Summary of contents:

RESISTANCE FORMULAE

With Doug Landau's reply to mine in MP 37½ this topic has changed to Steam Locomotive Resistance. There can be little debate about the Vehicle Resistance of the locomotive, so this letter is about the additional resistance, Machinery Resistance (MR).

A correct analysis of MR has to allow for resultants and offsets. One resultant occurs at Coupled Wheel Bearings (CWB), that of (a) static load vertically, and (b) piston thrusts, propulsive, compressive and dynamic, fore-and-aft, through the drive, at various angles near to horizontal. Item (a) is part of the Vehicle Resistance (VR). If (r) is the resultant of (a) and (b) at any point in a revolution, (r) – (a) is something additional to VR, and part of MR. That is simple geometry and arithmetic. If anyone wants to consider (r) alone, the same Locomotive Resistance (LR) will result, but proper analysis of MR per se will be prevented by some of the machinery effects being bound up in the resultant.

I do not understand why Doug sees a need to deduct cylinder frictional losses and what from. These are presumably of rings on cylinder walls. Such friction is positive and a component of MR. It does not depend on Piston Thrusts (PT) but on the pressure on the rings at each point of the piston stroke. Those pressures are the same as those determining the propulsive and compressive PTs, at the same points.

MR arises only after the effects of forces which oppose one another net out. MR is therefore MR, and net is superfluous. Doug thinks MR as a function of speed is more practical. He does not say than what or why, but presumably thinks thus because such would be simpler than a function which allows for the components of MR per se, (again presumably) so that it can be easily added to a VR to give an LR equation of the $a + bV + cV^2$ form. That seems not worth pursuing if LR is to be even reasonably soundly established, because the influences on MR are not dependent on weight, and the V^2 element in MR has to do with various masses, whereas the V^2 in VR depends on vehicle cross section area. In addition, the relevant masses differ considerably from engine class to class, on account of the differing extent to which reciprocating masses are balanced, the number of cylinders, and if more than two, the way they are arranged. Further, MR decreases or only slightly increases at higher speeds as VR increases (see further below on constancy of MR). True, the effort being developed at various speeds needs to be known (Doug's reference to an assumed IHP) to estimate the MR, but that problem can be overcome simply by iteration (described in my paper mentioned on p 213 of MP 34½, available on application to me at johnk.pb15@virgin.net). I have no practical problems dealing with MR separately from VR. Indeed, in arriving at the ITE of a steam locomotive I establish all other resistances first, those to the coupled wheel rims (rail tractive effort, RTE), and then add MR.

Many aspects of LR have only a modest sensitivity to the determinants (Doug's reference to sensitivity to effort). That is very likely in MR. Effort is high, friction coefficients low. The latter are mostly below .05 (a handful above), so that would be expected. A particular force (especially piston thrusts working through the drive) can act in full or part at several places where friction occurs, however, multiplying the rate of variation.

Knowing the fixed and slightly varying effects properly is as important as knowing those which vary strongly. A considerable proportion of MR is dependent on piston thrusts, especially at lower speeds. The extent of MR in total, its variation with effort for various efforts, and what proportion it is of ITE and LR for an LMS Class 5 can be appreciated from the following table for two levels of output at three speeds, estimated as shown in that paper. The first IHP at each speed represents about the best usually observed steaming rate at the speed, and the second half that rate. VR, MR, LR and ITE are in lbsf.

MR, VR and LR of LMS Class 5

	30 mph		50 mph		70 mph	
VR (still air)	770		1240		1900	
IHP	1375	688	1500	750	1550	775
MR	1420	940	1060	920	1100	900
LR*	2190	1710	2300	2160	3000	2800
MR as % ITE	8.2	11.0	9.4	16.4	8.3	6.7
MR as % LR	65	55	46	43	37	32

*LR = VR + MR

There are other influences, especially coupled wheel diameter and number of coupled wheels. The Queensland Railways C19 4-8-0 with 4ft diameter CWs working at full effort (about 67% cut off) with full load on a gradient at the usual 10 mph, had a VR of about 460 lbs and an MR of about 1420 lbs, MR 75% of LR. Dependence on V^2 is low in both VR and MR at 10 mph.

If an engine can be judged from the RTE to have been working hard, the speed is in the range of 200 – 350 rpm, and the engine carries 225 – 250 lbs working pressure, MR can be approximated satisfactorily by using 8 lbs mean effective pressure in the tractive effort formula, down to about 6½ lbs at 160 lbs pressure. This shortcut assumes an average percentage of reciprocating masses balanced – it is unsuited to low or zero reciprocating balance. The 30 mph column above, and the paragraph above show that at lower speeds, at which maximum ITEs are developed, an average or constant MR is unsatisfactory. Above 350 rpm or so, the same applies, because the V^2 element in MR becomes considerable.

Doug states that errors of 100lbs in MR or a part of it are tiny in horsepower terms. The statement requires a comparator of 100% accuracy to identify errors, and its import depends on the number of hundreds in the error.

Doug defends the mathematical fitting (trial and error basis decided by the analyst) of resistance curves, and excuses negative coefficients on terms which should from first principles be positive as refining the answer. I do not deny the likely calculation effort or decry the intention, certainly for the period, but deplore the claim. The desire should not be to best reproduce the data, but provide the best scientific (statistical) fit to it, together with the test statistics, which allow establishing the probability that the values of the coefficients and of the answers differ significantly from results of other analyses, or from zero. The same applies to the trendlines which EXCEL allows. These are chosen at will by the user, and might (often do) mean nothing. Doug should not be concerned about a proper regression line (rather than an EXCEL trendline) not passing through the actual data. A best fit will often not pass directly through any of the data. No method of analysis can make up for poorly measured/inaccurate/inconsistent data or improper specification of the equation to be fitted. If the equation is based on the proper physics of the problem, then any failure of the equation to live up to expectations is almost certainly the result of unsatisfactory data.

Par excellence, it is not hard to show (details on request) that the data of the pull on the Amsler dynamometer at the drawbar of locomotives tested on the Rugby Testing Station cannot be right. A series of articles *Locomotive Testing at the Rugby Plant BR*, appeared in the *Locomotive Railway Carriage and Wagon Review* in 1957. No author was named, but directly or indirectly, D R Carling, Superintending Engineer of the plant, was almost certainly the author. In the December issue, pp 233-4, it is said that despite all the favourable circumstances, it is *not* (his italics) possible to measure the internal friction (ie MR) of a locomotive accurately on a test plant, only to confine that value within comfortably wide upper and lower limits. As lower limits measured were negative, which is technically impossible, the comment on the lower limit at least is not helpful.

John Knowles

25th October 2016.

Reply from Doug Landau

The second paragraph positing the "correct analysis" of MR, sets out the salient forces, (a) and (b), producing resultant force (r). It then seeks to isolate and determine (a) (the static loading on the coupled axles), as a component of vehicle resistance (VR). This is utterly pointless, it needlessly complicates matters and leads to miscalculation. There is not the remotest possibility that a design office would treat the matter in this way: from the pistons via the motion to the coupled wheel rims, the losses would be treated as the power transmission system losses, in other words, MR. The outcome, after all, is exactly what is being measured by the test plant dynamometer. The additional resistance to determine the total LR is simply to add the VR of the uncoupled wheels plus the aerodynamic drag and the track resistance (the b term) for the whole locomotive and tender. Why is it thought necessary to isolate an intrinsic element of MR and treat it as VR? Nothing whatever is to be gained by doing so. The resultant of weight, traction and dynamic forces will be less than the mathematical sum of the parts, and cannot sensibly be isolated for the purpose of analysis.

The third paragraph appears not to understand (or has misunderstood my point), that piston frictional losses will reduce the connecting rod big end and coupled wheel journal loadings; not a huge quantity but finite nevertheless.

I'm unclear as to how the fourth paragraph was derived from what I actually said. I neither think nor say that MR is simply a function of speed, it can obviously be presented that way once MR has been determined, and as such is pertinent to the determination of LR HP. MR is clearly the product of manifold forces and elements; simple and dynamic, windage and frictional, weight and mass. Speed expressed as RPM is obviously relevant to the dynamic and windage elements of force. To say "the influences on MR are not dependent on weight" is pure nonsense, only propounded by an untenable view of the mechanical reality. Are we to suppose the losses attributable to axle load only kick in when the locomotive actually moves along a track and are absent on the rollers? Obviously not, John's pursuit of MR in

PMF terms (Pure Machinery friction,) is beyond logical comprehension. On what grounds is the numerical demarcation of simultaneous forces acting on a common point justified?

The MR values in the table are upwards of 50% higher than the values recorded on the Rugby test plant (of which more below). Likewise estimates using the suggested 8 lb MEP formula. The idea that a " scientific stastical" fit is axiomatically superior to the empirical evidence contradicts one of sciences basic tenets; repeatability, something the Rugby WRHP data amply demonstrates. The so-called statistical science is no such thing since it will involve many assumptions in regard to friction coefficients and so on.

John says (last para); "As lower limits measured were negative, which is technically impossible, the comment on the lower limit at least is not helpful." Technically impossible yes, but statistically quite probable. The problem is the relatively small difference between two large numbers which are subject to experimental error. Experiments with a random number generator, where notionally perfect ITE - WRTE data was entered (the answer was always 800 save for the fact the two inputs were randomly varied by up to +/- 2%), showed that negative values would occasionally occur. The programme was such that if a single entry was changed the whole data set of 70 entries was rescrambled, so it was possible to quickly generate numerous simulations of test data. The scatter patterns were very similar to those seen in the *later* Rugby test plant data. Unsurprisingly, when the difference was reduced to 600, the incidence of negative values increased. The +/-2% by the way was as the stated limitations of the test plant equipment and proceedures. These simulations were a simplification to the reality on the test plant, where variations in boiler pressure increased the natural scatter when plotting Willans Lines (Steam rate Vs IHP and WRHP). I said *the later* Rugby data because negative MR values were rampant in the early test data (70005/25, 35022 and 73008). For 3 test series from 1951/52 158 MR readings were recorded, of which no less than 95 (60%) were negative, and most of the remainder were improbably low. For the 12 test series 1953/59, of 572 MR readings 5 (<1%) were negative, in line with the simulation predictions, MR was averaging hundreds of pounds.

Clearly something changed post 1952. The recorded wheel rim horsepowers (WRHP) are consistent across these periods where the same locomotive or locomotive types were involved. The comparative data available is a bit random in the sense that the speeds adopted across the various test series varied somewhat. The BR5 tests with 73008 (1951/52) & 73030 (1953), when fitted with 5.125" blast pipe caps as first built, returned R squared values approaching unity for WRHP Willans lines (steam rate plotted against HP) at 20 mph (20 plots) and 35 mph (27 plots). When 73030's blast pipe caps were reduced to 5" and then 4.875" in the pursuit of improved steaming; the recorded WRHP reduced at each step. Later tests with 73031(1958), 4.875" cap, enabled comparisons with 73030 so fitted, again returning Willans lines of high consistency for 35 mph (10 plots) and 45 mph (12 plots). These comparisons were as for 73031 in standard condition in regard to the superheater arrangements. WRHP Willans lines for the various 9F test series again return R sq'd values approaching unity (>0.99): 92013 (1954) and 92050 Series 2 (1957) at 15 mph (18 plots); 92050 Series 1 & 2 (1955 & 58) at 30 mph (14 plots); etc,etc. The Crosti 9F 92023 was an exception, with higher machinery friction at all speeds, amounting to about 60 HP at 40 mph, a figure confirmed by the Crosti's reduced DBHP established on comparative road tests.

The measurement of WRHP was the simple product of drawbar pull and RPM, a process automatically recorded, monitored and controlled by a Mediating Gear under the control of a servo mechanism. For the benefit of readers unfamiliar with the Rugby test plant, the rollers were set with the coupled wheels sitting directly above set at top dead centre (TDC) using a special gauge. After a warm up period of some 40 to 60 minutes, stable running and steaming conditions having been reached, the test period began. The positioning of the coupled wheels relative to TDC was monitored by a differential gear box which measured the Mediating Gear inch seconds. Provided the fore and aft motion in the course of a revolution was equidistant about TDC, no inch seconds would be recorded, and the same inch seconds would be recorded at the beginning and end of the test period. A test sheet for 70025 at 30 mph registered a start/finish discrepancy of 3 inch seconds accumulated over 3618 seconds representing a negligible average shift from TDC of 0.0008" over the 1 hour test period. The WRHP was determined by the dynamometer integrator HPhrs over the whole test period, not spot readings. The amplitude of the fore and aft motion was moderate, on a demonstration run with 70025 working quite hard on 40% cut-off at 25 mph, it was within 1/8 of an inch. A nest of Bellville washers in the drawbar absorbed the disturbing forces, preventing any tendency for resonance to develop; in the words of test engineer Jim Jarvis, the Bellville washers "breathed". The differential gear box also operated the servo mechanism which automatically held the locomotive via the mediating gear at TDC.

The performance of the Farnboro indicator equipment at Rugby was somewhat chequered in the early years of operation. The "balanced pressure" sensors (for details see *From Shovels to CTs*, page 21 on the RPS website), that were key to the production of the indicator diagrams were mechanically and electrically unreliable and failure was frequent. Some correspondence with Ron Pocklington, who was involved with the operation and improvement of the equipment, spells out the various tribulations in detail: "I endeavoured to sort it out to become reliable and precise, including an accurate assessment of the dead centres as a reference and the compilation of the stroke diagram and its IHP assessment". In January 1953 some comparative tests were carried out between the Farnbro indicator and two mechanical types (Maihak and Dobbie McInnis) provided and operated by Swindon engineers. The initial results found the mechanical readings about 7% higher than the Farnbro, the resulting check found the Swindon calibrations to have been in error. After correcting for this the Maihak readings were consistently 2.3% higher than Rugby, the corrected D & M error averaged 3.9 % high but had the curious characteristic of being inversely proportional to steam rate; 7.2% high at the lowest rate falling to 0.7% at the highest. Some further comparative tests were carried out in early March 1953 between the Rugby and Derby versions of the Farnbro indicator. While both operated on the same basic principal the Derby model used a piston rather than a diaphragm as the balanced pressure interface. Vis a vis Rugby, the Derby results were scattered on, above and below, averaging 2% higher. In summary, the Rugby indicator was the lowest reading of the four indicators tested. Perhaps Carling and associates found this persuasive; the IHP curves in the Britannia test bulletin (Fig. 15) are measurably higher than the Rugby experimental data. Over time a process or trial and error achieved improved reliability and sensitivity, a modified diaphragm "produced the standard of diagram so long sought after". In 1955 some further comparative tests between the Rugby and Derby indicators on 9F 92050 showed closer agreement than previously, the Derby readings were 99.3% of the Rugby average, reversing the earlier result of Rugby being the lowest.

The tabled LR and MR values for the Black 5 are high relative to the empirical evidence. Report L116 reconciling 92050 road test steam rate anomalies includes a 9F LR curve, at 30 mph LR is 1680 lb, equating to steam rate 16,000 lb/hr, 1100 IHP. At this work rate (IHP), pro rata John's table, the Black 5 MF and LR works out at 1230 and 2000 Lb respectively, the latter *19% higher than the 9F*. At 5.6% the Black 5 MF sensitivity to Indicated Tractive Effort (ITE) is high relative to the test plant results for BR5 73031; 49 plots of WRTE v ITE at 30 mph return a sensitivity of 2.7%; at 1100 IHP the MF is 790lb. The R squ'd value was 0.9956 reflecting the low scatter. This is of particular interest since the 49 plots involved a wide range of superheat, with steam temperatures ranging from 450 to 750 Deg. F. At the lowest temperatures for a given IHP cut-offs were about 2.5% longer than at the highest.

The last paragraph citing Carling's observations regarding the uncertainty surrounding the determination of machinery friction omits his preamble. Here he dwells on the small remainder problem, setting out a numerical example. Writing in the in the Model Engineer, 7 November 1980, he again addresses the small remainder problem and gives a similar numerical example,a problem he describes as "very vexed" and "notorious", the difference is that on this occasion he was talking about locomotive resistance not machinery friction. Given the greater potential for variables, any upper and lower uncertainty limits for LR should be set wider than is the case with MR. Freed from the small remainder problem Carling regarded WRHP readings as a reliable bench mark of performance, and used them to monitor the before and after performance of the 9F 92015 regulator modifications, as published in *The Locomotive*, November 1958. The effect of the modifications proved insignificant, the minimal scatter of the before and after WRHP Willans Line plots was clearly evident.

Yours Sincerely,

Doug Landau.

Steam Locomotive Resistance

John Knowles

I comment on Doug Landau's letter of 2.12.16.

Doug's testy first paragraph remarks on my isolating the addition caused by piston thrusts to the coupled wheel bearing resistance (CWBR) of the vehicle resistance. He claims that this is utterly pointless, needlessly complicated, miscalculation, not what would be done by a Design Office, nothing to be gained, cannot be sensibly done. At least I have not been accused of treason! I disagree on all counts. His strong words are not accompanied by any examples of the terrible effects of my supposed error. I challenge him to show how there can be a miscalculation. Of course the resultant is less than the sum of the parts, but that does not mean that the effects of each part as additions to CWBR of the vehicle alone cannot be isolated. I obtain the addition – I do not make an arbitrary division of the resultant. I have found it useful to isolate the addition to CWBR in obtaining from first principles the parts of MR subject to piston thrusts, as in my process (1) below. In my terminology, MR excludes the CWBR of the vehicle resistance, but includes any extra loading thereon from mechanical effects. My approach cannot make any difference to LR. Further it has to it the logic that the locomotive is first a vehicle, and that without the vehicle the mechanical functions cannot be applied.

I know of no other analyst of the subject than Doug who considers that the whole of the resultant is part of MR. I can see that if he wants combined MR and the CWBR of the vehicle resistance for an LR and he is confident that the Rugby data is correct, he will do it his way. My difficulty is that I think the Rugby data poor/inadequate, only a handful of the world's locomotives were tested at Rugby, and I work at MR and LR more generally, for application to other locomotives. What is easy for him in principle for a handful of locomotives is only a tiny part of the need for well informed MR and LR.

My fourth paragraph was about statistical analysis by regression, which cannot have been a misinterpretation of a comment by Doug, because he has never used it, and appears not to understand it. He claims to have fitted an equation to some Rugby WRTE data (actually DP, dynamometer pull) and obtained values of r^2 of almost one, presumably as an indicator of how good the Rugby DP data is. What variable he chose to fit DP to, what form the equation took, and the results are not revealed, nor any statistical tests. I presume the work is really an Excel trendline of DP on Q, the steam rate, of a shape chosen by Doug, and not a regression at all. I suspect that if he compared ITE fitted to the same Q in the same way, he will have found that the difference between the two trendlines, an apparent MR, also varies strongly with Q, something which would never do for Doug, who advocates that MR is all but constant across the output and speed ranges. He can check that for himself. Such a way of using Rugby data to obtain MR is not valid, however; the direct ITE – DP data for each test is the source of MR plus CWBR (see below). (I know that Rugby used the term WRTE, but it was DP which was measured, and as will be considered below, a Damping Resistance could well intervene between the WRTE and the DP).

Doug's third paragraph responds to my point that a practical formula for LR, a simpler one than addition of a VR formula and an MR formula, which is difficult, for reasons I gave. Had he read my paper on Steam Locomotive Resistance, he would have seen that I have considered the subject. He says, without reference to my general point, that MR not being influenced by weight (of the locomotive) is pure nonsense, only propounded by an untenable view of mechanical reality, and that my pursuit of MR in terms of pure machine friction is beyond logical comprehension. On the last, if he searches the same paper, he will I use the term MR throughout, except to note in passing that Ell, a BR officer involved in the BR testing at Swindon, made an extremely low estimate of the MR of a Bulleid engine and called it PMF to distinguish it from other measures.

The rest of his remarks here are essentially the same as those in his first paragraph, except that these are richer in their insults. I have already explained that I see MR as the addition to VR. There is nothing wrong with that, it is capable of logical comprehension, so far as I know, by everyone interested except Doug Landau. The mechanical reality is not explained, but if it excludes what I do in isolating the addition to the CWBR of VR, it is not reality. Indeed, as Doug avoids the point of my remark about practical formulae for LR, that draws attention to the three term formulae for LR which he uses for calculations of steam locomotive output as far as IHP, which typically contain (so far as I can see) a constant MR in lbsf at all speeds and outputs.

Doug claims that my approximation to MR for certain circumstances is upwards of 50 per cent higher than Rugby. As Rugby MR is low (see below), I think worldwide evidence on MR, such as it is (a subject in itself), is on my side, and that Rugby offers no basis for comparison.

On his fourth paragraph and what follows, I have done three things which bear on the Rugby evidence, which will avoid Doug jumping to conclusions or at least enable him to sort out what I have done.

(1) I have established from first principles what MR might be expected to be, using various alternative assumptions in some cases for friction coefficients, a general approach for all steam locomotives, published in my paper. Doug is known to have attempted the same himself, but from what I know of it (it is not published), it omits some important influences. Such is not scientific statistics at all, as Doug believes (does he really call it that?) but applied mechanics. I have supplemented this by seeking empirical proof of the MR and LR from the literature. It is for this purpose that I isolate the incremental effect of different piston thrusts at the CWBs.

(2) I have analysed the TSMR (ITE - DP) data from Rugby to see how it compares with these principles, noting inter-observation consistency. I apply TS (Testing Station) to MR

because such data from a TS includes CWBR from VR. This I have done for all engines tested at Rugby after 1954 where there were at least a dozen observations at any one speed (V). One test is simply to graph TSMR against PTTE. This reveals tremendous ranges in TSMR for a given PTTE, and precious little of the repeatability Doug claims that the Rugby data possesses.
In this context, Doug says that for the 12 test series from 1953 until 1959 when steam testing ceased, of 572 MR readings only 5 (<1%) were measured negative, and in line with the simulation predictions, (TS)MR was averaging hundreds of pounds. As the simulation depends on the actual, it would be troubling if it did not predict the same. The Rugby figures are not the same as MR properly called, however. When the CWBR is removed to give MR per se, they become lower, and more become negative. As a further test, I have then excluded estimates of the resistance from the V^2 effects, and the constant of MR, leaving mostly the sources of resistance due to piston thrusts and rings. Almost all of these remainder observations are thereby reduced to values so low that they imply implausibly low friction coefficients, ie that Rugby data are generally low. Only 19 of the 158 observations in the constant speed data I examined could be said to show that they were the result of reasonable friction coefficients.

(3) I have analysed the same data as in (2) by statistical regressions, mostly all at one speed but in some cases across all speeds. This is where Doug makes some wild, sweeping and ill-informed statements. He claimed that this is a so-called statistical science, which requires so many assumptions such as friction coefficients to be no such thing. These remarks are quite wrong, an insult to the many people who apply statistical regressions in testing experimental data in all the sciences, and in establishing criteria for eg rejection of materials. The friction coefficients are used in (1) above, not (3), although they are used also in (2) as a criterion for the reasonableness of the Rugby data as just explained. Furthermore, Doug does not seem to be aware that such regressions are carried out on the observed empirical data.

In both (2) and (3) I too have analysed the effect of possible error ranges in the ITE and DP data. Doug's use of random numbers to show that these are what would be expected formalises that, but it makes no difference in the sense that the data are the data, and must be the basis of any analysis. Even knowing these ranges, the effects of the small difference between two large numbers problem could well prevent satisfactory data and analyses emerging.

I wrote about statistical regression in paragraphs 3 to 6, partly to avoid this knee-jerk reaction against it. Doug certainly did not seek to find out more about regression. There is a lot on the web about the subject, from simple to advanced, and many good books. The empirical data is used and tested in toto for its reliability. Regression provides a best fit to the data, and provides various tests which can be used to say how much confidence can be had in the results, in other words to say whether equations derived from the data can or should be sensibly used.

Regression of the ITE data (against Q and V) from Rugby is generally very good. Its consistency does not prove it to be right, however. Although I agree that the Farnboro' Indicator was eventually excellent, some of the ITE figures appear a bit low when tested by the Perform program. More apposite, I regressed TSMR (ITE – DP) against PTTE (piston thrusts expressed as tractive effort, this including propulsive, compressive and to and fro forces) at a particular speed where there are sufficient data at that speed. The logic is that an equation in TSMR should in those circumstances have a positive coefficient on PTTE and that the rest of TSMR should be included in a constant. The results for MR, however, are overwhelmingly disappointing, in terms of sense (ie behaviour and signs) and magnitudes, with wide standard errors of the estimate, low t scores on coefficients, high significance F values, and values of r^2 as low as 0.1. Neither the equation, nor the analysis is at fault, it is the poor, inconsistent data. Further, because the ITE data are generally good, the apparently erratic TSMR must be the result of the erratic DP data. With these results, no confidence can be placed in the Rugby ITE – DP data and results for obtaining MR.

I have also used Rugby data to apply the input/output approach to MR for a couple of classes, as used in obtaining the approximate MR of internal combustion engines. These yield MRs which are far too high. As in my last letter, all these results and a commentary thereon are available on request.

In his *Locomotive Testing Stations*, (IMech E and Newcomen Society (1973)), his last major statement about Rugby, D R Carling said that they ultimately got the (DP) answers right, but he did not say how that was done, nor how it was known the answers were correct. No mention is ever made, there or elsewhere, of using the proper dynamometer, that applying the braking on the plant, to check the DP measures. More important, however, and not mentioned by Doug, Carling was clear that they damped to protect the recording devices from the effects of resonance, not to perfect DP readings. Doug places a very favourable gloss on all of that. He omits mention of the dashpot, which after oil was removed from it, had air in it, and the frequency and magnitude of the forces affecting the apparatus. The to and fro forces came to an abrupt end at the ends of strokes several times per second (for a 9F, at 60 mph, this was 11 times per second at 60 mph, 3.7 times at 20 mph.), and could not be damped. It is impossible to dampen forces resulting from V^2 with a system operating in V.

Doug says that the measurement of WRHP (DP as a HP) was the simple product of drawbar pull and RPM, a process automatically recorded, monitored and controlled by a Mediating Gear under the control of a servo mechanism. The recording and calculation were separate from the mediating gear, which moved the engine as needed to keep the CWs on top of the rollers. To do that, it pumped oil into or from the hydraulic system which was the Amsler dynamometer used to record DP. How did it control (Doug's word) DP? How could the gear react several times per second to movements in both directions, ie was it capable of keeping up with the frequency of the sources of variation in DP?

The effect of the Belleville washers, air dashpot and mediating gear operating much more slowly that the fluctuating forces, must have regularly allowed the to and fro forces free rein, and at others resisted them. This would explain the large fluctuations recorded in DP at a given speed and PTTE and the erratic TSMR. If the Belleville washers and the air dashpot kept up with the fluctuations, there would have been frequent short hisses from both, rather than sighing. Indeed the delayed reaction could have added a damping resistance to the components of TSMR (as an extra positive item between ITE and DP, but not from the working per se of the mediating gear, the energy for which was outside the ITE – DP system).

Some comments on other remarks of Doug's. Of the many values in a considerable range of TSMR in the Rugby data for the various classes at any speed, which does Doug choose to be used as his TSMR, and why? Carling did not comment again about the plant being unsuitable for determining MR in the Model Engineer article in 1980, but the damping and measurement situation had not changed from his 1957 mention, so if he had commented, why would his opinion have changed? Avoiding damaging resonance was the prime function of the damping, and would have been first in his mind. Indeed, as he did not know that the DP results were right, he probably interpreted positive as being right. While I have read all of Carling's writings in the hope of guidance, I find little help from searching the runes.

