Investigations for the Development of Simulative Test Methods for the Durability of Thin Surface Course Systems

Sub-Task 1: Critical Review of DRaT Project Outputs

Project number: 60580090

6 November 2019

Quality information

Prepared by	Checked by	Verified by	Approved by
Giacomo D'Angelo	Chibuzor Ojum	Aswandaru Widyatmoko	Aswandaru Widyatmoko
Dr Giacomo D'Angelo Pavement Engineer	Dr Chibuzor Ojum Senior Engineer	Dr Daru Widyatmoko Technical Director	Dr Daru Widyatmoko Technical Director
Kieran Dhandwar			

Revision History

Revision	Revision date	Details	Authorized	Name	Position

Distribution List

# Hard Copies	PDF Required	Association / Company Name

Prepared for:

Matthew Wayman
Senior Pavement Advisor
Highways England (HE)

Prepared by:

Dr Giacomo D'Angelo
Pavement Engineer

AECOM Infrastructure & Environment UK Limited
12 Regan Way
Chetwynd Business Park
Nottingham NG9 6RZ
United Kingdom

T: +44 (115) 907 7000
aecom.com

Table of Contents

Figures

Tables

Executive Summary

In June 2018, Arup AECOM Consortium was commissioned by Highways England (HE) to conduct Task 1-614. The title of the project is “Investigations for the Development of Simulative Test Methods for Durability of Thin Surface Course Systems”. This project includes five sub-tasks. This report details the work undertaken under Sub-Task 1: Critical Review of DRaT Project Outputs.

The aim of this sub-task is to carry out a critical review of the outputs and results of the Conference of European Directors of Roads (CEDR) Development of Ravelling Test (DRaT) project. Based on this review, a method of measuring the ravelling characteristic of a Thin Surface Course System (TSCS) pavement in the United Kingdom (UK) is be proposed.

The main findings from the critical review of the DRaT project are the following:

- Ravelling mechanism is affected by several interdependent factors including: the presence of water, the aggregate type and gradation, the binder content and type, the mix design, the quality of construction and the weather conditions.
- Porous asphalts are more prone to ravelling than dense asphalts. For Thin Surface Course System (TSCS), it was found that aged asphalt has a lower resistance to ravelling whereas very low correlations were found with the binder content, binder type and aggregate size.
- With the aim of investigating four existing scuffing tests, the DRaT project tested three different asphalt mixtures (from open to dense), manufactured in their standard version and two lower quality variants. High quality and repeatability of the slab manufacturing and compaction was observed. Visual inspection was the most critical acceptance criterion, leading to the highest number of slab rejection.
- The main differences between the devices testing protocols are related to the sample dimensions and shape, the conditioning, the loading type and amplitude, the type of tyre used to apply the load (and consequent contact area), the test duration and the measured parameters. Harmonisation proposed by the DRaT project focussed mostly on conditioning of samples and measurement.
- Results were in terms of absolute weight loss. DRaT statistical analysis concluded that:
  - For the same material, the amount of material loss was not consistent between different tests (and between the same test used by different lab).
  - The overall weight loss for the Stone Mastic Asphalt (SMA) slabs was very low and hardly any physical ravelling was observed after completion of the tests.
  - The rate of material loss (slope of the curves) behaved differently, depending on the test device.
  - No correlations were found between mass loss and either density or texture.
  - The results showed large coefficients of variation, with the consequence that the discrimination potential strongly depends on the number of slabs tested.
  - The correlation study showed that scaling factors were not constant and depended on the material type. Thus, a single ‘universal’ scaling factor could not be obtained.

The suitability of these tests to discriminate between good and bad performing TSCS under English weather conditions, was assessed through an independent analysis, using normalised results. The analysis focussed on BBTM (very thin asphalt layer) and SMA, which are the most similar to TSCS. This analysis concluded that RSAT (Rotating Surface Abrasion Test) is the most suitable test (among those used in the DRaT project) to assess resistance to ravelling of TSCS. DSD (Darmstadt Scuffing Device) and ARTe (Aachener Ravelling Tester) are possible alternatives while TRD (Triboroute Device) was evaluated as least suitable for this purpose. Nevertheless, none of these testing methods take into account the effect of water on resistance to ravelling.

Alternative test methods that can better discriminate performance of TSCS in UK were explored, focussing on the possibility of including damage from water and/or using test methods available in the UK. The possibility of using the alternative test methods explored may be considered in future stage of the project subject to budget constraints.

Based on the critical review of the DRaT project carried out, the following recommendations are made for the next stage of the current project.

The air voids content was the selected as the key factor influencing ravelling resistance of TSCS. Therefore, two TSCS are proposed for this research and they will be manufactured to have 6% and 14% air voids content.

For the slab manufacturing, our recommendations include sieving, conditioning, mixing and compaction. A minimum of four slabs for each configuration will be manufactured. These will provide sufficient Detection Power as well as Discrimination Power for each device to assess different quality TSCS materials.

These samples should also meet the following acceptance criteria:

- Measurement of slab thickness using a calliper at eight locations for each slab. The maximum thickness difference should be 1 mm.
- Measurement of density to check that manufactured specimens are at the targeted air voids content ±1%.
- Visual inspection looking for greasy spots (>2 cm not allowed), scarce and lean spots (>50 cm^2 not allowed) and irregular distribution of the slab mix near edges.

Based on the independent evaluation of DRaT results, the test devices recommended for next stage are:

- RSAT
- DSD

To evaluate the reproducibility of the test methods, two laboratories per each device are recommended for carrying out tests.

The possibility of including water damage and/or accelerate ageing will be discussed with the respective laboratories.

1. Scope

In June 2018, Arup AECOM Consortium was commissioned by Highways England (HE) to conduct Task 1-614. The title of the project is “Investigations for the Development of Simulative Test Methods for Durability of Thin Surface Course Systems”. This project includes five sub-tasks:

- Sub-Task 1: Critical Review of DRaT Project Outputs
- Sub-Task 2: Devise the Test Programme
- Sub-Task 3: Execute the Test Programme
- Sub-Task 4: Results and Implementation
- Sub-Task 5: Project Management

This report details the work undertaken under Sub-Task 1. The aim of this sub-task was to carry out a critical review of the outputs and results of the Conference of European Directors of Roads (CEDR) Development of Ravelling Test (DRaT) project. Based on this review a method of measuring the ravelling characteristic of a Thin Surface Course System (TSCS) pavement in the United Kingdom (UK) will be proposed and/or developed.

The key questions to consider include: Can the test methods assessed in DRaT project be harmonised? Do they produce comparable results? Can the results be correlated from one equipment to another? Can any of the four methods discriminate between good and bad performing TSCS under English weather conditions? Is there an alternative test method that can better discriminate performance of TSCS?

Based on the stated aims, the following objectives were identified:

O1. Consider the usefulness of the findings in relation to the Highways England context and the materials that are specified for use in UK, with focus on TSCS

O2. Evaluate which of the test kits reported in the DRaT project might be useful to test TSCS, if any

O3. Explore alternative test methods available in UK which can be used and/or adapted to current specification

2. Review and evaluation of DRaT Project Deliverables and PD CEN/TS 12697-50

The following sections report on the critical review of the DRaT project framework, the material and methodology used and main findings. This review considers the usefulness of the DRaT Project Report findings to TSCS used on the Strategic Road Network.

2.1 Background of the DRaT project

The DRaT project was undertaken via the CEDR framework, under Call 2014, Asset Management and Maintenance (Wayman, 2015).

The partners of the project included: The Netherlands Organisation for Applied Scientific Research (TNO), Belgian Road Research Centre (BRRC), BAM Infra Asfalt bv (BAM), Heijmans Integrale Projecten (Heijmans), Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTTAR), Technische Universität Darmstadt (TUD) and Ingenieurgesellschaft für Straßenwesen (ISAC).

This project aimed at investigating four existing scuffing tests in order to identify (Wayman, 2015):

- The extent to which sample preparation needs to be standardised.
- The most effective method of measurement in terms of extent of differentiation, validity as a measure of ravelling and practicality.
- Whether the results from one or more scuffing machines can be validated from experience on site.
- Whether the results from different scuffing machines can be converted to a common measure.
- Estimates of the precision of the results with each scuffing machine or, if the results can be converted to a common measure, of the common measure.
- Whether the results from either pair of similar machines are comparable and their results are reproducible.
- A procedure to identify if other scuffing machines can be used for the standard test.

The outputs of the project were disseminated through 9 deliverables, which were reviewed as part of this sub-task.

2.2 The ravelling mechanism

The mechanism of ravelling (or scuffing or fretting) of asphalt pavement consists of the loss of fine and then coarse aggregate from the road surface by the passage of or weathering (Nicholls, De Visscher, Hammoum, & Jacobs, 2016).

This mechanism occurs when the bond between binder and aggregate reaches a critical point due to the shear forces due to traffic (De Visscher & Vanelstraete, 2017). This is due to an excessive deformation (and therefore stress) in the binder, which leads to a fracture (either cohesive, adhesive or mixed), and consequent loss of aggregate particles from the asphalt pavement. Once the mechanism initiated, the deterioration becomes progressively faster until failure. This failure mechanism is restricted to the surfacing, which is subjected to the scuffing forces from the vehicle tyres when changing direction and/or braking/accelerating (Nicholls et al., 2016). However, other causes/factors can initiate or accelerate the ravelling process. These are detailed in Section 2.3.

2.3 Main factors affecting ravelling in asphalt pavement surfacing

The potential causes for ravelling include lack of sufficient binder, inappropriate aggregate grading, poor adhesion between the binder and the aggregate, poor compaction, aggressive scuffing by the traffic, bitumen ageing and effect of climatic conditions, among others. These causes are often interdependent, making it difficult to assess the theoretical potential to ravel of an asphalt mixture in the design stage.

Deliverable D.2 of the DRaT project focussed on the literature review of the ravelling phenomenon, the causes/factors affecting it and the possible practices to prevent or mitigate ravelling. Table 1 summarises and reviews the findings from D.2 of DRaT.

