groups appear to be quite large at the asset group level (for example ±41% on metal underbridges). Given the level of expenditure associated with some of the asset groups it may be beneficial to conduct some additional sampling for metal and masonry underbridges, embankments and metal and masonry overbridges which have relatively large variances at the asset group level.

From Diagram 4 it is noted that there is a large increase in repairs expected during CP4 and it would be appropriate to understand the specific reasons for this increase. In broad terms it is understood that it is additional preventative maintenance that is driving up costs and but cost/benefit analysis would be useful to demonstrate the value for money from increasing up-front investment against a reduction in replacement of structures in future years.

Asset type	Population	Expenditure CP4 £m	Sample Size used in CECASE	Confidence Interval @ 95% Limit ± % (Sampling Theory)	Sample req'd for ± 10%@ 95% Limit (Sampling Theory)	Confidence Interval @ 95% Limit ± % (Modelled in CECASE)
Underbridges Deck, Metal	10,458	486.1	51	13.8	96	41
Underbridges Deck, Masonry	25,322	272.3	58	13.2	96	33
Underbridges Deck, Concrete	3,422	71.5	23	20.8	94	43
Underbridge Supports, Arched Masonry			92	21.3	93	
Underbridge Supports, Nonarched Masonry			121			
Earthworks Cuttings, Rock	1,600 km	60.9	6 (4.7 km)	50	96 km	74
Earthworks Cuttings, Soft	6,400 km	88.1	56 (24.9 km)	20.4	95 km	74
Earthworks Embankments, Soft	9,300 km	195.4	54 (23.2 km)	20.8	96 km	63
Overbridge Deck, Metal	3,173	145.7	22	21.3	94	34
Overbridge Deck, Masonry	6080	103.6	21	21.9	95	36
Overbridge Supports, Arched Masonry			28	21.3	93	
Overbridge Deck, Concrete	2,669	75.2	22	21.3	93	39
Overbridge Supports, Nonarched Masonry			73	21.3	93	
Culverts	23,000	49.4	20	22.4	96	36
Footbridges, Decks	3,200	43.4	16	25.2	94	49
Footbridges, Supports	*6,400*		25	20	95	
Retaining Walls	17,000	20.6	28 (8.3km)	18.8	96 km	63
Coastal & estuarial defences	300	20.6	24	19.6	73	63
Total		**£1632.8m**	**740**	**±8%**		**±17%**

Table 3. Sample Confidence Intervals (data from Network Rail, calculations undertaken by AMCL)[3]

[3] Some 1,805 and 2,072 over and underbridges do not have samples (classed as "other").

3.4 Analysis of Detailed Studies

As part of the detailed studies Network Rail sampled some 740 different assets using Policy A,B,C and E interventions (based on PR08 policies). Hence the detailed studies do not necessarily reflect the current practice which is underpinned by the ACR2003 policies and assumptions. Since the ARC2003 policies were not modelled the specific reasons for the increases are not transparent enough other than in general they are down to additional preventative maintenance works.

Network Rail provided an estimate of standard deviations of volumes of work for metal underbridges Policy B for the 30 year data are shown in the table below. Overall, the ABLE2 model had 'learned' from *prior* to the *updated* model from 57.5% down to 20.3% overall.

Metal underbridges account for only a part of the expenditure and when the estimates for other types of asset are added together the overall standard deviation reduces. Network Rail made the point that this was also observed in SACP model previously. Given the large sums of expenditure associated with certain asset groups, the level of variance at the asset group level should not be discounted though and further work may be necessary to understand the reasons for the uncertainty.

Intervention	Prior	Updated	All study
Replacement	120%	45%	26%
Strengthening	71%	49%	Not Available
Repair to Beams	239%	38%	14%
Repair to Decks	465%	48%	14%
Preventative	435%	53%	Not Available
Waterproofing	218%	56%	Not Available
All interventions	**57.5%**	**20.3%**	**-**
Replacement + Repair to Decks and Beams	110%	26%	10%

Table 4. *Prior* to *posterior* change due to learning

The 'all study' averages shown in table 4 are what would result from a 100% sample set as ABLE requires an estimate of this. In learning, ABLE reduces the prior variance towards the 'all study' variance and can go no lower. For the three intervention types where all data was available[4] ABLE has learned a substantial proportion of the difference from prior to updated

[4] To estimate this data it is necessary to modify an updated ABLE file. The process obtained results for three types of intervention but did not produce results for the other three as indicated in table 4.

model and additional samples would only result in a slow improvement in the standard deviation.

Prior intervention standard deviations were high mainly because there had been an underestimate of the frequency of interventions when creating the prior models as explained in the CECASE Functional Specification. It was also high because it multiplies together three quantities with fairly high variances. The prior variances had to be increased in order to cover this. This meant that ABLE had learned upwards and would have ended at less than the study averages.

In the CECASE study, the average cost data analysis was carried out in a separate model to the numbers of interventions. This did not cause a problem when estimating central values but to estimate variances the data needed to be in the same model. In order to produce the table above, Network Rail have had to extend the original models. This re-working did not change the data but brought it into a common model.

Network Rail did not use its own bottom-up work banks to calibrate the models. However, it is unlikely that the current work banks would reflect the PR08 policies C&B in any case. The current detailed studies largely reflect the PR08 policies which assume increased spending especially on preventative maintenance.

3.5 Unit Costs

The unit cost information is provided from an Access database which is exported to a spreadsheet for ease of handling. This tool is known as the Structures Benchmarking Tool (SBMT). Unit cost information is also used from the previous SACP work. SBMT was taken as the preferred source where available. The average cost per intervention and asset type can be scaled appropriately.

SBMT uses only schemes costing over £100,000 in its base data. It cannot be used for work categories of a lower value at present. SBMT is used for replacement, waterproofing and strengthening but not for other categories. In the detailed studies the costs are derived from first principles and then calibrated against SBMT data in order to ensure that costs are comparable on a like-for-like basis.

It is understood that unit costs are an area where there is significant uncertainty associated with developing the bottom-up estimates. Network Rail now has the opportunity to build up its knowledge of minor works and associated costs in order to better inform the CECASE modelling process for future updates. Network Rail is considering how at least some of the basic and repetitive cost information could be collected recognising the resourcing issues associated with collecting and storing the information.

