Smith, Andrew S.J., Kaushal, Aniruddha, Odolinski, Kristofer, Iwnicki, Simon and Wheat, Phill

Estimating the damage and marginal cost of different vehicle types on rail infrastructure: combining economic and engineering approaches

Estimating the damage and marginal cost of different vehicle types on rail infrastructure: combining economic and engineering approaches

Andrew S. J. Smith[a]*, Aniruddha Kaushal[b], Kristofer Odolinski[c], Simon Iwnicki[b], and Phill Wheat[a]

Abstract

EU legislation requires that European infrastructure managers set access charges based on the marginal cost of running trains on their networks. Two methods have been used in the literature for this purpose. Top-down methods relate actual costs to traffic volumes. Bottom-up methods use engineering models to simulate damage and then translate damage into costs based on assumptions about interventions and their unit costs. Whilst top down methods produce sensible results for marginal cost overall, they have struggled to differentiate between traffic types. The challenge for bottom-up approaches is how to translate damage into cost, with numerous assumptions being required which may be invalid.

This paper proposes a new, two stage approach to estimating the marginal cost of rail infrastructure usage. The first stage uses engineering models to simulate damage caused by vehicles on the network. The second stage seeks to establish a statistical relationship between actual costs and damage. It is thus possible to convert damage estimates into costs using actual cost data, rather than through a set of potentially invalid assumptions as in previous approaches.

Only the first stage is implemented in this paper. We show that it possible to produce total (annualised) damage measures for three damage mechanisms on five actual track sections in Sweden. Once extended, it will be possible to model the relationship between damage and actual costs for the first time; and thus better understand the relative costs of the different damage mechanisms and in turn inform the level and structure of track access charges.

[a] University of Leeds, UK

[b] University of Huddersfield, UK

[c] Swedish National Road and Transport Research Institute, Department of Transport Economics, Sweden

* Corresponding author. a.s.j.smith@its.leeds.ac.uk

1. Introduction

European policy since the mid-1990s has emphasised the promotion of within-mode competition as a way of revitalising the fortunes of Europe's railways. Progressively freight and international passenger services have been opened up to competition. The proposals contained in the European Commission's Fourth Railway package (European Commission, 2013) for further reforms of Europe's railways, mean that, once enacted, it will be compulsory to introduce competitive tendering for passenger services run under public service contracts and open access for commercial services across the whole of Europe. Vertical separation of infrastructure and operations or, at least, fair access to infrastructure and transparent prices for access, is seen by the Commission as a key enabler of competition in the sector.

The above developments mean that understanding the cost, and in particular, the marginal (infrastructure) cost, of running an extra service on the network has become more important than ever. Existing legislation requires that charges for access to the infrastructure must be based on "costs directly incurred as a result of operating the train service". This can be interpreted as what economists would call the short-run marginal (or incremental) cost imposed on the infrastructure by the service running on the network. This paper focuses on one element of short-run marginal cost, namely the additional maintenance and renewal cost required to rectify the incremental damage caused a train service (the marginal wear and tear cost). Of course, the need to estimate marginal cost of infrastructure use is not merely for the purpose of meeting EU legislation. It is important from the purpose of economic efficiency (in terms of making best use of the existing network) that train operators pay at least the short-run marginal cost of running trains on the network. Further, track access charges that vary according to the different damage and cost imposed by different vehicles, should ensure that the "right" vehicles are run on the network and potentially that new rolling stock designs are developed that reduce whole system costs (operator and infrastructure managers costs).

The previous literature, and practice by bodies responsible for charging, contains three approaches for estimating marginal costs of traffic and also different types of traffic: cost allocation methods; statistical approaches; bottom-up engineering methods. In practice the approaches used by charging setting bodies can be a hybrid of approaches, such as that used by the British Office of Rail Regulation (ORR). However, each of the approaches has significant drawbacks. The contribution of this paper is to propose an alternative, two-stage approach that combines engineering and statistical approaches in a way that seeks to combine the best features of both; and thus overcome some of the weaknesses of previous approaches.