Table 1. Factors affecting ravelling in asphalt pavement surfaces

Factor	Sub-Factor	Description	Prevention/ mitigation	Comments
Presence of water	**Permeability**	Water permeability influences moisture damage. Mixtures with higher air voids content are likely to be interconnected allowing water to travel through the mat, stripping the bitumen from the aggregate particles. This can result in a loss of bond that leads to ravelling.	Use of asphalt mixture with low air voids content.	As additional factor to be considered, hydraulic pumping is caused by the action of vehicle tyres on a saturated pavement surface, i.e. water is forced into surface voids in front of the vehicle tyre. Low voids mixtures are more prone to suffer this mechanism.
	Surface macro-texture	If water is removed from the surface by interconnected voids, the pressure is reduced and so damage is less. However, when negatively textured surfaces are filled with detritus damage occurs due to water retention.	Use of asphalt mixture with adequate macro-texture.	HRA[1] is less permeable than SMA[2] and BBTM[3].
Materials	**Aggregate-binder affinity**	Poor aggregate-binder affinity increases the likelihood of ravelling. The presence of water molecules reduces the affinity of aggregate-binder.	The type of aggregate influences the degree of affinity. Aggregates such as basalt and limestone are generally more generally more affine to bitumen than quartz and granite. Hydrophobic aggregates are preferred.	Silicates absorb water and reduce the affinity between the constituents.
	Aggregate cleanliness	When aggregates are dirty their adhesion ability is reduced due to the presence of dust of fine aggregate.	Aggregates should be cleaned before mixing.	Ability to do this depends on the site set up and where mixing occurs.
Mix Design	**Air Voids**	Ravelling is closely related to in-situ voids, with the higher the voids the greater the ravelling.	Design towards low air voids content.	This factor has to be balanced with other factors such as permeability and macro-texture.
	Binder content	Low binder content (1-2% below the optimum) results in a lack of "glue/bond" between constituents: low binder film thickness. Damage in asphalt mixtures can occur within the mastic (cohesive failure) or at the aggregate-mastic interface (adhesive failure). However, thicker bitumen films do not noticeably increase the resistance to ravelling.	Use of optimum bitumen content without negatively influencing rutting/bleeding.	Most likely, for a given asphalt mixture, there is a threshold film thickness (and binder content) below which the expected ravelling increases as thickness decreases.
	Binder grade	Typically, binders with lower stiffness (or viscosity) improve the resistance to ravelling, as reported in several studies. According to Hunter et al. (2015), ravelling is most likely to occur at low temperatures and at short loading times when the stiffness of the binder is high.	DRaT D.2 concluded that the use of more viscous binders will reduce the tendency for ravelling.	The conclusion of D.2 is possibly based on Van Loon and Butcher study. This study focused on asphalt including different % of RAP[4]. The correlation (relatively low) reported by Van Loon and Butcher refers to the initial modulus of the asphalt mixture – not necessarily

Factor	Sub-Factor	Description	Prevention/ mitigation	Comments
		Apparently in contrast, it has been reported by Van Loon and Butcher (Van Loon & Butcher, 2015) that along with decrease in stiffness of the asphalt mixture is associated an increased potential for ravelling (R-value of 0.38). See comments.		related to the binder stiffness (and therefore viscosity). In addition, the same study reported no correlation between viscosity and ravelling.
	Binder type (use of polymer modification)	Some studies indicate that the use of PMBs[4] can increase the resistance to ravelling whilst other studies did not find noticeable influence of PMBs on asphalt life extension.	Overall, the advantage of using PMBs is uncertain.	Most of the studies reported in DRaT D.2 focussed on porous asphalt. None of the studies found negative impact of PMBs on ravelling.
	Aggregate grading and filler content	Larger aggregate size and gap-graded/open-graded mixtures increase ravelling due to larger number of shear planes.	Use smaller aggregates and/or well-graded mixtures can give a higher number of resisting shear planes, with increased resistance to ravelling.	The more open asphalt mixtures tend to be more susceptible to ravelling because the aggregate particles are not "protected" by being embedded in the mortar on all sides. Therefore, rather than the aggregate size itself, the mixture grading is the key factor.
Construction quality	**Compaction**	Poor compaction results in higher air voids (%) thus reducing the adhesion of particles in the mat. It is claimed that this is the most important factor affecting ravelling.	Compaction must be completed to the specified range. A minimum of 92% of maximum density achieved on site it is claimed to mitigate ravelling and promote a durable pavement. The use of intelligent compaction technologies can assist ensuring quality during work execution.	Excessive compaction can lead to other undesired distress such as rutting. Thus, air voids should be optimised.
	Segregation	Segregation can result in areas with high air voids (%), more prone to ravelling.	Ensure segregation does not occur during construction.	Segregation affects also other types of distress.
	Layer thickness	Excessively thin layer does not provide sufficient room for the aggregate to reorient itself into a dense configuration. This increases the potential of ravelling.	Ensure the layer is at least two times thicker than the nominal aggregate size.	Pavement and mix design should take into account this factor to allow reorientation of constituents and correct compaction.
	Asphalt temperature	If asphalt is not sufficiently hot when laid, poor compaction can occur due to the bonds already formed. This occurs especially at the ends of loads.	Ensure asphalt is at a correct temperature when laying and compacting. Suggested minimum temperature is 145°C at mid-depth of asphalt layer. This ideal temperature depends on a number of factors.	Too hot temperature during construction can lead to excessive bitumen ageing. This factor should be considered during construction works.
	Wet weather	Laying in wet weather impacts the compaction quality and the adhesion, leading to increased potential of ravelling.	Laying should occur in dry conditions (no fog, rain or high humidity).	A warm enough asphalt may be able to evaporate moisture at low levels
	Joints	Ravelling often initiate where excessive/poor longitudinal joints have been cut.	Joints should be cut to the correct size with care, especially longitudinally.	Workmanship should be controlled. Avoid material segregation. Use pavement joint heater or joint sealant.
In Situ conditions	**Bitumen ageing**	Premature ravelling can result from overheating during mixing while long term ravelling can result from weathering/brittleness.	There should be no overheating (temperature typically above 165°C) during mixing as this can cause premature aging.	Overheating and its impact on the amount premature aging depends on the binder type.
	Weather	Cold weather can affect ravelling in different ways: the action of freeze-thaw mechanism can break the bonds in the mixture; low temperatures will make bitumen	Ensure that bitumen is adequately designed for the climatic in-situ conditions.	In DRaT D.2 it was reported a study finding that the maximum adhesive performance of porous asphalt was achieved at 0 °C, whereas the adhesion at -

Factor	Sub-Factor	Description	Prevention/ mitigation	Comments
		more brittle. These result in higher ravelling potential. Warm weather can affect ravelling to the extent that softening of the binder can reduce the adhesion strength between constituents.		10°C was about equal that at +10°C.
	Substrate	For layer thickness lower than two and a half times the nominal aggregate size disaggregation may propagate upwards from the bottom of the layer. Therefore, the substrate stability and bond coat efficiency can influence the ravelling.	Ensure adequate bond coat.	This factor is more relevant for very thin layers.
	Traffic loading	Vertical static or dynamic loading (unless associated with other factors) is not typically related to ravelling. Areas of braking, acceleration and cornering are more prone to ravelling.	If required, surfacing may need specific design.	Binders with elastic recovery may exhibit better performance in these areas.

Notes:

[1] HRA denotes Hot Rolled Asphalt.

[2] SMA denotes Stone Mastic Asphalt.

[3] BBTM (Béton Bitumineux Très Mince) denotes very thin asphalt layer.

[4] PMB denotes Polymer Modified Bitumen.

Based on this review, the main findings relevant to the Highways England context are the following:

- The presence of water affects ravelling mostly with two mechanisms: stripping of bitumen from the aggregates, mostly in asphalt mixtures with interconnected air voids; pumping effect, caused by the action of tyres on a saturated pavement surface (low voids mixtures). These two apparently contrasting needs could be balanced by using dense mixtures with adequate macro-texture.

- The use of aggregate with good affinity to bitumen reduce the likelihood of ravelling for adhesion failure.

- Open asphalt mixtures tend to be more susceptible to ravelling because the aggregate particles are not embedded in the mortar on all sides. The use of small size aggregate, well graded aggregate and low voids mixture contrast this mechanism.

- Binder characteristics affect ravelling in various respects: binders with low viscosity improve the resistance to ravelling; low binder content increase the likelihood of ravelling whereas the advantages of using PMBs is uncertain.

- The quality of construction works is essential to improve the resistance to ravelling: to ensure that air voids requirements are fulfilled, asphalt should be compacted to the designed density at the optimum temperature range, possibly in dry conditions; to ensure aggregate re-arrangement the layer thickness should be at least twice the nominal aggregate size; poorly cut joints are often a cause of ravelling initiation.

- The action of traffic loading is relevant to ravelling in areas of braking, acceleration and cornering.

- The above factors are often interdependent, making it difficult to assess the theoretical potential to ravel in the design stage.

2.4 Review of findings from previous studies

Deliverable D.3 of the DRaT project reviewed available data on the performance of various mixtures with respect to ravelling on site. Table 2 summarises and reviews the findings from D.3 of DRaT.