3.6 Sensitivity Analysis

As part of this review Network Rail was asked to consider the effect of increasing the sample size by 10% and what effect this would have on the model. This was answered as part of question 29 (see Appendix A). Network Rail have stated that the model had already learned significantly from the prior to the posterior and an increased sample size might move the confidence interval from 17.5% to 17% assuming the study samples were of the same quality.

Considering the large increase, particularly for repair work (see Diagram 4), with a variance of 38% and 48% for repairs to beams and decks respectively, it may well be worthwhile to either re-visit the assumptions at the prior stage or else consider additional studies to reduce the uncertainty as repairs, in particular, are the key cost driver in CP4.

It is understood that the preferred PR08 policy is modelled in the prior and updated models. The Policy Planning Tool is able to change the various policies easily on a deterministic basis outside of the statistical model. The ABLE2 modelling is concerned mainly with how many interventions are likely to occur over a period of time and is sensitive to changes in assumptions. If the detailed studies showed that more interventions were likely and Network Rail was confident that this was the case then the model would learn from the detailed studies.

When considering how sensitive the model is to policy application, AMCL has found that this is the key determinant of volumes and expenditure since this specifies the level of intervention and timing. Diagram 5 demonstrates the unconstrained model output of Policy C08 for Metal Underbridges. It shows that around CP13 to CP19 the volume of replacements is expected to increase to an unsustainable level if only Policy C08 is applied.

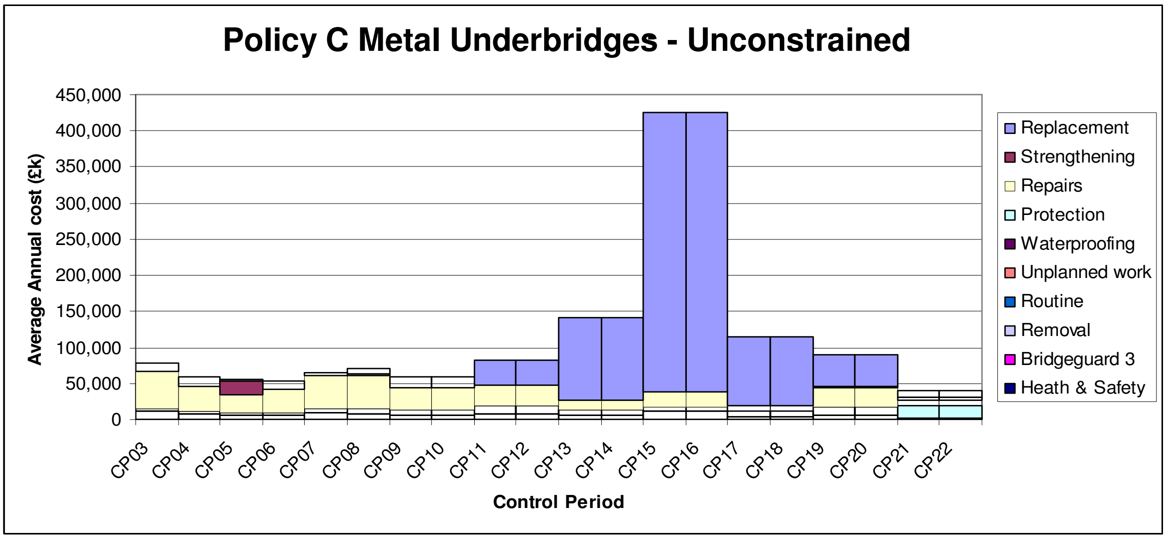


Diagram 5. Sensitivity to policy selection – Source Network Rail

Diagram 6 shows that even if the replacement programme were to be smoothed it still represents a significant volume of renewals work under Policy C08. In order to avoid a large amount of replacement in the future, Policy B08 demonstrates that more repair work should be undertaken up front in order to avoid unsustainable levels of replacement in the future. This was also shown in Diagram 4 earlier.

The Policy B workload would also need to be smoothed out over a number of control periods. Network Rail's long term strategy is to get the Civils assets into a "steady state" condition.

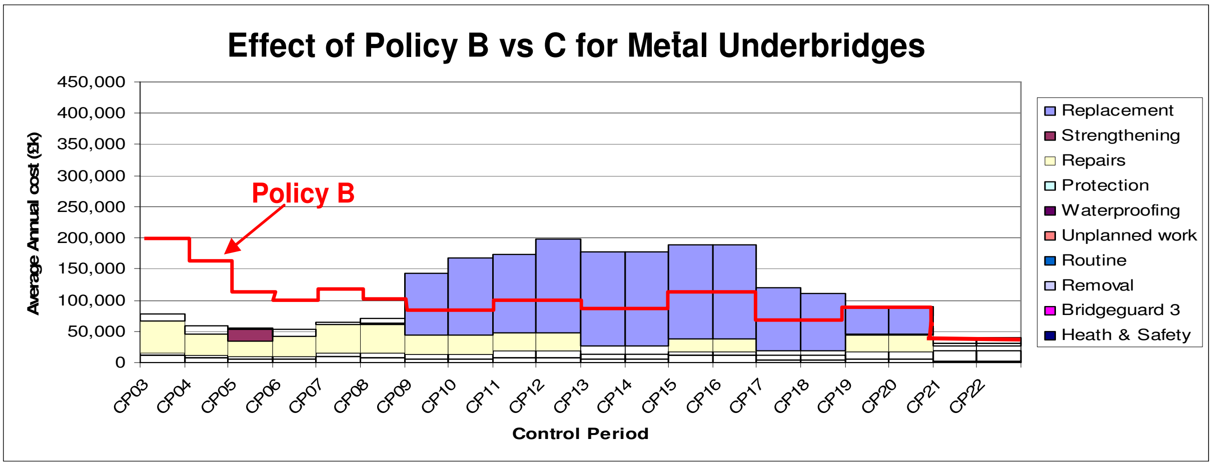


Diagram 6. Sensitivity to policy selection – Source Network Rail

3.7 Learning Process for STAMP

Network Rail discussed the possibility of a learning process from actual interventions with Mouchel at the outset of the CECASE project. At that time Network Rail had good knowledge of the use of STAMP and the development of a model (SACP) in using STAMP sample studies.