The remainder of the paper is structured as follows. In section 2 we explain the previous approaches used to estimate marginal rail infrastructure costs and their weaknesses. The proposed methodology is set out in section 3 and the data used for the empirical work outlined in section 4. Section 5 presents the results and section 6 concludes.

It should be noted that at this stage of the research we have only been able to implement stage 1 of the methodology. This is because we do not yet have sufficient data to implement the second stage, though this is anticipated soon. This conference paper therefore reports and discusses the methodology and the first stage results. The second stage will be implemented post submission of this conference paper.

2. Previous approaches

To summarise briefly, there are two methods for producing estimates of marginal costs (Wheat and Smith, 2008). Top-down methods relate actual costs to traffic volumes, controlling for characteristics of the infrastructure. Bottom-up methods can be characterised as using some form of engineering model to estimate the damage inflicted by different types of vehicle on the network. Then assumptions can be made about the intervention / remediation required to deal with that damage, combined with estimates of unit costs of that remediation activity, to give the marginal cost estimates. These approaches are summarised in Figure 1. A "third" method, the so-called cost allocation method, can be thought of as a hybrid that utilises engineering judgement and econometric evidence and other rules of thumb to establish the variability of different cost categories. This approach is therefore not discussed further.

Figure 1: Alternative approaches for estimating marginal costs

- Method 1: engineering approach
  - Simulate damage done by traffic (engineering model)
  - Determine action need to remedy damage (e.g. tamping)
  - Activity volume * Unit cost of activity = (marginal) Cost
- Method 2: top down statistical approach
  - Relate actual costs to passenger and freight tonne-km (regression)
  - E.g. Log Cost = a + b* Log Passtonne + c * Freight tonne
  - Compute marginal costs from the parameter estimates (the a and b) from that model

Both methods have strengths and weaknesses. The advantage of top-down methods is that they use actual cost data. Their weakness lies in the fact that it is likely to be very hard to capture the complexity of factors that will affect the relationship between traffic and cost, and in particular, it has proved difficult to get any sensible estimates of the relative cost of passenger and freight vehicles. The bottom-up method is very good at capturing complexity and it is possible to model and gain estimates of the relative damage of different vehicle types. The problem is how then to translate these damage estimates into cost.

It is worth noting, more precisely, why there is a difficulty in getting from damage to cost in bottom-up approaches. First, traffic results in different types of damage. In practice, one vehicle may cause more of one type of damage and less of another, thus meaning that information is needed on the relative cost of the different types of damage to obtain estimates of relative marginal cost. This leads to a second problem, namely that assumptions are needed on what type of activity and how much of it, are needed to rectify the damage done. This potentially requires a very detailed model or alternatively simplifying assumptions are needed which might be wrong.

Added complexities include the fact that the mix of damage types will affect what activity should be undertaken (e.g. some traffic types might cause damage but at the same time alleviate the need for other forms of remediation, such as rail grinding) and the fact that some damage mechanisms lead to more maintenance activity, whilst others result in more renewal (and the costs of these can be very different). Finally it is hard to estimate unit costs of activities as these will depend on, inter alia, the location, the nature of the job, the length of possession and the scale of the activity.

The research question therefore is as follows: how can we obtain better estimates of the relative cost of different damage mechanisms (which in turn can then help estimate the relative marginal cost of different vehicle types).

In terms of its relation to other approaches, in the past ORR has used a top-down cost allocation approach, based on engineering judgement, to determine the general level of cost variability, and then used an engineering formulae to allocate costs to vehicles based on their relative damage (vertical forces only). Since then Network Rail has developed its bottom-up cost modelling approaches (its infrastructure cost model (ICM) model, which in turn is based on the Vehicle Track Interaction Strategic Model (VTISM)) to estimate marginal costs from the bottom up). This approach measures the overall variability of costs with respect to traffic. Engineering approaches, based on both vertical and horizontal forces, are then used to allocate that element of costs that is deemed variable down to individual vehicles (see, ORR, 2013). This approach means that Britain, unlike other European countries, has highly differentiated charges, by vehicle, which should incentivise the use of more track friendly vehicles.