Table 2. Summary and review of findings from previous studies

Location and asphalt type	Findings	Comments
Netherlands – Porous Asphalt	• Extent of ravelling can vary within the same asphalt mixture (sometimes significantly). • High binder contents do reduce the likelihood of ravelling. • The use of PMBs does not reduce the likelihood of ravelling. However, adding modifying the bitumen was necessary to prevent binder drainage at for higher bitumen content. • Slag aggregate makes asphalt more susceptible to ravelling.	For Visual Condition Survey (VCS), distress was measured as a % lane length (lane length was 100 m) affected and then the % per m^2 stone loss. Increased bitumen requires use of fibres (cellulose or acrylic) or bitumen modification to inhibit binder drainage.
Belgium	• Twin-layer porous asphalt is more susceptible to ravelling than dense asphalts.	Sites had daily traffic of 10,000 vehicles (16% heavy on weekdays, 5% heavy on weekends). Site was contaminated with dirt from nearby agricultural machinery.
United Kingdom	• Ravelling increases with age. • Resistance to ravelling can be ranked as (best to worst): SMA[1] → BBTM[2] → AUTL[3] • Higher binder contents and larger aggregate sizes tend to reduce ravelling.	Variables included type of asphalt, aggregate size, binder type and age of asphalt. Visual condition was assessed by a panel of experts and ravelling categories were discreet. Therefore, trend lines reported were considered only descriptive. Ravelling correlation with the variables analysed varied from low to very low: R^2 = 0.29 with the age of asphalt R^2 = 0.06 with the type of asphalt R^2 = 0.02 with the aggregate size R^2 = 0.04 with the binder content

Notes:

[1] Surfacing was classified as SMA when the mixture had unmodified bitumen but incorporated fibres to carry relatively high binder content; generally had a polymer modified bitumen bond coat.

[2] Surfacing was classified as BBTM when the mixture had polymer modified bitumen and had unmodified binder coat.

[3] Surfacing was classified as AULT (Asphalt for ultra-thin layers) when the mixture had an unmodified binder but was laid on heavily polymer modified bond coat.

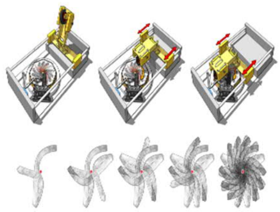
The above-reviewed studies confirmed that the air voids content play an essential role in the resistance to ravelling: porous asphalts are more prone to ravelling than dense asphalts.

Studies conducted on TSCS revealed that aged asphalt (more brittle bitumen) have a lower resistance to ravelling whereas very low correlations were found with the binder content, binder type and aggregate size.

2.5 Test methods used to evaluate ravelling

The DRaT project assessed four test devices for ravelling and proposed recommendations for revision of standard PR CEN/TS 12697-50. Table 3 provides a summary of these devices, their main features and lists the laboratories where each device was available.

Table 3. Summary of the test devices assessed in the DRaT project

Test Devise	Aachener Ravelling Tester (ARTe)	Darmstadt Scuffing Device (DSD)	Rotating Surface Abrasion Test (RSAT)	Triboroute Device (TRD)
Description	The specimen is fixed in a box which move forwards and backwards. During this movement, a set of two wheels with pneumatic tyres rotate over the loading table and the asphalt specimen, inducing shear stresses due to the combination of the lateral movement of the table and the rotation of the wheel set	The asphalt specimen is attached in a fixture oscillating 180°, which is mounted on a horizontal table moving forwards and backwards. During this movement, a vertical load is applied though a tyre	This test simulates the real condition of a tyre tread continuously deformed when in contact with the pavement. These deformations result in shear stress in the contact area. These stresses cause fatigue in the asphalt surface and lead, eventually, to aggregate loss	This test aims to measure the resistance of asphalt surfaces to tangential forces in a laboratory. It is composed of a braced vertical column supporting the load applicator and mounted on a classical hydraulic press and a roller-mounted horizontal table
				
Laboratories	TU Aachen BAM	TU Darmstadt BRRC	Heijmans	IFSTTAR
Sample dimensions (mm):				
Length	500 or 320	260	500	185
Width	500 or 260	260	500	247
Min. Thickness	30	25	-	-
Max. Thickness	80	60	-	-
Core dimensions	Core samples not explicitly covered but can be tested	Core samples not explicitly covered but can be tested[3]	Diameter of 150±1 mm and height between 30 mm and 60 mm – Three cores per test	300 mm in diameter
Slab dimensions (mm*mm)	500*500	260*260	500*500	400*600
Conditioning	(20±2)ºC for at least 4 h	(40±1)°C for 2.5 h	Test temperature for 14 h to 18 h - Preloaded with ≥ 20 kg for ≥ 1 h	(20±2)°C for 2 h to 3 h
Test temperature	18ºC to 25ºC	(40±1)°C In DRaT round robin BRRC also tested with the DSD at (20±1)°C	(-10±1)°C to (25±1)°C with standard (20±1)°C	(20±2)ºC
Other initial preparation	None specified	None specified	Removal of loose material	Removal of loose material
Initial measurements	Dimensions and mass; Photograph or 3-dimensional texture	Mass and photograph	No additional measurements to those in the main text required	Surface flatness; Macro-texture; Photograph; Dimensions and mass

Test Devise	Aachener Ravelling Tester (ARTe)	Darmstadt Scuffing Device (DSD)	Rotating Surface Abrasion Test (RSAT)	Triboroute Device (TRD)
Test loading[1]	(250±5) kg with a contact area of 108.7 cm^2 for an average contact pressure of 230 kPa	(1000±10) N with a contact area of 33.3 cm^2 for an average contact pressure of 300 kPa	(35.0±0.1) kg with a contact area of 5.7 cm^2 for an average pressure of 600 kPa	Average 2000 N with an amplitude of 500 N with a contact area of 11.2 cm^2 for an apparent contact pressure of 1330 kPa
Operation during test	Slab rotated 180° halfway through test	Vacuuming of loose grains and wiping off as required	Removal of all loose material by vacuum cleaner	Removal of all loose material by vacuum cleaner
Test duration[2]	600 cycles (100 minutes)	16 cycles (6 minutes)	86,600 passes (24 h)	10,000 cycles (4 h)
Final measurements	Visual, photograph and 3-dimensional texture (if available)	Photograph; Residue and loose grains from the asphalt specimen and the tyre	Aggregate loss after removal of rubber lost from tyre	Aggregate loss; Number of cycles to reach specified degree of degradation
Comments	Default test duration is designed for dense graded mixtures	Developed for the measurement of Porous Asphalt (PA). However, it has the capability of testing SMA and very thin surface courses; Binder type has a greater influence if testing is completed at 40°C rather than 20°C	Tests can be prematurely ended by high stone loss	Allows loading to be constant regardless of surface deformations; Temperature control is essential for repeatability

[1] The average pressure applied to the asphalt surface by each tyre of a typical tandem 10 tonnes axle load is 800 kPa (Hjort, Mattias, & Jansen, 2008) for an average contact area of 312.5 cm^2.
[2] This refers to the standard duration developed for porous bituminous mixtures, which have relatively short expected service life. DRaT project recommended to double the duration when testing longer-lasting asphalt mixtures.
[3] This is based on the additional information gathered from De Visscher and Vanelstraete (2017).

The main differences between devices are related to the sample dimensions and shape, the conditioning, the loading type and amplitude, the type of tyre used to apply the load (and consequent contact area), the test duration and the measured parameters.

The DRaT project conducted a study to harmonise these test methods, as reported in Table 4.

Table 4. DRaT proposed additional harmonisation and comments

Attribute	DRaT proposed additional harmonisation	Comments
Sample dimensions	Dependent on apparatus; harmonisation not considered appropriate	-
Core dimensions	Dependent on apparatus; harmonisation not considered appropriate	Not all devices are designed to allocate cores
Conditioning	(20±2)ºC for at least 4 h DSD: (40±2)°C for at least 4 h	-
Test temperature	(20±2)ºC other than DSD at (40±2)ºC	This parameter can be harmonised also for DSD (different loading conditions would be needed)
Other initial preparation	Removal of all loose material	-
Initial measurements	Visual inspection; At least one photograph; Sample dimensions;	-

Attribute	DRaT proposed additional harmonisation	Comments
	Sample mass; Macro-texture by patch method (possible with smaller volume); Dimensional texture by laser measurement (if equipment available)	
Test loading	Dependent on apparatus; harmonisation not considered appropriate	If the load amplitude can be varied for each device, it is recommended harmonisation in terms of applied pressure
Operation during test	No additional requirement	-
Test duration	Dependent on apparatus; harmonisation not considered appropriate	If test loading can be harmonised, it is recommended to harmonise test duration
Final measurements	As for initial measurement plus mass of loose material collected[1]	-

[1] Primary weight loss was used as main parameter; secondary weight loss was also recorded as alternative parameter

Further to the harmonisation proposed by the DRaT project, loading conditions, test duration and temperature could be varied to obtain a similar rate of damage for all the devices.

Results reviewed in the following sections can provide important data sets to improve future harmonisation of ravelling test devices.

2.6 Standard asphalt mixtures used and their variants

Deliverable D.5 of the DRaT project reports the methodology used to select materials to be tested. Three standard mixtures (to EN 13108) typically used for surface courses in Netherlands, France and Germany were selected. Each asphalt mixture had a standard mix design and two variants, as summarised in Table 5 and Table 6.

Table 5. Summary of the standard test mixtures used in the DRaT project

Asphalt Mixture	Country	According to	Binder	Compaction Temperature	Air Voids
Porous Asphalt (PA)	Netherlands	EN 13108-7	5.2% 70/100 bitumen	150°C	20%
Very Thin Layer Asphalt (BBTM)	France	EN 13108-5	5.6% 50/70 bitumen	160°C	12-19%
Stone Mastic Asphalt (SMA)	Germany	EN 13108-2	6.8% PMB 25/55 with 3% SBS[1]	155°C	3%

[1] SBS denotes Styrene Butadiene Styrene thermoplastic elastomer (elastic behaviour at room temperature and plastic behaviour when heated; not cross-linked allowing better workability).

Table 6. Summary of variant one and two for each of the standard mixtures reported in Table 5

Asphalt Mixture	Variant	Binder	Compaction	Air Voids
Porous Asphalt (PA)	One	5.2% 70/100 bitumen	**105°C**	20%
	Two	**4.2%** 70/100 bitumen	150°C	20%
Very Thin Layer Asphalt (BBTM)	One	5.6% 50/70 binder	**110°C**	12-19%
	Two	**4.6%** 50/70 binder	160°C	12-19%
Stone Mastic Asphalt (SMA)	One	6.8% PMB 25/55 with 3% SBS polymer	**105°C**	3%
	Two	**5.5%** PMB 25/55 with 3% SBS polymer	155°C	3%

The TSCSs typically used in the UK are proprietary materials having air voids content ranging from 4% to 19% (Khojinian, Parry, & Thom, 2016), intermediate between the SMA and BBTM used in the DRaT project. Therefore, assessment of results in relation to the Highways England context will focus on these two asphalt mixtures.