The development of CECASE was seen by Network Rail as ambitious in terms of the more complex modelling, the outputs required, the use of policy options and the short timescale for achieving it. It was decided that the introduction of a further unknown, in the form of feeding back the learning from the detailed studies to STAMP (and by implication the Markov Model), potentially put the principal objectives of the project at risk, particularly as the correlation between of the theoretical interventions sample studies had yet to be proven.

Now that the novel modelling issues are better understood, there are a number of developmental possibilities which Network Rail could consider. These will include the option of learning from actual interventions and refining the intervention and degradation assumptions used in the CECASE model.

As the Policies in effect drive the modelling outputs the timing of an intervention changes as a function of the prescribed policy. The implications of this are that a STAMP type approach would need to take account of these differences in scope and timing and there may need to be a number of STAMP model options which reflect the different policy options.

4 Conclusions

Overall AMCL considers that the CECASE approach is robust and fit for purpose in modelling long lived Civils assets and has a history of use in both the Rail and Water sectors. Its strength lies in its ability to learn from various data sources, its ability to refine imperfect or lacking information using relatively small samples. CECASE is more accurate over the longer term especially over the 30-100 year horizon. It is less accurate over the short term (i.e. 1-5 year period) and is not able to reconcile bottom-up estimates over CP4 owing to the limitations of this type of modelling and policy assumptions. The model is also progressively less accurate when analysing the asset group level and even less so at intervention level.

In order to answer the question of sufficiency of samples, AMCL has calculated the variances based on sampling theory and found that Network Rail had targeted the sampling effort appropriately at the most important asset groups with underbridges, earthworks and overbridges which make up over 75% of the expenditure for CP4. From AMCL's calculations on the basis of sampling theory alone, the overall variance which could be expected would be in the order of ±8% for the whole portfolio at the 95% confidence limit. The level of uncertainty at asset group level ranges from ±13.2% to ±50% on this basis. It is noted that the actual variances are much greater than this though.

AMCL's expectations at the outset were that the overall variance that could be achieved would be in the order of ±15-20%. Based on Network Rail's own calculations the overall variance is actually in the order of ±17% which is reasonable for this type of modelling. It should be recognised that at the asset class level the variances are much larger ranging from ±33 to ±74% and so there may be some scope for additional sampling for certain asset groups.

Even if a 100% sample was available then there would still be uncertainty associated with sources including engineers' judgement, likely interventions, timing, costs, etc. Network Rail has estimated that for metal underbridges as an example, the best level of accuracy that could be achieved is ±26% for replacements and ±14% even if a 100% sample was available.

The statistical model in CECASE implicitly takes into account the distribution of structures across the route types as it models a preferred policy for each asset element at the start of the process. However, the statistical model is not stratified by route level and would require additional sampling to enable modelling in this fashion. It is debateable whether it would be cost effective to set up the statistical model in this manner as the Policy Planning Tool is able to model down to route level. It is noted that the Policy Planning Tool is a deterministic model which simply manipulates the outputs from the statistical model to consider the effects of changing policies and scenarios.

Regarding the sensitivity of increasing the sample sizes on the resulting variances, Network Rail has indicated that increasing the sample size by 10% for example, would have a marginal impact on the overall variance. Network Rail has stated that since the learning from the detailed studies already undertaken was significant, that additional studies would have a limited effect on the overall variances.

AMCL's own calculations suggest that the detailed studies sample sizes would have to be doubled or trebled before further significant improvements in the variances would be achieved. There would be marginal benefit in additional sampling for all of the asset groups but a targeted approach for the high cost assets may improve the results in key areas. Certain asset groups such as metal and masonry underbridges and overbridges together with embankments may benefit from additional sampling given the relatively large uncertainties associated at asset group level which range from ±33% to ±63%.

The unit costs are an area of particular uncertainty and it appears that bottom-up estimates in the detailed studies are higher than the outturn costs that Network Rail normally experiences. It would be beneficial for Network Rail to build up its own library of minor works costs in order to better calibrate the CECASE model in the future.

In terms of sensitivity of the model to sample selection, the *prior* model is based on a sample of 16,000 structures which is derived from the SCMI studies. A set of assumptions is used to extrapolate the attributes to represent the total asset base. The detailed studies which consisted of 740 studies were then used to update the model. Network Rail has provided the basis for selecting the detailed studies which were stated to be representative of typical structures found on the network.

The outputs from the CECASE model are very dependent on the policy selection and it should be recognised that this is the most important factor in calculating the overall expenditure forecast. There is a limited amount of “learning” taking place in a purely “Bayesian” sense since Policies B&C08 have not been reconciled against actual current practice which is also reflected in the current work banks. As such, the learning effect is to an extent circular given that the detailed studies have been based on the PR08 policies not necessarily what happens on the ground. The additional expenditure is mainly down to additional preventative maintenance in CP4.

Network Rail decided not to model the ACR2003 policies citing problems with the number of different types of interventions which the model would have to contend with. Network Rail claims that minor works items for example are already inherently modelled within the PR08 policy C and B options. AMCL believes that the ACR2003 policies could have been modelled using Policy E (which was modelled in CECASE) to provide the minor works elements as a starting point. AMCL believes that Network Rail should reconcile the ACR2003 and PR08 policy assumptions, quantify the differences and justify the benefits of additional expenditure for preventative maintenance in CP4.

Network Rail considered the possibility of a learning process from actual interventions with Mouchel at the outset of the CECASE project. At that time they had good knowledge of the use of STAMP and the development of a model (SACP) in using STAMP sample studies. Owing to the short timescales in developing the CECASE model it was decided to review the learning from the detailed studies and feed back the knowledge into STAMP (and by implication the Markov Model) at a later date since the correlation between STAMP and the sample studies was unclear at the time of modelling. Network Rail intends to consider further development of CECASE tool including learning from actual interventions in due course.

5 Recommendations

It is recommended that Network Rail undertakes a further review to consider the potential benefits of reducing the variances of the expenditure forecasts associated with particularly metal and masonry underbridges and overbridges and earthworks where the variances range from ±33 to ±63%. This analysis should also consider the costs and benefits of stratified sampling and statistical analysis at route level.