There is little science about MR in Doug's letter, more criticism of my approach without troubling to read or understand it. I find it beyond belief that he feels so strongly about something that does not matter a scrap for correct LR.

STEAM LOCOMOTIVE RESISTANCE

DOUG LANDAU'S SPREADSHEET

I refer to this spreadsheet placed on the Society Website in January 2017.

General

Doug Landau is quite correct that any way of obtaining empirical evidence of steam locomotive MR, indeed of LR, is subject to the problem of that evidence being the small difference between two large numbers which are themselves subject to measurement errors, variations or defects in method. This problem is well known, not only in testing locomotives, and is dealt with in statistics textbooks. It is one of reasons why D R Carling thought the Rugby testing plant would not yield satisfactory figures for the internal resistance of the locomotive (MR + CWBR) (see *The Locomotive Railway Carriage and Wagon Review* December 1957 p 234).

What Doug Landau terms WRTE is DP, Dynamometer Pull, and what he terms MF is TSR, Testing Station Resistance, ie ITE – DP as measured on the station, Indicated Tractive Effort less DP. (ITE – DP) in turn equals MR + CWBR + DR, respectively Machinery Resistance, Coupled Wheels Bearing Resistance (as if part of vehicle resistance, but excluding enhancements due to resolving PTTE with it, and Damping Resistance if present). If any friction in the damping is built into achieving the damping, and the damping in perfect, ie any net to and fro (TF) forces in the drive are completely neutralised, then DR is simply the work done in achieving that neutralisation. Damping is very relevant in considering Rugby DP data, and is considered in its own right below. I convert all HPs to TEs for consistency, and abbreviate the Small Difference (between two large numbers) Effect to SDE. PTTE is followed if necessary by an S if the propulsive and compressive effects of steam on the pistons is the subject, and by V^2 if that from unbalanced reciprocating masses; if the sum of the two, then simply PTTE.

Randomised TSR Simulations

Nothing is said about the purpose of the exercise set forth in the spreadsheet, why it is necessary to simulate where there are actual TSR data, the reason for introducing randomness, and the extent to which the conclusions depend on the randomness or simulations. Indeed, demonstrating the existence of SDE does not require randomness or simulations. It can be shown by taking proportions of the average range in the data. To show the SDE is fine, but what then? Is SDE the only reason why Rugby TSR values are erratic? If so, how is that taken into account?

Is the purpose simply to say that the range in the Rugby TSR is what would be expected, under certain circumstances such as those assumed, that is also fine. If however the intention is to justify the terrible TSR and by implication DP values from Rugby, enlarge the sample, home in on the average of the enlarged sample, then it is not. The TSR data from the instruments at Rugby are the data which are to be analysed for what they reveal, not some corrected or improved version, or a much increased number of simulated observations. The TSR is assumed to be 800 lbs at all speeds and efforts, so it is not surprising that the average of many trials yields almost exactly 800 lbs. A lower assumption, say 600lbs, and a higher, say 1200 lbs, would do the same, although there would be an effect on the significance of the results of any analysis (size of sample and variation from average are major influences on significance, as that term is used in statistics). Further, the procedure does not treat the real area of uncertainty in the Rugby data, the DP. It works on ITE (which, see below, is generally consistent in Rugby data) and an assumed constant TSR, completely certain so far as the procedure is concerned, as a result of the assumed constant value. The treatment of SDE brings a range of uncertainty into the simulations, but that means the real source of uncertainty in the Rugby data, the DP, is ignored.

Steps in the Procedure

The typical ITEs are not given, but can, with some study be implied from the tables. It is not said what engine is the example. The constancy of TSR at all speeds and efforts is a further major assumption, a doubtful one. This results, with large numbers of simulations, in assuming what is hoped can be obtained from analysis of the data, even allowing for the SDE. It also assumes the nature and behaviour of TSR, variation with other sources ignored.

Carling, who did not analyse the Rugby DP results for the extent to which these sources of variation in DP applied, did not have a confidence level. Rather, he stated, on an impressionistic basis, how accurate he thought the measured results were, on which see below. (A confidence level in statistics is the end of a range over which certain conclusions can be drawn about probabilities of results occurring by chance, the levels and range suggested by test statistics).

Further, he had no way of knowing the true ITE resulting from tests at Rugby. He did not say that ITE measurements were +/- 2% accurate. Rather, he said that during a given test (constant boiler pressure, regulator setting, and cut off, hence speed also), results were typically in a 2% range. Indeed, the ITE readings for a test were averaged. Some comparisons were made with other indicators. The accuracy of ITE is however unknown. Similarly, he had no way of telling whether the DP measured at Rugby was accurate. He said that the manufacturer claimed the Amsler dynamometer was +/-1% accurate, and that when the instrument was statically tested at Rugby it was accurate to within +/-1%.

The accuracy of the Amsler in use, however, was unknown. The difference between the two items measured on the plant, ITE - DP, or TSR, in turn comprised MR + CWBR + $PTTEV^2$ + DR if any. The plant was not designed or operated to achieve accurate DP readings, but to avoid resonance damaging the plant and the equipment. Nor were the Amsler readings compared with the other dynamometer on the plant, that providing the braking of the rollers, which provided the resistance against which the locomotives under test worked. The DR is unknown, and was never tested or measured. On account of its importance for DP measurements, damping is considered further below in Seeing Sense in the Rugby Data.

It is said the scatter patterns look remarkably familiar compared with those in the TSR data. For that to have any meaning, the two patterns need to be compared, eg standard deviations, and correlations between actual and simulated values. No mention is made of that having been done. The creation of a larger sample than that given by the Rugby data amounts to creation of extra data to reinforce the actual data, reinforcing some preconceived idea of the best explanation of that data, assuming that there are no other considerations to take into account in explaining the behaviour of DP. On the same theme, there is mention of normal experimental error as an adequate explanation of some characteristics of the data. How much is normal? To what extent is it a real error or something inherent in the running of the locomotive on the plant? And, very important for considering the data, how random and large are the errors?

Such testing of the apparent similarity in scatter patterns is not a reason for accepting the data (actual figures and characteristics, especially the distribution) as adequate to explain anything. It is perfectly possible for the data to fit the SDE argument but not to be suited to explaining anything, especially finding a reliable TSR, its values and characteristics. No matter how many simulations are made, the Rugby data are not likely to reveal sound TSR, for reasons to emerge below.

Mention is made of 2 – 3% sensitivity to effort. What is the origin of that? If the subject is MR, MR displays considerable sensitivity to PTTES, not ITE, in the range of 5 to 7%, as I have mentioned here before. PTTE is a large number, so even 5% is considerable in MR.

An outlier envelope is introduced. Outliers are values which are considered to be out of place on account of their extremely high or low values. Outliers should not be discarded, but examined for reasons why they are so high or low. If there are good explanations, they should be left in. The outer lines so far as I can see were by means not stated fitted to the highest values. They are all very well, but what are the averages, and the one, two and perhaps three times standard deviations on each side, to indicate the distribution (fortunately, from other diagrams, it can be seen that lots of simulated values cluster close to the averages).

Conclusions on the Method

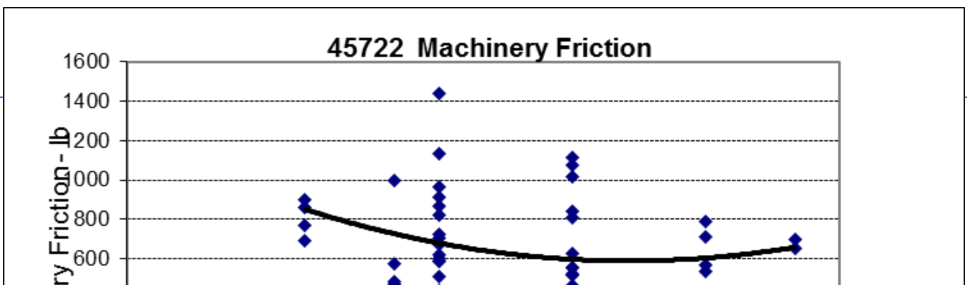
If it was the hope or intention that the randomised approach used by Doug Landau allows the Rugby TSR data to be corrected or improved so that it allows a supposedly sound TSR to emerge, that cannot be the case. For one thing, the way it was done means that the TSR is that assumed by the analyst, which cannot be correct. Most importantly, the method is not the correct way of analysing the data, including allowing for scatter, measurement error, SDE generally, and other influences on the elements of TSR..

The only way it can be shown how good TSR data are, is to regress TSR against its components, PTTES, PTTEV^2 and CWBR (a multiple regression). SDE will remain a problem. If the measurement errors of SDE are truly random and small relative to the TSR, and there is a high number of observations, the randomness will have little effect on the results. If those errors are large in magnitude relative to TSR, are not random but are biased or erratic, and the number of observations is small, then the components of TSR will not emerge with any (statistical) reliability or significance. Indeed, no sensible values of the components will emerge, ie the Rugby TSR data will not reveal anything at all about locomotive MR. The latter is the way things turn out, on which see below. It is possible that there are other influences than those so far determined from first principles which might affect the determination of TSR.

In that case, the residuals (data unexplained by the relationship so far fitted) are studied, often by graphing, also by further regressions, to see if that is the case. I have tested all likely explanations of TSR measured at Rugby post 1953, and all are wanting, likewise any residuals.

Part 3 of the Spreadsheet, Examples of Rugby TSR Data

Doug Landau presents this graph and the following sentence:



> Notwithstanding the scatter, the trendline shown reflects a speed/magnitude relationship roughly in line with theoretical expectations.

His Machinery Friction is of course TSR, or MR plus CWBR. Nothing is said about the form of the trendline (the equation to it) or how it was fitted, and there are of course no test statistics. From inspection of the graph, despite what Doug says, there is no speed/magnitude relationship. For there to be, the data at each of the six speed points would have to be tightly placed along the curve shown. Rather, there is observably much more variation in (his) MF at each of the speeds (about 380 to 1400 lbs for example at 35 mph) than there is, in highly averaged terms, in speed alone (circa 600 to 800 lbs along his trendline). Doug has no idea of how such data might be interpreted and analysed. He should be trying to analyse what causes the variation at those speeds. There are sufficient points of data at each of those speeds to test any hypothesis he might have, for example how hard the engine is working, and he believes the Rugby TSR data to be good. Notably, as no equation is given for the trendline, so there is no guidance on how the approach can be applied to the vast majority of locomotives which were not tested at Rugby, or anywhere else.

To test his MF/speed relationship, I fitted a regression equation to the very same data for 45722, $TSR = cV^n$, in logs $\ln TSR = \ln c + n \ln V$, V speed, c and n constants, in ln terms in order that there was least constraint from the form of the equation. Being a regression, my equation emerges with test statistics.

The result is $\ln TSR = 1857 - 0.29 \ln V$, or $TSR = 1857/V^{0.29}$. That relationship has an odd form. What, in terms of TSR, does the constant mean? What does the low power of speed in the denominator mean (its value is 2.38 at 20mph, 2.92 at 40, and 3.56 at 80 mph)? The test statistics show that no empirical relationship at all exists between TSR and V in Doug's trendline (r^2 is .06, Significance F is .05 and the ranges in the results at which they are significant at reasonable levels of probability very wide). Nor is there a theoretical expectation that TSR varies with V alone. TSR = ITE – DP = MR + CWBR. $MR = C + aPTTES + bPTTEV^2$.

(To obtain TSR requires the addition of CWBR, taking care that all relevant forces are resolved as necessary.) The line is continually decreasing from 800 lbs at 20 mph to 530 lbs at 75 mph, ie there is no turnup or shallow U as speed increases. How could such an equation, however valid for engine 45722 be made useful for other locomotives?

Further, his trendline and my fitted equation suffer from an error, which results from the data. The constant of both is at least 1000 lbs. The constant of MR is less than 100 lbs, and the constant of the CWBR of a Jubilee about 150 lbs, in total only a quarter of that figure. That emphasises that his MF/speed relationship does not exist and that there are at least eccentricities in the data. In addition, and of course, my equation like Doug's has enormous spread of data above and below the trendline (his) and fitted equation (mine).

He also says that notwithstanding the scatter, the trendline reflects a speed/TSR relationship roughly in line with theoretical expectations. Elsewhere in the spreadsheet document, reference is made to a shallow U shape for this curve, which probably influenced the undeclared shape chosen for the trendline. He does not say what those theoretical expectations are. There is also no connection between TSR and the dimensions and masses of the engine. The trendline is of no use for estimation of TSR without some characteristics of the locomotive and how it is being worked. Speed enters MR through characteristics of the terms. The propulsive forces tend to fall with speed, the compressive to increase, and the TF forces to increase with V^2. A great deal depends on the masses of the reciprocating parts, and the extent to which they are balanced in the mechanism. TSR however does not vary with V per se, for engine 45722 or any other.

Testing the Rugby Data

Before research is attempted on any data, that data should be examined closely for its characteristics, and the way it was gathered, measured and presented. In the case of the TSR data, three sensible and useful things can and should be done.

i) **Examining the Damping at the Drawbar/Dynamometer connection**

R C Bond in his autobiography *A Lifetime with Locomotives* (1975) shows (pp 120-1), that as the first Superintending Engineer of the Rugby plant, responsible for the design, he was well aware of the TF forces from the unbalanced reciprocating masses, and variation in steam pressure on the pistons during a stroke. He relates how on the French plant at Vitry, the frequency of those forces often coincided with the frequency of the plant, which led to resonance being set up, and violent oscillation of the locomotive under test and the plant. The TF forces concerned reached a maximum once in each direction per revolution and formed a resultant with the unidirectional force from the application of steam to the pistons. The Research Department of the LMS Railway was given the task of analysing the problem. The Rugby plant was therefore designed to dampen these forces, to ensure suppression of resonance for any tests likely to be done there. In Carling's 1957 article mentioned in the first paragraph, it is said that it was assumed in the design of the plant that the pull would vary with Simple Harmonic Motion, but it was found in practice that that the pull varied, not in SHM but in a highly irregular and unsymmetrical way, on account of play in the axleboxes and other bearings, the unsymmetrical variation being ascribed to the 90° spacing of the thrusts.

Damping the pull to eliminate the fluctuations falsified the results. Nothing is said about what the damping was, how it was known that the fluctuations were actually eliminated, how the results were falsified, and to what extent. Considering the surviving information, judging from the large number of low and negative values of TSR the falsification of the results continued until 1953. The intention was to damp these forces, presumably either to eliminate them, or to absorb forces in one direction and release them in the other. Until 1953 at least, the damping was poorly designed, and led to most observations of TSR being negative, by several hundreds of pounds in many cases, at least as measured.

It was not the play in the axleboxes and other bearings which caused the highly irregular and unsymmetrical pull, but the TF forces – the effects at the axleboxes and other bearings were a result of those TF forces. Their fluctuation was the result of their movement being interrupted forcibly by the end of the stroke occurring while their value was still high, ie by the TF forces continuing in one direction when the piston changed direction.

After the modifications to lessen the value of DR about 1953, the damping was the result of:

a) air being sucked into a dashpot, compressed, and exhausted; this could in principle damp TF forces as they occurred. If the orifices were much the same as when oil was placed in the dashpot, it probably provided little damping, but if the air pressure built up before any release, it would have resulted in erratic effects.

b) Belleville washers (sixteen pairs) which could dampen only at a constant rate, and were therefore unsuited to damping the forces and their pattern.

It was not simply a matter of what these devices did, but how well they could keep up with the reciprocation of the locomotives, which at the fastest the engines were run on the plant approximated one stroke per .09 second.

Further, proper damping must balance or neutralise the net forces in one direction with equal and simultaneous forces in the other, ie exactly the same pattern at exactly the same time, and for any friction in the damping per se to be part of the damping, for all four strokes occurring together. The damping which remained at Rugby after 1953 could not do that. What was wanted was opposing the TF forces as they occurred. In each stroke of a two cylinder locomotive, the TF forces changed from assisting the propulsive forces to opposing them, those in one stroke being balanced by those in another, but were still in progress as opposing forces as each stroke ended, the reason for the jerk effect, which was not balanced or opposed. The dashpot with air in it was not capable of dealing with these variations. In any case the TF forces had to be calculated in advance to design proper damping.

While Carling referred to getting the Rugby numbers right after the modifications of 1953, presumably the DP numbers, he did not say how that was achieved, nor could he have known they were right. He emphasised that the main function of damping continued to be prevention of damaging resonance to the plant, rather than satisfactory DP values. Indeed he acknowledged that avoiding the effects of the inappropriate damping would have required complete redesign of the plant. That was not done, so Carling admitted in effect that the damping was not right after 1953, which in turn means the values of DP were not right even then. Because it was not correct in form, damping must have in itself absorbed energy, which would have reduced DP and in turn increased TSR. Even so, as the TSR values are low by comparison with MR + CWBR from other sources, it would seem that the errors from pre 1953 must have persisted, which could well have been in inappropriate measurement.

Keeping the engine on top of the rollers so that that there was no reduction in TSR when it was running downhill and vice versa was achieved by the mediating gear adding to or subtracting oil from the Amsler dynamometer. That was a slow process, but the effect of deviations from the correct were registered, and the recorded DP figures adjusted for them.

In all the analyses I have done of Rugby data, ITE regressed on Q and V, gives good mutually consistent results, and DP very poor results. It is possible the ITE figures are consistent, but all wrong, perhaps all too low. Those I have examined with the Perform program, appear a little low but not a great deal. The problem is therefore with DP, or with one of the constituents of TSR. CWBR should not be in error, hence the PTTE is the problem, not surprising when it is considered that the damping cannot be correct.

(ii) Seeing Sense in the Data

The second approach is to test the data for its sense, a normal practice before conducting any further analysis of it. I used three approaches.

a) Graphing TSR against PTTE

To do the three tests in this exercise I considered the data for every engine tested at Rugby where there were at least 12 observations at any one speed (13 engine/speed combinations), and graphed TSR against PTTE (both sources). The spread of data in all cases was discouraging – what should have been a near straight line of TSR figures from a constant on the vertical axis (see (c) below) spreading upwards and outwards was a confusion of such points, with, in most cases no such pattern.

b) Implied friction coefficient of PTTES induced by steam effects, propulsive and compressive.

From the TSR of the 13 sets of data mentioned in (a), I deducted my estimates of CWBR and the $PTTEV^2$ effects. For any engine class, the sum of CWBR and $PTTEV^2$ should have been constant at each of the speeds considered. That left resistance data varying with PTTES as a residual, which residual I compared with PTTES data. That residual is such a small ratio of PTTES that the data imply improbably low Cfs (coefficients of friction) in the mechanism from steam effects, often less than half the lower set of Cfs I used when assessing MR from first principles, and (by examining what data there are on LR, and by elimination of other sources of resistance, MR. I can also report from having done the above, that that TSRs are erratic at a speed/output combination. I admit that in this exercise I introduce an SDE even more acute than that which occurs in TSR, but the results are very clear. Data on LR and MR from elsewhere in the world tends to justify the figures for MR, hence TSR, that I use, so I consider this exercise shows Rugby TSR to be decidedly on the low side and erratic.

c) Test Equations for Each Engine Class Tested at Rugby where there are at least 12 observations at any one speed.

Where speed is constant, $PTTEV^2$ is constant, as is CWBR. That leaves PTTES as the only component of TSR which at any one speed should show variation with TSR, ie
TSR = PTTES + $PTTEV^2$ + CWBR + constants in any of these variables, ie
TSR = Constants + b PTTES

Note that this equation for TSR will include CWBR. It will also include any net DR. This is a simple relationship, easily established if the data are any good. That was found not to be the case, however, not surprising considering (a) above. The constant should be positive, as should the coefficient on PTTES. The equations for most engines have at least one negative.

The t ratios on both constants and coefficients on PTTES are low, the Standard Errors of the Estimate wide, and the values of r^2 low, many less than 0.1. Results from two engines share some outwardly apparently redeeming features. That for the Duchess at 50 mph gives 522 +.015PTTES. The .015 is to low by far, and the r^2 is only 0.11, ie there is really no relationship after all.

Much the same remarks apply to 9F 92250, the last steam engine tested at Rugby, the data for which gives 227 + .02PTTES at 20 mph. At 30, 40 and 50 mph, the constant turns appreciably more negative, as in:.

1 Speed mph	2 Obser-vations	3 Equation for DP	4 Value of r^2	5 t on constant	6 t on coeffic-ient	7 Standard Error of the Estimate
20	15	227+.02PTTES	.11	2.56	1.24	291
30	17	-436 + .05PTTES	.23	-0.9	2.11	299
40	12	-1207 +.12 PTTES	.55	-1.94	3.55	195
50	16	-2774 + .22 PTTES	.14	-1.66	2.08	277

The large negative constants in the equations for 30, 40 and 50 mph, and the very high coefficients on PTTES at 40 and 50 mph, show that the plant did not produce reliable DP results. Very little of the data is explained by the form of the equations. Many of the t values are such that little confidence can be placed on the equations occurring other than by chance, reinforced by the large SEEs. The equations provide the best fit to the data, which says nothing for the data.

Further, as the interest is in MR, or in the case of the Rugby data, TSR, the data for Q, ITE and DP, all high numbers, are inter-correlated, and for that reason, should not be used together in attempts to find a relationship for TSR.

All constants for the Jubilee and Royal Scot are negative. The data for 9F 92166 at 30 mph, however, gave 281 + .047PTTES, with a good t value on the PTTES coefficient, and r^2 0.49. The coefficient on PTTES is encouraging, and the constant exceeds CWBR. If every equation for the set of 13 were like this, then the Rugby data might have been redeemed, but so many other 9F results say otherwise.

Could the values of CWBR and PTTEV^2 deduced in my analysis of MR from first principles be too high, rendering the PTTES too low, bias the above results? That is possible, but Cf of the CWBR is fairly well established, and PTTEV^2 is modest at low speeds.

More likely is that fluctuating DR is present. There is no way of isolating that.

I also tried the input/output (Willans line) approach to obtaining TSR (MR + CWBR). The article by S J Pacherness, *A Closer Look at the Willans Line*, in paper 690182, Society of Automotive Engineers, International Automotive Engineering Congress, January 1969 explains the underlying idea. ITE is regressed on DP, the opposite of the usual cause and effect representation. The resulting regression line is projected back until it intersects the DP line in the negative range, ie left of the ITE line. That negative section with sign changed gives TSR, which minus estimated CWBR leaves MR. For 92250, all relationships linear, this gave 333 lbs TSR at 20 mph, of which 229 lbs is estimated CWBR, leaving 104 lbs MR, and 247 lbs at 40 mph, which after the same CWBR leaves 18 lbs for MR. These MR values are obviously far too low. At 30 and 50 mph, the TSRs are too low to give any MR at all.

For Duchess 46225, a linear equation gave an MR of 370 lbs at all rates of working at 50 mph. I also fitted a curve to the same 50 mph data (ln ITE on ln DP), differentiated it, and found the slope at various values of DP, all within the data range. For a DP of 7000 lbs, MR is 228 lbs, for 10,000, 419 lbs, and for 16,000 lbs, 813 lbs. These MR values are certainly too low at DPs of 7000 and 10,000 lbs. All these equations for the input/output approach had good test statistics except for the constants, on which the t test measures were poor, in turn leading to large standard errors of the estimate, and considerable uncertainty in the values of TSR.

These results all point to the low values of the Rugby TSR data for analysis of that subject.

Conclusions

The Rugby TSR data are so poor that sound values of TSR will not emerge from them. The equations are not to be blamed for these results; they are the result of the unsatisfactory data. Or, for those not aware of the niceties of fitting relationships to data, the data are such that sound relationships, or relationships that might be expected, cannot emerge.

From all points of view, I would consider the Rugby TSR data of post 1953 to be too low and too erratic to be credible, let alone useful. SDE is only partly responsible for those conclusions. I would ascribe much of the reason for that to be the improper damping and measurement of DP. Doug Landau's approach to the Rugby TSR data is in my view one of wishful thinking about its soundness and hopes of using it, and playing with figures to defend it. As previously related (first paragraph above), it was the view of D R Carling, Superintendent of the Rugby plant during its operating life, that the plant was not suited to obtaining the internal resistance of locomotives. In saying that he referred to the SDE, but he also pointed out that the damping provided was to prevent resonance developing, not to provide accurate TSR; indeed, it could not. Rugby TSR data should not be used for deriving TSR or MR, indeed for anything. It is strange that Doug Landau should defend the Rugby results so stoutly. If the data are not satisfactory, no good can come of playing with it.

Full results of any of my analyses mentioned are, as previously, available on request. I will also make the relevant Rugby data available to anyone who wants to investigate the subject.

John Knowles

Letter from Doug Landau

7th March 2017

Locomotive Resistance

This is in response to John Knowles letter 2nd December 2016. The many points raised are not necessarily taken up in chronological order. Words in quotation marks are John's own words unless otherwise stated. This letter is longer than was perhaps essential because it contains information that may be helpful to readers unfamiliar with this topic.

You say "I know of no other analyst of the subject other than Doug who considers that the whole of the resultant is part of MR". Who are these other analysts on the subject? As one mechanical engineer well versed in the ways of steam recently opined regarding your procedure; "As you correctly point out, WRTE and the pull recorded on the Amsler dynamometer were one and the same. Also, WRTE has to be net of all the machine friction inherent in driving the locomotive. Axlebox friction forms part of MR, it does not appear in

WRTE; it represents part of the difference between indicated power and WRHP. It cannot somehow escape to be part of the WRTE only to be absorbed later, I do not see the logic of that." The relevant force diagrams can be found in Lomonossoff's *Introduction to Railway Mechanics.*

Direct studies of MR as opposed to the subject of LR are a distinct rarity, largely because experimental data on the former is scarce, and such as it is generally unsatisfactory. It is only in recent years that researches at the NRM have brought to light the wealth of relevant experimental data from the Rugby test plant. The available data from the Vitry test plant is very limited in this regard, and amount to some data for the EST 241 – 004 4 cylinder compound.

The *Railway Mechanical Engineer* (USA) for May 1943 featured an article by Lawford H Fry analysing locomotive test plant MF data for 10 locomotive types dating back to the tests at Purdue University about the turn of the 20th century to the 1930s, involving four, six, eight and ten coupled locomotives. Most of the tests were carried out on the Altoona test Plant. Notwithstanding the extent and diversity of scatter present in the various data sets, Fry sought to reconcile the data notwithstanding "given an uncertainty of 40 per cent", into a formula for machinery friction. The outcome was a function of coupled axle loading, driving wheel diameter and the number of coupled axles. In the event the latter factor was given undue significance, but was perhaps the best he could do with the data available. Curiously Fry was using the 'small remainder' (ITE – WRTE) data for the exercise; a problem he fully acknowledged. Perhaps he thought this was the best way to compare different data sets, or possibly the data available to him was incomplete. The scatter magnitude was uniformly much greater than present in the Rugby data. As far as I know this was the only published study specifically focusing on locomotive machinery friction based on experimental data, or from a purely theoretical standpoint.

It's surprising you cited Ell's comments on the locomotive resistance curve for the Rebuilt Merchant Navy in Test Bulletin No. 20. The curve itself is the same as appeared in Bulletin 15 for 71000. Given the very similar basic architecture of the two types this was not unreasonable, and must have assumed any frictional differences for the different valve gears would be too small to be of practical significance. Ell assessed the resultant frictional augment over the standing coupled axle load losses at about 300 lb, a long way short of the magnitudes you ascribe to your statistical exercises. A constant 300 lb was a bit of a simplification, but in magnitude was not dissimilar to what the WRHP data recorded at Rugby indicates. At an ITE sensitivity of 2.5% (typical value) and say 1850 IHP at 60 mph (a typical express work rate), it works out at out at 290 lb.