2.7 Asphalt mixture manufacturing method

All specimens were produced in a single laboratory (i.e. BAM) to reduce the variability between specimens. The method to prepare each batch of samples is described below:

- Sand and coarse aggregate were sieved according to BS EN 12697-2 (2007) in the following fractions: <2 mm, 2 – 5.6 mm, 5.6 – 8 mm, 8 – 11.2 mm, 11.2 – 16 mm, >16 mm. These fractions were then used to be dosed in the asphalt mix.
- The constituent components were pre-heated, and the asphalt was mixed in a Bear Varimixer mixer (type AR60/MK1). About 45 kg of asphalt mixture can be effectively mixed in about 90 seconds.
- The mixtures were poured into boxes of the required size. The boxes had wooden bottoms and steel side walls to reduce slipping between the slabs and the box. The amount poured was calculated taking into account the desired density and the density compensation for shrinkage during the cool-down phase. Temperature was constantly measured during the production and compaction process.
- A standard compaction roller (HAMM HD10 VV, a tandem roller with a mass of 2.5 mton) was used according to EN 12697-33. This roller is regularly used in practice and allowed for a surface texture like asphalt in the field. The slab was turned by 90° during compaction to simulate the compaction which occurs in practice.
- Once compacted, for the determination of quality various measurement were performed: dimensions, flatness, surface texture, density. If a check produced a large variation, then new slabs were prepared.

A total of 177 slabs were compacted to provide the necessary number of samples. It was decided that each laboratory shall test four samples per mix variant. Three labs need a whole slab per test sample, the other three labs only need a quarter slab. To ensure that the variance between the four samples of each variant is identical for all, the three labs using quarter slabs received four quarters coming from four different slabs. BRRC used the spare quarters to perform tests at two different temperatures (20°C and 40°C).

Based on this method, recommendations for next stage of the project are:

- To sieve sand coarse aggregate according to BS EN 12697-2 in suitable nominal fraction sizes. Weigh the required amount of each fraction into metal buckets using UKAS calibrated scales which combined make up all constituents of the slab.
- Place the metal buckets in a thermostatically controlled oven and heat the aggregates at the established mixing temperature ±5°C for a minimum of 4 hours. Mix the asphalt in mixer for the period required for the TSCS selected. The amount of asphalt should be calculated taking into account the desired density and the density compensation for shrinkage during the cool-down phase. Temperature has to be constantly measured during the production and compaction process.
- A standard compaction roller should be used according to EN 12697-33. Two A3 sheets are to be sprayed with silicone lubricant and placed over the top of the mould, this will prevent the compaction arm from sticking to the asphalt in the mix creating voids in the slab.

2.8 Acceptance criteria

To ensure that the samples were uniform and acceptable, the following criteria were applied:

- The difference in the thickness of the slabs must not exceed one millimetre. This was measured using a calliper at eight locations for each slab. The maximum thickness difference measured was 0.9 mm hence this was below the limit and all slabs passed this criterion.

- Initially each slab was to be measured in four locations with a nuclear measuring device. After initial checking (repeatability and reproducibility trial) it was determined that the nuclear measuring device did not allow for measuring density below the required precision of ±15 kg/m^3. Therefore, density was measured using the volume of the slab and the mass of the slab. This method has a disadvantage such that measurements in different areas of the slab could not be taken. From 9 different mix designs, only one mix did not comply with ±15 kg/m^3.

- Texture measurement comprised of two methods. For the PA a laser scanner was used and for the SMA & BBTM the sand patch method (EN 13036-1) was used. The laser scanner has good repeatability, but poor reproducibility and the measurements recorded did not allow quantifying of the surface as good or bad. The sand patch method was completed in four locations for each slab. The results of these tests suggested identical texture depths for the same mix and hence no slabs were rejected.

- Finally, visual inspection occurred looking for greasy spots (>2 cm not allowed), scarce and lean spots (>50 cm^2 not allowed) and irregular distribution of the slab mix near edges. SMA had no rejections, BBTM has 3 rejections while PA had 6 rejections.

Based on this, the acceptance criteria proposed for next stage of the project are:

- Measurement of slab thickness using a calliper at eight locations for each slab. The maximum thickness difference should be 1 mm.

- Measurement of density to check that manufactured specimens are at the targeted air voids content ±1% (see Section 6).

- Visual inspection looking for greasy spots (>2 cm not allowed), scarce and lean spots (>50 cm^2 not allowed) and irregular distribution of the slab mix near edges.

2.9 Test results and statistical analysis

Deliverable D.7 of the DRaT project details results from the four devices and the 9 configurations of asphalt mixtures.

Statistical analysis was carried out to evaluate the following:

- Discrimination – establishing significant differences between the test results for each mixture and the ability of the equipment to discriminate between good and bad mixtures.

- Precision – focussing on repeatability (variation of slab measurements under near-homogeneous conditions) using statistical testing and using the results from the two ARTe & DSD devices to assess reproducibility.

- Correlation – looking at the rate of ravelling between the test equipment to see if there was a similarity in general or for each specific asphalt type.

A summary of the issues, observations and patterns from testing is summarised in Table 7.

Table 7. Summary of the DRaT project test results

Laboratory (equipment)	Deviations from Instructions	Observations	Comments
TU Aachen (ARTe)	Room temperature of 23°C for SMA tests; Surface temperature at beginning not always within 18-22°C[1]	Resistance to ravelling (standard mix)[3]: BBTM → SMA → PA Discriminations (from standard mix)[4]: PA: 1 BBTM: 0 SMA: 0 No correlation between density and mass loss possible; No correlation between MTD[2] and mass loss	The differences between standard materials in terms of resistance to ravelling presented inconsistencies (TU Aachen *v.* BAM); The results obtained strongly differed among the two laboratories; The ability to discriminate between standard mixes and their variants was similar for PA and BBTM but presented inconsistencies for SMA; Correlation between mass loss and density including all material was not considered
BAM (ARTe)	All test conditions were as specified	Resistance to ravelling (standard mix): BBTM ~ SMA → PA Discriminations (from standard mix): PA: 2 BBTM: 1 SMA: 1 No correlation between density and mass loss or MTD and mass loss	
BRRC at 20 °C (DSD)	Using a test load of 1000 N for 16 cycles, there was very little damage for any of the variants. The test load was therefore increased to 2000 N and the number of load cycles to 50; Room temperature >22 °C during testing, surface temperature not always within 18-22°C; Delay of 12 months for PA, 11 months for the BBTM and 9 months for the SMA	Resistance to ravelling (standard mix): SMA → BBTM → PA Discriminations (from standard mix): PA: 2 BBTM: 0 SMA: 0 No correlation between density and mass loss or MTD and mass loss	The differences between standard materials in terms of resistance to ravelling presented few inconsistencies, for the same test temperature (BRRC *v.* TU Darmstadt); The ability to discriminate between standard mixes and their variants was similar for SMA and BBTM but presented inconsistencies for PA; Correlation between mass loss and density including all material was not considered
BRRC at 40 °C (DSD)	All test conditions were as specified	Resistance to ravelling (standard mix): SMA → BBTM → PA Discriminations (from standard mix): PA: 1 BBTM: 0 SMA: 0 No correlation between density and mass loss or MTD and mass loss	Results obtained using the same equipment from different laboratories showed relatively good consistency
TU Darmstadt at 40 °C (DSD)	Few surface measurement > 42°C (max 43°C) at the start	Resistance to ravelling (standard mix): SMA → BBTM → PA Discriminations (from standard mix): PA: 1 BBTM: 0	

Laboratory (equipment)	Deviations from Instructions	Observations	Comments
		SMA: 0 No correlation between density and mass loss or MTD and mass loss; Slight correlation for M3-1, M3-2, M3-3 for MTD against mass loss where an increase in MTD produces an increase in mass loss	
Heijmans (RSAT)	Conditioning temperature was 5°C for SMA & BBTM; Delays of 1 months for PA and 2 months for BBTM & SMA	Resistance to ravelling (standard mix): SMA → BBTM → PA Discriminations (from standard mix): PA: 2 BBTM: 0 SMA: 2 No correlation between density and mass loss or MTD and mass loss	The detection of the 3 asphalt mixtures was consistent; Correlation between mass loss and density including all material was not considered
IFSTTAR (TRD)	Room temperature >22°C for SMA; Delay of 1 month for BBTM and 2 months for SMA	Resistance to ravelling (standard mix): SMA ~ PA → BBTM Discriminations (from standard mix): PA: 0 BBTM: 1 SMA: 2 No correlation between density and mass loss or MTD and mass loss	Correlation between mass loss and density including all material was not considered

[1] For this set of test results, no systematic relation was observed between mass loos in a test interval and the temperature at the start of the test interval. Therefore, there is no strong reason to reject the measurements made outside the specified temperature range for the further statistical analysis.

[2] MTD denotes Mean Texture Depth.

[3] Ranking reported in this column is based on mean values.

[4] Discrimination reported in this column is based on statistically significant differences in the round robin tests reported in Deliverable D.8 of the DRaT project.

Deviations from instructions

It was reported that maintaining the test temperature within the specified range was the most difficult test condition to satisfy, since there is significant heating of the plate surface due to friction (De Visscher, 2017).

In a few cases, the time between manufacturing and testing was exceeded by one to two months. Since the specimens were all stored correctly at a temperature below 20°C, it was not expected that there could have been any significant ageing in that period which would have an impact on the resistance to ravelling (De Visscher, 2017).

Changing the test temperature from 40°C to 20°C (for DSD at BRRC) had a great impact on the ravelling resistance. Results showed that at 20°C, the test load had to be doubled and the number of load cycles tripled, in order to measure a sufficiently high amount of material loss (De Visscher, 2017).