It is recommended that the ACR2003 and PR08 polices should be reconciled to provide a recognisable baseline for the ORR to assess the changes. The future work bank should then be generated on a consistent basis which will help to inform the front end (i.e. years 1-3) of the CECASE model using real data. The additional expenditure on preventative maintenance should then be justified on a cost/benefit basis, where the trade-off is additional up-front repair work against costly replacements in the future.

It is recommended that Network Rail ensures that the differences between the ACR2003 and PR08 polices are understood at the Territory level in order that the detailed studies which informs the modelling are consistent and produce reliable results. This will improve the assumptions and estimates used in the forecasts, and will also give the ORR confidence that the actual work undertaken in the Territories is broadly in line with that predicted by the CECASE modelling process.

Network Rail should build up its own library of unit cost information for minor works (i.e. less than £100,000) in order to reduce the uncertainty of the bottom-up cost estimates. Over time this should reduce the uncertainty of the overall forecasts as theoretical unit cost estimates are more closely tied to actual outturn costs. A review of the difference between the bottom-up estimates and the SBMT costs should be conducted to assess the uncertainty. Network Rail is considering the best way to approach this issue given the cost considerations of storing and managing this information.

It is recommended that the lessons learned from the detailed studies should be fed back into the STAMP model (and by implication the Markov Model) in order to improve the *prior* assumptions. The model should continue to be updated on a periodic basis with a calibration made against the 1-5 year bottom-up work bank.

In addition, a number of recommendations are made below that are aimed at reducing the possible errors and uncertainty in the modelling process as follows:

1. Given that the 16,000 SCMI structures (around 40% of the total) have been used to extrapolate the total population for the prior model, there will be some uncertainty as to the actual starting condition of the structure and what interventions will actually be required in real life. As more structures are assessed using the SCMI process this should reduce the uncertainty over time.

2. In addition to the uncertainty surrounding the actual starting point, there will also be some uncertainty concerning past interventions. There may be some occasions when the prior model assumes an intervention is necessary when in reality that intervention, such as strengthening or waterproofing, may already have been carried out. This could lead to some inefficiency or double counting in the model. Network Rail commented that given the relatively small amount of structures that have had significant interventions over the study period that the impact is likely to be small. It would be useful to conduct an analysis to confirm the assumptions.

3. The rate of degradation is the key determinant of the timing for an intervention. In the statistical model, Network Rail has used a straight-line deterioration rate when in reality the rate is more curvilinear (see Diagram 3). Network Rail claims that their approach is conservative and assumes a later intervention timing than that of a curved function. Given the uncertainty surrounding the exact timing of an intervention, it would be useful to calibrate the degradation model using the SCMI scores which are starting to come into the second round of scoring based on the 6 yearly inspection cycle.

4. Given that Civils structures are predisposed to the effects of rainfall in particular, it would be useful to consider whether climate change, with the possibility of hotter and drier summers and warmer wetter winters will have a significant effect in the long term given that Network Rail have modelled a 100 year horizon. There are a number of initiatives in place to better understand long range climate scenarios (e.g. www.ukcip.org.uk) which could inform the CECASE modelling work.

5. The approach taken to detailed studies should be reviewed. Whereas, currently, all PR08 policy options are costed, there is now an opportunity to focus only on the preferred applied policies. This should reduce the time involved as well as the cost of the detailed studies which can then be used to increase the sample sizes and reduce uncertainty both at asset group level and overall.

Appendix A – Information Request to Network Rail

Information Request to Network Rail

General

1. **Provide an overview presentation by the CECASE Development Team of how the model has been constructed to take account of the different types of structure, and the different types of route, etc.**

   A presentation and meeting on 12 March 2007 outlined how the CECASE results were derived and how policy planning tool was used to allocate different Policies to different assets across the network.

Model setup and assumptions

2. **Is there a list of the key assumptions under which the model operates?**

   There are two sets of assumptions that need to be distinguished, one for the studies and one for the prior models. The final answer can be regarded as a weighted average of the two. However, for some estimates the study assumptions are dominant and for others the weighting tends more to the prior. This arises from the difficulty in handling time.

   Time presents a particular problem for the statistical modelling of interventions because the probabilities of carrying out activities in adjacent time periods tend to be negatively correlated. For example, if a repair is carried out to an element in one period of time it is much less likely that a repair will be required in the following period. Instead of modelling these variables directly they were transformed into a set of quantities that were independent of each other which could be recombined through the model equations into the output quantities required.

   As a simple example assume that there are only three periods, X, Y and Z. Define the following variables:

   Pt = The probability of an intervention in period t for t = X, Y, Z.

   PT = The probability of an intervention in the whole time period (X plus Y plus Z)

   Define

   pX = PX /(PX + PY.+ PZ)

   pY = PY /(PY.+ PZ), i.e. the probability of an intervention in a period as a proportion of the residual probability to the end of the period.

   The three new variables pX , pY and PT are statistically independent.

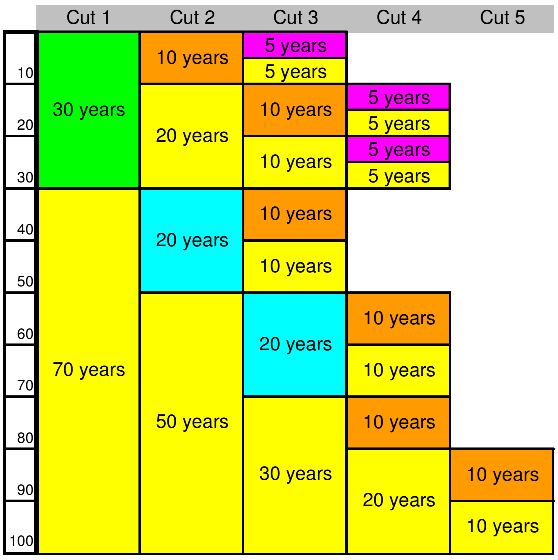
   Then,

   PX = PT * px

   PY = PT * (1- px) * pY

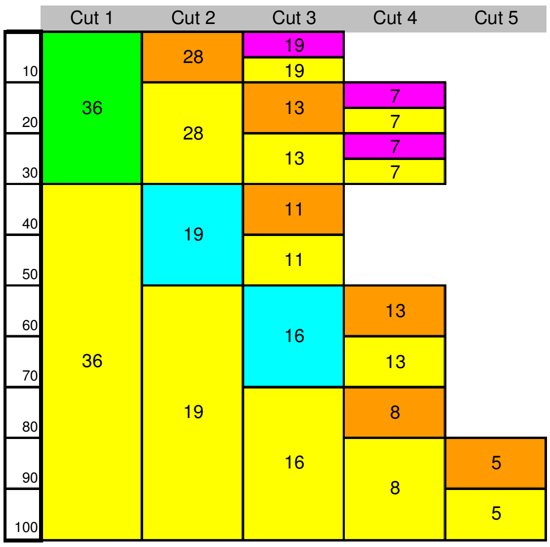
   PZ = PT * (1- px) * (1- pY)