In 1944 E S Cox presented a paper on Locomotive Axleboxes to the I.Loc.E, it included an analysis and diagrams of the forces encountered by the coupled axle boxes of a Midland 4F working in 30% cut-off at 15 mph. The exercise was essentially the resolution of two forces, the net axle load which was a constant defined as the 'vertical load', and the combination of piston thrusts, a variable. The resultant axlesbox loads in the course of a revolution were quite nuanced, a situation involving the cross couples between the RH and LH pistons phased at 90^0 and the axleboxes. The resultant peak RH and LH loadings were about 80% or slightly less of the summed forces. The overall "work factor" for the RH box was 18% higher than the LH. Details are also given for an outside cylinder arrangement, this equalised the RH and LH workloads. This exercise and its modelling is indicative of how MF would have been tackled had there been an interest in estimating it.

You say "How could the gear (mediating) react several times per second to movements in both directions, i.e. was it capable of keeping up with the frequency of the sources of variation in DP? "

The mediating gear and servo mechanism did not operate in the way you describe. Firstly the dynamometer, of the hydraulic type was exactly the same as fitted to the LMS dynamometer car No.3 commissioned in 1948. Such dynamometers are more than capable of absorbing, measuring and integrating the variations in drawbar pull during the course of a revolution. The mediating gear and servo mechanism played no part in adjusting to these transient forces; this was not its function. Variations in drawbar pull could be quite severe relative to the mean value. The offical report on the 1948 Locomotive exchanges contains a number of drawbar pull (DP) traces; some trace a sharp zigzag profile in the course of a revolution. The WD 2-8-0 for example, not blessed with any reciprocating balance, delivered a very spiky trace, with an amplitude of +/- 7% about mean pull.. The GW 28XX 2-8-0 and the LNER O1 were not much better. These traces were in the 17 - 26 mph range. The Stanier 8F was much smoother, about +/- 2%, but there were some random intermittent spikes about double this. The LMS Class 5 trace showed minimal ripple at 55 mph, but the mean pull was undulating. The B1 at 53 mph was not quite as smooth as the LM 5, and again delivered an undulating trace. The GW Hall was notably uneven in one example which is captioned "Increased oscillations encountered at 31 to 38 mph (+/- 10%). All the multi cylinder engines delivered smooth traces with undulations. A slight exception here was the GW King, with intermittent periods of zigzag present in the trace. These undulations likely reflected local changes in gradient, curvature and track condition. The various dynamometers used in these trials were evidently sensitive to all the locomotives could throw at them.

In the normal way of things any disturbing forces resulting from transient changes in drawbar pull were dissipated in parasitic motion (swaying, rolling, hunting) of the rolling stock. Some of this behaviour, as clearly observable from inside an underground train, is down to ride characteristics and track imperfections. This situation is also sensitive to the drawgear arrangements. As first built, Britannia hauled trains were soon receiving complaints of "shaking effects" from passengers. After mathematical analysis the solution proved quite simple; a reduction in the initial compression of the tender drawbar spring. "The rogue W.R. two-cylinder engines were found to be just as amenable to this arrangement as were the BR engines themselves "(E S Cox). The situation on the test plant with the dynamometer anchored solid is rather different, any potentially resonant forces have nowhere to go. In the absence of any damping equipment as first built, the French test plant at Vitry dating from the 1930s, encountered severe resonance problems with 2 cylinder locomotives, a situation largely resolved by the addition of Bellville Washers (springs) to the test plant drawbar. These achieved a satisfactory damping effect. This lesson was well understood when the Rugby Test Plant was in the planning stage. Jim Jarvis commented that the more solid anchorage of the Rugby dynamometer brought further improvement (written communication). The damping deflections involved were slight, within $^{1}/_{8}$”,

The function of the mediating servo mechanism was solely maintaining the locomotive at top dead centre (TDC) and correcting any drift from this situation. It was insensitive to any drawbar pull variations of shifts from TDC in the course of a revolution or even many revolutions. Key to its function was a differential gearbox, its two wheels rested on a disc and were friction driven by its rotation at constant speed as a function of time. Provided the wheels were equidistant about the disc centre the gearbox output shaft was stationary. The faces of the wheels were transverse to the fore and aft shift. The set up was such that the gearbox moved back and forth about the rotating disc centre line in equal magnitude to, and in synch with the fore and aft motion of the locomotive. Provided the motion was equidistant about TDC the fluctuations of output shaft cancelled out to zero, and no “inch seconds” would be recorded. Should TDC not obtain the “inch seconds” would be added to or deducted from the recorded value dependent on whether the TDC shift was fore or aft. The second function of the differential gearbox was to drive two moveable electrical finger contacts. These were interposed by a third contact that moved back and forth between these contacts at the identical amplitude (typically less than 1/8”) of the locomotive’s fore and aft motion. This shuffling contact was fixed to the mediating control rod connected to the locomotive; this rod was not subject to any stress or stretch and incorporated positional adjustment provision to suit any locomotive type. The distance between these outer contacts was such that in the TDC situation no contact would be made between the two differential controlled contacts and the shifting middle contact. The two outer contacts swung, too and fro pendulum fashion in the course of a revolution as the differential gearbox picked up the fore and aft shifts. In the TDC situation the deflection would be equidistant about the zero datum line, the swing per revolution remaining equidistant left and right. In the event of a shift from TDC, the finger contacts swing would be biased to increased swing in one direction, building up the swing bias such as to eventually make contact with the centre contact, initiating remedial plus or minus action by the dynamometer dependent on the initiating contact, the other contact will have become more distant. I have no details of the Amsler circuitry, but this transient contact will have closed a control relay with a time delayed drop-off, in other words the dynamometer was given a nudge for a finite period of time. At this period of history such timed relays were dashpot controlled and adjustable, so the optimum timing could be fine tuned during the commissioning phase. Such nudges would occur at intervals, reducing the rate of swing bias until the "inch seconds" reading stabilised, remaining constant. This situation was probably well in hand by the time the warm up period was complete and the test period commenced. A continuous paper trace plot of shifts about TDC was recorded. The operation of the mediating gear can be summed up in one word – ‘measured’. The test sheets also included a provision for mathematical correction should there be an “inch seconds” discrepancy. As first supplied the mediating gear was over reponsive, the differential gearbox ratio was reduced as a consequence.

Amsler's conditions of contract included performance guarantees. The dynamometer was guaranteed to within 1% in regard to pull. Carling believed it was well within the guarantee and that it was consistent to even finer limits. Work done was guaranteed to within $1^1/_2$%.

You say: "My difficulty is that I think the Rugby data poor/inadequate, only a handful of the world’s locomotives were tested at Rugby, and I work at MR and LR more generally, for application to other locomotives. What is easy for him in principle for a handful of locomotives is only a tiny part of the need for well informed MR and LR."

In regard to MR, quite where this body of alternative of "well informed" MR data comes from I cannot think, a problem as alluded to above. Regarding locomotive resistance there is certainly plenty of data, world wide if you want it, but such as it is can fairly be described as a minefield of disparity. LR is after all a variable, subject changes in effort, wind speed and direction and track condition. Regarding the latter some tests in the USA using the same set of rolling stock on three different railways found significant changes in rolling resistance which was attributed to differences in track and track bed formations. Locomotives would be similarly affected.

Specific resistance USA Coaching Stock 1938 – Lb/ton Index

MPH	PRR	C & NW RR	UP RR	Davis Formula
60	89	100	109	100
70	86	101	104	100
80	89	103	100	100
90	90	102	96	100

100	87	100

Some of the French compounds displayed extraordinarily high locomotive resistance compared to other continental types. The disparity is too high to be explained by attributing a few percentage points inaccuracy by the indicting equipment. The recorded MRHP for the EST 241 mentioned above is likewise high.

It is ironic that I am accused of insulting scientists when your letter is riddled with attempts to disparage Carling and associates at every supposed opportunity; no matter how speculative, or ill informed. Likewise, Amsler, at the time a world leader (if not the world leader), in the field of scientific instrumentation metrology, are, by implication, similarly rubbished. This dubious collection of non sequesters does not survive scrutiny (your text in inverted commas).

1. "The Rugby figures are not the same as MR properly called, however. When the CWBR is removed to give MR per se, they become lower, and more become negative. As a further test, I have then excluded estimates of the resistance from the V^2 effects, and the constant of MR, leaving mostly the sources of resistance due to piston thrusts and rings. Almost all of these remainder observations are thereby reduced to values so low that they imply implausibly low friction coefficients, ie that Rugby data are generally low."

"Properly called"? Only by a definition of your own creation, the absurdity of contriving a number that is incapable of verification by actual measurement, and discarding one that was is proper? Treating the *driving* wheels as *passive* objects is a fundamental conceptual error; I can see that this approach might help you find the number you first thought of. "My difficulty is that I think the Rugby data poor/inadequate," Your comments on the higher incidence of negative MR outcomes, having reduced same by deducting your assessment of CWBR seeks to undermine the Rugby data by implication. The reality is, as my *Experimental Error* paper shows: the lower the actual remainder between two given quantities at stated limits of uncertainty, the higher the statistical incidence of negative outcomes. As to; "the remainder observations are thereby reduced to values so low that they imply implausibly low friction coefficients": the corollary of this is that your own frictional assessments are too high. It is apparent from the Rugby WRHP data that most of the track ride losses incorporated in LR formulae (the B term, are absent when running on the test plant rollers. This is no surprise; when running on more solid foundations the track deflections, track bed deflections, and rail joint percussive losses are not encountered. "Axlebox heating was a very serious problem on the Vitry plant; it was greatly reduced when the roller-pedestals were mounted on large rubber pads, such as were incorporated at Rugby from the start, probably due to Vitry's experience. This was in fact giving the plant some of the elasticity of the track *without its irregularities*, which if kept small were not entirely harmful.' – Carling (my italics).

2. "D R Carling said that they ultimately got the (DP) answers right, but he did not say how that was done, nor how it was known the answers were correct. No mention is ever made, there or elsewhere, of using the proper dynamometer, that applying the braking on the plant, to check the DP measures."

The first sentence appears to imply Carling was hiding something. He was not appearing in a court of law. Absence of statement does not prove absence of action or deception. As far as I can see his only omission was a detail description of the mediating gear servo control mechanism and how a differential gearbox was fundamental to its function. In the Rugby Test Plant publicity brochure it is simply referred to as 'a special device'. This was an elegantly simple solution; mention of it could only have reinforced his case. Your second sentence again infers guilt by omission of action, and is nothing more than speculation that something never happened. Having designed and commissioned a number of control schemes for a variety of industrial processes, and control and protection schemes for high voltage generation and distribution networks, I can assure you that when trouble shooting, no stones are left unturned. Most problems prove routine; some can be quite challenging. You seem to regard the Rugby staff and the Amsler test engineers as a bunch of incompetents.

The metering of the hydraulic brakes (speed and torque) was primarily to facilitate the equalisation of work between the coupled wheels and the detection of slipping. Obviously it would be a useful, though approximate cross check with the Amsler dynamometer behaviour during the commissioning phase. However to describe the hydraulic brakes were the “proper dynamometer” to verify drawbar pull is optimistic. I don’t have the figures for Rugby, but the similar Heenan and Froude brakes at Vitry were guaranteed at +/- 5%, actual performance was assessed at +/- 3%, someway short of the Amsler dynamometer performance inside the guaranteed +/- 1%. I have heard it suggested drawbar power could have been determined from the temperature rise and mass flow of the brake units cooling water. This is unlikely to have proved very accurate, aside from the obvious problems of thermometry and mass flow assessment; there would have been significant radiation losses from the brake unit bodies.

3. “Important, however, and not mentioned by Doug, Carling was clear that they damped to protect the recording devices from the effects of resonance, not to perfect DP readings. Doug places a very favourable gloss on all of that. He omits mention of the dashpot, which after oil was removed from it, had air in it, and the frequency and magnitude of the forces affecting the apparatus.”

The Amsler dynamometer was a recording device, and the only one that potentially could be damaged by resonance. Such damage is unlikely to have left accuracy unaffected, thus there was every reason to protect it. For readers unfamiliar with the history of the Rugby test plant, the resonance problems at Vitry were cured by incorporating Bellville washers in the drawbar, as specified for the Rugby plant based on that experience. At the suggestion of LMS research department an oil filled dashpot, was added to provide additional damping; the proverbial belt and braces solution. The dashpot incorporated a controllable by-pass to regulate the damping effect. It was operated by a bell crank arrangement connected to the drawbar. For the commissioning tests a WD 2-10-0 was selected, two were used, the first having proved ‘an old bag of bones’. Having no reciprocating balance, from the resonance standpoint, it was the severest test the plant was to encounter. The LMS research department had assumed the drawbar pull waveform was sinusoidal; the reality proved otherwise, the fore and aft wave forms of drawbar pull proving dissimilar in shape, and amplitude. As a consequence the dashpot, falsified the drawbar pull. Note that the falsification of drawbar pull was clearly apparent from the available instrumentation. Modifications to the dashpot, considerably enlarging the bypass capacity proved no solution. In his Newcomen Society paper Carling explained ‘In the end they simply took the oil out of the damping dashpot and left it with air in it, which damped sufficiently to prevent any damage, had resonance ever occurred. Afterwards no trouble of that kind had arisen and they got their results right. Being wise after the event he considered that, had the whole dashpot system been suspended on the drawbar, not fixed to the foundations, it would have acted as an inertia damper, there could have been no falsification of mean pull. It would have involved a major engineering modification as was not justified.’ Writing in 2005 Jim Jarvis recalled: “…. The oil was drained from the dashpot and care taken to check that no untoward effects arose. In the event, the revised drawbar & dynamometer etc, characteristics avoided any significant disturbance even when the dashpot was made ineffective. It was considered that the plant-drawbar pull figures were accurate after the dashpot problem had been settled."

4. "The effect of the Belleville washers, air dashpot and mediating gear operating much more slowly that the fluctuating forces, must have regularly allowed the to and fro forces free rein, and at others resisted them."

The functioning of the mediating gear as explained above, operated in a 'measured' way; it was not compromised in any way by "fluctuating forces". The deflection of the Bellville washers was consistent as a function of load and instantaneous (Hooke's Law). The amplitudes were slight, typically within $^1/_8$". Given the dashpot was filed with a compressible medium, air (a sealed unit –no hisses), it will have behaved in a similar fashion (Boyle's Law). In this form as a pneumatic damper with very small deflection and a huge clearance volume (in the relative sense), it probably achieved very little if anything. "I have a feeling that sometime subsequently the dashpot equipment was disconnected." - Jim Jarvis

Given the uncertainty of exactly what is being suggested above, perhaps I should point out that the drawbar pull and the dynamometer reaction are always equal and instantaneous.

5. "If the Belleville washers and the air dashpot kept up with the fluctuations, there would have been frequent short hisses from both, rather than sighing."

So that clinches it: the Bellville washers were not making the right noises!

"Of the many values in a considerable range of TSMR in the Rugby data for the various classes at any speed, which does Doug choose to be used as his TSMR, and why?"

I do not understand the question; obviously TSMR will vary according to speed, effort and locomotive type, such matters have to be determined on a case by case basis. The consistency I describe refers to the WRHP Willans lines at given speeds across separate test series with the same locomotive, or with different locomotives of the same type.

"Doug says that the measurement of WRHP (DP as a HP) was the simple product of drawbar pull and RPM, a process automatically recorded, monitored and controlled by a Mediating Gear under the control of a servo mechanism."

I am not saying the mediating gear controls the drawbar pull as you go on to infer. The mediating gear regulates the integrity of the measurement process by sustaining TDC, it measures "inch seconds" to self monitor its performance. Speed (rather than RPM), was determined by distance travelled over the test period.. The work done was measured by the dynamometer integrator

"One test is simply to graph TSMR against PTTE. This reveals tremendous ranges in TSME for a given PTTE, and precious little repeatability Doug claims that the Rugby data possesses."

This is just a restatement of the small remainder problem in modified guise, so the scatter described is no surprise. In any event why, by inference, does the measured WRTE take all the blame? The TSMR scatter is the product of two uncertainties, not one, a compound error; moreover, Carling rated the dynamometer accuracy higher than the Farnbro indicating equipment which he put within +/- 3% . The consistency I describe is in the form of Willans lines – plots against steam rate. Carling thought this could be determined to within 1%, although it must be said that within this there were slight variations in pressure and temperature over the course of a test series, introducing an additional source of scatter to both IHP and WRHP.

Additional to Willans line however, where plots ranging from 29 to 45 are available of WRTE against ITE, as in the case of 73031, scatter is very low for 20, 30 45 and 65 mph. Given these data sets embrace enhanced superheat, de-superheated and part regulator working, thus involving variations in steam volume and cut-off , it is apparent that cut-off is of little significance, ITE is. Given that between 15 and 40 % cut-off, an increase of 260% occurs for an increase in valve travel of only 16%, this is no surprise.

From this data set the MF at 30mph, 688 and 1375 HP for a Black 5as tabled in your earlier letter works out at 600lb and 890 lb as against the 940 and 1420 lb shown. Given these differences are as high as 56 and 60%, the slight dimensional differences between a Black 5 and BR5 are immaterial. Since however your values represent an emasculated definition of MR the true discrepancies are even larger. Further evidence of a failure to match the empirical evidence can be found is Report L116, which includes a resistance curve for a 9F at 16,000 lb/hr steam rate as derived from constant speed road tests. At 30 mph, 1090 IHP, the LR was 134 HP, 1675 lb, as against the 1710 lb given for the Black 5 at 688 IHP, and by extrapolation at 1090 IHP the Black 5 LR works out at 1990 Lb, 19% higher than a 9F, notwithstanding the latter’s 5 coupled axles as opposed to 3 and a coupled axle load 19% higher. As previously mentioned the plant test MF differences between the Crosti and Std 9F were confirmed in road tests. When asked to adjudicate on the test data for the standard and Crosti 9F, Chapelon concluded it was the most accurate locomotive test data he had seen.

In summary the supposed shortcomings of the Rugby Test plant, its designers and operators are groundless. The available experimental data demonstrates consistent repeatability over time and circumstance. Repeatability is a key indicator of metrological integrity. That is not to say everything is perfect and falls in place in place like a jig saw. Given the understood limits of experimental error, however small, and the random nature of scatter, the real world is more complicated. Exactly the same problems obtain when reconciling the data from road tests. Road tests have however confirmed the differences in test plant MR in the case of the Crosti and standard 9Fs. In other words the empirical evidence derived by different methods remains consistent. A key test of scientific proof is that its claims are consistent with the empirical evidence. The powers of the regression statistical process used by John Knowles fails the empirical test significantly and is thus unsound, supposed statistical integrity notwithstanding.

Yours sincerely,

Doug Landau

PS; I have only just seen John Knowles letter 21 February 2017 on the website, as at 7th March, and have not had time to study it as yet. – *see below .*

---

14 April 2017

Locomotive Resistance

This is in response to John Knowles letter 21st February. As in my previous responses the many points raised are not necessarily taken up in chronological order. Words in emboldened quotation marks are John’s own, unless otherwise stated and with regular quote marks.

The first point is that in stating **"**D R Carling thought the Rugby testing plant would not yield satisfactory figures for the internal resistance of the locomotive", John omits to mention that Carling took exactly the same view of locomotive resistance (Model Engineer 17 November 1980), as I have previously pointed out. This does not mean he mistrusted the WRHP data any more than for the IHP, rather less in fact: he was simply stating the inherent uncertainty of the small remainder problem and experimental error.

Carling's stated uncertainty for the IHP data was put higher than as for the WRHP: "Practically every instrument used at Rugby was checked in one way or another. A special calibrating device was used for the Amsler, of a kind used for testing large materials testing-machines, and the device itself was tested by the National Physical Laboratory. Amsler's guaranteed the measurement of pull within 1%; he had reason to believe it was it was well within the guarantee and that it was consistent to even finer limits. Work done was guaranteed to within $1^1/_2$%, and the indication of power to within $2^1/_2$%, but the derivation of power from the recording was to considerably closer limits. More difficult to quantify would be the accuracy of indicating; but, generally, the scatter of values for several sets of diagrams for any one test fell within, or very little outside, 3%." (Carling - Locomotive Testing Stations Part I; Newcomen Society Paper.).

The second paragraph again propounds his ideas on how the damping measures supposedly sent the dynamometer into some kind of a spin. I dealt with his ideas on the damping equipment and mediating gear comprehensively in my letter 7th March, no need to repeat my observations on how the damping equipment etc actually functioned here. I see he still thinks that coupled wheels are not part of the propulsive machinery; it's a wonder trains ever managed to move.

"After the modifications to lessen the value of DR about 1953, the damping was the result of:

a) Air being sucked into a dashpot, compressed, and exhausted; this could in principle damp TF forces as they occurred. If the orifices were much the same as when oil was placed in the dashpot, it probably provided little damping, but if the air pressure built up before any release, it would have resulted in erratic effects.

b) Belleville washers (sixteen pairs) which could dampen only at a constant rate, and were therefore unsuited to damping the forces and their pattern.**"**

I don't know where John gets the idea that the abandonment of the oil damping dashpot, replacing it with air, did not occur until as late as 1953. Perhaps he seeks to use this date to correspond with the time when negative MF values became a rarity. The idea that it took 4 years to reach this decision is absurd; had it been so, many heads would surely have rolled in the meantime. The reality was that the problem was treated with some urgency during the tests with WD 2-10-0 73788 in 1949. As Jim Jarvis[1] recalled; "We all worked well into the night on at least one occasion. After waiting for stable conditions to exist, the damping by-pass setting was altered, accompanied by a significant change in the

recorded pull on the Amsler table. In consequence the oil was drained from the dashpot, and care was taken to check that no untoward effects occurred." (Perhaps I should point out that my many Jim Jarvis quotations are taken from letters addressed to John, likewise citations given in respect of Ron Pocklington[2]).

At no point in JJ's correspondence can I find any reference the to a dashpot modification opening it to the atmosphere. It cannot have been built that way for obvious reasons.

The tests with WD 2-10-0 73788 took place in four episodes between 22.4 and 19.12.1949, amounting to 59 days and 46 test runs (the previous choice, 73799 having been declared unfit as an 'old bag of bones). The intervals were occupied by D49 62764 for indicator tests of the Reidinger Poppet Valve gear. It is clear from JJ's comments above, that the damper was air filled by the time these WD tests were concluded.

"Indeed he (Carling) acknowledged that avoiding the effects of inappropriate damping would have required a complete redesign of the plant. That was not done, so Carling admitted in effect that the damping was not right after 1953**"** (note the spurious date).

He said no such thing; this is just a crude attempt to put words into Carling's mouth. What he actually said was (repeating my previous letter I'm afraid); "In the end they simply took the oil out of the damping dashpot and left it with air in it, which damped sufficiently to prevent any damage, had resonance ever occurred. Afterwards no trouble of that kind had arisen *and they got their results right* (my italics). Being wise after the event he considered that, had the whole dashpot system been suspended on the drawbar, not fixed to the foundations, it would have acted as an inertia damper, there could have been no falsification of mean pull. It would have involved a major engineering modification as was not justified."

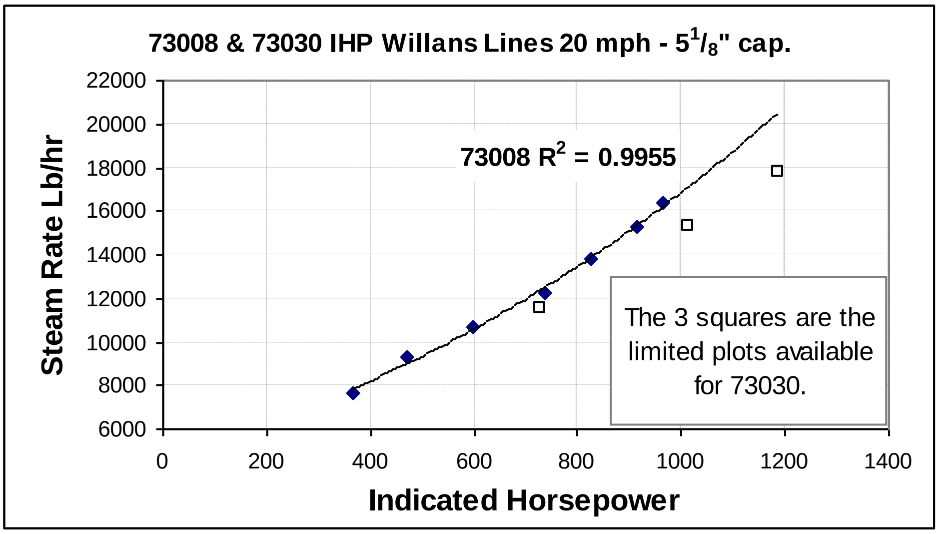
The tests with B1 61353 involved three spells at Rugby, 1950/1. No indicating was carried out, comprehensive WRHP readings were recorded. A couple of Rugby Test Station drawings dated 6.4.1951 show a family WRHP curves Vs speed for steam rate, cut-off, and curves for WRHP (estimated) and WRTE at 18.000 lb/hr steam rate, plus IHP (estimated) and WRHP curves. Such drawings would hardly have been prepared with the damper problem unresolved.

On point **"**b)**"**, the nest of Bellville washers was adjustable to suit the test programme , as *Engineering* 19 November 1948 reported: "A crosshead at the front end of the dynamometer, Fig. 20, is pulled by a forked member which passes loosely through it; the fork passes through it a number of Bellville washers which act as "smoothing" springs. Some of these washers may be replaced by plain washers if it is desired to alter the compression modulus to suit locomotives of different masses, etc."

The assertion that the Rugby IHP data is **"**generally consistent**"** and survives his statistical rigours, whereas in contrast we are told, the WRHP/WRTE data fails it seems, on all counts- **"**very poor**"**, is not a situation I am able to recognise from the available test data.

On the first count, consistency, the opposite is true. When IHP and WRHP or ITE and WRTE are plotted as Willans Lines, the R^2 values are uniformly high, typically approaching unity. This value is an index of scatter, a perfect outcome (no scatter) returning a value of 1. Various curve fitting options are provided by the Excel programme. The general shape of Willans lines is known from first principles; the chosen option for Willans Lines is a polynomial. Twenty five plots of IHP Willans lines randomly selected involving 46225, 70005, 73008/30 and 92013/250 returned an average R^2 value of 0.9853; the same exercise for WRHP Willans Lines returns a marginally higher value of 0.9888. Clearly, on this test, the WRHP data holds its own. Carling put the steam rate accuracy within 1%; even so there was some true scatter for given steam rates from test to test because there were slight variations in steam chest pressure and temperature. In a given circumstance of speed and cut-off, a reduction in pressure of 2lb might reduce IHP by about 1% (function of absolute pressure ratio). These high R^2 values are not in them- selves proof of accuracy, it is a measure of low scatter, more telling is the test of repeatability. Fixed or systematic calibration errors would not disturb the R^2 values.