Detection results

The ranking of asphalt mixtures in terms of resistance to ravelling varied significantly depending on the test method used. From experience (De Visscher, 2017; Schoen, Van Vliet, Mookhoek, & Meinen, 2016), the tested mixtures, under the same loading conditions, would demonstrate an order in resistance to ravelling as: SMA → BBTM → PA. This was in agreement with the obtained results from RSAT and DSD.

Overall, for the same material, the amount of material loss was not consistent between different tests (and between the same test used by different lab). Furthermore, the test results demonstrate that the overall weight loss for the SMA slabs was very low and hardly any physical ravelling was observed after completion of the tests. The weight losses for almost all laboratories did not exceed the 25 g limit after the full testing procedure. This was associated to the fact that current test procedures were designed to test more open materials and/or the designed SMA material is quite unsusceptible to ravelling damage in general (De Visscher, 2017; Schoen et al., 2016).

Considering each material separately, no correlations between mass loss and either density or texture was found. The variations in density and MTD within each series of samples of the same mix were very small. Correlation between mass loss and density including all material was not considered.

Discrimination results

The results showed large coefficients of variation, with the consequence that the discrimination potential strongly depends on the number of slabs tested. Considering each material separately, results in terms of discrimination were following:

- PA: All scuffing devices excepting IFSTTAR's TRD were able to detect differences between the low-bitumen mixture and the standard mixture. Results from RSAT showed that a statistically significant increase in weight loss with respect to the standard was established for both low-bitumen and low-temperature mixtures.
- BBTM: The resistance to ravelling of low-temperature mixtures was not greater than for the standard mixtures, for any of the laboratories.
- SMA: Results from RSAT showed that a statistically significant increase in weight loss with respect to the standard was established for both low-bitumen and low-temperature mixtures.

Correlation results

Results obtained from different devices have been compared to establish whether correlation/scaling factors could be used, as reported in Deliverable D.8 of the DRaT project (Schoen et al., 2016).

For each asphalt mixture, scaling factors were calculated based on the final weight loss obtained from each device.

Results showed that, for given material, scaling factors strongly differed mutually (ranging from 0.2 to 388.8). Furthermore, when comparing scaling factors between devices for different materials, a large variability was observed. This indicated that scaling factors were not constant and depended on the material type. Therefore, a single 'universal' scaling factor could not be obtained.

It could be argued that absolute values were used for this comparison. Normalising the weight to the original weight of the sample could represent an alternative method for correlation purposes.

In addition, the weight loss at the end of the test was used. Further harmonisation of test loading conditions and durations could make these results more comparable and/or allow to establish a duration for each device which should produce similar amount of damage to the other devices.

3. Evaluation of test methods reported in PD CEN/TS 12697-50

A key question to be considered in the current work was: Can any of the four methods discriminate between good and bad performing TSCS under English weather conditions?

To answer this question, this stage focusses on ascertaining which of the four test devices in PD CEN/TS 12697-50 might be useful to test TSCS considering climatic and loading conditions in England.

An independent analysis of results obtained in DRaT (Appendix 1 of Deliverable D.8) project was carried out, as follows:

- Normalisation of results.
- Evaluation of Detection Power (differentiate between materials) of each device.
- Evaluation of Discrimination Power (differentiate between different variants of the same material) of each device.
- Calculation of the average coefficient of variation obtained for each device.
- Building of a correlation matrix between devices for each material.
- Ranking to assess these tests and propose the most suitable alternatives to the purpose of this project.

3.1 Normalisation of results

To allow for a more homogeneous comparison and correlation between devices, materials, and variants, results have been normalised as follows:

$$mass\ loss\ (\%) = mass\ loss\ (g)/initial\ mass\ (g)$$

The normalisation was carried out on the primary weight loss only, because the results in terms of secondary weight loss (optional parameter) were not available for all the devices.

Table 8 shows an example of normalisation of DRaT results and the statistics for each variant.

Table 8. Example of DRaT results normalisation

Device	Lab.	Material Type	Variant[1]	Initial Mass (g)	Mass loss (g)	Mass loss (%)	ID	Mean	Standard Deviation	Coefficient of Variation
ARTe	TU Aachen	PA	S_1	15000	35	0.23%	ARTe, TUA, PA, S	0.15%	0.08%	54.12%
			S_2	14683	12	0.08%				
			S_3	14630	12	0.08%				
			S_4	14693	32	0.22%				
			T_1	15259	52	0.34%	ARTe, TUA, PA, T	0.26%	0.06%	23.05%
			T_2	14768	31	0.21%				
			T_3	14758	33	0.22%				
			T_4	14892	37	0.25%				
			B_1	14620	76	0.52%	ARTe, TUA, PA, B	0.48%	0.21%	44.49%
			B_2	14643	102	0.70%				
			B_3	15041	28	0.19%				
			B_4	14713	75	0.51%				

[1] S denotes the Standard mixture; T denotes the lower compaction temperature variant; B denotes the lower bitumen content variant; the subscripts from 1 to 4 denotes the four repetitions for each variant.

From this example it can be noted the relatively high coefficients of variation obtained for each configuration. These statistic parameters were evaluated for all the devices, materials, and variants, as reported in Table 14 (Appendix A).

3.2 Evaluation of Detection Power and Discrimination Power

The ability of each device to differentiate between materials and between qualities of the same material was evaluated on the basis on the normalised results.

The two-sample t-test (α = 0.05) was used to determine whether two population means are different (Minitab Inc., 2010). This method helped evaluating whether there was sufficient evidence to conclude that different materials/variants had a different resistance to ravelling. This analysis was undertaken for each test method.

The detection power and discrimination power for each device are summarised in Table 9.

Table 9. Summary of Analysis carried out by AECOM based on DRaT results

		Test Devise	Aachener Ravelling Tester (ARTe)		Darmstadt Scuffing Device (DSD)		Rotating Surface Abrasion Test (RSAT)	Triboroute Device (TRD)
		Laboratories	TU Aachen	BAM	TU Darmstadt	BRRC	Hejimans	IFSTTAR
Detection Power		PA - BBTM	1	1	1	1	1	1
		PA - SMA	1	1	1	1	1	1
		BBTM - SMA	0	0	1	1	1	1
Discrimination Power	PA	Standard - Low Temp.	0	1	0	0	0	0
		Standard - Low Bitum.	1	1	1	1	1	0
	BBTM	Standard - Low Temp.	0	0	0	0	0	1[1]
		Standard - Low Bitum.	0	1	0	0	0	0
	SMA	Standard - Low Temp.	1[1]	1	0	0	1	0
		Standard - Low Bitum.	0	0	0	0	1	0

[1] In these cases it was observed that the variant had a better performance than the standard mixture. Temp and Bitum denote compaction temperature and bitumen content, respectively.

Based on this analysis, all the devices were able to differentiate between different materials, except for ARTe (from both TU Aachen and BAM laboratories). This is considered to be a key attribute, as experience show that the analysed materials have marked variations in terms of resistance to ravelling.

The analysed devices had also different discrimination power. BAM ARTe was able to discriminate 4 variants. However, the same device from TU Aachen obtained a lower score: 2 discriminations, one of which stating a better performance for the variant (which is not expected in on-site behaviour). RSAT discriminated 3 variants from standard materials, while the other devices 1 variant. Results, in terms of detection and discrimination, obtained by the two laboratories using DSD were consistent to each other.

3.3 Correlation matrix

Scaling factors have been calculated based on the normalised results as reported in Table 10.

Table 10. Mutual scaling factors between devices for PA, BBTM and SMA

PA	ARTe TU Aachen	ARTe BAM	DSD TU Darmstadt	DSD BRRC	RSAT Heijmans	TRD IFSTTAR
ARTe TU Aachen	1.00	0.90	0.04	0.04	0.52	7.17
ARTe BAM		1.00	0.04	0.04	0.58	7.97
DSD TU Darmstadt			1.00	1.06	14.34	196.17
DSD BRRC				1.00	13.56	185.42
RSAT Hejimans					1.00	13.68
TRD IFSTTAR						1.00
BBTM	ARTe TU Aachen	ARTe BAM	DSD TU Darmstadt	DSD BRRC	RSAT Heijmans	TRD IFSTTAR
ARTe TU Aachen	1.00	0.67	0.03	0.03	0.20	0.16
ARTe BAM		1.00	0.04	0.05	0.30	0.25
DSD TU Darmstadt			1.00	1.26	8.13	6.57
DSD BRRC				1.00	6.46	5.22
RSAT Hejimans					1.00	0.81
TRD IFSTTAR						1.00
SMA	ARTe TU Aachen	ARTe BAM	DSD TU Darmstadt	DSD BRRC	RSAT Heijmans	TRD IFSTTAR
ARTe TU Aachen	1.00	0.77	0.74	0.97	0.41	0.93
ARTe BAM		1.00	0.97	1.27	0.53	1.21
DSD TU Darmstadt			1.00	1.31	0.55	1.25
DSD BRRC				1.00	0.42	0.95
RSAT Hejimans					1.00	2.26
TRD IFSTTAR						1.00

Overall, these scaling factors strongly differed mutually were dependent on the material type, in agreement with DRaT conclusions. However, the following additional considerations could be made:

- TU Darmstadt DSD correlates well with BRRC DSD, independently of the material testes: a universal scaling factor can be proposed between these devices.
- TU Aachen ARTe correlates well with BAM ARTe, independently of the material tested: a universal scaling factor can be proposed between these devices.
- RSAT device had a fair correlation with both ARTe devices, independently of the material tested: a universal scaling factor can be proposed between these devices.
- Scaling factors between ARTe devices and DSD devices are very similar for PA and BBTM: a unique scaling factor can be proposed for asphalt mixtures with characteristic similar to PA and BBTM.

- TRD presented the highest scaling factors variations, ranging from 0.16 to 196.17. These factors were the most affected by the material type. This could be attributed to the different loading type used for this test method.
- Scaling factors for SMA ranged from 0.41 to 2.26, indicating that ravelling results obtained using different devices for this type of materials are fairly similar to each other. This is attributed to the relatively high resistance to ravelling offered by the SMA analysed.