At the same time, top-down econometric methods relating actual costs to traffic volumes, controlling for other factors have been extensively used and the results used by the European Commission (see, for example, Johansson and Nilsson (2004), Wheat et. al. 2009 for a summary, and Andersson et. al., 2012 for subsequent developments in modelling of renewals costs). These studies have covered a range of European countries, and suggest that the marginal cost of rail infrastructure maintenance is in the region of 20-35% of maintenance costs (or up to 45% for heavily used sections). Wheat et. al. (2009) found that the available evidence was much less strong for renewals, though suggested an indicative overall cost variability proportion of around 35% of renewal costs. More recent evidence has put this at a higher level;

at approximately 55% (Andersson et. al., 2009). As noted above, these methods have been useful in determining the extent of cost variability with traffic in general, but less effective at allocating to types of traffic or vehicle. It is worth further noting that the engineering based bottom-up approach used by Network Rail puts the variability proportion at less than 10% which is out of line with the top-down econometric evidence from across Europe.

Our proposed approach is therefore positioned within the existing literature and would particularly enhance ORR's approach to allocating variable costs to vehicle by providing new evidence on the relative cost of different damage mechanisms, in turn leading to better estimates of the relative marginal cost of different vehicle types. The approach could also be used to determine absolute marginal cost and cost variability levels to compare against the results of top-down and other engineering models.