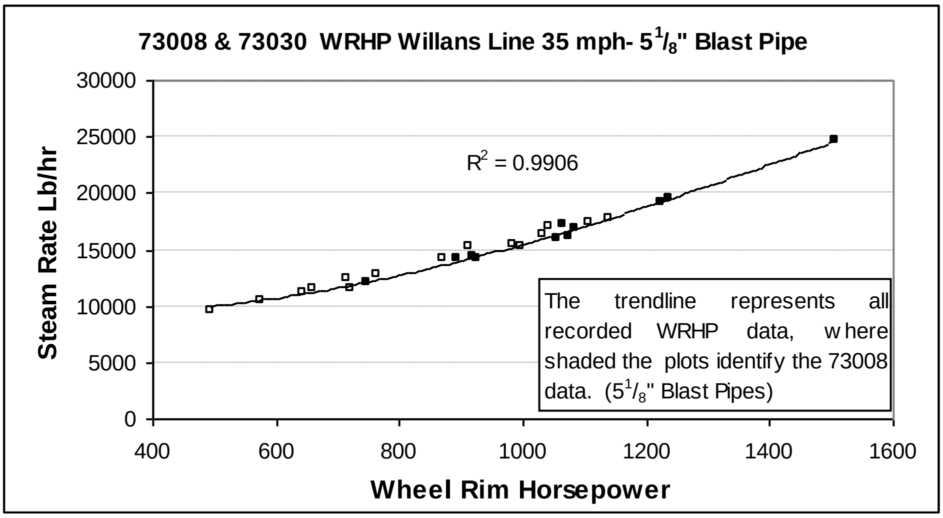
When it comes to what might be the called ‘handshake’ test, plotting combigned data sets from tests separated by time, it is the IHP data that lacks consistency; the WRHP data consistently passes this test. The scope for such tests is constrained by the data available. Suitable IHP data pre and post the improvements to the Farnbro indicator introduced by Ron Pocklington is confined the tests with BR5s 73008 and 73030. In the latter case the suitable data is confined to the tests with the $5^1/_8$” blastpipe. The scope for 73030 IHP data at a given speed is confined to 3 test runs at 20 mph. There is more adequate WRHP data for both engines at 35 mph.



The 3 plots for 73030 (1953) trace a distinctly separate path to the earlier tests with 73008 (1951/52). Note the increased IHP for a given steam rate. A trend line for 73030 has not been fitted since the default resolution with only 3 plots is to return an optimistic R^2 value of 1.

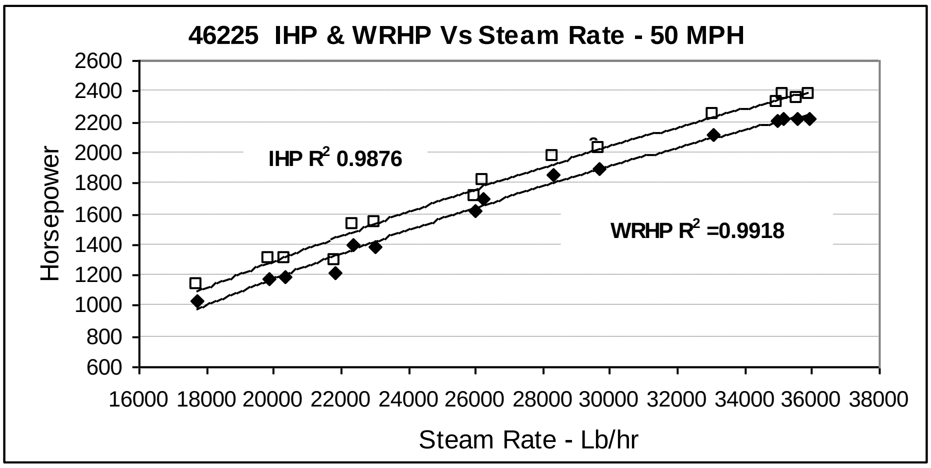
As mentioned in my previous letter, the test bulletin IHP data for the Britannia was uplifted relative to the actual experimental data. Apparently in recognition that the early IHP results where low: hardly an endorsement of **“**consistent**”** data.

The chart below demonstrates a firm WRHP “handshake” of consistency between separate test sequences for the same locomotive type at 35 mph. It spans the same time frame as the chart above. This is in clear contrast to the disparate IHP data.



In summary, the recorded WRHP data was consistent over time; many other examples could be given. This was not the case with the IHP data pre/post early 1953. The change here was clearly the outcome of improvements in the indicating equipment, the dashpot problems having been sorted long since by the end of 1949. Clearly repeatability is not in itself proof of accuracy, but it is an essential first step.

Plots of simultaneous IHP and WRHP data against the steam rate base demonstrate a clearly visible 'Master/Slave' relationship between the paired IHP & WRHP plots (see Chart below).



The WRHP plots are clearly sensitive to upward and downwards movements of the IHP plots. Note in this regard the obedience to this rule of the outlying set 4th from the left. This is an example of simple Boolean logic where B = A-x, where x is a variable as a function of effort; in practice both A and B are subject to random experimental error, hence the elasticity of uncertainty as manifested by the varying separations of the paired plots.

"Testing the Rugby Data – I) Examining the Damping at the Drawbar/Dynamometer Connection**"**

This section covering 1100 words is a further attempt to disparage the Rugby test plant set-up and its operators. I have already dealt with the various misconceptions and inaccuracies on offer either above or in my earlier letter 7th March. I therefore see no need to cover this ground in detail again. I will just add that the claim in the last paragraph of "consistent" ITE data is curious given its demonstrated inconsistency. If such consistency is deemed the case with 73008 and other pre 1953 ITE data, the applied statistical tests are clearly unsound. In contrast the pre 1953 ITE minus WRTE plots return negative MF values; afterwards when positive values emerge, the WRTE data has not shifted, unlike the ITE.

"II) Seeing Sense in the Data"

The arguments stated under this sub heading are not easy to follow given the opacity of the presentation and the surfeit of acronyms. The three steps, a, b, & c set out in an attempt to "test the data" are unsound. The difference between two measured quantities, ITE & WRTE, is reduced by the subtraction of two estimated quantities; the coupled wheel bearing resistance (CWBR) and the plant test tractive effort V^2 dynamic losses ($PTTEV^2$), this being the losses attributable to rotating and reciprocating mass dynamic forces. This process effectively reduces the remainder from a measurement to the status of an estimate. There is no indication that and how the mitigation provided by competing force vector resultants that are less than their mathematical sum has been taken into account. The actual measured WRTE relative to ITE is discarded.

"That residual (the remainder) is such a small ratio of PTTES that the data imply improbably low Cfs (coefficients of friction) in the mechanism from the steam effects, often less than the lower set of Cfs."

The simple answer is that the two estimates and process in the exercise were wrong. The analysis of dynamic force effects is a complex matter. Far from being a Eureka moment, these results were a case for back to the drawing board.

"I consider this exercise shows Rugby TSR to be decidedly on the low side and erratic.
Data on LR and MR from the rest of the world tends to justify the figures for MR hence TSR that I use, so I consider this exercise shows the Rugby TSR to be decidedly on the low side and erratic."

Given that the measured WRTE data from Rugby demonstrates high repeatability, the supposed erratic behaviour only emerges when subjected to the deduction of estimates.
By implication said deductions are erratic and inaccurate. It is not clear what TSR is defining here: it surely cannot be the small remainder data; that, inevitably, is erratic.

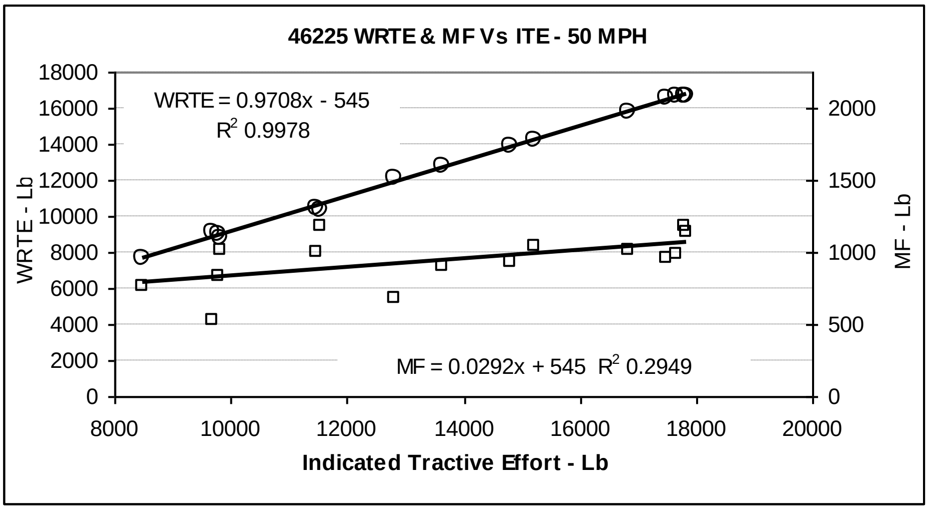
Citing international LR data as back-up is unconvincing, said data is a minefield of disparity. If any threads can be found sufficient to detect a trend, it would just be one amongst many alternative trends available for use. Take your pick.

Below a plot of the Rugby data for 46225 at 50 mph; the only speed for which sufficient simultaneous values of IHP/ITE and WRHP/WRTE data are available (15 pairs).

Beyond the small remainder outcome (MF) this chart does not display the erratic nature of WRTE claimed by John, but then the values plotted are as measured, not the emasculated estimated values created in the pursuit of an untenable concept.

It is tempting to assume the negative value of 545 lb notionally represents the MR when coasting without steam. However in that situation some compression losses will occur in the cylinders, disturbing the projected mathematical trend from when under power. In the absence of said coasting losses, the projected constant would embrace the coupled wheel journal and windage losses and all the cylinder and motion frictional and dynamic losses.

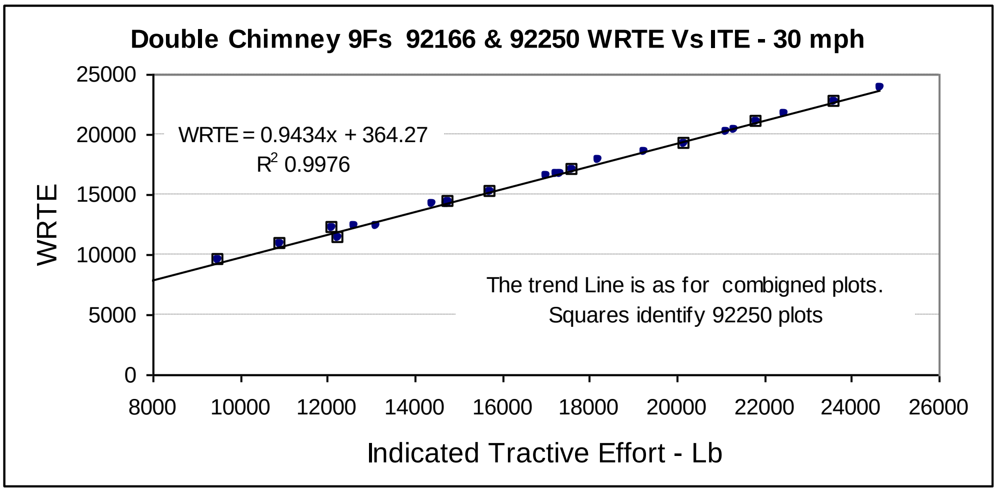
The sensitivity to effort implied in this example is 3%, about mid range of typical Rugby values yielded by similar plots; 1 to 5%. This sensitivity occurs on two basic counts; firstly real effects down to piston thrust on bearings and motion according to work rate, and likewise piston and valve ring pressurisation. Under the notional conditions of zero tractive effort, significant residual losses would remain; for example piston and valve friction would not fall to zero, likewise the dynamic losses. The losses attributable at given speeds are therefore x%ITE + constant n. Secondly, potentially false sensitivity and anomalous outcomes down to random scatter patterns, which in effect, falsify what might be dubbed the 'compass setting' by a degree or two.



The lower trend line represents the machinery friction. Remarkably, notwithstanding the low R^2 value, the formula is the exact inverse of the WRTE formula, returning the same MR values. This, however, is wholly exceptional; typically there is some mismatch between the formulae outcomes of such derivations. I can only think the scatter pattern of the MR plots in this instance is fortuitously balanced. This is far from the usual case, the raw MR plots (ITE – WRTE) are generally not suitable for the direct determination of MR, which in addition are often too limited in number for given speeds to obtain sensible relationships between ITE and WRTE. With just a few plots over a limited power range the scatter produce slopes in the wrong direction; MF seemingly an inverse function of effort.

Other problems are the sensitivity of trend lines to the plots coincident with the lowest and highest abscissa coordinates. This sensitivity can be examined by experimentally removing plots. In the case of the 46225 chart above, removing two plots from the low end increases the residual from 545 to 768 lb. The curve fitting programme and formulae so generated are a mathematical smoothing exercises, *and therefore hostage to the randomness of the scatter pattern.* Solitary plots at the start and finish of trend lines are strongly trend setting, especially if the nearest adjacent plot is somewhat distant. This can still occur with high plot numbers overall, especially in the case of Willans line polynomials. It is apparent that any formulae fitted by the excel programme that approximately coincide with theoretical expectations, as is the case of 46225 above, are fortuitous. Some more insight into this problem is examined below.

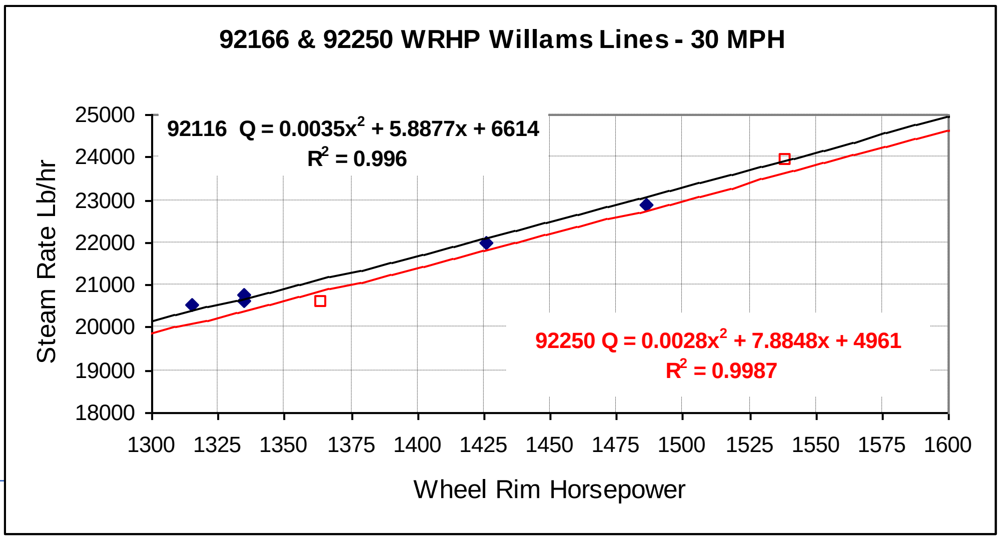
It is curious that John Knowles cites the data for 9F 92166 as encouraging, but condemns same for 92250; the plots of WRTE against ITE for both engines are effectively identical. Both were double chimney engines; 92166 was fitted with a mechanical stoker and $3^7/_8$" blast pipe caps, 92250 4" caps.



The results for 92166 and 92250 were also plotted separately yielding the results tabled below. If any proof were needed that the Excel curve fitting programme formulae bear no relationship to the causal reality this is surely it. Note nevertheless the trivial difference in outcome these diverse coefficients and constants deliver.

Double Chimney 9Fs 92166 & 92250 WRTE @ 20,000 Lb ITE - 30mph					
Engine	Plots	R^2	Formula	20K ITE MR	MF HP
92166	14	0.9978	WRTE = 0.9525x + 192.69	757	60.6
92250	10	0.9974	WRTE = 0.9373x + 476.91	777	62.2
92166/92250	24	0.9976	WRTE = 0.9434x + 364.27	768	61.4

Results for 73030 showed a fall in WRHP against steam rate as blastpipe diameter was progressively reduced in the pursuit of free steaming on Grade 2B coal: $5^1/_8$", 5" and $4^7/_8$" diameter. Given this phenomenon the outcome on this count was examined for 92166 and 92250. Plotted as separate WRHP Willans Lines over the full working range, the curves are so close as to appear as a single curve. Hence it was therefore necessary to focus on an enlargement as below to reveal the effect of reduced blastpipe caps as below. The penalty here for 92166 over the range shown is about 20 HP. The outcome for 73030 was similar. WRTE is a linear function of ITE; this is consistent with the Rugby data generally.



"Nothing is said about the purpose of the exercise set forth in the spreadsheet**"** (Experimental Error)

The simple answer is to inform. In that regard I believe that a few charts demonstrating the graphic outcome of the small remainder problem to be far more informative than its mathematical explanation. John Knowles seems unhappy that I have put this up for scrutiny, hence over 2000 words of general irrelevance seeking to pick holes in it. The spreadsheet as presented is straightforward enough, with clear caveats regarding its scope and simplification relative to actual test circumstances, so I will spend no time addressing these comments, other than those referring to the chart for 45722 plotting the machinery friction data recorded at Rugby. To refer to **"**more variation in (his) MF at each of the speeds" seems to imply the trend line is some kind of concoction on my part and questions the absence of a formula. The trend line is simply the product the excel curve programme, so is presumably the product of the least squares method for the data available. The formula is not necessarily accurate, given the uneven scatter, so is irrelevant. It does however, as the caption says; 'Notwithstanding the scatter, the trendline shown reflects a speed/ magnitude relationship roughly in linewith theoretical expectations.'

Note the word 'theoretical'. Back in 2004 I undertook a theoretical examination of the various elements contributing to locomotive machinery friction and the resulting outcome. The exercise was broken down into nine elements variously contributing to force, friction, dynamic effects, windage, simple harmonic motion etc. The forces were a matter of calculation, the masses known, but obviously the friction coefficients had to be assumed based on published data sheets, technical manuals and some rolling stock empirical data. The values adopted and method erred on the pessimistic. There was no input to this exercise from the Rugby test data or any other similar data. So it was coincidental when the first such exercise was of a similar magnitude to the Rugby data and dish shaped, further exercises for various locomotive types followed this similarity.

John Knowles is fully familiar with this work, so for him to say; **"**Doug has no idea of how such data might be interpreted and analysed. He should be trying to analyse what causes the variation at those speeds.**"** is wholly disingenuous.

"Doug Landau's approach to the Rugby TSR data is in my view one of wishful thinking about its soundness and hopes of using it, and playing with figures to defend it.**"** It con- tinues later on with great irony; **"**If the data are not satisfactory, no good can come playing with it.**"**

Really? This is incongruous; throughout this correspondence I have simply reported and plotted the Rugby data as it exists, at no point have I 'played' with it, in direct contrast to the processes set out in **"**Seeing sense in the data.**"**

"It was the view of D R Carling, Superintendent of the Rugby plant during its operating life during that the plant was not suitable to obtaining the internal resistance of locomotives. In saying that he referred to the SDE, but he also pointed out that the damping provided was to prevent resonance developing, not to provide accurate TSR; indeed it could not.**"**

This bowdlerization of what Carling actually said and thought is not without its absurdity. If the dynamometer was damaged it wouldn't work accurately or even not at all would it? What Carling was talking about was the small remainder problem, not the dynamometer performance, of which he said (I repeat): *"they got their results right*". As previously cited, Carling considered the determination of locomotive resistance equally problematical because of the small remainder problem. If the scatter patterns of MR and LR data are considered as statistical crime scenes they share a common felon; Indicated Horsepower. John seems unable to acknowledge that IHP played any part in the Rugby MR data scatter.

> **"**My difficulty is that I think the Rugby data poor/inadequate.**"**
>
> In summary, this view has not been supported by the arguments submitted.

1. The several supposed shortcomings of the Rugby Test plant set-up in regard to the Amsler dynamometer, have, one by one, been shown as inaccurate and often ill informed.

2. The inaccurate attributions to what Carling actually said, wrote and clearly thought can be dismissed as 'spin'

3. The various players in the design, manufacture, construction and operation of the Rugby test plant were not incompetent.

4. The suggested timescale for de-commissioning the damping dashpot is inaccurate.

5. The treatment of the coupled wheels as part of vehicle resistance is pointless, unsound, and degrades a measured quantity to the status of an estimate. This compromises any statistical analysis.

6. The consistency of the measured WRHP over time, in given circunstaces, sometimes with different locomotives of the same class, appears to have been disregarded.

7. The consistency of the IHP data has been overstated, and does not hold over the timescale involved.

8. **"**Seeing sense in the data**":** The procedures as described have manifestly sown chaos in places where it did not previously exist. Measurements of high consistency are usurped by a feast of needless, and by implication inaccurate estimates. No wonder improbable results follow.

9. Given the controlled environment, the Rugby test station was better placed for the determination of MR than was the case with road tests in regard to LR. The test plant was not subject to the vagaries of wind, track condition and curvature.

Doug Landau

1. Jim Jarvis, as his elder brother Ron, were both LMS Derby engineering apprentices. Under BR Ron was promoted to Chief Technical, CM&E, Southern Region. He was in charge of all design work throughout the region, Based at Brighton, this involved the leading design work on the BR 4MT 4-6-0, the 4MT 2-6-4T and the 9F 2-10-0. He was later responsible for the Bulleid pacifics' rebuild design. Jim was assigned to the Rugby test plant from its earliest days, he is present in a photograph of the ceremonial opening and demonstration run with 60007 in 0ctober 1948. By 1951 one he was in the USA serving a two year scholarship with the Norfolk and Western,

and attending Illinois University where he gained an MSc in mechanical engineering. On return to the UK he undertook the very successful design of the 9F balancing arrangements.

2. Brighton trained engineer Ron Pocklington was in charge of the Fanrbro indicator operation and development at Rugby. In the early days sensitivity and mechanical reliability was poor, and the electrical circuitry was troublesome in various ways. Progressively, improvements were introduced and problems eliminated. In its final state the indicator pressure diaphragm was sensitive to "the slightest breath applied to the steam inlet could make and break the contact." Exact date unknown.

LOCOMOTIVE RESISTANCE FORMULAE - 4th July 2017

Reply by John Knowles to Letter from Doug Landau of 7th March

This is the first stage of my reply to Doug Landau's letter of 7th March. As usual Doug's criticisms are laced with at least as many insults as science, plus in this case calling on several great men most of whom had nothing to do with the subject of the Rugby test plant or LR. In addition he calls on repeatability as a criterion for acceptability or accuracy of data, when all the repeated data can all be wrong. The matters he presents require a great deal of answering. I intend to do that in three parts – first, here, (i) the accuracy of data, statistics and regression, and the form of argument he has adopted, (ii) the great men, and (iii) other matters, including the Rugby plant.

A list of abbreviations used is given at the end.

1 What I am accused of and Regression Analysis

In his final paragraph, he says:

In summary the supposed shortcomings of the Rugby Test plant, its designers and operators are groundless. The available experimental data demonstrates consistent repeatability over time and circumstance. Repeatability is a key indicator of metrological integrity. That is not to say everything is perfect and falls in place in place like a jig saw. Given the understood limits of experimental error, however small, and the random nature of scatter, the real world is more complicated. Exactly the same problems obtain when reconciling the data from road tests. Road tests have however confirmed the differences in test plant MR in the case of the Crosti and standard 9Fs. In other words the empirical evidence derived by different methods remains consistent. A key test of scientific proof is that its claims are consistent with the empirical evidence. The powers of the regression statistical process used by John Knowles fails the empirical test significantly and is thus unsound, supposed statistical integrity notwithstanding.

He has not shown any of his claims made in this conclusion, ie the conclusions come out of the air unsupported by the content of the paper. He has not shown anything to be wrong with regression, and what criterion he has employed to reach his astonishing conclusion about it. He does not appreciate that repeatability is an insufficient criterion for acceptability of experimental data – the repeated data can be all wrong. He does not show repeatability to exist in the Rugby data – I find precious little of it. He gives no reference for the claimed confirmation of TSR by road tests for the Crosti and standard 9Fs, nor explained how he reconcilied what are essentially different measurements – TSR given on the test plant and LR on the road. Given the lack of repeatability in the Rugby data, he does not say which 9F data among the non-repeating 9F data he picked for his own use as the resistance of the 9Fs. The doubts about the test station results are far from groundless, his assertion notwithstanding.

I have answered much the same points in my previous letters on the Society webpage on this subject. As he pronounces further on the subject with no more evidence of knowing much about scientific analysis, and in particular about testing data and regression, there will be repetition in this reply.

He has not explained what he means by his statement that is not to say everything is perfect and falls in place in place like a jig saw, and that given the understood limits of experimental error, however small, and the random nature of scatter, the real world is more complicated. It is all very well to claim there is scatter in data, that is random and that it cannot be avoided, but scatter is lack of repeatability, and its extent and pattern gives the probability of the data yielding sound results. Indeed, what appears to be scatter could be "good" in revealing important aspects of behaviour, which were not previously appreciated. Randomness, in the sense of absence of bias, is an essential feature in experimentation and in analysis of data.

Does he mean that if the data do not fit precisely what he is looking for, the random scatter has to be treated in some way to make it amenable? That is precisely where statistics, as a science accepted by millions of practitioners worldwide, has its place. Simply drawing a line through data, or fitting an equation to data by trial and error, with a self-chosen criterion of acceptability of the relationship implied by the line is no proof that accuracy or acceptability of data has been established, quite the contrary. Further, where there are two or more determining variables, or the relationship posited is complex (eg it changes over the range of the data, or there is variation with powers, including fractional powers, in one or more of the determining variables, it is impossible to fit a relationship to data without regression. The supposed deficiencies of regression are mostly the result of Doug Landau's lack of knowledge of the process and what it can achieve. He is decrying regression because it can show deficiencies in data and/or methods and/or relationships which he wants to claim are satisfactory, that the Rugby data in his hands can be declared to be satisfactory, and is declaring often, apparently in the hope that if the declarations are made often enough, they will eventually be accepted, especially if he can deprecate my explanations and remarks sufficiently. I say that because he has done nothing to show the data to be satisfactory. As for deprecating, see the net paragraph also.

Whatever is the basis of his claim that the powers of the regression statistical process I used fails the empirical test significantly and is thus unsound, supposed statistical integrity notwithstanding? This conclusion is not even discussed, ie he gives no basis for it. The conclusions are therefore not based on a scientific approach or discussion. There is no reference to the small difference problem (SDP). Nor any appreciation that data can exist but can be not good enough for any sound result to emerge; or that any analysis or conclusions require testing the data, choosing the right form of analysis, ie the right form of equation, and applying well known and easily available tests of the probability of the results being acceptable. In other words, the nearness to fitting the jigsaw or some other criterion says whether the data really say anything worthwhile.

Conclusions of a paper follow from its content. In this case they do not. Doug Landau's supposed conclusions do not follow from the content. These are broad statements of his beliefs not supported by the content of the paper, and without any references to other literature which do support them. His approach amounts to false argumentation, false accusation, especially in relation to things I have said. In other words, anyone quickly reading the conclusions could be led to believing the paper had cogent argument about regression and the soundness of the Rugby data (among other things) whereas it does not even remotely do that. What are his motives for such action? Is he hiding that he has no supporting arguments, or trying to put readers off what I have said?