   This concept can be extended to any number of periods. For longer time periods with more subdivisions the overall period can be divided in many ways. The division of 100 years into 20 five year periods used in CECASE is shown in the diagram below. All colours apart from yellow represent a defined cut of a larger block with the yellow blocks representing the residual of the period. The first cut was a 30 year block with a 70 year residual. The 30 year block was cut into a 10 year block with a 20 year residual block and so on.



The reason for doing this is that the cuts on the left tend to have more sample points for the update. Columns to the right have a greater chance that the proportion associated with a block will be zero divided by zero which is indeterminate. All cuts that derive from this block will also be indeterminate. The cut with the most data will be cut 1 making the 30 year total likely to be the most accurate, although it does depend upon other factors in individual cases. The cut of years 80 to 100 into two ten year periods is likely to be the least accurate relying heavily upon the prior estimate. In this way we can select the periods for which we require the data to be generally the most accurate.

An example of the way the sample numbers reduce is shown below for ‘repairs’ to metal underbridge beams[5].



Total Sample Size 51

[5] Please note that this is an example and doesn’t reflect all the interventions identified on the detailed studies for ‘metal underbridge beams’ as equivalent interventions were identified and modelled for the other 8 categories of work (replacement, strengthening etc..). Additional interventions were also identified for each work category for different underbridge ‘element groups’.

Fifteen samples had no intervention over the 100 year time period and so no data was available for these assets on the split within the 100 years.

The extent of the learning within ABLE depends significantly upon the number observations available for the update. As this increases, the estimate moves in general terms from the prior average towards the study average and the confidence level narrows. This is also affected by outlier values however as these are given a lesser weight by ABLE.

The most significant estimate in the exercise is for the 30 year period as it is the 30 year average that was used as the main basis of the submission. The more detailed profile indicated a probable need for a high volume of work early in the period so it was decided to phase this more evenly over the period. It is the 30 year average or total that is the most accurate and closest to the study values. It is consequently the study assumptions that influence this most strongly. This is also the case for estimates for the first ten years.

The study assumptions are set out in some detail in Appendix K of the functional specification. The general assumptions for the prior model are set out in Appendix K. There are many detailed assumptions in the individual models which were derived in the ways described in Appendix K.

3. **Can NR provide an explanation of the engineering model and how this operates? How have the probabilities of faults and failures been derived and validated?**

   See section 7.4 and Appendix A of the Functional Specification.

4. **How does Network Rail overcome the issue of age, condition and remaining life of a structure i.e. to what extent has past intervention been modelled or what assumptions have been made about past interventions?**

   Past activity has only been considered in as far as it has affected the starting condition, strength and characteristics of the asset. The engineering team has used most recent data available for a study asset. For each Policy, future work is based on known condition and strength of an asset, on engineering views of deterioration and on the discounted cost of adopting alternative activities.

   Age is considered where this affects the material characteristics of the asset (for example whether material is wrought iron or steel, or concrete composition and expected detailing).

   Historical problems may also be considered if they affected future performance of the asset.

5. **How complete and accurate is the base data?**

   This varied across the asset types and by type of data as explained in Appendix K. Bridge data was generally the most complete. Models were used only where there was sufficient data to populate the prior model in a reasonable way. Where individual assets did not have sufficient data for this, their intervention requirements and costs were estimated by association with other similar assets. For example, bridges were included in the ABLE models only if there was condition data. Other bridges were assumed to have probabilities of intervention equal to the average of modelled bridges of the same type.

6. **How many SCMI assessments have been carried out?**

Refer to Appendix K, Annex E – SCMI data for details of how SCMI was used in the model.

7. Can NR explain the basis for firstly selecting "special" or exceptional structures and secondly how the ones that are modelled are treated so that they do not skew the results?

This is set out in the executive summary paper for major structures, produced by Network Rail.

8. Does the model only consider steady state investment? How is enhancements to structures treated e.g. the need to increase capacity of a structure above the design capacity?

The models made no allowance for changes in capacity or capability as this is funded separately

9. How has Bridgeguard 3 been modelled?

NR provided a bottom up estimate of the numbers and costs for Bridgeguard 3 (based on the current state of this project) which were inserted into the results. Refer to Appendix K, Annex E – Bridgeguard 3 data.

10. Has accidental damage e.g. bridge strikes etc been modelled?

This was excluded from the CECASE model. Network Rail has not included any additional costs within the results submitted to ORR on the grounds that the total costs are relatively small and the majority of such costs are recoverable under insurance arrangements.

11. Has the effects of weather been included for structures e.g. sea defences, bridge scour, etc?

Some of the sample assets were affected by scour and this was taken into account in determining the management regime that should be applied to them. Sea defences were assessed separately. See Appendix K, Annex C – Engineering Assumptions for Sample Coastal and Estuarial defences – Key Exclusions.

Sample Sizes

12. On what basis were the 580 samples of data selected? Is this all of the information available? Please provide a breakdown of the samples by type of structure, route etc.

In SACP we calculated the variances derived from the ABLE model. We tabulated the results against the number of samples and estimated the benefit we would derive from adding more samples to each asset type. This was used to determine where additional samples would be most beneficial in terms of reducing the overall variance and a plan was developed to use the sample time and budget effectively. The cost of samples was substantial so the number of possible samples was constrained.