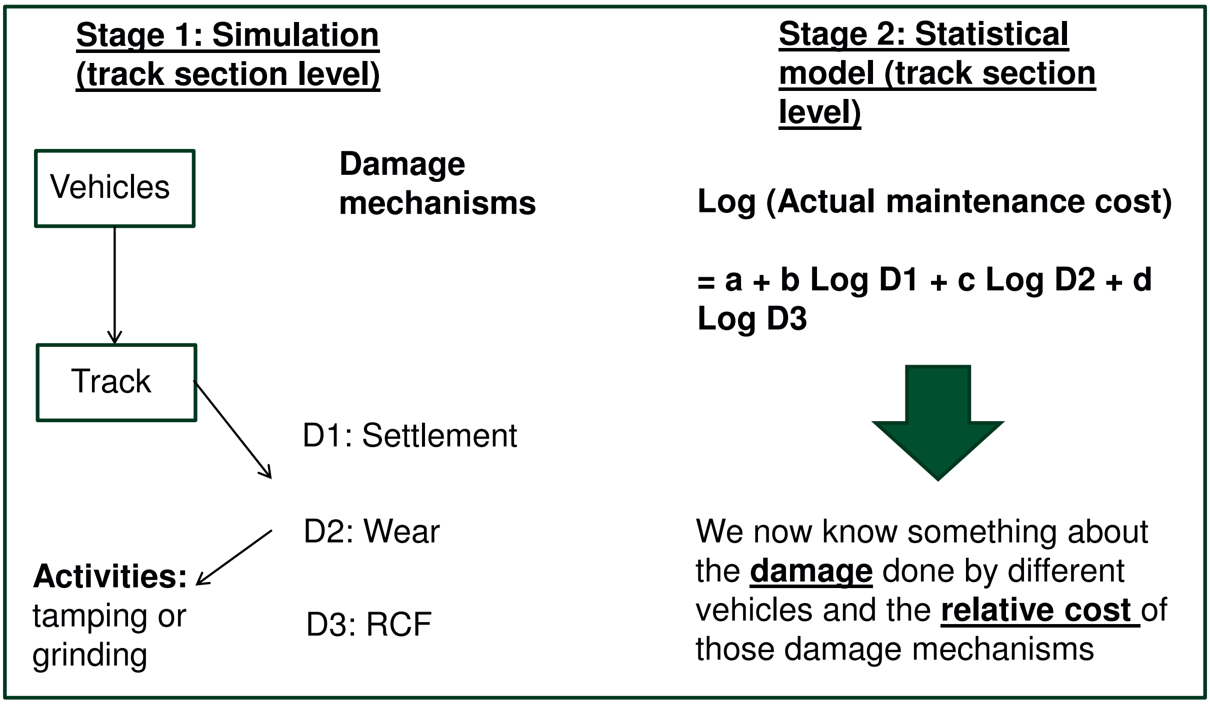
3. Methodology

Our proposed method essentially consists of two stages. The first stage involves an engineering simulation exercise in which traffic (of certain vehicles and mixes of vehicles) is run down a network of known characteristics, to produce estimates of the resulting damage (denote these D1, D2, and D3, to represent the three main damage mechanisms: settlement, wear and rolling contact fatigue). For this exercise we choose actual track sections from Sweden where we have data on the maintenance and renewal costs, the traffic volumes and the infrastructure characteristics. The second stage involves establishing a statistical relationship between actual costs (maintenance and or renewal) for actual track sections on a network (in our case the Swedish network) and damage. The approach is summarised in Figure 2 below.

The three damage mechanisms are defined as follows. Wear of the rail is a natural process in which material is removed from the head and/or the gauge corner of the rail when railway vehicles run. The rate of removal is affected by the forces and contact conditions. Severe wear can change the cross sectional profile of the rail, resulting in a change of the running surface seen by the wheel. Rolling contact fatigue (RCF) occurs if the rail surface is subjected to repeated plastic deformation as is often caused by repeated wheel passages.

The 'Tgamma' value provides a measure of the wear and RCF performance of the vehicle. Tgamma is the product of the tangential or creep forces and the slippage or creepage in the contact patch between wheel and rail. Tgamma combined with a non-linear damage function produces a RCF damage index as shown in 3. This index is then used to interpret whether the vehicle is damaging the track due to wear, RCF or a combination of both.