Further, I should say Doug Landau is not in a position to judge on the matters just mentioned, or the conclusions he drew. Consider two examples of "analyses" he performed, which are simply not right. First, he wanted to establish the TSR for 9F 92050 at 30 mph. He chose seven observations from a Rugby test of that engine, and obtained a trend line from a computer program (Excel) in the form of a quadratic equation ($aX^2 + bX + c$) for each of IHP and WRHP (at Rugby this was DPHP) against Q, the steam rate. The results were:

$IHP = -1Q^2/10^6 + .1148\ Q - 463.45$
$WRHP = -9Q^2/10^7 + .1064Q - 440.41$ (this WRHP is DPHP)

From these trend lines, it follows that
$IHP - DPHP\ (= TSRHP) = -Q^2/10^7 + .0084Q - 23.04$ by subtraction,
And $TSR = -12.5Q^2/10^7 + .105Q - 288$, multiplying by 12.5 to convert HP at 30 mph to a force. From that,

For Q of 14,000, TSR = -245 + 1470 – 292 = 933
For Q of 21,000 (ie plus 50%), TSR = -551 + 2205 – 292 = 1362 (plus 46%)
For Q of 28,000 (ie plus 33%), TSR = -980 + 2940 – 292 = 1668 (plus 22%)

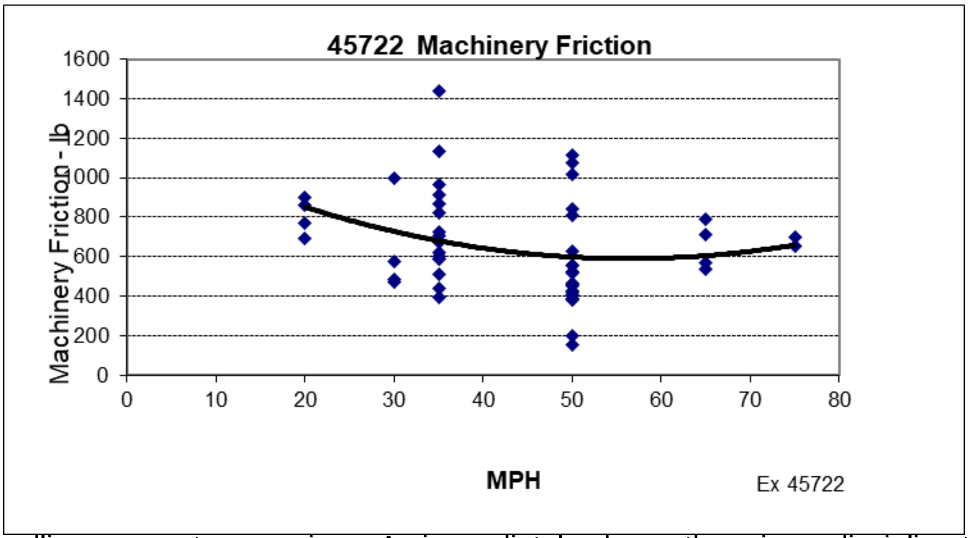
This exercise was supposed to show that TSR was constant at 30 mph (like a dog following its master on a lead he claimed – see *Backtrack*, April 2014, p 253). It does the exact opposite. It shows TSR supposedly varying with Q, but not as fast, and at a declining rate, to high levels.

But this is inappropriate analysis. There are only seven observations, out of 191 for all non-Crosti 9Fs tested. It is unscientific to select only some data from the total without a good scientific reason. Why were not all observations at 30 mph pooled, or indeed all 191, and the effect of speed tested as well? With only seven observations, the chance of finding sound results is much reduced. With the considerable range usually found in Rubgy TSR values under similar circumstances (as exemplified below) that is a considerable failing – it is not known how reliable the answers are. Nor is there any examination of the data and these results in relation to the Small Difference Problem (SDP), nor any testing of the data, to see if it is sensible.

Why was a quadratic chosen? Q has its effect on ITE (not in direct proportion, because SSC varies across the range of Q). Q^2 however is not known to have an effect on ITE, especially when its value is in millions (steam rate Q is expressed in lbs/hr, which occurs in thousands). Presumably the idea was to obtain something resembling the quadratic form of the VR element of LR, in the hope that the TSR and VR could be added together. That results in a minute coefficient on Q^2 as would be expected, but as the values of Q^2 are in millions, they are still large. In any case, the unit squared, Q, is not the same as the unit squared in the VR, ie V. No statistical tests are available, a considerable failing, for they would have shown the fallibility of the reasoning and analysis.

The basis of the analysis is incorrect in using Q at all. IHP is dependent on Q, but not as a straight line (as is clear from any curve of SSC). But DP is not dependent on Q. It is dependent on ITE and TSR (and the components of TSR), not on Q or Q^2.

Second, he is in the habit of using inappropriate trend lines to draw conclusions. See my previous post, in which I pointed out that a trendline of TSR against speed, and only speed, cannot be the right relationship to examine. The six vertical lines obviously contain the real determinant of MR, with speed a lesser factor. The proper approach would have been to use the data at each speed separately (look at the number of observations at both 35 and 50 mph), and test the various possible explanations, of which PTTE is likely to be the best, because it is the major source by far of MR, and to fit regressions rather than trend lines.



These trendlines are not regressions. As immediately above, there is no discipline to them – Doug Landau has used them here to obtain relationships which do not exist in physics or mechanics. They can be done without any of the tests possible with regressions.

Doug Landau's statement that a key test of scientific proof is that its claims are consistent with the empirical evidence is certainly not satisfied by either of these cases, by observation. In the graph above, the line claimed by the relationship ignores most of the data, because the supposed relationship is not valid. At each speed, TSR (his vertical axis) is shown dependent on speed. But TSR is little dependent on speed, which is why his supposed relationship ignores most of the data. TSR varies mostly with other things, on which see below.

The usual logic applied in scientific investigation is formulating hypotheses which from first principles might be relevant to the subject in hand, gathering data which enables the hypotheses to be tested and new ones to emerge (ie almost everything which can be measured about the subject should be measured), testing the data through physical and statistical tests, forming relationships from the tested data to show whether the hypotheses can/should be accepted, including to what degree the acceptability applies. The data has to agree with the theoretical, scientific and/or common-sense expectations, there has to be enough of it, and it has to be sufficiently exact. The empiricism is only part of the process.

For the kinds of claims he makes, he should appreciate that things have moved on since he was a boy, that for decades the data used in deriving a relationship is tested in advance for its soundness, and subject to various forms of analysis, of which regression is the most common, that analysis subject to tests of goodness of fit, whether it differs sensibly from alternative values (including zero), and tests of alternative explanations. With some education in the subject, he would learn that regression is often *the* empirical test, or the most important and useful empirical test – ie part of testing the data for soundness, for formulating explanations of the data, and saying how sound any explanations tested by regressions are. That would save him having to offer weak excuses, such as, to quote, the understood limits of experimental error, however small, and the random nature of scatter, and the real world being more complicated.

Further, on his idea that a key test of scientific proof is that its claims are consistent with the empirical evidence. This puts the cart before the horse. The empirical evidence might be wrong, very poor in itself, subject to the SDP, or untested for its reliability. Then he has to test the relationships, ie establish scientific proof. Doug seems to believe the data are sacrosanct, apparently perfect, or if not perfect (a real world situation?) they are as good as can be obtained in the real world, and are not to be questioned. Not so, as should be clear from almost everything I have written so far. He should be aware of a good example in locomotive testing in this country. The overall BR testing system was badly flawed in the principles guiding it because it depended on an unjustified assumption that a constant blast pipe pressure (BPP) ensured constant Q, at all speeds, and on the plant and on the road. That is why, in general, it is not possible to take the ITE from the plant (where it was usually measured), and deduct EDBTE from road tests for the same Q and V, EDBTE corrected for ind conditions, and to claim that the difference between ITE and EDBTE (as shown in the BR Test Bulletins) gives LR. Only late in the testing was it discovered by simple consideration of the data, that for LR in this case, that such was not correct, that for a given pressure Q varied with speed (as seems obvious). Further, the Q provided by the boiler for a given BPP was different on the road from that on the plant, so my question to him about the 9Fs is crucial.

It is difficult to prove conclusively that experimental data are correct. As above, sheer repeatability is insufficient – all the data can be wrong. Doug uses Carling's belief that because the ITE results for the same test circumstances fall in a narrow band, the ITE data are acceptable, even accurate. Carling also believed that the results from the Farnborough indicator used at Rugby were much the same as those from mechanical indicators available to BR. Mechanical indicators were susceptible to lags and incorrect readings, however, on account of the multiplier in the working, and the small size of the indicator cards being difficult to measure. No proof there. Inserting the input data (pressure, Q, cut off, steam temperature) into the Perform program gives results a little higher than those from Rugby. Perform is by far the best way of approximating cylinder outputs, but itself requires some approximations to inputs, especially cylinder temperature at the beginning of a stroke. Very persuasive, but not absolutely a proof. The Rugby indicator results are highly consistent for a given engine when regressed against Q and V (which themselves determine cut off and steam temperature) in an equation of the form $ITE = cQ^aV^b$, a, b and c being constants, giving good equations and good test statistics. Again, not absolute proof, because the data could all be wrong.

Doug is a great advocate of the accuracy of the instrumentation proving something, eg the Amsler dynamometer, claimed to be accurate to within +/- 1%. That too says little, nay can be completely misleading, if what pull reaching the dynamometer is itself distorted or other factors he has not allowed for, or the SDP is present. (See equation below for the passage of energy from ITE to DP.) [The same Amsler was the source of the DP readings in the first two years of the operation of the Rugby plant, when DP typically exceeded ITE, ie that energy was added to TSR (ITE – DP) by processes in TSR which should all have absorbed energy, ie what was measured by the DP was impossible. This was said to have been cured, by taking oil out of the dashpot in the chain between ITE and DP and replacing it with air, ie replacing a high resistance (oil in the dashpot) in the chain by a lower one (air in the dashpot) resulted in energy being absorbed between ITE and DP, as it should have been. If the change of the medium in the dashpot is all that was done to the system, it is not an explanation for the change in the relativity of ITE and DP, and DP readings remain suspicious. If of course, other things never reported were done, that could well be different.]

The major test to use if there are none available for the data as data is to fit the relationships to which the data should conform, decided either from past research, or from first principles, as used in formulating hypotheses about the subject before the research started.

And if data fail tests, or no tests are possible, then no more use can be made of it. It cannot be used to prove anything, except how not to specify and conduct experiments, and whether it is possible to obtain TSR at all.

Doug Landau does not appreciate that the data are the real world, (see his remark above about the “real world” and things not fitting together like a jigsaw puzzle). Whether he likes it or not, in science, he cannot interfere with data. He might, with some statistical and technical analysis, show that is probable (even to a degree of probability) that the data would be useful for finding MR or TSR if such and such had been or not been done (I do some of this below), but he cannot impose anything on the real world.

Last, be it remembered that it was said in the *Locomotive Railway Carriage and Wagon Review* for December 1957, pp 233-4, in one of a series of articles in that journal during the second half of 1957 on *Locomotive Testing on the Rugby Plant, BR*, that it is *not* possible to measure the internal friction of a locomotive accurately on a test plant, only to confine its value within comparatively wide upper and lower limits. (As the data are so unsatisfactory, the confidence with which any declared upper or lower limit can be held must be low.) The articles were unattributed, but were almost certainly prepared by D R Carling, Superintendent of the Rugby Testing Station during its operations. Certainly, Carling did not refute the point. It is therefore extraordinary that Doug Landau, after all these years, claims to be able to judge the Rugby data better than Carling, and to want to do so without explaining how. That is the same as setting his face against regression results – nothing declaring against the Rugby results, specially by me, is to be tolerated.
I suspect too that he believes that scatter is evenly distributed and that the true answer lies in some sort of average of all the data. I fear not. The testing and consideration of the data requires consideration of the scatter, its extent and an examination for biases.

Simply declaring that the Rugby data are fit for providing TSR values avoids crucial steps in showing that it is fit. Declarations are empty if the steps have not been taken. Doug Landau has never shown that he has considered the data, so it follows his declarations are empty.

I have therefore turned to testing the data for their soundness. This involves going back to the first principles of the mechanics involved, analysing the forces involved, and considering from acceptable references the likely friction coefficients involved. I have found the data lacking.

2 Are the Data Sensible?

I have considered their "soundness" in four ways. First, they have been graphed against PTTE, for their consistency or repeatability. This has been done for every engine tested on the plant where there were at least a dozen observations at one speed. In some cases, more than one speed was available, with up to four speeds suited to this analysis. In no case were the data consistent or repeating. [Graphing is mostly sufficient to show this, but in one case (Duchess 46225) it was shown in addition by painstakingly listing and ordering the observations which are inconsistent with one another.]

Second, I considered the values of TSR obtained from ITE – DP (the experimental results) for their magnitude. Using the same data from the cases where there are at least a dozen observations at a single speed, from each TSR observation were deducted the CWBR and the items varying with speed squared (where relevant), both of which items should be constant at the speed concerned, to leave a residual, which ought to be the value of all items varying with piston thrusts. In analyses and comparisons of mine, these were found to be a ratio of .05 to .07 of PTTE (details available on request). In these Rugby TSR data, the ratio is much lower than .05 to .07. For the twelve engine-class/speed combinations considered, the vast majority result in ratios which on average are less than .025. Only the Jubilee at both speeds (40 and 50 mph) could be said to demonstrate coefficients approximately those expected, but still on the low side, but the Jubilee data are problematic in other respects. Some are very low indeed, and the value of the ratio is generally erratic.

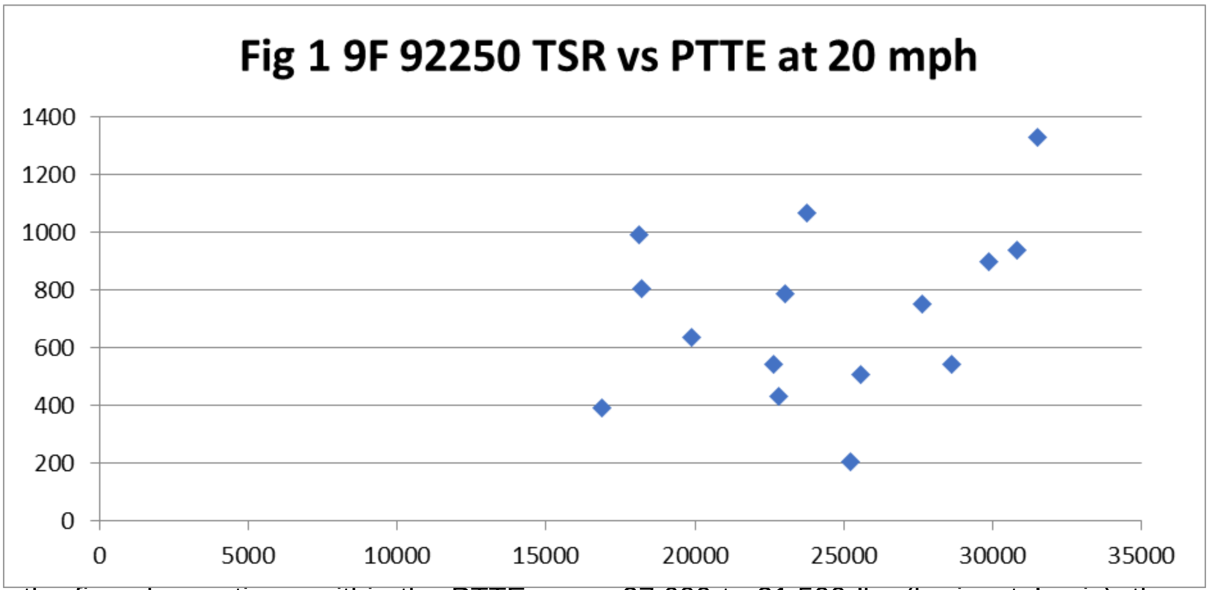
Third, TSR was regressed against PTTE for the same twelve class/speed combinations, for each speed/class combination. The logic is that an equation in TSR at each speed should in those circumstances have a positive constant covering all items constant at that speed, and a positive coefficient on PTTE covering all items varying with PTTE, ie constant + xPTTE at each speed.

Fourth, Rugby data were also used to apply the input/output approach to MR for a couple of classes, as used in obtaining the approximate MR of internal combustion engines. These yield MRs which are far too high. This is consistent with the low values of TSR. This however is incidental to the previous three approaches.

3 Consistency/Repeatability of Rugby Data

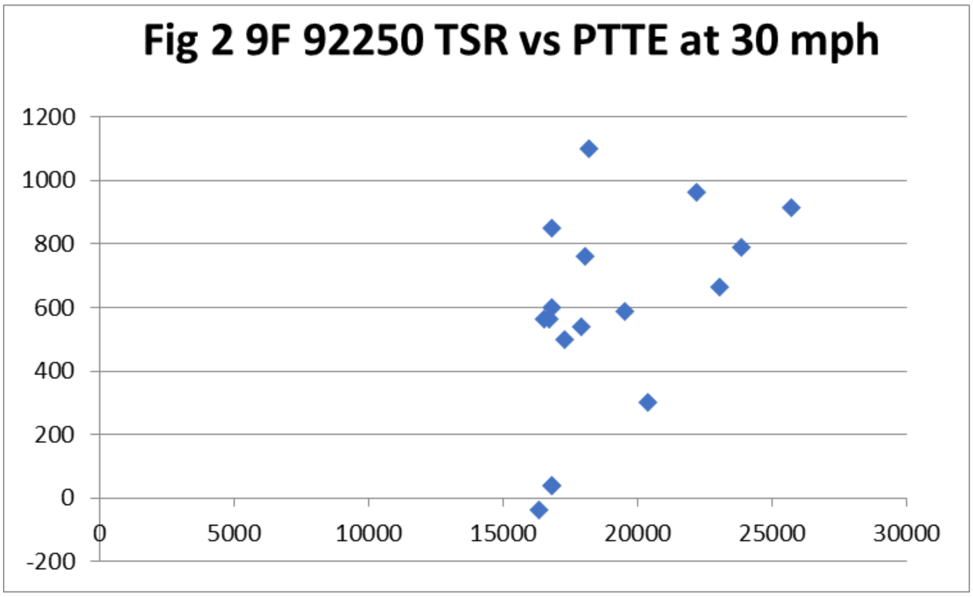
To exemplify the point about non-repeatability of the Rugby TSR data, I have chosen the data from 9F 92250 , the last steam engine tested on the plant. By then, practice on plant should have been as good as it ever was. In this case, the data are available for at least 12 observations for four speeds, 20, 30, 40 and 50 mph.

In all the figures TSR is on the vertical axis, PTTE on the horizontal.



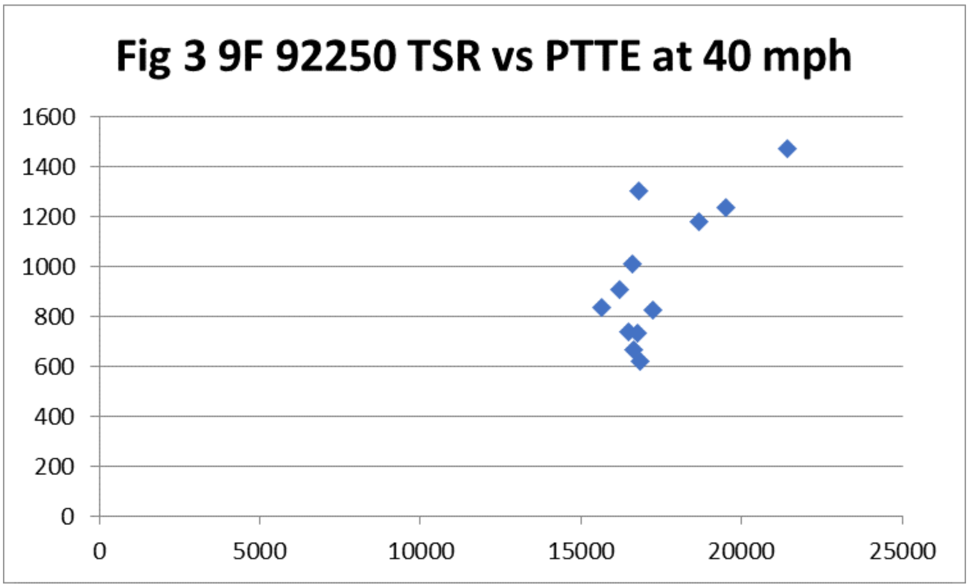
For the five observations, within the PTTE range 27,600 to 31,500 lbs (horizontal axis), the TSR range is 544 to 1331, the average TSR is 844, and its Standard Deviation 290.

30 mph



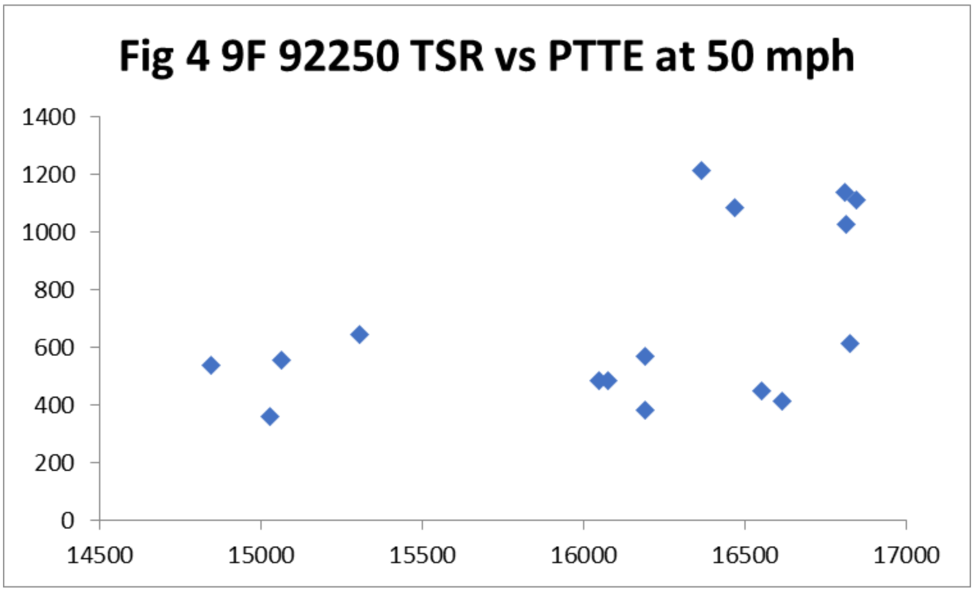
Twelve of the 19 observations fall in the PTTE range of 16,300 to 19,500 lbs, in which the TSR range is -38 to 1100 lbs. The average TSR of these 12 observations is 508, and their standard deviation 343.

40 mph



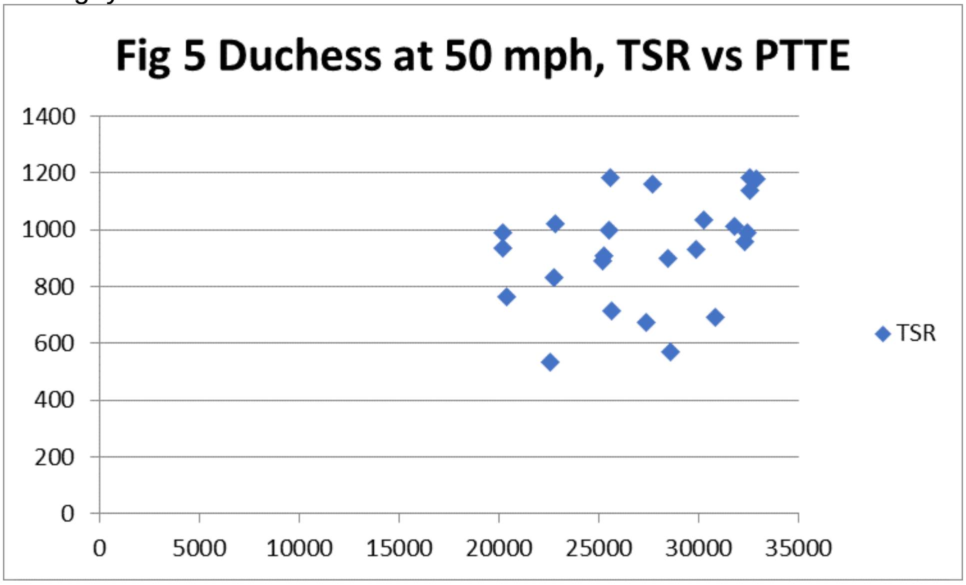
Of the 12 observations, nine are within the PTTE range of 15,600 lbs to 17,200 lbs. The TSR range of those observations is 619 to 1303 lbs, the average 849 and the standard deviation 209.

50 mph



The four observations at about 16,800 lbs PTTE contain TSR in the range 615 to 1140, for which the average is 973 and a standard deviation of 243. The four observations at about 16,500 lbs contain TSR in the range 615 to 1140, for which the average is 823. Given the circumstances of their origin (and the SDP), the three observations in the far top left of Fig 4 are as good as could be expected, but the fourth observation at 16,800 lbs demonstrates the lack of consistency, or repeatability.

In addition, Fig 5 gives the TSR and PTTE data for Duchess 46225 at 50 mph, for which here are 24 observations, the greatest number at any one speed for any single engine tested at Rugby.



At a PTTE of close to 25,000lbs PTTE, the five TSR values vary from 713 lbs to 1185 lbs, with an average of 939 lbs. At a PTTE in the range of 28,000 to 30,000 lbs, TSR varies from 570 lbs to 1163 lbs, with an average of 881 lbs. At a PTTE of 32,000 to 33,000 lbs, the six values of TSR vary from 960 to 1185 lbs, with an average of 1083 lbs, this being the only case of TSR values being even remotely close of all the PTTE ranges discussed here, there being two groups of three observations which could even be said to demonstrate repeatability, even though the two groups of three are about 200 lbs or 20% apart.

In all five cases, the spread of data is much greater than modest variations about what Doug Landau seems to consider the right value of TSR derived from the Rugby data, these modest variations being what he terms scatter, something he regards as unavoidable, but perhaps excusable. TSR is of course the subject of interest. The variation is in most cases indeed modest in terms of ITE or DP or PTTE, but in terms of TSR it is large, on account of the SDP. Far from showing that TSR is constant at a wide range of PTTE, the data characteristics show the opposite, that TSR varies a lot to a degree to which mechanics provides no basis, seen also in the large standard deviation in TSR. Further, considering the variability in relation to DP is not sound, because DP is simply a measurement of ITE less TSR, ie DP is a result of those other two items; or DP is the result of the effect of TSR. Furthermore, scatter is not something to be judged according to the ideas of Doug Landau. Statistics has methods for making this judgement in relation to the best fit to the data recorded, and the size and regularity of the deviations from the best fit, ie whether even the small amount of repeatability occurs by chance.

I did the same for every engine tested at Rugby for which there are at least a dozen observations at a given speed. It all shows similar characteristics. The data are available on application.

It is obvious that there is almost no sensible repeatability in most of these data. No doubt this will draw forth the cry that strict repeatability is impossible in most experimentation, and that there are some observations that are close enough to be regarded as the same. Where the observations are close, that is indeed what I expect. But I have considered narrow ranges of PTTE above and found a wide variation in associated TSR, in each case detailed under each Figure. The TSR data can be said to be no better than erratic. Further, a considerable number of observations are low, which raises the question of what value they should have. On that see the next two sections.

With the wide spread of TSR data at a given rate of working, given his criticisms of my remarks, it would be of interest to know what Doug Landau would consider to be the TSR of 92250 in the range of 20 to 50 mph based on Rugby data. Given his defence of these data, that seems a fair question to ask him to answer.

4 Implied Value of (TSR – CWBR – MR – (resistances varying with V^2))/PTTE

In this exercise, it is considered that TSR comprises CWBR, MR, resistances (friction and work) varying with V^2, DR and heat. The value of these constituents of TSR is not separately measured, but any DR for example will be included in TSR. If heat is lost, it is not included in TSR.