3.4 Ranking

Based on the present analysis which considers all the three materials, the different devices have been ranked according on the following criteria, from 1 (best) to 4 (worst), as reported in Table 11:

- Detection Power
- Discrimination Power
- Repeatability
- Reproducibility
- Correlation

Table 11. Ranking Matrix for the different test devices analysed, considering results from PA, BBTM and SMA

Device	ARTe	DSD	RSAT	TRD
Detection Power	4	1	1	1
Discrimination Power	1[5]	3	2	4
Repeatability[1]	3	1	2	4
Reproducibility[2]	2	1	NA[2]	NA[2]
Correlation[3]	2	3	1	4
Overall Ranking[4]	3rd	2nd	1st	4th

[1] Ranking based on average coefficient of variation for each test device (Table 14– Appendix A).
[2] Reproducibility ranking could be established only for ARTe and DSD devices, available in two laboratories.
[3] Ranking based on average coefficients of variations from Table 15 (Appendix A).
[4] Reproducibility, unavailable for two devices, was not considered.
[5] ARTe ranked 1st considering only BAM results. However, inconstancies with results obtained in TU Aachen were found.

3.5 Discussion

The analysis undertaken aimed at evaluating which of the test kits reported in the DRaT project might be useful to test TSCS, if any. For this purpose, the same criteria used in Section 3.4 have been used to rank the test devices based only on results from BBTM and SMA, which are the most similar to TSCS. This ranking is reported in Table 12.

Table 12. Ranking Matrix for the different test devices analysed, considering results from BBTM and SMA

Device	ARTe	DSD	RSAT	TRD
Detection Power	4	1	1	1
Discrimination Power	1[5]	4	1	4
Repeatability[1]	3	2	1	4
Reproducibility[2]	2	1	NA[2]	NA[2]
Correlation[3]	2	4	1	3
Overall Ranking[4]	3rd	2nd	1st	4th

[1] Ranking based on average coefficient of variation for each test device (Table 14 – Appendix A).

[2] Reproducibility ranking could be established only for ARTe and DSD devices, available in two laboratories.

[3] Ranking based on average coefficients of variations from Table 16 (Appendix A).

[4] Reproducibility, unavailable for two devices, was not considered.

[5] ARTe ranked 1st considering only BAM results. However, inconstancies with results obtained in TU Aachen were found.

The first factor considered was the ability of each device to detect materials with different air voids content. All the devices, except for ARTe (from both TUA and BAM), were all able to distinguish (with a statistical significance) BBTM from SMA (see Table 9).

The ability of discrimination good and bad variants of the same material was assessed through a statistical analysis. Considering only BBTM and SMA:

- RSAT was able to discriminate 2 variants.
- ARTe results different between BAM (2 discriminations) and TU Aachen (1 discrimination but the variant had a better performance than the standard mixture – which contradicts experience findings).
- DSD was not able to discriminate between variants from standard BBTM and SMA.
- TRD provided 1 discrimination but the variant had a better performance than the standard mixture – which contradicts experience findings.

Repeatability of testing methods was evaluated by means of the average coefficients of variation calculated for BBTM and SMA results (based on Table 14 – Appendix A). It was found that:

- RSAT obtained the most consistent results (17.97%).
- DSD ranked second and it was consistent between TU Darmstadt (25.74%) and BRRC (25.63%).
- ARTe ranked third, considering the average between TU Aachen (41.45%) and BAM (30.79%).
- TRD produced the less consistent results (39.06%).

Reproducibility could be assessed only for ARTe and DSD, which were available in two laboratories: DSD demonstrated a very good consistency between results obtained from TU Darmstadt and BRRC; ARTe results presented not negligible variations in results obtained from BAM and TU Aachen.

Correlation, as additional factor considered for ranking purposes, was taken into account as follows: For each material, mutual scaling factors were calculated to correlate results from different devices (Table 10); the variability of these scaling factors with the material tested was evaluated as coefficient of variation of BBTM and SMA values (Table 16 – Appendix A); for each device, these coefficient of variations have been averaged to assess which device is more prone to correlate with the other devices independently of the material tested. Results show that RSAT ranked first, followed by ARTe, TRD and DSD. This result reflects the fact that DSD was the most aggressive test in terms of mass loss produced, having therefore the poorest correlation with the other devices (considering only BBTM and SMA results).

Overall, based on the analysis undertaken, RSAT is proposed as most suitable test (among those used in the DRaT project) to assess resistance to ravelling of TSCS. ARTe and DSD are possible alternatives while TRD was evaluated as least suitable for this purpose.

None of these testing methods take into account the effect of water on resistance to ravelling. This factor is discussed in the following section.

4. Explore alternative test methods for PD CEN/TS 12697-50

This section details the work undertaken on exploring alternative test methods that can better discriminate performance of TSCS in UK. The options explored focussed on the possibility of including damage from water and/or using test methods available in the UK.

4.1 TRL Scuffing Test

The scuffing-wheel apparatus consists of a loaded wheel which bears on a specimen held in a moving table (Figure 1). The table moves to and from beneath the wheel with the axle of the wheel held at an angle of (20±1)° to the vertical plane perpendicular to the direction of travel (TRL Report 176, 1997).

Figure 1. TRL Scuffing test (Ojum, 2016)

The sample size is 305 mm by 305 mm. The vertical load applied to the wheel is (520±5) N and the tyre pressure is (310±10) kPa. The centre of contact area of the tyre describes simple harmonic motion with respect to the centre of the top surface of the specimen with a frequency of 42 passes per 60 seconds (TRL Report 176, 1997).

According to TRL procedure, samples should be conditioned at (45±1)°C for a period of 4 to 6 hours and during the test, the temperature should be maintained at (45±1)°C (TRL Report 176, 1997).

This test is being currently used for the erosion measurements on high friction surfacings. Although not being currently described in a EN standard, TRL scuffing test method follows similar principles as those tests stated in last version (2018) of PD CEN/TS 12697-50.

This test would require modification to account for the water damage mechanism.

4.2 Wheel-Tracking Device

Wheel-Tracking tests are typically used to assess susceptibility of bituminous materials to deform. This susceptibility is assessed by measuring the rut depth formed by repeated passes of a loaded wheel at a fixed temperature. Different devices can be used for this purpose. In large-size devices the specimens are conditioned in air during testing. With small-size devices, specimens are conditioned, in either air or water (BS EN 12697-22:2003, 2003). This last feature makes the small-size devices a possible option to be explored for the present work.

Small-size devices combine the effects of rutting and moisture damage by rolling a steel wheel across the surface of an asphalt concrete specimen that is immersed in hot water. Therefore, they are widely used to evaluate the resistance to rutting and moisture susceptibility of asphalt mixtures. The Hamburg Wheel Tracking Device (HWTD), developed in Germany, is one example which meets BS EN 12697-22 (Figure 2). The test is typically conducted on 260 mm by 320 mm by 40 mm slabs. The 47 mm wide wheel is tracked across a submerged (underwater) sample for 20,000 cycles using a 705 N load (Pavement Interactive, 2011).

Figure 2. Hamburg Wheel Tracking Device (Pavement Interactive, 2011)

For the purpose of this work, this device has the advantage to be designed to test immersed specimen. The load is applied along the same direction, providing a relatively low scuffing effect, but often sufficient to induce material loss on poor quality asphalt specimens.

4.3 Model Mobile Load Simulator (MLS)

The MLS system is an accelerated pavement testing device simulating the action of traffic on asphalt pavement. MLS system can vary the test conditions to conduct tests both in the laboratory and on site (Ojum, 2016). There are different versions of MLS, varying in dimensions, maximum load, loading type and frequency, type pressure, and maximum rut depth, among others. The MLS has been applied as an accelerated pavement testing system to evaluate failure mechanisms that include rutting, moisture damage, cracking and ravelling.

MLS11, formerly known as MMLS3, is a third scale model (Figure 3). This test is covered under South African National Standard (SANS) 3001-PD1:2006 and has been used in a number of countries in Europe, South Africa, USA and Australia (details can be found in http://www.mlstestsystems.com/.) This device consists of four recirculating axles, each with a single 300 mm diameter wheel. The tyres may be inflated up to a pressure of 800 kPa. The axle loads ranges between 2.1 kN and 2.7 kN. The axle loads are automatically kept constant at a predetermined value by the special suspension system. Nominal wheel speed is 2.5 m/s, applying about 7200 loads per hour (Ojum, 2016).

Test can also be done in dry or wet condition. An environmental chamber can be utilized to maintain temperature of samples or pavement during testing. The wheels can be laterally displaced across 150 mm in a normal distribution to simulate traffic wandering. The lateral wander could be applied to evaluate the stripping distress (Figure 3) (Ojum, 2016).

Figure 3. MLS 11 Equipment and Example of Stripping of Porous Asphalt tested with MMLS2 (Ojum, 2016)

4.4 Modified Brushing Test

The standard brushing test procedure involves initial soaking of a test specimen made in the laboratory or cored from a pavement in a fuel, followed by a brushing period with a steel brush mounted in a mixer. The material loss of the specimen is a measured (BS EN 12697-43:2005, 2005).

A modified version (without the pre-conditioning by immersion in fuel) of this test was used in other UK studies to measure the influence of asphalt preservatives on asphalt mixture resistance to ravelling (Thameside Test & Research Ltd, 2015).

The specimen is fixed in a cylindrical mould and a crimped wire cup brush 70 mm external diameter and 30 mm internal diameter fixed to a Hobart N50 type mixer is applied to the surface of the specimen (Figure 4), with a maximum rotation speed of 8,500 rpm.

At the bottom of the specimen a pneumatic cylinder pushes the specimen onto the steel brush with a constant force. In the modified version, the force on the brush was reduced from 60 N to 30 N.

Figure 4. Modified Brushing Test (Thameside Test & Research Ltd, 2015)

The specimens have a diameter of (150±2) mm and a height between 40 mm and 60 mm.

The brush moves in epicycloids passages over the surface. After 30 s the brushing stops and the specimen is removed from the mould. Then the loss of mass is measured and the brushed surface is visually inspected. After that the specimen is put back in the mould and the same procedure is carried out again after each 30 s brushing period, up to 120 s in total.