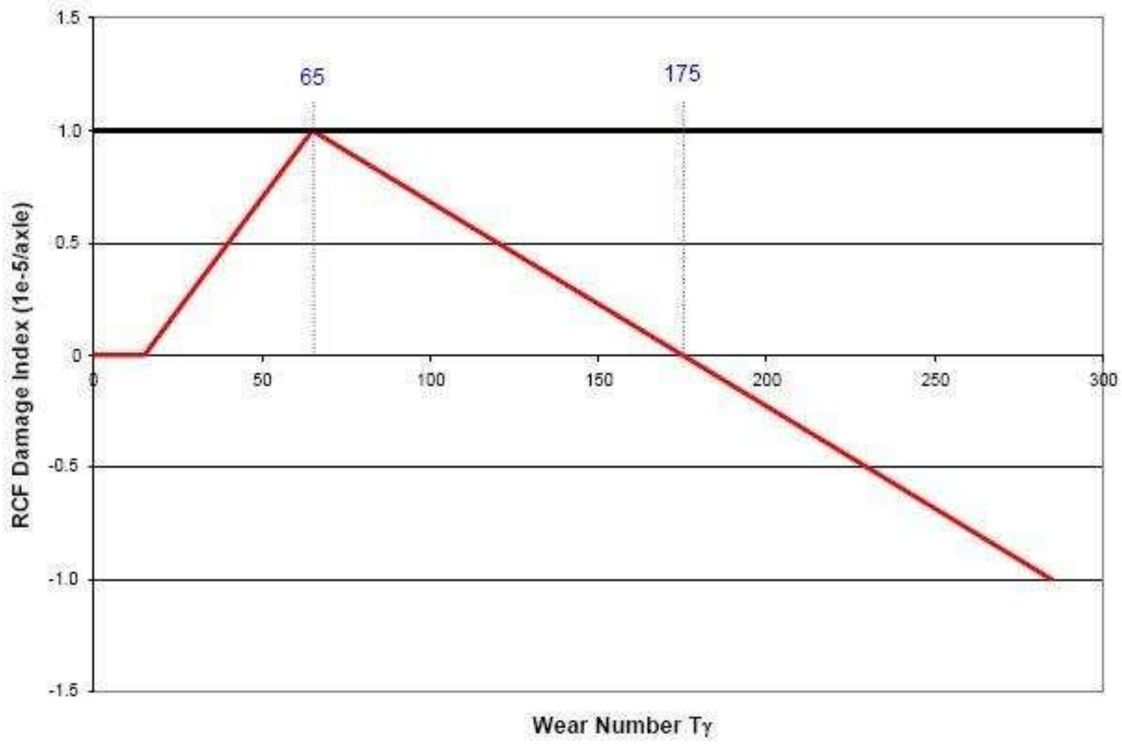
Figure 2: Overview of the methodology



With reference to 3, as Tgamma increase from 0 to 15 N, no RCF damage is generated as there is insufficient energy to initiate RCF cracks. As Tgamma increase from 15 to 65 N, the probability of RCF incitation increase, to a maximum of 1 at a Tgamma value of 65 N. As Tgamma increase further from 65 to 175 N, the level of energy is such that the dominant form of surface damage is wear (rather than crack initiation), therefore the probability of RCF damage decreases as wear increases. Negative values of RCF damage indicate values of Tgamma greater than 175 N, resulting in wear and no RCF initiation. The units of the RCF damage index are 10-5 per axle. This indicates that for a damage index of 1, 100000 (One hundred thousand) axle passes would result in RCF initiation.

Finally, track settlement may be defined as the sinking of the track (in the vertical plane) into the ballast under a variety of conditions. A number of models have historically been used to predict track settlement. Initially the Technical University of Munich (TUM) model and the Sato model were considered. It is assumed that the settlement is going to be proportional to the vertical force on the wheelset.

Figure 3: Relationship between wear number (Tγ) and RCF damage index



Estimates of the relative cost of different damage mechanisms can be estimated from the parameters in this second stage regression (the b, c and d parameters in Box 2), which in turn allows us to estimate the relative cost of different vehicles. Of course, a more complex relationship could be assumed in the second stage statistical model, for example to include interactions between damage types (e.g. D1*D2). The detailed assumptions for stage 1 of the approach, as they apply in practice, are set out in section 5, after the description of the data. As noted in the introduction, stage 2 of the approach is not implemented in this paper due to lack of data at this time.

4. Data

There are four main aspects to the dataset, which has been provided by the Swedish Transport Administration (*Trafikverket*). These are, for the damage simulations, data on track alignment and vehicle characteristics. Secondly, for the second stage modelling, data on costs (maintenance and renewals), traffic volumes and infrastructure characteristics.

So far the analysis has focussed on five track sections. Data for another one hundred sections has been requested but has not been received in time for inclusion in this conference paper. The additional data points will be used in subsequent analysis to implement the second stage.

4.1 Vehicle data

Three vehicle types have been modelled running down the selected five sections (see Table 1). These include two types of freight locomotives (with associated wagons) and one EMU. The data provided by Trafikverket revealed the generic type of vehicles and their usage on the network. However, the information was not sufficiently detailed to identify the exact vehicle. Hence generic freight and EMU vehicle types from the EU funded research project, INNOTRACK[1], were used. These types are considered to represent typical freight and EMU vehicles running on the EU network. In future work it is hoped to obtain more detailed information on the exact vehicles running on the Swedish network.

The vehicle models were described in terms of mass properties, geometry, axle load, unsprung mass, wheel radius and suspension characteristics. The VAMPIRE simulation package uses the model description to generate equations of motion for the vehicle which it then solves with the track input. Once the simulations are completed, outputs are available for wheel-rail forces vehicle motions and can be used to estimate damage.