Using the same data from the cases where there are at least a dozen observations at a single speed, from each TSR observation were deducted the CWBR and the items varying with speed squared ($PTTEV^2$) (where relevant), both of which items should be constant at the speed concerned, to leave a residual, which ought to be the value of all items varying with piston thrusts. The deductions for CWBR and $PTTEV^2$ were obtained in my earlier analysis of MR from first principles (available on request), and are very reasonable values (the CWBR uses Cfs consistent with rolling stock resistances which emerged from Ell's researches into British rolling stock resistances (Ell was an officer in the locomotive testing on BR). In that analysis, the value of this ratio was found to be .05 (low) to .07 (high) of PTTE. Note that this .05 to .07 is not a coefficient of friction, but the proportion of the friction to the net forces involved in PTTE both at a common point, the CW rims. The actual Cfs occur at many locations (piston rings, glands, crosshead, and its guides, gudgeon pin, rod pins and the addition to the vehicle only CWBR from the PTTE forces); Cfs at particular points vary from .012 to 0.14. Amalgamated, these yield the ratio of .07. Lower illustrative values in some cases yield the .05.

The following tables are the results of applying this approach to 9F 92250.

In Tables 1 to 4, (a) represents net friction of rods on pins and work done working on unbalanced reciprocating masses; and residual (b) is column 3 – column 4 – column 5.

20 mph

1 Run	2 PTTE	3 TSR	4 CWBR	5 V sqd items (a)	6 Residual (b)	7 Residual/PTTE (c)
2237	16875	393.75	228	38	127.75	0.008
2251	18116	993.75	228	38	727.75	0.040
2168	18200	806.25	228	38	540.25	0.030
2229	19879	637.5	228	38	371.5	0.019
2243	22626	543.75	228	38	277.75	0.012
2249	23016	787.5	228	38	521.5	0.023
2164	22831	431.25	228	38	165.25	0.007
2226	23774	1068.75	228	38	802.75	0.034
2250	25240	206.25	228	38	-59.75	-0.002
2167	25585	506.25	228	38	240.25	0.009
2230	27644	750	228	38	484	0.018
2255	28613	543.75	228	38	277.75	0.010

2235	29851	900	228	38	634	0.021
2233	31496	1331.25	228	38	1065.25	0.034
2170	30822	937.5	228	38	671.5	0.022

Table 1 Ratio of residual (see text) to PTTES in Rugby Data for 9F 92250 at 20 mph
Average value of column 7, .019.

30 mph

1 Run	2 PTTE	3 TSR	4 CWBR	5 V sqd items (a)	6 Residual (b)	7 Residual/PTTE (c)
2146	16324	-37.5	228	86	-351.5	-0.021
2238	16515	562.5	228	86	248.5	0.015
2150	16801	37.5	228	86	-276.5	-0.016
2228	16825	600	228	86	286	0.017
2155	16790	850	228	86	536	0.032
2147	16808	37.5	228	86	-276.5	-0.016
2227	16703	562.5	228	86	248.5	0.015
2144	17302	500	228	86	186	0.011
2252	18073	762.5	228	86	448.5	0.025
2156	17908	537.5	228	86	223.5	0.012
2225	18179	1100	228	86	786	0.043
2145	19518	587.5	228	86	273.5	0.014
2231	20358	300	228	86	-14	-0.0006
2157	22200	962.5	228	86	648.5	0.029
2234	23049	662.5	228	86	348.5	0.015
2148	23877	787.5	228	86	473.5	0.020
2149	25718	912.5	228	86	598.5	0.023

Table 2 Ratio of residual (see text) to PTTES in Rugby Data for 9F 92250 at 30 mph
Average value of column 7, .024

40 mph

1 Run	2 PTTE	3 TSR	4 CWBR	5 V sqd items (a)	6 Residual (b)	7 Residual/PTTE (c)
2177	15639	834	228	153	453	0.029
2176	16189	909	228	153	528	0.033
2162	16474	741	228	153	360	0.022
2253	16657	666	228	153	285	0.017
2239	16748	731	228	153	350	0.021
2174	16795	1303	228	153	922	0.055
2163	16841	619	228	153	238	0.014
2175	16586	1013	228	153	632	0.038
2161	17232	825	228	153	444	0.026
2180	19504	1238	228	153	857	0.044
2160	18683	1181	228	153	800	0.043
2186	21428	1472	228	153	1091	0.051

Table 3 Ratio of residual (see text) to PTTES in Rugby Data for 9F 92250 at 40 mph
Average value of ratio (c) in column 7 .033

50 mph

1 Run	2 PTTE	3 TSR	4 CWBR	5 V sqd items (a)	6 Residual (b)	7 Residual/PTTE (c)
2244	14846	540	228	239	73	0.005
2183	15307	645	228	239	178	0.012
2241	15030	360	228	239	-107	-0.007
2169	15066	555	228	239	88	0.006
2246	16077	487.5	228	239	20.5	0.0013
2248	16190	382.5	228	239	-84.5	-0.005
2240	16048	487.5	228	239	20.5	0.0013
2165	16190	570	228	239	103	0.006
2247	16613	412.5	228	239	-54.5	-0.003
2242	16552	450	228	239	-17	-0.001
2182	16842	1110	228	239	643	0.038
2166	16807	1140	228	239	673	0.04
2245	16812	1027.5	228	239	560.5	0.033
2257	16823	615	228	239	148	0.009
2181	16366	1215	228	239	748	0.046

Table 4 Ratio of residual (see text) to PTTES in Rugby Data for 9F 92250 at 50 mph
Average value of ratio (c) in column 7 .013

The residual is often negative or very low. The value of the ratio in Col 7 in each table is far too low, meaning TSR is too low subject to the items in cols 3, 4 and 5 being correct. A ratio of .05 to .07 expected, for low and expected coefficients of friction. Only two runs in the 92250 data, in table 3, 40 mph, nos 2174 and 2186, satisfy this criterion, and then only at the lower expected value.

This approach relies for its conclusions on other analyses I have made in other contexts. I do not claim that the data could not be tested for this purpose in other ways. I do not accept the judgemental comment (made with no exemplification) of my critic that I have selected friction coefficients to justify my conclusions on this second or any other approach. I defend the values I chose from several sources.

5 Regressions of TSR data

TSR was regressed against PTTE for the same twelve class/speed combinations used in analyses above. Each regression was made at a particular speed. The logic is that an equation in TSR should in those circumstances have a positive coefficient on PTTES and that the rest of TSR should be included in a constant, the values of that coefficient and the constant emerging from the data, not imposed. The question then arises, what relationship should be sought?

Between ITE and DP are the components of TSR, plus BR and DPP. BR is braking resistance at the braked rollers and equal to WRTE. It is transmitted through the frames and CW bearings to the locomotive drawbar, where it emerges as drawbar pull DBP, equal to BR (although as a couple causing oscillation in a vertical plane about a horizontal axis, resulting from the differing heights above rail of the locomotive drawbar and the CW centres). The resistance of CWs rolling on the rollers of the braking dynamometer has been considered as part of the resistance against which the engine was working.

ITE – DP = TSR.

ITE – CWVBR – MR = WRTE, all terms measured at the CW rims

WRTE = BR = DBP

DBP – DR – Heat – fV^2 = DP, whence

DP = ITE – CWVBR – MR – WRTE + DBP – DR – Heat – fV^2

= ITE – CWVBR – MR– DR – Heat – fV^2
(Remember that in my approach, after resolution of the weight borne static load on the bearings and the forces of the mechanism on those bearings, the CWVBR is deducted from the resolved sum and the remainder (the extra resulting from mechanical action) is part of MR.) [Doug Landau appears to be unaware of the convention applying to the term static axle or bearing load. He thinks it means without the wheels turning. It applies to both circumstances. There are plenty of examples of the term static in the sense in which I have used it – see for example the paper by Cox on locomotive axleboxes, which he quoted, with the flavour that Cox's paper proves I am wrong in some way. If this still offends him, he can ignore the word static].
The only source of energy in this system is ITE. DP has no independent existence of its own. It is merely a pressure measuring device giving the pull resulting from ITE – TSR. If any ITE observation is wrong in fact, the unintended error will affect the MR, DR and Heat elements of TSR, and pass to DP. If the ITE is correct, but measured wrongly, then MR, DR, Heat and DP will be correct, or at least as correct as if there were no error in ITE, but TSR will be wrong. It is therefore highly probable that DP will be wrong and TSR wrong in consequence. The equality between WRTE and DBP must remain. The fV^2 term applies to any net forces and net work associated with the revolving masses on pins, and work done revolving unbalanced masses. Heat arises at the dampening, (the dashpot when filled with oil was water cooled), and to any other loss of heat between the CW rims and the DP. Multiplying throughout by (-1), dividing some terms into fixed (constant) and variable portions, and rearranging:
at any given speed, CWVBR and fV^2 will be constant, as will any constant in MR, which means
TSR = constants + bPTTE
There are no data of DR per se. The Belleville washers and dashpot will have reacted in proportion to the forces involved (the dashpot) or be fixed for the speed concerned (the Belleville washers), and partly in proportion to the effort, which effects should divide into constant and variable in a regression. Heat from any effect (the Belleville washers and dashpot) will be lost from measurement, so that measured DP will have been too low and measured TSR too high.
TSR = constants + bPTTE is therefore what is to be regressed. That would be followed by examining the results and the residuals for any sensible conclusions which can be drawn about the effect of heat, even from calculating its value from first principles. Alternatively, if the results can be obtained for several speeds, and are very good, they can themselves be analysed for the approximate values by elimination. It will be noticed that the relationships are a result of the data speaking for themselves – nothing is imposed.

It would be wrong to regress DP against Q. Q has already influenced ITE, at a rate varying with Q per se and V, and as seen in the Specific Steam Consumption. The same applies to regressing DP against ITE. That would not provide any relationship of any value, on account of the big number which each represents, ie that DP will be close to ITE. The difference between the two large numbers will be small, and it is to be expected that the two will be highly correlated, which can distort the results. Further, any such regression will as a result give a high value of r^2, which to many unpractised analysts is the be all and end all of regression or other approaches to obtaining relationships. But regression will also give how high are the probabilities that the terms and coefficients on them are close to being correct (or significant, meaning significantly different from something in the relationships (eg zero, or a close value; significant does not mean large). The coefficients on ITE would have definite high values of the t ratio, ie that the slopes of the relationship are high, indeed very high. The relationships however give very low t values for the constants, which means that it is not possible to fix the relationships with any certainty. TSR is the thing to regress. It would be expected that if the data are good, a well-established constant and coefficient on PTTE would emerge.

It would also be wrong to regress the equation DP = ITE – constants – bPTTE. The three terms PTTE, ITE and DP are close in magnitude and highly correlated, which can affect the answer. Secondly, unless the data are very good, it will be impossible to separate ITE and DP statistically and obtain a sensible coefficient, especially when DP is so erratically related to ITE (as the data in Table 1 shows). Further, any estimation should use actual data and then as parsimoniously as possible, ie without needless complication. TSR is the subject of interest, the matter under investigation. TSR is available in its own right, ie as (ITE – DP). It is not acceptable to “smooth out” ITE and DP, by fitting an equation to both separately. Even if such is done, it is necessary to show that ITE and DP are statistically separable with various degrees of confidence, ie are significantly different at some level of probability accepted for experimentation of this kind from zero. As such, it is not possible to show that, because the values of both ITE and DP have wide confidence intervals of their own.

6 Regressions of the TSR data.

6.1 Duchess 46225

Equations for TSR at 50 mph

There are 24 observations for this engine at 50 mph, the greatest number at one speed for any engine tested at Rugby, the number a useful characteristic in obtaining good results.
The result of fitting TSR = constants + bPTTE is
TSR = 522 + .015 PTTE. [1]
The observed data are very dispersed, as shown in Figure 5, with consequential low significance of the results. Equation [1] is the fitted equation of Fig 5, the best fit to the data using the above form of equation. The SEE is 183, Sigf F 0.114, t on constant 2.12, and on the coefficient 1.66, with R^2 0.111. On all possible grounds, this is unacceptable. Despite having the right signs, the coefficient on PTTE is double or more that expected, ie the slope of the relationship is much too high. The SEE puts a range of $^+/-$ 183 to give an answer 68% significant, and $^+/_-$ 365 one significant at 95%, as might be expected from Fig 2 below. At 68% the range on the coefficient on PTTE is .006 to .02. Everything about the fitted equation, the best fit to the data, is the uncertainty of the results, and their low value, ie that DP recorded high. The extent of the high reading being unknown, that is no help in obtaining MR generally.

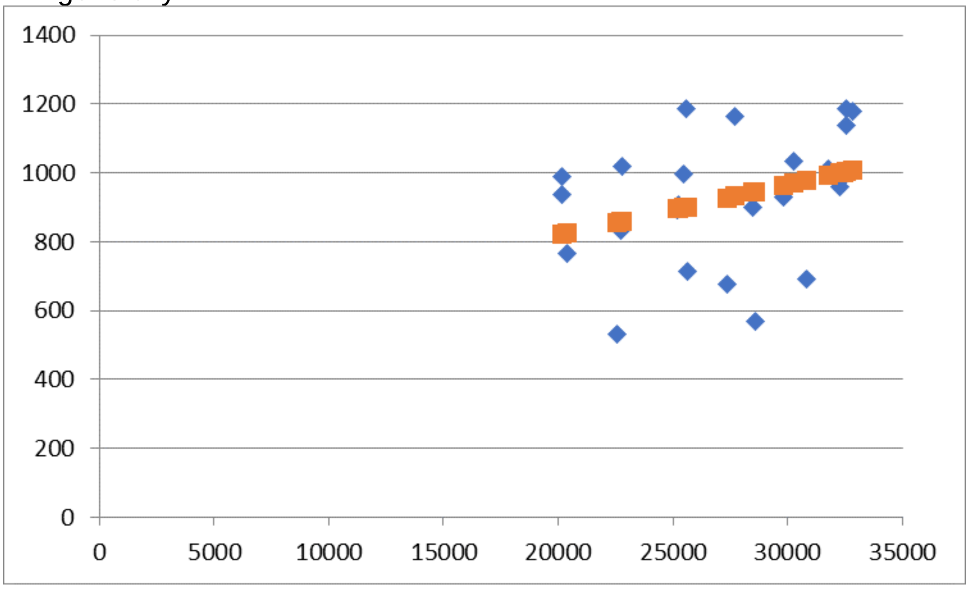


Fig 6 Observed and Fitted TSR (vertical axis) on PTTE (horizontal axis), blue observed, brown fitted by regression

From previous analysis, the expected constants in TSR for a Duchess would be some 228 for CWBVR, $.22V^2$ for speed related items or 550lbs at 50 mph, normal constant of MR 120, total constant about 900, to which .07PTTE has to be added. (Of this, the 228 of CWBR constant is not MR), all much higher than given by equation [1]. the data occur only at high values of PTTE. It is in that range that there is interest in TSR, but use of [1] to give them must be of even less reliability than and there is no professional way of formulating an equation in TSR which gives TSR values for lower values of PTTE.
See also Relating Input to Output, Willans Line approach to Determining MR directly for this engine, below.

6.2 9F 92250

This was the last steam engine tested at Rugby, in 1959. It could be said that procedures should by then have been such that the results were as good as they were going to be. On the other hand, the DP results were still problematical as shown below. The tests of this engine include both a double chimney arrangement and a Giesl ejector exhaust. Both these fittings should have resulted in lower back pressure, and slightly lower MR. Some of the tests involved use of slack coal, to test the ability of the Giesl ejector to allow satisfactory steaming with such. That should not of itself have affected MR. There is the considerable advantage in using the data for this engine because there are 60 observations, 15 at 20 mph, 17 at 30, 12 at 40 and 16 at 50 mph, a reasonable number at each speed for analysis, and for obtaining the effect of speed, although 12 observations at 40 mph is just sufficient. The equations in TSR are:
[2] 227 + .02PTTE at 20 mph.
15 observations, SEE 291, Signf F .02356, t values 0.56 and 1.24, r^2 0.106
[3] -436 + .053 PTTE at 30 mph.
17 observations, 299, 0.523, -0.9, 2.11, 0.23
[4] -1207 + .1246 PTTE at 40 mph.
12 observations, 195, .0058, -1.94, 3.50, 0.55
[5] -2774 + 0.215PTTE at 50 mph.
16 observations, 277, 0.277, 2.23, 1.51, 0.24
The data do not allow sensible explanations of TSR. The constant cannot be negative. The negative constants are compensating for the unduly high coefficients on the PTTE terms, at least at most values.
At all speeds together, ie all 60 observations, for TSR
[6] 433 +.0149 PTTE.
324, .136, 2.22, 1.51 and .038
[7] including V, -422 + .0373 PTTE + 12.17 V.
310, .017, -1.08, 2.87, 2.50 and 0.13
[8] including V^2 as well as V, -1292 + .044 PTTE + 56.9V - 0.61 V^2.
307, 0.174, -1.81, 3.21, 1,81, -1.44, 0.164
[9] $2.62 \times 10^{-10} \times PTTE^{2.64} \times V^{1.35}$.
2.17, 2.69, 2.40 and .12
By solving [6] to [9] inclusive in turn for the four values of V, those equations can be converted to equations similar to [2] to [5]. They are, for TSR, first based on [7]
[2a] at 20 mph -179 + .037 PTTE,
[3a] at 30 mph -57 +.037 PTTE,
[4a] at 40mph 65 + .037 PTTE,
[5a] at 50 mph 187 +.037PTTE
Then based on [8]
[2b] at 20 mph, -130 + .044PTTE
[3b] at 30 mph, 360 + .044PTTE
[4b] at 40 mph 886 + .044PTTE
[5b] at 50 mph 1400 + .044PTTE
Then based on [9]

[2c] at 20 mph 110 $x10^{-10}$ x $PTTE^{2.64}$
[3c] at 30mph 258.5 $x10^{-10}$ x $PTTE^{2.64}$
[4c] at 40 mph 381 $x10^{-10}$ x $PTTE^{2.64}$
[5c] at 50 mph 515 $x10^{-10}$ x $PTTE^{2.64}$
It is obvious that the TSR data do not lead to any sensible explanations of TSR. Even [8] with satisfactory t values, has poor F and r^2 tests. This is not the fault of regression, but of the data The TSR must have a positive constant, at least the 228 lbs or so expected value of the CWBR. [2], [3] and [4]) not only have unacceptable negative constants, but [3] and [4] also have coefficients on PTTE far too high, so the data have characteristics which by having these high coefficients, throw an increasing negative value on to the constants. The (c) set of equations require raising the PTTE to a power of 2.64, then multiplying it by a very small value coefficient, which is not sensible in principle. The t values on the coefficients are in many cases so low that the probability of the values given is too low to be acceptable. The values of r^2 are too low for the equations to be said to explain the data, that after all the relevant forms of analysis have been tried. All the coefficients on PTTE imply variation with PTTE well below the .05 which would be expected from engineering data, and there is no other term in which that friction appears. The lowest PTTE is about 15,000 lbsf, and the highest about 32,000 lbsf, as seen in the Figures, but that should not render the constant negative, and in some cases considerably so. Whatever, the negative coefficients in TSR equations, indicates an extra resistance between ITE and DP of at least 600 lbs. Equations [3]) and [4] for 40 and 50 mph respectively are not sensible at all, the high negative constants and the high coefficients on the PTTE not being credible.
If ITE is regarded as sensible, then the DP is too high generally, and behaves erratically with those features with which it should vary, in engineering terms. In other words, what was recorded at Rugby for DP was not worth recording, even at the end.

5.2 Other 9F

There is sufficient data at least at one speed to analyse the results for some other 9Fs, as follow:
[10] 92013, 1954, 14 observations at 25 mph.
TSR = 639 - .005PTTE.
241, 0.78, 1.79, -0.28, .007
[11] 92166, 1958-59, 15 observations at 30 mph:
TSR = 281 + .047PTTE.
175, .00387, -1.05, 3.5, 0.49
The ITE data for the 9F as a class in all tests combined are consistent, satisfying ITE = $13.24Q^{1.011}V^{-0.84935}$ with a r^2 of .99 and excellent statistical tests. The self-consistency does not mean they are perfectly measured. The poor TSR results have to result from odd behaviour of the constituents of the TSR, with, perhaps, low values of ITE.

The data cannot provide a sensible TSR for any engine of this class. Only [11] approximates what might be expected, and then with such low t value on the constant that it is clear that the vertical location of the curve (ie above zero PTTE line) cannot be fixed. Nothing consistent or conclusive results, the signs are completely inconsistent (those on the constant and the coefficient must both be positive, the statistical tests are almost all very poor, and most of the coefficients on PTTE which are positive are far too low.

5.3 Royal Scot 46165

In the 61 observations for this engine, 13 were at 40 mph and 20 at 50 mph. The remaining observations were at speeds from 20 to 80 mph, in small numbers at each speed. The modest numbers at 40 and larger number at 50 mph were regressed in the same way as for the Duchess and the 9Fs.
[13] At 40 mph, TSR = - 9386 + 0.702 PTTE.
The SEE is 345, the Sigf F 0.418, the t values -0.79 and 0.84, and r^2 0.06.

The high negative constant and high coefficient on PTTE are equally exaggerated, and the constant has the wrong sign. Such an equation simply shows that the data are so poor that an explanation of TSR is not possible.
[14] At 50 mph, TSR = -412 + 0.071 PTTE. The SEE is 309, Signf F .533, the t values -0.25 and 0.63, and r^2 0.02.

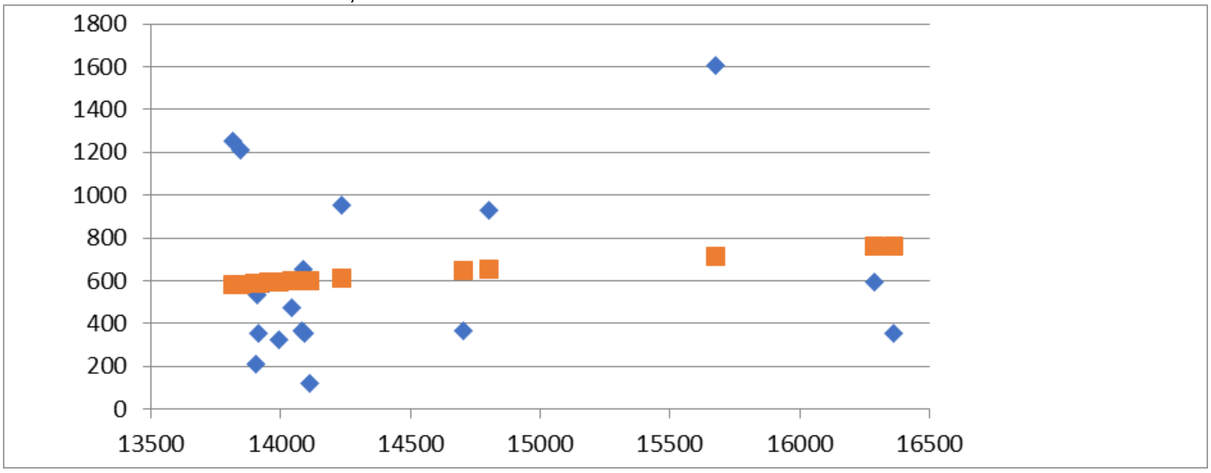


Fig 6 Observed and Fitted TSR (vertical axis) against PTTE, 46165 at 50 mph, blue observed, brown fitted, best fit

The PTTE data occur only in the range of ca 14,000 to 16,500 lbs. It is obvious why no satisfactory equation can be fitted to these data, given the wide dispersion. At ca 14,000 lbs PTTE and one speed (at which so many items are constant), the TSR should be close to constant, yet it is distributed from ca 100 lbs to 1200 lbs. It is low, given that TSR includes the constant for CWBR, and the value at 50 mph of the terms in MR which vary with V^2. It is slightly increasing with PTTE, at about the expected rate (but here with such a low t value that the value of that rate is not at all certain). That slope should continue back to zero PTTE, where it should have a positive constant. The result here of a negative is further indication that the values are all low. It cannot be argued that "something" would cause the fitted TSMR line to rise as it is projected back to zero PTTE, something not present in the data, or allowed for in the equation. That cannot be: as above there should be a considerable positive constant, all V and V^2 effects should be in the constant, and it is logical for the rate of variation of TSR with PTTE to be much the same at lower PTTEs as at higher.
Across all speeds, a regression of TSR against PTTE and V^2 gives a result of

[15] TSR = -376 + .074 PTTE – 0.165 V^2,

with the coefficient on PTTE significant at the 95% level of confidence. This is a better equation than those at 40 and 50 mph, SigF .0003, SEE 633, t values -0.2, 2.29 and -1.02 but with an r^2 of only 0.29. But a negative constant when the CWBR constant is 150 lbsf, and the negative coefficient on the V^2 term show that no relationships based on the technical first principles of MR emerge from these data.

There is no obvious pattern to the residuals. No interpretation can be placed on these results. The even effect of TF forces in a three cylinder engine with cylinders in line does not exist on this engine because the outside cylinders drive on to the second coupled axle, while the inside cylinder is forward, and drives on to the leading coupled axle, but that cannot explain the enormous negative constant.

5.4 Jubilee 45722

This engine was tested in 1956-57. There were 18 tests at 35 mph and 25 at 50 mph. The regression results were:

[16] 35 mph, TSR = -193 + .068 PTTE.
271, 0.143, -0.3, 1.54 and .13
[17] 50 mph, TSR = - 866 + .112 PTTE.
254, 0.316, -0.63, 1.02, .04

The same conclusions as drawn for the 9F apply in this case.

5.5 Standard 5 73030

There were 12 observations at 55 mph. The regression result was:
[18] TSR = -523 + .097PTTE.
343, 0.473, -0.34, 0.74 and .05.
The same conclusions as drawn for the 9F apply in this case.

Of course, it is possible to say that a negative constants are impossible, it is he absence of data below the observed values which are the reason for both the negative constants and high coefficients on the PTTE. That could of course be true had tests been conducted at lower efforts, but such data do not exist, and imposing values which make the data appear better, and at the same time removing the above deficiencies is not scientific. Further, the composition of the constant and PTTE terms are such that they should capture variation right down to low but positive values, ie the data should have such behaviour in it if the data were satisfactory. Further, it is at high values of PTTE that data will be observed because the experiments were conducted at outputs of interest to those testing the engines, and it is for values n about the same range for which TSR and LR will be needed. Whatever might be thought about the constants, the coefficients on PTTE cannot be judged other than being far too high.

Is the assumption that the relationship with PTTE is linear justified? I have not yet tested that, but do not expect any change in the conclusions.

7 Other Notes on Rugby Results

Some effort was devoted to the data for all classes across the whole speed range. Apart from finding consistency in the basis of ITE, no results of use emerged. In addition, the average MR of each class which emerged, MR here being TSR less CWBR, was analysed, with the following results:

Class	Average TSR lbs (a)	Calculated Constant of CWVBR lbs	Average recorded TSR lbs	In speed range, mph
9F	542	228	314	36 – 60
Duchess	953	227	726	50 - 85
Standard 5	640	151	489	45 – 75
Jubilee	681	150	531	50 - 85
Royal Scot	586	150	436	50 – 85
Crab	642	169	473	40 - 70

Fig 8 Comparison of Observed Average Apparent Resistances at Rugby for Five Classes
average TSR hides any variation with V^2, or more generally $(rpm)^2$. It differs from MR by CWBR

The Jubilee and Royal Scot differ mechanically essentially only in cylinder diameter. The latter has the larger diameter, with more circumference of piston rings to slide on the cylinder walls. Yet the average TSR of the Scot in the Rugby data is 18% lower than that of the Jubilee. The Crab and Standard 5 TSRs are also out of line. The Crab should have a higher average resistance than the 5, partly on account of its smaller CWs, partly on account of its bigger cylinder diameter. In that case, however, the lower pressures on the rings of the Crab will affect the comparison.