The sample numbers are set out in the following table:

Asset type	Sample numbers
Culverts	20
Earthworks Cuttings, Rock	6
Earthworks Cuttings, Soft	56
Earthworks Embankments, Soft	54

Asset type	Sample numbers
Retaining Walls	28
Footbridges, Decks	16
Footbridges, Supports	25
Overbridge Supports, Arched Masonry	28
Overbridges Deck, Masonry	21
Overbridges Deck, Concrete	22
Overbridges Deck, Metal	22
Overbridges Supports, Nonarched Masonry	73
Underbridge Supports, Arched Masonry	92
Underbridges Deck, Masonry	58
Underbridges Deck, Concrete	23
Underbridges Deck, Metal	51
Underbridge Supports, Nonarched Masonry	121
Coastal & estuarial defences	24

We have not analysed the samples by route, but they were spread around the zones fairly evenly.

The lengths of Earthwork samples are tabulated below:

MdRef	Model	Number of studies	Total Length (km)
ECR	Earthworks Cuttings, Rock	6	4.674
ECS	Earthworks Cuttings, Soft	56	24.877
EES	Earthworks Embankments, Soft	54	23.115
ERW	Retaining Walls	28	8.304

13. How were the actual samples chosen?

Most of sample structures were identified early in the SACP study. A series of attributes were determined and various combinations were selected to ensure a reasonable spread. Each zone was asked to identify structures meeting the combinations of attributes they were given. They were asked to avoid structures that were very untypical of their general stock. The same assets were used as samples in CECASE including some that had been identified but not used in SACP. The sample assets were re-assessed by the engineers using updated information to determine the appropriate management regimes for the different policies. The specification of the first 63 assets is set out in Appendix 1 as an example.

14. **How does Network Rail ensure that the samples used were representative and not a skewed sample?**

    In a Bayesian process they do not need to be fully representative, as the prior model helps to determine the frequency with which different attributes are present. The aim of the sampling was to provide reasonable coverage of the range of variation in the assets.

15. **Was consideration given to possible assessor's bias in relation to the samples that were used? If so, what was the process for correcting the bias or verifying the validity of the data? Were the samples subject to a peer review?**

    The sample results were determined in many cases by a panel of experts drawn from Mouchel and Network Rail. In other cases the estimates were produced by engineers who had been involved in a number of panel meetings. All results were checked for consistency of policy application by a Mouchel engineer (Hari Indran).

***Prior* Model**

16. **What is the basis of the degradation assumptions and how have these been validated?**

    The degradation assumptions originally came from panel meetings of engineers as part of SACP. These were reviewed with Mouchel engineers in the light of new policies. These initial assumptions were used for masonry and metal and concrete underbridges which are by far the most dominant groups. For other types of bridge where the sample was small, the rates of deterioration implied by the studies were compared with the original assumptions used for the prior models in SACP. Where the difference was significant, the prior assumptions were adjusted by 50% to 75% of the gap. Care was taken to ensure that the adjustments did not cause ABLE to learn from the opposite direction as there is a discontinuity at this point.

    For earth structures the assumptions were developed by Geotechnics engineers in NR and Mouchel working together.

17. **How are the prior volumes of work calculated?**

    Explained in the presentation by a modelling process relying on a Markov approach.

18. **How have the unit costs been calculated? How are the unit costs scaled for the different sizes of structure? What is the process for validating the unit costs against outturn costs?**

    NR will explain the long process to ensure costs of samples were consistent with their current costs. *prior* estimates were derived from SBMT where appropriate or from output of SACP updated for inflation and efficiency. As dimension data was not available for most types of structure, costs were scaled by span. Linear assets were scaled by length, allowing for the fact that only a part of an asset length might be worked on in carrying out an intervention.

19. **Is the model updated with fresh data on a regular basis? How often is the forecast updated?**

    This is not currently undertaken, but ways of updating are under consideration. This update would be infrequent as the process is very labour intensive.

ABLE Model

20. **Has the ABLE model been stratified to reflect the different types of structure and route etc?**

    A separate model has been used for each asset type and each major element. No stratification is carried out by route as this would introduce a substantial number of new parameters and could weaken the learning process considerably.

21. **Has the model been set up to learn from all of the posterior samples or only from the relevant samples? For example for masonry underbridges does the model only learn from masonry underbridges samples or all of the bridge samples?**

    The model learns only from the assets of the type relating to the model (e.g. masonry decks).

22. **What is the process for entering and running the ABLE update?**

    The standard update process is used to run the model using variance learning on the first pass. Where this fails non variance learning has to be used. There are some variables where the study estimates are nearly all zero and these tend to produce instability in the variance learning. For the most important variables the value of the t parameter (the learning factor) is compared to average movement from prior to study. If this appears to be at odds after taking account of outliers the variances would be checked and if justified changed. In some cases non variance learning was used to deal with difficult sets. The assumption of normality in ABLE can lead to difficulties in updating, as it will sometimes reduce the weighting attached to data that is considered to be valid for inclusion.

23. **What is the basis for assigning the random and systematic variances and what processes are in place to ensure integrity?**

    The study variances were derived in SACP by group meetings. The most important variance of prior minus study is estimated from the data to ensure that the learning is appropriate.

24. **On what basis is the degree of learning from the posterior model setup?**

    This is produced automatically by ABLE once the variances are reasonable. It is a function of the systematic and random variances (prior minus study variances) and of the number of studies. It is output in the .aur file as the value of the parameter t (with sign reversed).

25. **What peer review processes are in place to ensure consistency and to elicit the variances?**

    See question 24

26. **What is the overall accuracy of the ABLE output?**

    Data requested at meeting of 12/03/08 for metal underbridges will be sent through as soon as possible.

27. **Can Network Rail provide a list of volumes of work in terms of minor works, intermediate works and renewals volumes?**

    Output is available by standard NR categories as used in SBMT.

Sensitivity Analysis

28. What sensitivity analysis has been undertaken?

A sensitivity analysis was undertaken (under an earlier stage of the project) to explore the impact on the overall cost profile of changes to possession costs and to discount rates for different policies. The impact of changes was analysed within the detailed engineering studies and the overall impact on the overall population was then estimated in a qualitative way. The same broad conclusions would apply to the most recent CECASE results.