4.2 Track data

Track alignment data were obtained using a track geometry coach which measures the curvature, cross level, vertical irregularity, lateral irregularity and gauge at fixed intervals of 25cm. This data was then used to create VAMPIRE track input files for the chosen track sections. An example of the track data for section 629 is shown in Figure 4. The five graphs represent:

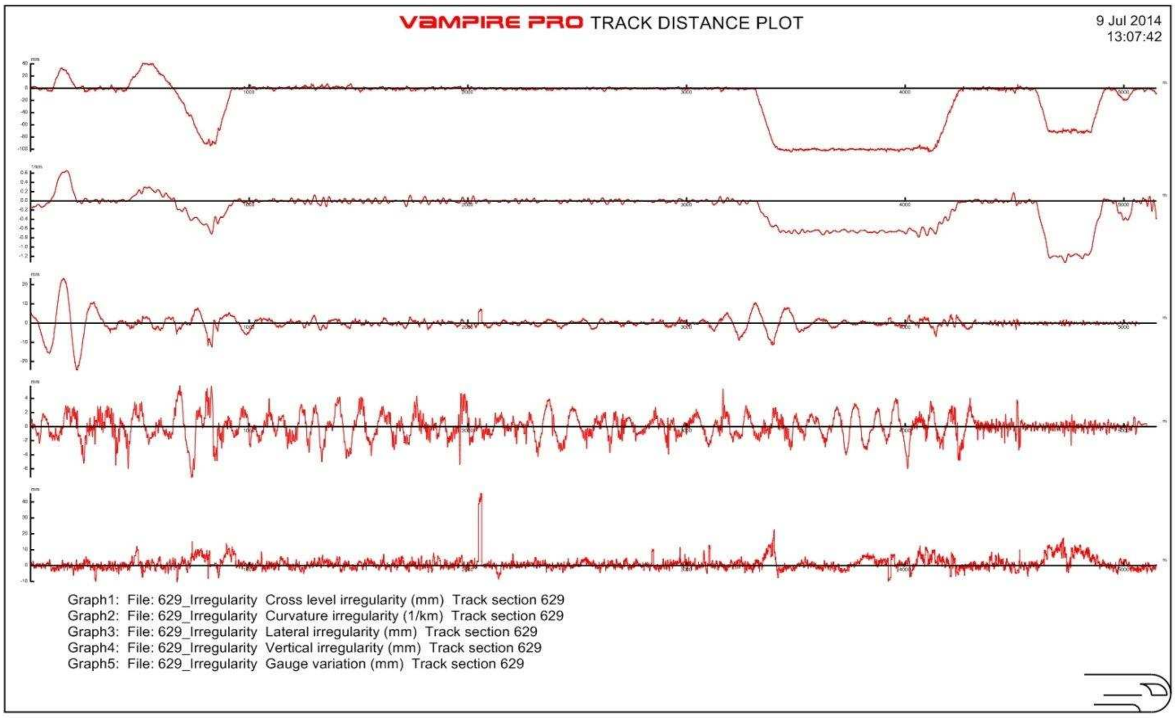
1. Cross level versus Distance along the track
2. Curvature versus Distance along the track
3. Lateral Irregularity versus Distance along the track
4. Vertical Irregularity versus Distance along the track
5. Gauge Variation versus Distance along the track

[1] http://www.innotrack.net/article/presentation

Table 1: The vehicles running on each track section

Track Sections	Vehicle ID	Type
326	RC2	Loco
	RC4	Loco
	4-axle Wagon	Wagon
	2-axle Wagon	Wagon
629	RC2	Loco
	RC4	Loco
	4-axle Wagon	Wagon
	2-axle Wagon	Wagon
652	X12	EMU
	X14	EMU
654	X12	EMU
	X14	EMU
821	X11	EMU
	X31	EMU

Figure 4: Track alignment data for track section 629



VAMPIRE Plot

4.3 Second stage data: costs, traffic and track characteristics

Costs for maintenance and renewals are collected in Sweden at the track section level. Included in these costs are activities conducted to maintain the railway assets as well as replacement and refurbishments to restore the assets to original condition. Snow removal is defined as maintenance in Sweden, though these costs are excluded in our data set. Renewal activities are not as frequent as maintenance, and a renewal cost is only observed for one of the track sections (see Table 2).