This test does not simulate the wheel scuffing effect. However, it can provide a relative quick measure of the quality of aggregate/mortar bond, useful for comparative analysis.

A possible modification would be the pre-conditioning of samples by immersion in water.

4.5 Modified Saturation Ageing Tensile Stiffness

The standard Saturation Ageing Tensile Stiffness is detailed in BS EN 12697-45 (2012). This test consists of initial saturation under vacuum prior to placing compacted asphalt core samples in a high temperature and pressure environment in the presence of moisture for an extended period of time.

SATS conditioning regime is used to age the specimens in the presence of water. In the standard version, this test method is limited to bituminous specimens with consistent air voids contents (6-10%) and hard binder. The test is intended to be used as a screening test for the assessment of a combination of aggregate, filler and additives with respect to the retained adhesion properties after simulated ageing in a moist atmosphere for lean/stiff base and binder course mixtures. The stiffness modulus measured after the test divided by the stiffness modulus measured before the test (retained stiffness modulus), and the specimen saturation after the test (retained saturation), are used as an indication of the sensitivity of the compacted mixture to combined ageing/moisture effects.

Figure 5. Modified SATS Test Equipment and Set-up (Ojum, Widyatmoko, Heslop, & Khojinian, 2017)

The standard SATS test protocol was quite successful for the asphalt mixtures tested which comprised low bitumen content, high air voids and high stiffness. An alternative test protocol (Modified SATS or Immersion Ageing Test) incorporates improvements to the EN 12697-45 SATS to widen the scope and applicability ensuring different types of asphalt materials can be tested to ascertain effects of moisture susceptibility and ageing (Ojum et al., 2017). The modified SATS test protocol follows the same procedure with the standard SATS test protocol except for a lower pressure (from 2.1 to 0.5 MPa), a lower duration (from 72 to 24) hours and an higher number (from one to three) of asphalt specimens (100 mm diameter) fully immersed in water (Figure 5).

This test method does not take into account the impact of traffic, such as the shear forces induced by accelerating, braking and turning of heavy vehicles which can accelerate wear. However, for the purpose of this work, it is considered as a possible accelerated ageing test to assess the effect of scuffing before and after SATS protocol, particular on core samples. The feasibility of using this test as sample conditioning method prior to scuffing test can be assessed via sending a questionnaire to laboratories who perform scuffing tests on cores (e.g. Heijmans - RSAT).

5. Discussion and Conclusions

This report details the work undertaken under Sub-Task 1: Critical Review of DRaT Project Outputs. It is part of an overarching project commissioned by Highways England (HE): Task 1-614.

The aim of this sub-task was to carry out a critical review of the outputs and results of the Conference of European Directors of Roads (CEDR) Development of Ravelling Test (DRaT) project. Based on this review, a method of measuring the ravelling characteristic of a Thin Surface Course System (TSCS) pavement in the United Kingdom (UK) was to be proposed.

To achieve the declared objectives, a review and evaluation of DRaT Project Reports and PD CEN/TS 12697-50 was carried out. The test methods reported in PD CEN/TS 12697-50 were evaluated and ranked with focus on TSCS, considering climatic and loading conditions in England. Furthermore, alternative test methods that can better discriminate performance of TSCS were explored, considering the possibility of including damage from water and/or using test methods available in the UK.

Main findings from the critical review of the DRaT project are described below.

- Ravelling mechanism occurs to asphalt surfacing when the bond between binder and aggregate reaches a critical point. This mechanism is typically present in areas exposed to changing direction and/or braking/accelerating. However, other causes/factors can initiate or accelerate the ravelling process. The most relevant factors are: the presence of water, the aggregate type and gradation, the binder content and type, the mix design, the quality of construction and the weather conditions. These factors are often interdependent, making it difficult to assess the theoretical potential to ravel in the design stage.

- The experiences reviewed by DRaT show that the air voids content plays an essential role in the resistance to ravelling: porous asphalts are more prone to ravelling than dense asphalts. The previous studies on TSCS reviewed show that aged asphalt (more brittle bitumen) have a lower resistance to ravelling whereas very low correlations were found with the binder content, binder type and aggregate size.

- With the aim of investigating four existing scuffing tests, the DRaT project tested three different asphalt mixtures (PA, BBTM and SMA), manufactured in their standard version and two lower quality variants. BBTM and SMA are the most relevant mixtures for the purpose of evaluating which test is the most suitable for the TSCSs typically used in the UK.

- The protocol used to manufacture slabs produced very consistent samples in terms of density and texture, reflecting the high quality and repeatability of the slab manufacturing and compaction. Visual inspection was the most critical acceptance criterion, leading to the highest number of slab rejection.

- The main differences between the devices testing protocols are related to the sample dimensions and shape, the conditioning, the loading type and amplitude, the type of tyre used to apply the load (and consequent contact area), the test duration and the measured parameters. The DRaT project conducted a study to harmonise these test methods, focussing mostly on conditioning of samples and parameter to be measure before and after each test. Further harmonisation could focus on the loading conditions, test duration and temperature, to obtain a similar rate of damage for all the devices.

- Results used in DRaT were in terms of absolute weight loss. DRaT statistical analysis concluded that:

  - For the same material, the amount of material loss was not consistent between different tests (and between the same test used by different lab).

  - The overall weight loss for the SMA slabs was very low and hardly any physical ravelling was observed after completion of the tests.

- The rate of material loss (slope of the curves) behaved differently, depending on the test device.
- The plots showed no correlations between mass loss and either density or texture.
- The results showed large coefficients of variation, with the consequence that the discrimination potential strongly depends on the number of slabs tested.
- The correlation study showed that scaling factors were not constant and depended on the material type. Thus, a single ‘universal’ scaling factor could not be obtained.

To evaluate which of the four test methods would be the most suitable to discriminate between good and bad performing TSCS under English weather conditions, an independent analysis was carried out using normalised results. The analysis focussed on BBTM and SMA, leading to the following main findings:

- All the devices, except for ARTe (from both TUA and BAM), were able to distinguish (with a statistical significance) BBTM from SMA.
- Considering the ability of discrimination (with a statistical significance) between good and bad variants: RSAT produced the most consistent result followed by ARTe, DSD and TRD.
- Repeatability results show that: RSAT was the most consistent test (17.97%); DSD ranked second and it was consistent between TU Darmstadt (25.74%) and BRRC (25.63%); ARTe ranked third, considering the average between TU Aachen (41.45%) and BAM (30.79%); TRD produced the less consistent results (39.06%).
- In terms of reproducibility, DSD show a very good consistency between results obtained from TU Darmstadt and BRRC whereas ARTe results presented not negligible variations in results obtained from BAM and TU Aachen.
- Correlation results show that RSAT ranked first, followed by ARTe, TRD and DSD. This is attributed to the fact that DSD was the most aggressive test in terms of mass loss produced, having, therefore, the poorest correlation with the other devices (considering only BBTM and SMA results).
- Overall, based on the analysis undertaken, RSAT is considered the most suitable test (among those used in the DRaT project) to assess resistance to ravelling of TSCS.
- Nevertheless, none of these testing methods take into account the effect of water on resistance to ravelling.

Alternative test methods to discriminate performance of TSCS in UK were explored, focussing on the possibility of including damage from water and/or using test methods available in the UK. The reviewed test methods include:

- TRL Scuffing Test, which has been used in the UK for evaluating asphalt resistance to scuffing. Although not being currently described in a EN standard, TRL scuffing test follows similar principles to those stated in last version (2018) of PD CEN/TS 12697-50. However, the effect of water damage is not included in this test.
- The Hamburg Wheel Tracking Device is a standard test (EN 12697-22) with the advantage of testing immersed specimen. In this test, the load is applied along the same direction, providing a relatively low scuffing effect although sufficient enough to induce material loss on poor quality asphalt specimens.
- The MLS system is an accelerated pavement testing device simulating the action of traffic on asphalt pavement. The main advantages are: the test can be carried out in dry or wet condition; the wheels can be laterally displaced across 150 mm in a normal distribution to simulate traffic

wandering, which can produce stripping distress. Whilst the test is not an EN method, it is standardised under SANS 3001-PD1:2006.

- The Modified Brushing Test is based on the EN 12697-43 but without conditioning in fuel. This test does not simulate the wheel scuffing effect. However, it can provide a relative quick measure of the quality of aggregate/mortar bond, useful for comparative analysis. A possible modification would be the pre-conditioning of samples by immersion in water.
- The Modified Saturation Ageing Test Tensile Stiffness (SATS) is based on the EN 12697-45. For the purpose of this work, this test has the potential to be used as accelerated ageing test to assess the effect of scuffing before and after SATS protocol. However, the feasibility of this procedure should to be discussed with the relevant laboratories carrying out scuffing tests on cores.

6. Recommendations for next stages

Based on the critical review of the DRaT project carried out, the following recommendations are made for the next stages of the current project.

The air voids content was the selected as the key factor influencing ravelling resistance of TSCS. Therefore, two asphalt mixtures are proposed. These mixtures will have 6% and 14% air voids content, to represent good and poor compaction, respectively.

For the slab manufacturing, the recommendations for next stage of this project are:

- To sieve sand coarse aggregate according to BS EN 12697-2 in suitable nominal fraction sizes. Weigh the required amount of each fraction into metal buckets using UKAS calibrated scales which combined make up all constituents of the slab.
- Place the metal buckets in a thermostatically controlled oven and heat the aggregates at the established mixing temperature ±5°C for a minimum of 4 hours. Mix the asphalt in mixer for the period required for the TSCS selected. The amount of asphalt should be calculated taking into account the desired density and the density compensation for shrinkage during the cool-down phase. Temperature has to be constantly measured during the production and compaction process.
- A standard compaction roller should be used according to EN 12697-33. Two A3 sheets are to be sprayed with silicone lubricant and placed over the top of the mould, this will prevent the compaction arm from sticking to the asphalt in the mix creating voids in the slab.
- A minimum of four slabs per each configuration should be manufactured to keep sufficient discrimination power of the devices to detect lower quality materials.