Further sensitivity analysis was undertaken when developing CECASE results to establish the preferred management approach for each study asset for each Policy. Where 2 or more solutions were identified for a particular asset with similar discounted costs, the sensitivity of the solution to assumptions about deterioration and intervention costs was undertaken. This helped inform the engineers of the preferred management solution which then fed into the final results.

Please note that the full model process does not lend itself to traditional sensitivity analysis as the work involved would be very substantial for even a limited analysis. This was discussed at the meeting with ORR on 12/03/08.

29. If volume of samples was increased by 10% say, to what extent would Network Rail expect to increase the accuracy of the output?

There would be a limited benefit to overall model results by increasing the sample size by 10%. The increase in overall accuracy will depend upon how many additional samples of each type were included and this was considered when the numbers and distribution of samples was originally chosen.

The posterior variance parameters in ABLE indicate that learning had moved substantially towards the study average so that additional studies will have only a limited effect on variances (this is a general observation as there are several thousand of these parameters in the models). The original estimate was that the 90% limit would reduce from 21% in SACP to 17.5% in CECASE assuming that the studies were of the same quality. On the same basis an extra 10% of samples spread across asset types proportionately would reduce the limit to 17%.

If the extra samples were concentrated on earth structures and underbridges the limit would reduce by further 0.25 percentage points. The studies are however better, particularly the costing and this would bring the variances down further.

30. If unit costs are increased by for example 20% or so, what effect does this have on the model?

If the sample costs were increased by 20% it would raise overall costs by between 15% and 18%. If prior estimates were raised by 20% is would probably add 2% to 3% to costs.

31. If the base data say the age of all bridges was 20 years older or younger what effect would this have on the output?

Model is based on ‘as found’ condition rather than age.

Validation of the Outputs

32. What is the process for validating the outputs and what corrections are made?

Overall checks by Mouchel engineers and NR to ensure results acceptable. Main checks are on study results and ABLE learning.

33. What correlation analysis has been carried out against SCMI information?

The prior model is driven off SCMI data and hence the estimates are related to SCMI scores. No data was available showing the relationship of SCMI to actual schemes in the workbank so the trigger for interventions within the prior model was set on an assumption that work would be carried once an element score fell to the 40% to 50% band. This was described in greater detail at the meeting of 12/03/08.

34. What correlation has been carried out against historical volumes of work?

NR are looking at this.

35. How closely does the CECASE model correlate to bottom up forecasts over 1, 3 and 5 years?

Model is not accurate at that timescale. It is intended to show trends of different policies over 30 to 100 years. Data for early years indicates a need for early investment at a level that is not practical so bottom up forecasts are constrained to provide a practical build up to higher longer term levels.

36. When an actual intervention has taken place how is this information used to update the model and/or the samples used to refine the model?

Currently not updated but Mouchel are thinking about this problem for the future. It will be for NR to decide how they wish to proceed.

37. How are actual interventions used to update the STAMP (Structures Asset Management Plan) model and accuracy of the decision tree?

STAMP has only been used to look forward against a policy view. Existing problems are known and taken into account but NR response may not yet be consistent with any particular policy as this is developing.

Information Request to Network Rail No.2

Questions related to the Functional Specification (Version 1.9 w)

Statistical Model (page 24)

38. **It appears from the text that the prior and posterior models did not agree (para 7.5 page 24). Normally when the prior and posterior models do not agree, then additional sampling is required in order to reduce the uncertainty for the particular groups. Was additional sampling sought in order to reduce the uncertainty?**

    In some cases the study cycle times differed substantially from the initial assumptions being used to run the prior model. These assumptions were based on SACP and updating discussions with the engineers. The reason for the difference was that the estimates produced by engineers were based more on a minor element view than on the major elements included in the model. If several minor elements require repair at different frequencies it leads to a higher frequency overall and this is difficult to judge subjectively particularly as there is little experience of working with the new policies. Where the differences were substantial the assumptions in the prior model were changed to move closer to study actuals. The update process was run on the amended prior data. The corrections would not have affected final results greatly for longer term averages but would have affected profiling because of the loss of observations on short period data (see response to question 2). In practice these changes affected only the overbridge models as the sample numbers for these were lower than for underbridges. They account for a relatively small part of the expenditure. No adjustments were made for underbridges.

39. **Further to question 12, could NR please provide a breakdown of the samples in terms of their SCMI scores? Where the CECASE samples chosen have had a previous SCMI assessment(s) then could NR provide the corresponding year and previous score(s) also.**

    This is being extracted. It is Unlikely that any previous data is available.

40. **Before the correction was applied (paragraph 7.5 page 24), did the prior model show that the assets were better or worse than the sampling had suggested?**

    The learning was nearly always upwards. In other words the prior models underestimated the need because cycle rates were too long. The extent of this was reduced by the adjustments made for overbridges.

41. **Regarding the correction for bias, could NR please explain the following:**

    - **How was the bias calculated?**

      See above

    - **What was the magnitude of the correction?**

      Very little effect on updated estimates for longer term averages. Some effect on profiles within first 30 years and last 70 years.

    - **How was the bias correction factor applied to the non-sample base?**

      It was used for the whole prior estimation process

    - **What was the uncertainty of the overall prior model?**

      Sample results being calculated.

- **Was the prior uncertainty calculated by asset group (e.g. bridges, earth structures etc)? If so, can you please provide the prior breakdown?**

  It will not be possible to determine results for the whole model in the available time.

- **What was the improvement in the confidence level from prior to posterior?**

  Being produced for sample.

42. Under paragraph 7.6 it states that there was substantial gaps in the data, could NR please quantify the size of the gaps in terms of quantity and quality? How were the gaps filled?

The only gap filled at the input stage to modelling was RA. Where this was not available but SCMI was available, it was replaced with a probability of being under strength, up to strength or unknown. Probabilities were estimated from assets for which data was available.

43. Para 7.6 goes on to say that "Where little data was available…the asset was omitted from the modelling process." How were the "omitted" assets accounted for in the model?

For assets where significant data was not available estimates were taken at the end of the process by taking averages of intervention probabilities from modelled assets of the same type. This process is incorporated in the PPT.

44. Please explain briefly how the "Markov process" is applied? Is this a Monte Carlo simulation?

Described in presentation

Calibration of the condition model (page 39)

45. It states that the calibration of the model was carried out manually. What was the quantity and quality of the historical information which underpinned the condition, deterioration rate, and intervention point analysis? How was a "reasonable range" defined?