The average MRs for these engines are very low for the sizes of the engines, generally. Whatever might be considered about anything I have calculated, the correct average MR of the 5 of 489 is very low. The standard 5 should have much the same average MR as the Black 5 – its slightly larger cylinders are roughly balanced by its slightly larger CWs – instead of less than half. For an engine with such small CWs, the TSR of the 9F is very low. The third is that it cannot reasonably be expected that the MR should be constant over all outputs and speeds.

The results for the TSR regressions, however, are overwhelmingly disappointing, in terms of sense (ie behaviour and signs) and magnitudes, with wide standard errors of the estimate, low t scores on coefficients, high significance F values, and values of r^2 as low as 0.1. Neither the equation chosen, nor the basis of the analysis (regression) nor its application in this case, is at fault, it is the poor, inconsistent data. Further, given the remarks above about the ITE data being generally consistent when regressed against Q and V in ln form, while not necessarily accurate, (they appear a bit low when tested by the Perform program), the erratic TSR must therefore be the result of the erratic components of TSR or TSR as a whole (and that accepts that the DP measurement is accurate). With these results, no confidence can be placed in the Rugby ITE – DP (TSR) data and results for obtaining MR. Even where the constant and the coefficient are sensible, by sign and magnitude, the standard errors of the estimate are so high that the mean value is reduced to negative if two SDs are deducted from the mean.

The hypothesis can be put forward that the rapid to and fro movement on the Rugby plant distorted the results even after 1955. That fits with Chapelon's view that two-cylinder simple engines needed to be balanced to some 95% of the reciprocating masses to give acceptable results. At Rugby, a little extra reciprocating balance was added to a couple of classes where the proportion of reciprocating masses balanced was lower than average on some engines, but not all, and not to the extent of 95% suggested by Chapelon. Chapelon did not remark so far as I am aware about the balance of three and four-cylinder simple engines, but given different connecting rod lengths and drive on to different axles, they would have required reciprocating balance (GWR four cylinder engines had such), leaving some on a particular axle well below 100%, and subject to the same considerations as two cylinder engines. Or an hypothesis might be put forward that the to and fro forces were having a distorting effect, as implied in Chapelon's writings, but the origin thereof needs further thinking. Whatever, any TSR value will be subject to the SDP.

[Chapelon said quite clearly in five places that two cylinder engines did not give satisfactory results on testing stations on account of the recoil effect of the two and fro forces. (The sources for that are the Chapelon and Sauvage book *La Locomotive à Vapeur*, 1979 reprint, Section 77; his own book *La Locomotive à Vapeur*, 1935 edition, p 832; his 1952 paper *Conférences sur la Locomotive à Vapeur prononcées en Amérique du Sud* in 1952; and his comment p 137 of the Carling 1972/3 paper on Locomotive Testing Stations (Newcomen Society, Institution of Mechanical Engineers). He states that accurate answers for such locomotives on test plants required them to have 95% of the reciprocating masses balanced, which did not happen at Rugby. Note too that Carling did not explain why alterations to the plant in 1953 made the answers correct. I do not know any more about Chapelon's experience leading to these views. Further, the real problem which design and practice at testing stations in both France and the UK was avoiding resonant forces damaging the plant or its components, rather than achieving accuracy.]

Adrian Tester, who wrote a series of articles in *Backtrack* Vol 27 2013, about stationary testing plants, has informed me (personal communication) that Carling, superintendent of the plant, noted that the Amsler could record to +/- 1% for pull, and provided data within a +/- 1½% range for work done and +/- 2½% range for power (these are presumably at its own recording table, as might be expected from what these terms represent and the accuracy of the components. Only the pull, however, was recorded.

7 Relating Input to Output, Willans Line approach to Determining MR Directly, ITE made dependent on DP for 9F 92250 and Duchess 46225

Some Rugby data have been further analysed to test the idea that relating input to output can reveal the internal resistance between the input and output, in this case ITE to DP. This is not in terms of Q to DP, because on a steam locomotive, Q is first converted to ITE, and it is the relationship of ITE to DP which reveals TSR as used in this paper. As ITE is the independent variable in a relationship between ITE and DP, this case, performing a regression of ITE on DP is "back to front" in terms of the usual analysis based on cause and effect. The result is TSR, from which CWBR has to be deducted to give MR. For a 9F, CWBR by calculation is about 229 lbs.

The article by S J Pacherness, *A Closer Look at the Willans Line*, paper 690182, Society of Automotive Engineers, International Automotive Engineering Congress, January 1969, explains the underlying idea. If fuel is graphed as dependent linear variable against brake output of an internal combustion engine at a particular speed, as an increasing function, and projected back beyond the fuel line, the point where the graph line cuts the DP line, at zero fuel consumption, which occurs in the negative range of DP, represents, with the sign changed to positive, an approximation to the internal resistance of the motor. The slope of the line at any point is the specific rate of conversion of fuel to DP. If the graphed line at a particular speed is clearly a curve, ie Q is an increasing function or power function of DP, the tangent to the curve at any point projected back in the same way as the linear graph gives an approximation to the internal resistance of the motor at that speed and rate of working, and the slope to the tangent gives the specific fuel consumption at that speed and rate of working. Consistent derivatives can also be graphed. The fitting of the graph should be a regression in each case, but that is not said. In the automotive engine, the "friction" will include pumping losses and blowby. To result in correct MR, the engine must be working as it would be in use, and not be turned over by an external device. Numerous tests are said in the paper to give internal combustion MR of 6 to 8 psi.

For 9F 92250, using linear equations for each speed, this method yields an MR at 20 mph of 104, and at 40 mph of 18. At 30 and 50 mph, the constants in the relationship between ITE and DP are negative, which makes the method inoperative. All four equations, those for each speed, have excellent test statistics except that all have a low t score on the constant, which in turn leads to a high SEE, and inability to fix the location of the curve with any certainty.

For Duchess 46225, the equation to test this has been estimated in both linear and curved (power) forms (lnITE on lnDP).

A linear equation of ITE on DP is good statistically, ITE = 683.4 + 1.0199DP, signf F 3.08E-29, t on constant 4.51 and on coefficient 85.4, r^2 .997, standard error of the estimate 186.5. This results in a negative DP of –670 when ITE is zero. As there is a constant slope to the fitted line, that means MR + CWBR is 675 at all outputs at 50 mph, or MR alone is 446 lbs. Such constancy at all outputs should not be the case. The linear fit is based on observations of DP between 7373 and 17,085.

The curved form is ln ITE = c + b(lnDP), regressed on the 50mph data in ln form,
ln ITE = .650182 + 0.938868 ln DP, or ITE = $1.9159DP^{0.938868}$ (a)
This is statistically a good equation, signf F 2.08E-28, t on constant 5.776 and on coefficient 78.3, r^2 .996. When DP is 0, ITE is 5, reflecting the problem of ln for 0 and 1. The differentiation of the curve to give the slope (dITE/dDP) reduces to 1.9159 x 0.938868 DP^-.061132, or 1.7988/DP^-.061132. The following shows the steps in obtaining MR for three trial values of DP within the data range at 50 mph:

DP lbs	Equival-ent ITE lbs (a)	DP^-0.061132	Slope dITE/dDP (b)	Equivalent horiznl of DP = ITE/slope	MR+CWBR =(equiv hornzl of DP – DP) lbs	MR lbs
7,000	7806.7	0.5820	1.047	7456	456	228
10,000	10,910.6	0.5695	1.0248	10,647	647	419
16,000	16,962.3	0.5533	.9954	17,041	1041	813

by above equation (a), ITE = $1.9159DP^{0.938868}$
previous column x 1.7988
This method, although it can be applied, yields MR values which, by other analysis, are too low, by a considerable margin, in this case because DP measurement is not satisfactory. The slope of the curve of ITE on DP is too steep. In the range in which DP measurements occur, they are therefore too low, consistent with the conclusions above about the high values of DP and consequent low TSR.

8 Discussion

The Rugby data have been analysed in various ways, indeed all possible ways which reveal whether it is satisfactory, and what relationships are present in it, and in all cases, they are found wanting, being erratic and low for the circumstances. The bases of these conclusions have already been explained.

They only possible explanation of why that is so is the measurement system at Rugby. Scatter is unavoidable in investigations like these. But that does not mean that any old scatter is acceptable. The equations fitted are best fits – in probabilistic terms, nothing better is possible. If data are close to an expected or sensible relationship in physical terms, there will be a high probability that the relationships found are acceptable. Widely dispersed data do not do that. Scatter may be inevitable, but the more of it, or the less rational it is, the worse for the investigation. Just because data exist does not mean that they will be useful, and if they diverge widely from the expected relationship, and that relationship is correct, the worse for the data. All figures above show by observation that the data are far from conforming to the expected relationship; indeed, it has been a waste examining them as far as has been done here. They cannot be tampered with to “raise” them, by saying they are low only by reason of scatter. The experimental data have to be the basis of the investigation, not tampered figures.

To reemphasise, only the second approach depends at all on other analyses I have made, hence on my judgements of friction coefficients. All other of the four approaches rely on the data speaking for themselves.

There are several consequential comments.

Before I learned of the problems with the Rugby data, by examination and analysis, I earnestly hoped to obtain MR/TSR data from the tests done there, which data could be analysed to give reliable values for MR. I spent time at the NRM in 1988 extracting Rugby test data and reading files on the operations of the plant. I spent time since fruitlessly analysing those data, and from time to time testing out some new relationship or analytical approach to redeem those data, to no avail.

Doug Landau resolutely refused to declare is his letter what he does himself with Rugby data to convert it to MR, or LR, both for locomotives tested at Rugby, and for engines not tested there. Why the refusal, the sidestepping of the issue? What is he hiding? For this discussion to have been useful, other readers and I need to be informed how he drew conclusions on the Rugby data. Further, did he test the data, examine its soundness? How did he treat the SDP?

I shall write further comments on his letter, the great men approach and other, in due course.

10 Conclusion

Much of what Doug Landau has said about my previous letter amounts to unsupported declaration without analysis, tests or support. By example of what analyses he does, he is obviously not in a position to make these declarations. I consider his approach unscientific. I also consider his writing conclusions to his paper which have no relationship to the content of the paper to be dishonest. His motives for doing that are obviously not the purest. But I suspect they are to impress readers that his (unexplained) approach to obtaining TSR from the Rugby data is the correct one, and that such TSR values are good, and to deter readers from considering the matters put forward by John Knowles to be right.

Doug Landau's comments on regression are unsound, the fears expressed about consequence of its use groundless. They were made without any explanation of what are the supposed consequences of its use. Rather, regression is an essential tool in the analysis and explanation of experimental data. No one would now present a scientific paper relating to matters numerical, even less have it accepted for publication, without sections on examining or testing the data, and analysing the data for possible relationships in them, which analysis would be performed by regression (or similar). There are internationally accepted bases for drawing conclusions about the soundness of the results of analyses. It is not acceptable to deviate from them, and prefer Doug Landau's own (again unexplained), his attempt to justify use of data which to anyone else is not useful.

Given Doug Landau's stout defence of the Rugby data, his unwillingness to say how he uses the Rugby data to obtain MR and LR of locomotives generally is inexcusable in a scientific context.

What is said about repeatability of the Rugby test data is wrong and misleading. As is his comment that a key test of scientific proof is that its claims are consistent with the empirical evidence, when he accepts the empirical evidence without question or test.

From tests of the apparent soundness of the data, and relationships fitted the Rugby results on TSR are erratic, incapable of explaining the origin of TSR and on the low side.

I did and do not say that everything done at Rugby, its designers and operators, had shortcomings. That is Doug Landau putting words into my mouth. I certainly think that the measurement of the DP had shortcomings – it is hard to see otherwise. It might not have been possible for them to do better. By Carling's admission, elimination of the problems at Rugby would have required complete redesign of the plant, which was not done. At least Carling was clear that TSR at Rugby was not a sound measure of the internal resistance of the locomotive. It is also noteworthy that the DP data at Rugby were never published. What is known was obtained by me and a few others taking out the data at the NRM. I also think there were shortcomings in the whole philosophy and system of testing, even to having a Testing Station when there were the Mobile Testing Units, but especially the way of apparently or supposedly ensuring Q was a certain level on both the testing station and on the road. This is not the place to go into detail on these matters, but the same designers and operators, especially he senior ones, who Doug Landau is reluctant to see criticised, were involved in both the plant and road tests to some extent. But that criticism of Rugby is groundless is not right.

Last, this drawn out, often bad tempered, discussion on steam locomotive resistance has followed from a letter I wrote about the correctness of the second term in the usual formulae for railway vehicle resistance. It was Doug Landau who changed the subject to Steam Locomotive Resistance. Why did he do that? I find his motives questionable. In one sense, answering his erroneous notions is a waste of time, in another, it is useful if I can correct some of his ideas (as above). But no more than that. In my view he has not advanced the subject of steam locomotive resistance in these letters one jot. Overall, it would be better if this discussion were conducted in a peer reviewed scientific journal. For that to happen, he would need to learn about testing numerical data and scientific methods of analysing it.

Abbreviations

BPP	Blast Pipe (or Nozzle) Pressure.
BR	Braking Resistance
Cf	Coefficient of Friction
CO	Cut Off
CWVBR	Coupled Wheel Vehicle Bearing Resistance, without the wheels being powered
DBP	Drawbar pull (ontesting station)
DP	Dynamometer Pull
DR	Damping Resistance
EDBTE	Equivalent (to running on level track) Drawbar Tractive Effort
ITE	Indicated Tractive Effort
IHP	Indicated Horsepower
ln	In terms of Naperian logarithms
LR	Locomotive Resistance, basically VR plus MR
MR	Machinery Resistance, including the addition to CWVBR from the CWs being powered
PTTES	Piston Thrust Tractive Effort propulsive and compressive
$PTTEV^2$	Piston Thrust Tractive Effort forces from unbalanced reciprocating masses, dependent on speed squared
PTTE	The (net) sum of PTTES and $PTTEV^2$
Q	Steam Rate lbs per hour
SSC	Specific Steam Consumption, Q per Indicated Horsepower Hour
SDP	Small Difference Problem, as exists between two large numbers often or usually preventing exact measurement of the difference
SHM	Simple Harmonic Motion
SSC	Specific Steam Consumption (lbs per IHP hour)
TF	To and Fro (or Fore-and-Aft) Forces

TSR	Testing Station Resistance (ITE – DP)
V	Speed, mph
VR	Vehicle resistance
WRTE	Tractive effort (normal definition, cf PTTE) at coupled wheel rims
WRHP	WRTE as a HP

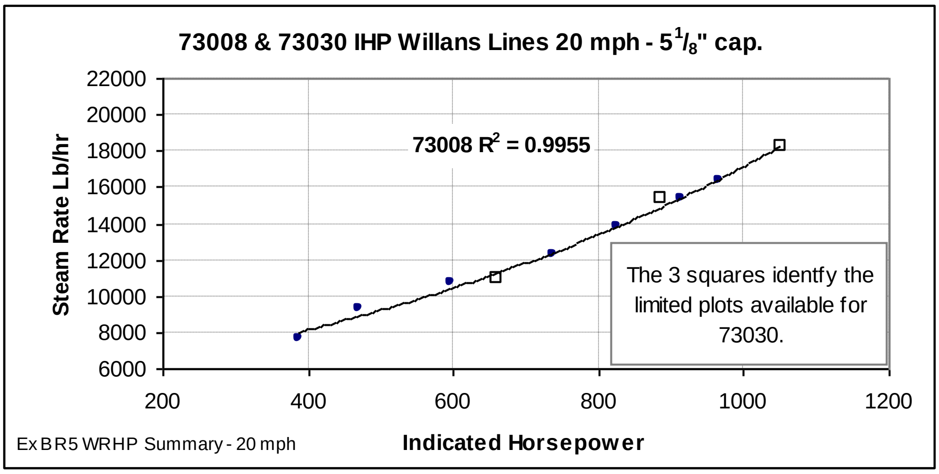
Descriptions of statistical tests are not given. (Standard Error of the Estimate, Significance F, t, r^2, Standard Deviation) can be found in Statistics texts.

John Knowles

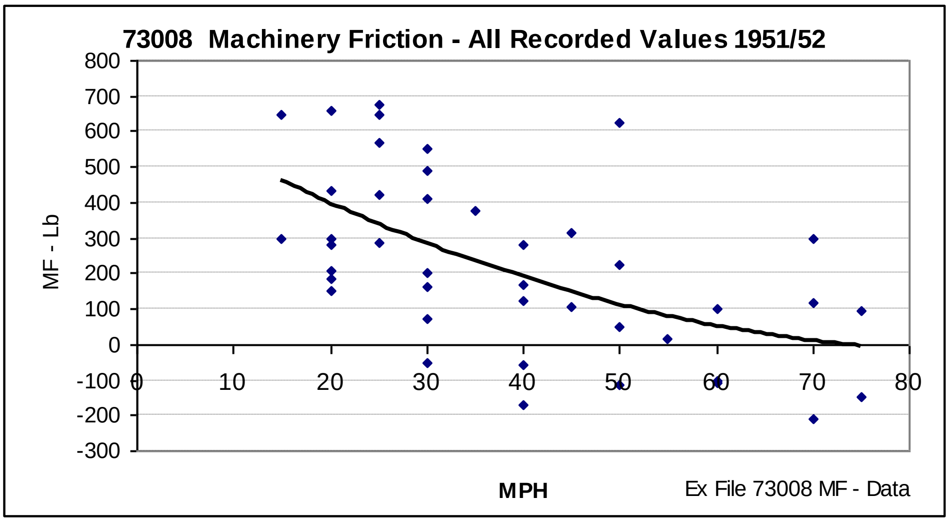
4th July 2017

Locomotive Resistance - 7 July

I've recently identified a serious plotting error in my letter 14th April. This concerned the graph comparing the indicated horsepower data plots for BR5s 73008 and 73030 at 20 mph when fitted with $5^1/_8$" blastpipe caps. Two of the three plots shown for 73030 were erroneous, misidentified data having been entered. I should have been suspicious at the time since the separation of the two data sets was more than might be expected. Entering the corrected data, as below, and contrary to the original outcome, it shows no separation of the two data sets beyond normal scatter.

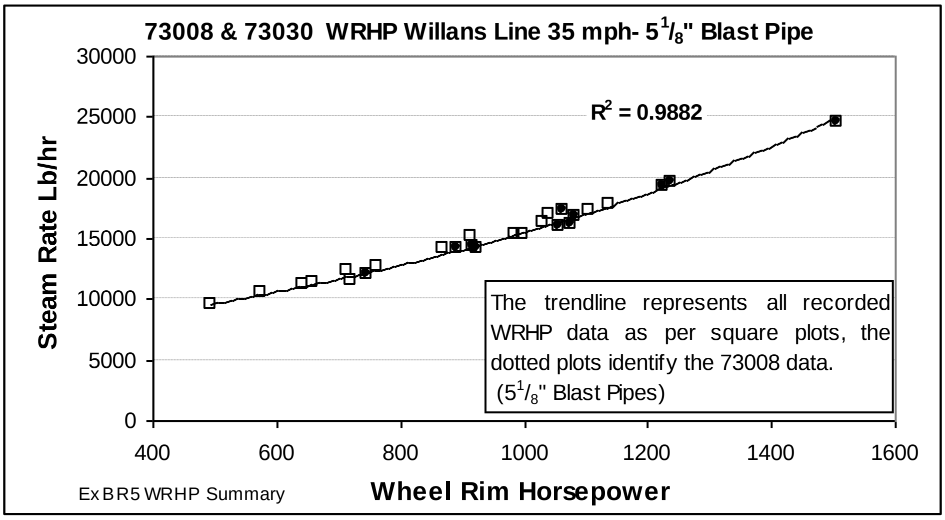


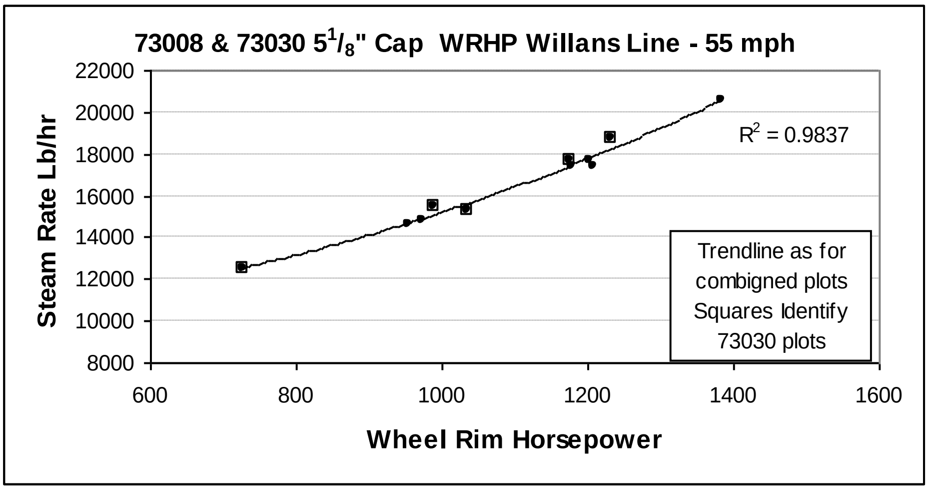
The available IHP data at higher speeds for 73008 and 73030 when fitted with a $5^1/8$" blastpipe was only coincident at 35, 55, and 70 mph, and such it is was very meagre, respectively amounting to no more than 4, 4, and 5 IHP plots *in total for the two engines*: insufficient to support any comparative plots. The 73008 tests took place when negative MF data was still being encountered with undue frequency. This tendency increased markedly with rising speed as plotted below.



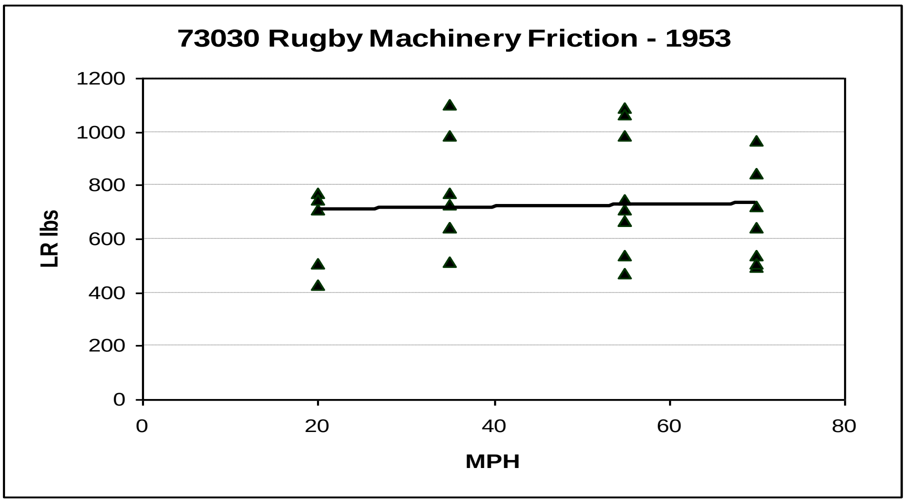
The incidence of negative MF outcomes clearly increases as a function of speed. Merchant Navy class 35022 showed similar traits, although the slope was less marked, the magnitude and frequency of negative outcomes was greater.

The available WRHP data at higher speeds for 73008 and fitted with the $5^1/_8$" blastpipe cap is sufficient for plotting Willans Lines, as in the two examples below for 35 and 55 mph. The recorded data is consistent across the two-test series.





The MF data scatter diagram for 73030, as below shows a dramatic improvement; negative MF values have been wholly eliminated. The plots shown include the data for all three blast pipe caps tested. The trend line shown is virtually constant, at about 725 lb. Such an outcome compares with the shallow dish shaped trend lines generated by 42725, 45722, 46165, and 46225. Such outcomes are to some extent down to the chance influence of the scatter pattern. As the example below shows, the speed groupings may develop an upward or downward bias, in this instance the latter at 20 mph.



This concludes what is essentially a corrective note, plus little supplementary information. I see John Knowles has submitted another letter a few days ago, 4th July. In due course I will have a look at it, but it will be some time before I do so. Among other projects, I am currently busy putting together, what will, inter alia, form a definitive vindication of the Amsler dynamometer at the Rugby test plant.

Regards,

Doug Landau

From Doug Landau – October 2017

Locomotive Resistance

This is just an interim note to report research on the Rugby Test Station NRM archive in late September. The programme I set myself for the day proved over ambitious, and much of the material I had requested went untouched.

My key interest was the chronology and record of events during the commissioning and early working up phases of the test plant1949- 50. It was not until quite late in the afternoon that some key material sufficient for the objective was discovered, but much important material not related to the Amsler dynamometer had to be skipped over as time ran out. Certain key dates were however established. Below is a brief summary of the record.

The initial commissioning of plant with WD 2-10-0 73799 commenced on 26 November 1948. Initially only 10 test runs were completed. It is unlikely any serious testing occurred during this phase, more a case of finding out how and if everything worked, so I did not trace this far back in the record. Some indicating tests with Caprotti Black 5 44752 followed before 73799 returned for a further 20 tests, bringing the plant test runs total to 50 on 13th April 1949. The replacement for the “old bag of bones” was another WD 2-10-0, 73788, making its first test run on 22nd April 1949, completing just three test runs before the first of three interruptions for D49 4-4-0 62764 indicating tests of the Reidinger poppet valve gear. These breaks were probably to undertake modifications of the dashpot damper system, of which there were many. Eventually 73788 completed 46 test runs on the plant, the last, run 144, was on 19th December 1949. The two intermediate test sequences both lasted for only 3 test runs, as had the initial tests. It seems probable that on all three occasions it was quickly established that it was a case of “back to the drawing board” in regard to the damper modifications.

At this period Carling was writing progress reports to the railway executive on a weekly basis, and the ‘Damping Dashpot Investigation’ was a hot topic; because of pending modifications he sometimes had to report “in abeyance”. In a letter 21 March 1949, which coincides with 73799’s final stint on the test plant, Carling reports; “the dashpot can increase drawbar pull 100%.” By the time 73788 was on the plant, some modifications to the dashpot appear to have met with a modicum of success; writing on 27 April 1949, Carling was able to report “error approximately halved.” Not good enough however, it was probably the last of the three tests completed in 10 working days. The dashpot was first tested drained of oil on 4th November 1949, details of the run notes: ”Run made with dashpots drained of oil (Run 126), in order to investigate amount of oscillation and to obtain values of drawbar pull unaffected by dashpots.” Writing to the Railway Executive on the 7th November, Carling reports; “There is now no reasonable doubt that differences of oil pressure in the dashpots account for the whole of the falsification of the record of drawbar pull on the Amsler table. A special test was carried out on Friday afternoon when the dashpots had been emptied of oil preparatory to fitting the new type of damping control, which is promised for delivery on the 7th November. This test was intended to explore the possibility of in the manner believed to be used at Vitry, i.e. with no dashpots in action. It was found the locomotive oscillations were very severe at 3 or 4 miles per hour, but became quite reasonable at high speeds of 45, 40, 35 and 30 miles per hour. The locomotive was behaving quite satisfactorily as far as oscillation was concerned at 25 miles per hour but before a test could be finished slipping occurred and before the speed could be steadied the blowing of a fuse in the electrical control circuits prevented completion of the test.”