The simulation models produce an overall estimate of damage per tonne-km for each vehicle type. It is therefore necessary to scale up these measures based on the volume and mix of traffic that is actually running on the network. Traffic data include the declared train kilometres, weight (gross tonnes), number of wagons, and train length for different types of locomotives or EMUs/DMUs. The data is summarised in Table 3.

Table 2 – Costs, SEK

Track section no.	Maintenance costs	Renewal costs
326	5 757 077	132 707
629	323 503	0
652	15 609 488	0
654	8 735 644	0
821	3 566 033	0

In the second stage it is also important to recognise that the cost of addressing different levels of damage will depend on the characteristics of the infrastructure. Additional data has therefore also been collected on measures such as rail weight and linespeed (these are not shown to keep the discussion tractable).

Table 3 – Traffic data

	Train-km			*Tonne-km*		
Track section no.	Passenger trains	Freight trains	Empty trains and service trains	Passenger trains	Freight trains	Empty trains and service trains
326	8 150	212 521	15 247	1 676 878	192 837 314	1 208 129
629	18	2 983	95	2 810	684 166	5 502
652	285 707	15 728	22 586	38 554 466	15 262 308	3 442 515
654	312 434	5 368	3 404	33 160 000	5 327 958	338 685
821	325 744	16 927	53 390	53 677 213	14 306 152	8 709 709

5. Results and discussion

The output from the simulation runs are the vertical track force and 'Tgamma', which has been described earlier. The vertical track force was used to calculate a damage index, which is the vertical force raised to the power 1.21. This damage index is then used to calculate track settlement. Tgamma was used to calculate wear and RCF damage. All damages for each track section were measured per gross tonne-km.

It should be noted that damage is estimated at a number of points along the section (that is, each section is split up into segments covering 40 or 200 metres for wear/RCF and settlement respectively). Thus it is necessary to produce a measure of overall damage for each of the three damage types over the whole section. In the case of wear and RCF these are the summations of the damage over each of the segments. In the case of settlement, an average measure is taken, as a summation would not make sense from an engineering perspective.

The summary damage measures per gross-tonne km are then scaled up based on the total annual gross tonne-km run by each of the vehicle types in Table 1 in order to produce an annual total estimated damage. It is this measure that can then be related to annual costs in the second stage of the modelling. However, a mix of other vehicle types has also run on these track sections accounting for a small proportion of the total gross tonne-km. The damage, wear and RCF from these vehicle types are assumed to be proportionate to the damage measures caused by the vehicle types in Table 1. We have therefore scaled up the measures with respect to the total tonne-km run on the sections. The summary, total damage measures are shown in Table 4 below.

Table 4 – Total damage, wear and RCF

Track section no.	Vehicle	Loading condition	Damage	Wear	RCF
629	RC2	Laden	9 235 914	274 163 841	886 093
	RC4	Laden	13 895 739	408 431 752	1 285 242
	Wagon 1	Tare	96 482	2 001 759 475	783 707
		Laden	58 945 934	916 350 922	7 014 644
	Wagon 2	Tare	7 371	6 614 805	70 194
		Laden	673 991	5 098 989	57 969
326	RC2	Laden	1 175 899 756	47 387 842 433	68 569 622
	RC4	Laden	1 200 996 700	48 100 608 862	68 487 534
	Wagon 1	Tare	375 216 324	3 136 983 770	33 453 667
		Laden	23 018 827 311	366 396 576 247	2 788 072 442
	Wagon 2	Tare	1 217 851 938	4 364 299 435	14 699 661
		Laden	4 079 811 306	9 770 716 020	67 225 065
821	X11	Tare	9 496 796	16 774 577 118	47 132 486
	X11	Laden	42 756 976	59 680 115 333	150 847 091
	X31	Tare	56 019 027	100 016 274 153	279 023 469
	X31	Laden	334 761 323	471 596 050 478	1 187 786 786
654	X12/14	Tare	0	0	0
	X12/14	Laden	1 165 046 538	14 559 357 588 564	544 327 218
652	X12/14	Tare	27 638 191	2 740 211 270	20 480 736
	X12/14	Laden	2 277 950 252	132 724 585 837	1 371 949 647