No historical data is available. Engineers made best predictions for sampled assets looking at available historical data. This was converted to a deterioration rate for each SCMI element for each sampled asset. An average value for the asset type was taken into the condition model.

***prior* estimates for earth structures (page 70)**

46. The sixth paragraph mentions that the whole length of an earth structure is subject to a repair whereas in reality only a small section of the structure may have or will fail in the future. How was the model calibrated to reflect actual practice? What learning took place from the earth structure samples?

The costs applied to the interventions were calculated as cost per km of structure even if work was only on a small section. This would clearly produce a small cost factor if only a part of the structure was worked on. When multiplied by the length of the assets worked on this gave correct results. Output measures are however for kms of structure considered to need work. A structure is defined in terms of a natural structure.

Statistical Update Process (page 72)

47. **What was the extent of the systematic error observed between the prior and study estimates?**

    There were thousands of base quantities involved in the model. A sample for metal underbridges is being extracted from the model and will be sent to ORR early next week.

48. **What were the criteria or limits of accuracy that determined the size of the study samples?**

    The information in question 29 provides the basis for this. We examined the distribution of the overall variance and estimated the improvement we would obtain from increasing the number of samples for each asset type. The final decision was based on a balance and number that provided an economic improvement. We had reached the point where the improvement available from additional samples appeared to be relatively small. Improvement in base data and sample quality was felt to be likely to provide more improvement.

49. **The third paragraph states that final estimates of interventions, costs and numbers of possessions are produced for individual assets. How are interventions graded or grouped (e.g. minor, medium, renewal) or other? Can NR provide a breakdown of the interventions, costs and numbers of possessions over the next ten years?**

    Interventions are categorised using NR current convention. The outputs are available from the PPT.

50. **Can the model be used to look back in time? If so can the model produce a list of interventions for bridges from 2002/03 up to 2006/07 broken down into interventions (e.g. minor, medium and renewals)?**

    No, and the output is probabilistic (probability per five year period in general) which makes such comparisons very difficult over the short term.

Annex D Data Assumptions (page 102)

51. **Under the section dealing with SCMI data it states that where an asset had multiple examination dates the most recent was used? What analysis was carried out to ensure that there was consistency between the previous and current inspections? How did the changes in condition from one examination to the next inform the degradation assumptions?**

    SCMI data (comprising of 400,000 entries at element) was processed and used as input data to represent starting condition within the prior model. It was assumed that most recent SCMI data for an element was the best reflection of its current condition. There were few instances of repeat data (at that time) so they were not of practical value for the modelling exercise or to inform degradation.

Appendix 1

SACP – Structure type for which data is required for STAMP input					
ZONE	**Material**	**Span**	**Old Condition**	**SCMI**	**AGE**
East Anglia	Masonry – Stone Under Bridge	4 – 5m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with Jack Arch Deck Under Bridge	13 – 20m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	8 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
Southern	Masonry - Brick Under Bridge	5 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with steel deck Under Bridge	8 – 12 m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	10 – 15m	Good Fair Poor	70 – 100 40 – 69 <40	Various
Great Western	Masonry – Brick Under Bridge	4 – 5m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with Steel deck Under Bridge	>20m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	8 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
Midlands	Masonry – Brick Under Bridge	5 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with Jack Arch Deck Under Bridge	13 – 20m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	10 –15m	Good Fair Poor	70 – 100 40 – 69 <40	Various
North West	Masonry – Stone Under Bridge	4 – 5m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with steel deck Under Bridge	>20	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	8 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
London North East	Masonry – Brick Under Bridge	5 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with concrete deck Under Bridge	8 – 12m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	10 – 15m	Good Fair Poor	70 – 100 40 – 69 <40	Various
Scotland	Masonry – Brick Under Bridge	4 – 5m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Steel with Jack Arch Deck Under Bridge	13 – 20m	Good Fair Poor	70 – 100 40 – 69 <40	Various
	Concrete Under Bridge	8 – 10m	Good Fair Poor	70 – 100 40 – 69 <40	Various

Railtrack / Mouchel SACP Project

Following the meeting Birmingham, and it was good to see so many there, sorry that some of you were not available, we attach the sample requirements for the first 63 samples.

The broad outlines are as follows:-

1. Each zone is asked to identify 9 Structures according to the table attached, you will note that some of the spans and materials vary slightly according to the zone. It would be helpful if you could accommodate these requirements to enable us to gain a first representative sample.

(Further sample requirements for samples across a broader range of structures with some very specific specifications will be outlined as the second phase of the Structures sample for STAMP.)

2. The Zone co-ordinating Engineer should now identify the actual structure that is to be assessed and collect and collate the information on the structure from the files and records available in the office and in line with the detailed requirements for data collection outlined at Birmingham.

3. The zone co-ordinating Engineer should liaise with their nominated Mouchel Engineer and agree a programme of when the data will be available and a check list of what detail will be available.

4 The plan referred to in 3 above, should be complete by 22 MAY '01 and sent to Nigel Stapleton.

5 Railtrack engineers will be required to consider what is the maintenance strategy required for each structure and in liaisons with the Mouchel "STAMP input engineer" will decide which options should be entered for evaluation.

6 It is NOT envisaged that the Contractors will be present for the first set of samples.

The above requires that staff are available for finding this information in the zone offices and for engineers to make the necessary engineering judgements about the individual structures.

Following discussions within Railtrack we are able to assist on a call – off basis assistants who could help to recoup data under Railtrack guidance and/or Engineers with necessary experience. Nigel Stapleton will be co-ordinating these requests and liaising with various suppliers.

His contact numbers are: Office; Tel. 01932 337155 Fax 01932 356122 Mobile 07785 510774 Midweek evening 01932 840947 Weekend Tel 01453 872687 Fax 01453 873114

At the Birmingham meeting we mentioned that we would meet on May 25th,this will not be required for all staff, Mouchel STAMP inputting engineers will be sent a separate note of a date to meet around that time.

In the meantime Nigel Stapleton hopes to visit zones to ensure the set up has taken place and we would ask that you afford him a short time to ensure the project is taking off in the right direction.