The acceptance criteria proposed are:

- Measurement of slab thickness using a calliper at eight locations for each slab. The maximum thickness difference should be 1 mm.
- Measurement of density to check that manufactured specimens are at the targeted air voids content ±1%.
- Visual inspection looking for greasy spots (>2 cm not allowed), scarce and lean spots (>50 cm^2 not allowed) and irregular distribution of the slab mix near edges.

Based on the independent evaluation of DRaT results, the test devices recommended for next stage are:

- RSAT
- DSD

To evaluate the reproducibility of the test methods, two laboratories for each device are recommended for carrying out tests.

The possibility of including water damage and/or accelerate ageing will be discussed with the relevant laboratories. A summarising matrix of the proposed testing plan is given in Table 13.

Table 13. Matrix of proposed testing plan for next stage

		Test Devise	RSAT		DSD	
Material	**Variant**	Laboratory	Heijmans	TBC[1]	TU Darmstadt	BRRC
TSCS	**6% Air Voids**	Samples	3 cores	3 cores	4 slabs	4 slabs
	14% Air Voids		3 cores	3 cores	4 slabs	4 slabs

[1] To be confirmed. Currently seeking the availability of a second laboratory

The possibility of using the alternative test methods explored may be considered in future stages of the project subject to budget constraints.

7. References

BS EN 12697-2:2002+A1. Bituminous mixtures : Test methods for hot mix asphalt — Part 2: Determination of particle size distribution (2007).

BS EN 12697-22:2003. Bituminous Mixtures - Test Methods for Hot Mix Asphalt. Part 22: Wheel tracking, European Committee for Standardization § (2003).

BS EN 12697-43:2005. (2005). *Bituminous mixtures: Test methods for hot mix asphalt - Part 43: Resistance to fuel*.

BS EN 12697-45:2012. Bituminous mixtures — Test methods for hot mix asphalt - Part 45: Saturation Ageing Tensile Stiffness (SATS) conditioning test (2012).

De Visscher, J. (2017). *DRaT – Development of the Ravelling Test. Deliverable D.7 Factual report on test results*. CEDR.

De Visscher, J., & Vanelstraete, A. (2017). Ravelling by traffic: Performance testing and field validation. *International Journal of Pavement Research and Technology*, *10*(1), 54–61. https://doi.org/10.1016/j.ijprt.2016.12.004

Hjort, M., Mattias, H., & Jansen, J. M. (2008). *Road wear from Heavy Vehicles*.

Hunter, R. N., Self, A., & Reid, J. (2015). *The Shell bitumen handbook* (Sixth edit). London: ICE Publishing.

Khojinian, A., Parry, T., & Thom, N. (2016). Strategy Management and Maintenance for Thin Surface Course Systems. In *E&E Congress 2016*. Prague, Czech Republic.

Minitab Inc. (2010). Statistical Inference and and t-Tests. *Minitab Inc.*

Nicholls, J. C., De Visscher, J., Hammoum, F., & Jacobs, M. (2016). *DRaT – Development of the Ravelling Test. Deliverable D.2 Review of parameters influencing the propensity of asphalt to ravel*. CEDR.

Ojum, C. K. (2016). *Task 451 : Developing a Simulative Laboratory Ageing Testing Method for Thin Surfacings*. AECOM.

Ojum, C. K., Widyatmoko, D., Heslop, M., & Khojinian, A. (2017). Accelerated durability testing using the immersion ageing test for thin asphalt surfacings. In *Asphalt Professional* (p. September). Retrieved from www.instituteofasphalt.org

Pavement Interactive. (2011). Laboratory Wheel Tracking Devices. Retrieved from https://www.pavementinteractive.org/reference-desk/testing/asphalt-tests/laboratory-wheel-tracking-devices/

PD CEN/TS 12697‐50:2018. Bituminous mixtures - Test methods. Part 50: Resistance to scuffing (2018).

Schoen, E., Van Vliet, D., Mookhoek, S., & Meinen, N. (2016). *DRaT – Development of the Ravelling Test. Deliverable D.8 Report on Analysis of Results*. CEDR.

Thameside Test & Research Ltd. (2015). *Effect of Armaseal on durability of Micro Asphalt Surfacing Thameside Test & Research Ltd*.

TRL Report 176. (1997). *Laboratory tests on high-friction surfaces for highways*.

Van Loon, H., & Butcher, M. J. (2015). Analysis of a Thirteen Year Old Rap Site in South Australia. In *21st ARRB Transport Research Conference, Cairns*. Melbourne: ARRB.

Wayman, M. (2015). DRaT – Development of Ravelling Test - Project Overview. CEDR.

Appendix A – Statistical Analysis

Table 14. Normalised DRaT results

ID	Mean	SD	Coefficient of Variation
ARTe, TUA, PA, S	0.15%	0.08%	54.12%
ARTe, TUA, PA, T	0.26%	0.06%	23.05%
ARTe, TUA, PA, B	0.48%	0.21%	44.49%
ARTe, TUA, BBTM, S	0.05%	0.02%	42.91%
ARTe, TUA, BBTM, T	0.11%	0.09%	79.25%
ARTe, TUA, BBTM, B	0.04%	0.03%	71.83%
ARTe, TUA, SMA, S	0.08%	0.01%	13.37%
ARTe, TUA, SMA, T	0.06%	0.01%	18.96%
ARTe, TUA, SMA, B	0.07%	0.02%	22.38%
ARTe, BAM, PA, S	0.20%	0.03%	17.36%
ARTe, BAM, PA, T	0.32%	0.09%	27.80%
ARTe, BAM, PA, B	0.46%	0.12%	25.71%
ARTe, BAM, BBTM, S	0.08%	0.04%	52.82%
ARTe, BAM, BBTM, T	0.08%	0.02%	32.10%
ARTe, BAM, BBTM, B	0.14%	0.03%	21.44%
ARTe, BAM, SMA, S	0.08%	0.02%	25.47%
ARTe, BAM, SMA, T	0.14%	0.03%	24.59%
ARTe, BAM, SMA, B	0.07%	0.02%	28.31%
DSD, TUD, PA, S	7.41%	1.62%	21.81%
DSD, TUD, PA, T	7.47%	1.00%	13.43%
DSD, TUD, PA, B	9.82%	0.45%	4.61%
DSD, TUD, BBTM, S	2.60%	0.20%	7.56%
DSD, TUD, BBTM, T	2.91%	0.56%	19.20%
DSD, TUD, BBTM, B	2.61%	0.46%	17.77%
DSD, TUD, SMA, S	0.12%	0.03%	29.80%
DSD, TUD, SMA, T	0.09%	0.05%	52.50%
DSD, TUD, SMA, B	0.09%	0.02%	27.63%
DSD, BRRC, PA, S	7.20%	1.19%	16.55%
DSD, BRRC, PA, T	6.16%	1.14%	18.45%
DSD, BRRC, PA, B	9.59%	1.56%	16.31%
DSD, BRRC, BBTM, S	2.40%	0.36%	15.08%
DSD, BRRC, BBTM, T	2.11%	0.54%	25.63%
DSD, BRRC, BBTM, B	1.94%	0.44%	22.56%
DSD, BRRC, SMA, S	0.07%	0.02%	33.16%
DSD, BRRC, SMA, T	0.08%	0.03%	33.44%
DSD, BRRC, SMA, B	0.07%	0.02%	23.94%
RSAT, Heijmans, PA, S	0.38%	0.16%	40.67%
RSAT, Heijmans, PA, T	0.45%	0.21%	47.64%
RSAT, Heijmans, PA, B	0.81%	0.21%	26.17%
RSAT, Heijmans, BBTM, S	0.32%	0.02%	6.79%
RSAT, Heijmans, BBTM, T	0.26%	0.06%	22.25%
RSAT, Heijmans, BBTM, B	0.42%	0.19%	45.19%
RSAT, Heijmans, SMA, S	0.15%	0.02%	15.33%
RSAT, Heijmans, SMA, T	0.18%	0.02%	12.74%
RSAT, Heijmans, SMA, B	0.20%	0.01%	5.52%
TRD, IFSTTAR, PA, S	0.03%	0.01%	41.71%
TRD, IFSTTAR, PA, T	0.05%	0.02%	37.82%

ID	Mean	SD	Coefficient of Variation
TRD, IFSTTAR, PA, B	0.04%	0.03%	74.60%
TRD, IFSTTAR, BBTM, S	0.29%	0.11%	38.30%
TRD, IFSTTAR, BBTM, T	0.08%	0.01%	17.65%
TRD, IFSTTAR, BBTM, B	0.78%	0.58%	74.61%
TRD, IFSTTAR, SMA, S	0.07%	0.03%	41.92%
TRD, IFSTTAR, SMA, T	0.11%	0.03%	32.19%
TRD, IFSTTAR, SMA, B	0.06%	0.02%	29.70%

Table 15. Coefficients of variation of mutual scaling factors between devices (considering PA, BBTM and SMA)

Coeff. of Variation (considering all the materials)	ARTe TU Aachen	ARTe BAM	DSD TU Darmstadt	DSD BRRC	RSAT Hejimans	TRD IFSTTAR
ARTe TU Aachen		15%	153%	156%	43%	140%
ARTe BAM			154%	156%	31%	134%
DSD TU Darmstadt				11%	90%	163%
DSD BRRC					97%	165%
RSAT Hejimans						126%
TRD IFSTTAR						

Table 16. Coefficients of variation of mutual scaling factors between devices (considering only BBTM and SMA)

Coeff. Of Variation (considering only BBTM and SMA)	ARTe TU Aachen	ARTe BAM	DSD TU Darmstadt	DSD BRRC	RSAT Hejimans	TRD IFSTTAR
ARTe TU Aachen		9%	132%	133%	47%	99%
ARTe BAM			189%	131%	39%	94%
DSD TU Darmstadt				3%	123%	96%
DSD BRRC					124%	98%
RSAT Hejimans						67%
TRD IFSTTAR						