The damage measures in Table 4 are thus summary, total simulated damage measures per section, per year. It should therefore now be possible, once the same modelling has been carried out for more track sections, to estimate a relationship between these damage measures and actual costs on those sections. This second stage will enable us to understand the relative cost of different damage mechanisms and also how the different damage types and the cost of correction interact with each other to affect overall cost. This new information on the relative cost of damage, based on actual reported cost measures, will be an important contribution to the literature, where currently the link between damage and costs relies on a set of assumptions that may not be valid.

6. Conclusions

This paper proposes and implements the first stage of a new approach to estimating the marginal cost of rail infrastructure usage. The approach fills an important gap between the modelling of damage – which is well understood – to obtaining measures of marginal cost, where currently the relative cost of different damage mechanisms is not well understood. The approach combines engineering and statistical approaches in a way that seeks to combine the best features of both; and thus overcome some of the weaknesses of previous approaches.

A new dataset is utilised, comprising detailed information on individual track sections of the Swedish network. This data includes information, at the section level, on track alignment and actual maintenance and renewal costs, as well as information on the vehicle volumes and types running on those sections.

Through testing this approach out on five sections, we conclude that, as expected, it is possible and indeed relatively straight forward to model the damage resulting from running vehicles on the network. We have further shown, however, that is possible to produce summary, total (annualised) damage measures for each of the three damage mechanisms. Once the approach is extended to approximately one hundred sections, as planned, it will then be possible to explore the relationship between these damage measures and actual costs for the first time; and thus better understand the relative costs of the different damage mechanisms. Such information will be highly valuable to academics, the industry and policy makers as it will allow the cost implications of different technologies to be more clearly assessed. Further, from a track access charge perspective, the approach will permit improved estimation of the relative cost of different vehicle types, allowing more cost reflective charges for different vehicles on the network. In turn, more cost reflective access charges should incentivise the development and use of more track friendly vehicles.

There are a number of aspects to the research where assumptions have been used, for example, concerning the precise nature of the vehicles running on the network (we could identify passenger versus freight vehicles, and locomotive versus EMU, but had to use generic types for these rather than the actual vehicle characteristics). It has also been assumed that the damage caused by one vehicle is independent of the damage caused by other vehicles, and that the damage measures from one vehicle run can be scaled up in a simply manner. These are limiting but pragmatic and sensible assumptions; however, relaxing them would be interesting and useful for future research. Validation of the damage estimates, as compared to actual measurements would also be a useful addition if the necessary information can be obtained.

References

Andersson, M., Smith, A.S.J., Wikberg, A., and Wheat, P.E. (2012), 'Estimating the marginal cost of railway track renewals using corner solution models', Transportation Research Part A, 46 (6), 954-964.

European Commission (2013): The Fourth Railway Package – Completing the Single European Railway Area to Foster European Competitiveness and Growth.

Johansson, P., and J-E. Nilsson (2004): 'An Economic Analysis of Track Maintenance Costs', Transport Policy, 11(3), 277-286.

ORR (2013), Final determination of Network Rail's outputs and funding for 2014-19.

Wheat, P.E. and Smith, A.S.J. (2008), 'Assessing the marginal infrastructure maintenance wear and tear costs for Britain's railway network', *Journal of Transport Economics and Policy,* vol. 42 (2), pp. 189-224.

Wheat, P., A.S.J. Smith, and C. Nash (2009): 'Deliverable 8 - Rail Cost Allocations for Europe', CATRIN (Cost Allocation of TRransport INfrastructure cost).