“It had been expected that it would have been possible to run the locomotive at a speed as low as 20 miles per hour, but not much below this figure, as the calculated critical speed with the present number of Bellville washers in the drawgear is 12 miles per hour.”

“The outcome of this test is an indication that it should be quite feasible to run a Class 5 4-6-0 on the plant without using dashpots at speeds of 25 miles per hour and upwards. It is possible that, by reducing the number Belleville washers, a run at a speed below the critical for that locomotive and spring combination could be achieved, thus completing the speed range down to slightly below15 miles per hour, which is the slowest speed at which this class of locomotive can be run on the plant at full power.”

The next locomotive on the plant was Black 5 45218. Writing on 23rd January 1950, Carling was able to report:

“Tests with 4-6-0 L.M.R Class 5 Locomotive 45218”

“It has been definitely established that this locomotive can be run on the test plant at all speeds without oil in the damping dashpots. The locomotive has now been thoroughly run in and testing up to any speed desired will commence next week.”

By the time of this development, the dashpot problem had been passed to the research department at Derby, while some of the modifications and correcting some imbalance in the system had brought about a reduction in amplification of the drawbar pull, it seemed impossible to eliminate. Experiments with different types of oil and reducing the friction had no effect. The dashpot was manufactured by Heenan and Froude; I was surprised to find it incorporated a pump, having previously imagined is was a simple displacement device. The pump pressurisation was adjustable, in the examples seen it was ‘set’ at 15lb/sq.in (‘nominal’). On Run No. 130 11th November 1949, the pump was shut off for the 40 and 45 mph tests, resulting in an increase in the drawbar pull discrepancy.

Other points of interest gleaned from the NRM are listed below.

The mediating mechanism gear ratio was reduced by a factor of about 3 sometime in 1950. As first installed it was overactive, and subject to excessive wear. It was further reduced in 1953 by a similar amount, bringing the ratio down to about one 10th of the original provision.

The dynamometer integrating mechanism was refurbished at the back end of 1953.

The ‘Summary of Improvements to Plant Equipment in 1953’ lists 13 items ranging from a milling machine safety guard to a Marine type clock for the firing platform. The changes to the mediating gear referred to above are listed along with improvements to thermocouples, the manometer bank, and the Farnboro’ Indicator diagram converter. The Amsler pump motor was replaced.

The summary list for improvements in 1954 could only be briefly examined. Of the 20 or so items listed, many, such as improved mess room facilities and data storage racks, were not relevant to technical matters. Of interest were roller scrapers to stop slipping; a new improved spark generator “much improved” Farnbro’ indicator elements (July); an exhaust injector flow meter installed; and dead weight testing for pressure gauges;

The files contained many original worksheets, such as a plot of Bellville washer deflection and hysteresis characteristics; the latter effect was low, the washers being arranged in a set of opposing single pairs. The results of a routine static dynamometer load test on the 36.000lb scale in 1953 found errors ranging from -0.34 to - 0.7%, averaging -0.57%. On the 12,000lb scale there was 1.87% error (112lb) at a pull of 6,000 lb; at a pull of 12,000 lb the error had fallen to 15lb. 0.125%.

It was apparent the test plant underwent continuous development and improve- ment.

My promised “simple proof” of the Amsler dynamometer is almost finished, but completion will have to wait a while yet, pending attention to some late running commitments. The time taken so far is not for the basis of the proof, which is very simple, but extracting supporting empirical evidence from the highly suspect DBHP data contained in the BR test bulletins for the locomotives tested at Rugby is another matter. These suspicions are not my invention, for as Report L116 clearly states: “In all cases where locomotive trials at Rugby have been followed by road tests carried out with the LMR Mobile Test Plant there has been a lack of reconciliation of the results to the extent that values of locomotive resistance obtained by subtracting Drawbar TE from Rugby Cylinder TE have not been acceptable.” These shortcomings were attributable to a failure to control steam rates to the nominal values set for the road tests. L116 report gives some guidance in regard to correcting the drawbar data for the 9F, but none whatever for the BR5 and Britannia. Only report R13 for the Duchess has corrected DBHP data as derived from Report L109. In this instance the ‘simple proof’ and the empirical evidence are in close accord.

Doug Landau

A DEFECTIVE APPROACH IN UK STEAM LOCOMOTIVE TESTING

A RECONCILIATION OF TEST RESULTS WHICH DID NOT SUCCEED, INTERNAL REPORT L116 OF DECEMBER 1957

John Knowles

In his letter to Milepost of 17.3.17, Doug Landau claimed that those BR officers conducting measurements and research into locomotive efficiency and outputs were scientists. I believe that those who worked at Derby, the Testing Section of the London Midland Region, were anything but scientists, that their work was anything but scientific, a lot mistaken. It was their function to conduct Controlled Road Tests to obtain figures for EDBTE consistent with the Rugby ITE results, and for years on end they produced erroneous results.

Abbreviations and Explanations:

ITE	Indicated Tractive Effort
EDBTE	Drawbar Tractive Effort made Equivalent to the Train Running on Level Track by correction for effects of gravity (gradient), and acc/deceleration. Omission of the E implies Drawbar Tractive Effort, that measured at the drawbar without the rendition of the figures to allow for gradient and acc/deceleration
MTU	Mobile Testing Unit, a vehicle with rheostatic brakes and control over the extent of the braking effect, and control to keep speed constant.
CRT	Controlled Road Testing of Locomotives on the Road, in contrast to the stationary Testing as on the Rugby plant, with devices to measure coal and water consumption, and instrumentation to advise the BPP to the driver, who can alter BPP by altering the CO of the locomotive. The locomotive can be equipped with indicating gear, but the advice given to the driver of the BPP is meant to avoid indicating. See S O Ell, Developments in Locomotive Testing, JILE, Paper 527, 1953 p 561.
BPP	Pressure of steam as steam is exhausted at the Blast Pipe, referred to Pressure absolute (14.6 lbs/sq inch higher than atmospheric.)
Q	flow of steam at a certain temperature and pressure, lbs per hour

LR	Locomotive Resistance

Despite Doug Landau's staunch defence of British testing and its numerical results, there was a large scale defect in one aspect of the UK approach to testing which led to incorrect EDBTE results being declared for many locomotives. It existed throughout the period of testing. Several Test Bulletins have incorrect EDBTE results and were never corrected. The defects, present during the whole period of testing, were eventually acknowledged in an internal report, L116, by the testing officers themselves. Many defects in procedure which probably led to the defective answers were also pointed out by various testing officers.

Background
For some seven years or so, some locomotives were jointly tested by Rugby and Derby, ie by the Testing Station at Rugby and by the Testing Section of the London Midland Region at Derby. The latter conducted Controlled Road Tests intended to obtain figures for EDBTE consistent with the Rugby tests, which were entirely in ITE. From the inception of these tests, it was found that the results of these road tests were inconsistent with the Rugby results, a problem of method, and/or measurement, and/or calculation of results from the measurements. This inconsistency was observed through LRs of the wrong shape, indeed impossible shapes. That meant there were errors in production and/or measurement of ITE and/or EDBTE, or calculation of EDBTE.

Only at the very end of steam testing was "something" done about this defect, and a method devised intended to correct the recent results. This depended on inserting speed terms into the relationship between Q and BPP, even though there was no dependence on V in the relationship between Q and BPP (as I show below), and deriving a correction equation, which was in fact an erroneous method of relating Q to BPP. The correction method was therefore muddle headed thinking¸ with no science in it. Having such a correction system would or might appear to make the answers they gave at the end of the testing "all right then", while leaving a defect still present in the data published in the Test Bulletins which were the joint responsibility of Rugby and Derby, without any public admission of the defect or correction of the results of the testing. As importantly, there was really no valid correction system at all.
The internal documents concerned which are the basis of my conclusions (L109, R13 and L116), all prepared at the very end of steam testing, claim, however, that the correction system did convert LRs of the wrong shape into the correct shape. It is impossible to draw that conclusion from the fullest description of the correction mechanism or process. No data were given on the cases where the supposed correction led to the correct LR, ie the original data, the basis of correction, and the results of the supposed correction, and it not obvious that the corrections can be checked. Further, the judgement made, that the system worked, required a comparator, ie consistent ITE and EDBTE at various speeds for the locomotive under consideration. For the locomotives concerned, no such comparator LRs are known. The conclusion that the system worked is therefore without foundation.

This paper

This paper first considers the large number of wrong results, admitted in internal report L116. It then goes on to consider how those incorrect results could have arisen, and the modest research conducted with the intention of allowing the incorrect results to be corrected, research which was extremely poorly applied. The officers concerned considered that their results were wrong because they had not taken into account the effect of speed on the use of the Blast Pipe Pressure on the metering of steam. In that they were mistaken, for there was no such speed effect. The correcting mechanism and equation thy devised did not fit the data available, which led to wrong conclusions. They believed that they could conduct

desktop corrections of results, but in that they were mistaken also, and no explained corrections of results was given. Nor did they perfect the testing and measurement, and to the end the Derby measurements of ITE proved defective, including that of a Duchess. Although Derby thought it had a system which could correct LR, it never explained where the comparator locomotive came from. Checks were made of the apparatus and procedure, but the Derby errors were never corrected. This is surprising because testing procedure with similar intentions took place at Swindon and seemed to operate satisfactorily – it was Derby which did not succeed in measuring properly, and which devised a supposed correcting mechanism which was not a logical explanation for the mismeasurement which occurred.
The data available is analysed herein much more soundly than done for L116. See below.
Derby did not run its side of the joint Rugby – Derby testing soundly.

The Intended Measurement System

LR is the difference between ITE and EDBTE. If the testing method was to reveal LR those two items are needed, correctly measured. (Lots of locomotive testing in the world, probably most, did not seek to reveal LR.) The BR Test Bulletins all include data on ITE and EDBTE separately, the ITE being the end product of the boiler and cylinder outputs, and the EDBTE the work the engine can do at its drawbar, measured there by the dynamometer car.
Testing should preferably be done on the road, where the engine will operate, and where the draft on the fire and the escape of the exhaust are those of the open air and where there can be evidence of the inevitable variations in atmospheric conditions. Postwar, the British testing system had the Mobile Testing Units, which could apply a rheostatic brake to achieve constant speed, vastly better than use of steam locomotives with the cylinders acting as counterpressure brakes.
The aims and methods of the BR testing system were given in a paper by S O Ell, *Developments in Locomotive Testing*, read to the Institution of Locomotive Engineers, paper 527, 1953-54, and in BR Test Bulletin No. 1, 1951, on the testing of the Western Region Hall.
Whatever the system and measurement, LR even at a continuous output, can be expected to show uncertainty and inaccuracy on account of the Small Difference Effect I have discussed previously, the result of ITE and EDBTE being both large numbers, the values of which cannot be measured precisely, resulting in considerable imprecision in the LR, the modest difference between them. The problems I discuss here concern conceptual errors mostly, errors in approach.

To minimise measurement difficulties, it would be usual to measure ITE and EDBTE on the same test train, and simultaneously, the engine running continuously at the same output for sufficiently long for information on the stability of coal and water consumption to be available. To obtain LR and nothing else, the stability of the output is the important consideration. The coal and water consumption are more important if efficiency is being established as well. Constancy of output can be obtained by having the boiler develop sufficient steam at full pressure to provide the output, then setting CO at a given rate, and having the MTUs operate at a given speed to give a constancy of V, Q, ITE and IHP. Indeed that very method was used to test the WD 2-8-0 and 2-10-0 engines, mostly the latter, including the more important boiler outputs, and was not regarded as defective (see Test Bulletin 7).

The testing officers of the day decided on using a mixture of the stationary Testing Station at Rugby and separate test trains containing the MTUs on the road, operating in a controlled way, termed Controlled Road Tests.

Errors and Oddly Shaped LR

The intention was to duplicate the Rugby ITE on the road, at various Q and V, in CRTs conducted by Derby. A BPP and V combination was conveyed to the driver, who was to duplicate it during the CRT by adjusting the CO. It was necessary to get the Rugby ITE and V combination right if the EDBTE corresponding to it was to be correct. During the test, the load on the drawbar and speed were regulated by the MTUs, while the DBTE was measured by a dynamometer car. All going well, this system was to provide a consistent set of Q, ITE, EDBTE and V at sufficient points to map EDBTE as appeared in the Test Bulletins. (The intention for the tests done at Swindon was similar).

Rugby was used to establish the boiler conditions and efficiencies and the ITE, and the test trains the EDBTEs. A constant Q and V can be obtained by setting CO at a given figure, and the MTUs to give a constant speed, with full boiler pressure applying throughout, the CO and V being chosen to duplicate a test at Rugby, which gave the ITE for the same CO and V. Instead of doing that, however, the blast pipe was used as a steam flow meter. The Blast Pipe Pressure (BPP) was the basis of measuring Q. A given BPP was measured during a given stationary test at Rugby by a mercury manometer. The aim was to reproduce the same Q on the same locomotive on the road by giving the driver a similar manometer to measure the same BPP. The same manometer and piping could even have been used in both cases – after all, the blast pipe and a location near the driver were needed in both cases on the same locomotive. The driverin the CRT aimed to achieve the same BPP and V as in the Rugby test by varying the CO.

It did not Work Out That Way

The intended system worked well, as a procedure, for testing done at Swindon. The accuracy of the Swindon indicator is a different matter, not discussed here. What follows applies to the system conducted jointly by Rugby and Derby.
In L116, a diagram is given for the LR of a Crosti 9F following the Rugby/Derby testing practice from 1950 until L116 was issued about December 1957. By then testing of steam power had ceased at Rugby as had associated road tests intended to complement the Rugby work, the two together becoming the content of several of the Test Bulletins. The content of L116 therefore admitted and exemplified the defect of method and measurement. That is the major contribution of L116, that it admitted large errors in the shape and presumably value of LR. (The intended correcting procedure is discussed later.)

Fig 1 in L116 shows that in the range 20 to 50 mph, the LR of a Crosti fitted 9F using the testing method used by Rugby/Derby throughout the whole of the testing period, was of completely the wrong shape, indeed an absurd shape. See line 1 of Table 1 below. LR declined as speed increased. That cannot have been. A correct LR rose with speed (the resistance from passing through the air, from revolving rods, and from the revolving masses associated with partly balancing the reciprocating masses all rose with speed, indeed very much with speed squared, while those from the application of steam to the mechanism fell with speed as ITE fell, as it had to if Q was constant for a test, the usual practice during BR tests. When particular tests are gathered together in a summary table or graph, any LR extracted therefrom should not be expected to be constant at any speed – it should be expected to vary with the effort as well. Fig 1 of L116 implied that at up to 39 mph, EDBTE of a Crosti 9F exceeded ITE, which is technically impossible, because ITE exceeds EDBTE by the LR at every speed, and LR is always positive. (Reason for giving this at 39 mph is given below).

It was admitted in L116 that this shape of LR in Fig 1, applying solely to the 9F Crosti, was wrong. But the further admission is crucial, that this was not a one off problem, that it had occurred in all the Rugby/Derby tests, since the inception of the testing procedure, and that it had been known to exist throughout the period. It does not say a great deal for the scientific acuteness and ability of the testing officers that it had not been cured at that inception.

Table 1
Data Given in Figs 1 to 3 of L116:

	20	30	39	50 mph
1 “incorrect” LR of Crosti 9F from Fig 1 of L116, lbs	2985	2631	2518	2461
2 “correct” (a) LR of Crosti 9F from Fig 2 of L116, lbs	1895	2164	2518	3027
3 Apparent Error, (1) – (2), lbs	1088	467	0	-566
4 “correct” (a) LR of standard 9F from Fig 3 of L116, lbs	1448	1643	2060	2659
5 Higher resistance of Crosti 9F compared with that of standard 9F, (2) – (4), lbs, both declared “correct” (a)	447	521	458	368

Here I follow the wording of the authors of L116. I do not believe that the Derby testing officers or the authors of L116 ever knew the correct LRs.
Procedures set out in L116 were supposed to correct for the errors, and give the correct LR for both standard and Crosti 9Fs, as in lines 4 and 2. Exactly how that operated, how it yielded the appropriate EDBTE and with that LR, is not explained in L116. All that is said is that correct answers were obtained. It is definitely not scientific to fail to describe and explain the principles of the correction. I discuss that below. But using the correcting mechanism devised by the testing officers, the correct and incorrect LR intersect at 39 mph, lines 1 and 2.
The “incorrect” LR of the Crosti from 20 to 50 mph as declared in Fig 1 and line 1 of Table 1 above was approximately 2985 – 17.4(actual mph - 20) lbs, ie declining with speed (a straight line effect, used for illustration)[1]

As declared in Fig 2, the supposedly “corrected LR” was approximately 1895 + 37.7(actual mph – 20) lbs, ie increasing with speed[2]

A crude correction without any basis for the correction to [1] to give [2] is obtained by [1] –[2] or

{2985 – 17.4(actual mph - 20)} – {1895 + 37.7(actual mph – 20)} lbs

=1090 -55.1 (actual mph -20), which is the equivalent of the correction given in line 3 of Table 1.

Thus are simple correction mechanisms devised. That given here provides no explanation for why or how the error in EDBTE arose, and there is no basis in them for claiming that the correction is correct.

How the Defective Measurements Occurred

Report L116 does not give the EDBTE figures applying to the supposed corrected figures. But it is possible to use the data in Test Bulletin 13 on the 9Fs, Figure 11, and in Figure 11 of L116 to obtain some comparison. As an error in EDBTE requires an error in ITE of a slightly greater magnitude, the error in ITE in the road tests for the Crosti 9F was about 4.5% at 20 mph, 2.7% at 30, nil at 39 mph, and 5.1% in the opposite direction at 50 mph. The reason for using ITE in this comparison will emerge shortly.

It is also to be asked why a Crosti 9F was used in the identification and presentation of the problem. It was a peculiar engine from the LR point of view, and there were no other Crosti engines on the system. It was also a poor choice when the LR of the Crosti was untypically high at any rate of working. The Crosti engines had a higher LR than the standard 9F, on account of the high back pressure resulting from the highly restricted and primitive blast nozzles, the result of the need to draw the combustion gases through the boiler and the preheater. The higher LR accords well with the back pressure, as shown by the Perform program. The frequently quoted idea that the resistance of the Crostis was high because they had weak frames is unsubstantiated; those quoting it as the reason for the high LR need to consider where the effects of the higher back pressure were felt, and the lack of detection of the effect of weak frames, also whether weak frames increase LR. The back pressure effect did not disappear. In L116, the LR of the Crosti is higher than the standard 9F by 450 to 500 lbs at 20, 30 and 39 mph but only about 370 lbs at 50 mph - see line 5 in Table 1 above. (That weak frames were even suggested at Rugby for the higher LR is another reason for my doubting the scientific competence of the officers concerned; at least they noticed the higher LR of the Crostis, before they declared that all LRs were wrong).

L116 does not say how the erroneous LR and by implication, erroneous EDBTE arose in all these joint Rugby/Derby tests, even the data for the individual tests where ITE showed the same absurd characteristics as those for the Crosti 9F (as in Fig 1 of L116 and Table 1 above). Although it was EDBTE which was the immediate or arithmetical cause of the erroneous LR, it was wrong because an ITE was wrong, and that ITE was wrong because the instruction to the driver at what speed and cut off to run was wrong, or the arrangements for interpreting the BPP differed between the observation at Rugby and that on the locomotive on the road in the CRT. The intended speed for the test was also advised to the operator of the MTUs in the Dynamometer Car.
Indeed, as itemised in L116, the testing officers took steps to check whatever might have led to the absurd answers. The Rugby and Derby indicators were checked and found to give identical powers (powers is the word used in L116, but it is TE which is given by a dynamometer). The dynamometer was checked. The steam rate measurement was considered. The officers found that Q could differ with speed both on the road and in tests at Rugby, but there was no proof that such was the case in anything they did. Their analysis of this data was defective and biased the results of their thinking towards the idea that there was a speed effect. This defective analysis is discussed below, because it led to an erroneous method of amending (or intending to amend) the historic data to produce accurate EDBTE and LR (or intended method – it is not clear that such desk-top corrections ever took place). The ideas put forward prove nothing of consequence, and variation in Q could not be detected from the water rates (although the difficulty of detecting small changes in water rates is emphasised). A concomitant problem is not mentioned. It is assumed that the driver could alter the CO as needed to achieve a certain BPP. A few simulations using the Perform program show that the necessary adjustments to CO to maintain a BPP were minute, not physically possible. (The report R13 on testing Duchess 46225 admits, however, that on the Plant, the CO was moved each time to a definite notch and the speed adjusted to give the correct Q, presumably by regulator adjustment, but on the line a definite speed was used for each step and the CO adjusted accordingly; presumably in drawing out the results for Report R13, considerable interpolation was needed to draw the ITE and EDBTE relationships at the usual tens of MPH and thousands of lbs of Q. That well may have been necessary in reporting results for all Test Bulletins). Presumably where the regulator was used to make the adjustments mentioned, the effect on Steam Chest Pressure would have been very small.

More importantly, however, the data recorded specially to show the relationship among Q, BPP and V was wrongly analysed and interpreted. No speed effect on the relationship between Q and BPP was present in the data for a 9F, nor in data with the same items for a Royal Scot. The relationship between Q and BPP was unaffected by speed, as should have been expected from first principles. See below.

Indicating the CRTs

In L116 it is said (as above) that the indicator used on the CRTs gave much the same readings of ITE as did those given by the Rugby indicator. That leads to the question, if the locomotive was indicated on the CRT, why was the BPPabs of any importance, why was the practice continued of trying to replicate the Rugby BPP in the CRT? The only reason which occurs to me is to connect absolutely the Rugby ITE and Q values with those on the road, to ensure that the Q and V for ITE measured at Rugby were exactly the same as those measured on the road, thereby allowing EDBTE to be measured with the same Q, V, CO etc as was the ITE, as is usual in the BR Test Bulletins. Such perfect correspondence, if the reason, is an extreme action - if the road test ITEs and EDBTEs were made at different Qs from those at Rugby, it is always possible to interpolate. Indeed, ITEs obtained on the road must be superior, in view of draft and exhaust effects, to those on a Stationary Plant. In that case, the Rugby results could have been put to one side.

It is remarked in L116 that a comparison was made between Rugby ITE and Derby DBTE by running the engine at equal V and CO, which gave LR without reference to Q. That tells anyone checking what Derby did almost nothing because identical V and CO mean identical Q. If it is thought that Derby needed to explain a V effect, then experiments would have been needed at each speed separately. Indeed there was some of that – see below.

It is obvious, however, that if ITE and EDBTE led to LR results which were obviously wrong (as in Table 1 and by admission, many other tests), then the various ITE results were not compatible, a problem of method and measurement

Test Bulletins Left Uncorrected

The Test Bulletins recording the joint work of Derby and Rugby are listed below. These were invalidated by the problems revealed in L116. Of course L116 contains the following paragraph in the Foreword:

With regard to previously published curves, however (presumably ITE and EDBTE curves in Test Bulletins below) it is considered that the discrepancy is not sufficient to invalidate their use for train timing and similar purposes. No information is given on the size and nature of the discrepancy anywhere in L116 (but see my rough estimates above).

The following Test Bulletins were undermined, those based on Rugby work and Derby CRTs:

Bulletin 2, B1 61353 1950
Bulletin 5, Standard 7 1953
Bulletin 6, Standard 5MT 1952
Bulletin 10, SR 8P 35022
Bulletin 13, Standard 9F 1959, work done up to 1957
(Last steam testing at Rugby 92250, 9F Giesl, no Bulletin
LMR 8P 46225, no Bulletin, but reports R13 and L109 mention the L116 method of adjustment (see below)
[Equipment dismantled 1970, plant demolished 1984]

No attempt to correct these reports is known to have been performed. Despite Fig 16 in L116 attempting to show how the earlier work could be corrected, nothing in L116 shows how a corrected EDBTE could be obtained, even what the error was.

The Proposed Correction and Underlying Research

Some background from L116:

It was concluded that the problem arose from using blast pipe pressure as a steam flow meter without compensating for varying road speed. In the revelation of the odd LR shape problem, it is said early in L116 that the difference between ITE from the LTS and the EDBTE obtained from road tests, which is the Locomotive Resistance (LR), had not been acceptable in shape, that the discrepancy was large and consistent. It was said that it was believed (ie not shown to be the case) that the DBTE resulting from the procedures used was correct in the middle speed range, too low at low speeds and too high at high speeds. It is not stated how this was known, indeed, given the problem, how it was possible to know it. Similarly, in point 9 in the report, it is concluded that the steam rate for a test applied only at the mean speed for the test. This is not sensible if things worked properly. How does the instrumentation know what range of speeds will be tested and how many tests conducted at each speed, ie that the results can be correct for the mean? The mean will vary with the tests conducted.

At Swindon, ITE was measured on both the plant and the road (see Bulletin 1 p 5). Although the Bulletins claim that there were no significant differences in boiler and cylinder performance between the plant and road tests, it is generally considered that the plant tests were undertaken to determine boiler characteristics, and that both ITE and DBTE data used in reports prepared by Swindon were obtained on the road. As they were both subject to the same effect of V on P where P was used as the steam flow meter, they should give reasonable LR.

Discovering the effect of V on BPP at a given steam rate from plant tests to adjust the results of road tests requires correspondence between plant and road in all circumstances. It is doubtful that such correspondence could be achieved. The ability of a given BPP to bring about a given evaporation can be expected to differ on the plant and on the road in ways not considered in the report. There are at least two reasons for this. The first is that the scooping of air into the front damper and under the fire will reduce the need for draft for a given evaporation rate compared with a stationary locomotive on the plant. The same will apply to air drawn from the sides of the ashpan. (If the front damper is closed and underfire air is drawn solely through the rear damper, the draft requirement on the moving locomotive will be increased, to overcome the slight vacuum behind the rear of the ashpan.) The second is that the moving locomotive will create a small vacuum at the chimney top, which will provide a little draft, compared with a stationary locomotive. Both of these effects can be expected to increase with speed. Tests on the plant to establish the effect of speed on evaporation for a given BPP will not detect these two effects. The third possible consideration is that the resistance of the fire cannot be expected to be necessarily identical on the plant and on the road at a given steam rate, on account of firebed depth differing on account of fire management requirements and duration of the run, and different packing down of the burning coal. A given BPP on the road could lead to higher or lower evaporation than on the plant, even if all other factors were made identical.

Surprisingly, it was believed that the incorrect results could be corrected, as a desktop mathematical exercise. To correct something known to be wrong, it is necessary to discover what was wrong and why, and to know the correct answers. None of this applies in this case. Usually, it will not be possible to undo what has been done.

Derby therefore put forward a method for correcting the defective measurements of EDBTE in all the published Bulletins applying to Rugby/Derby tests, or new tests to be done, incorporating these modifications. (Bulletin 13 on the 9Fs was published in 1959, but was based on data gathered before 1957, and Bulletin 20, published in 1960, on the rebuilt Merchant Navy engines) included road test data only, was not tested on the pre-1957 testing system. See the extent of the effect of their correction for the Crosti 9F in Table 1 above. The modification was to develop a process and formula which changes the Q data.

To have any hope of making such a correction, the reason for the error has to be known. As above, it had to be a question of method and measurement. As these factors are likely to differ in effect from test to test, the correction task would seem hopeless. They considered three possible bases – adiabatic heat drop, compensation for change in density, and compensation for speed effect on the BPP/Q relationship. They could not find any thermodynamic reason, which probably meant there was none, and picked, in speed effect, something which did not exist, as I show below. It is true that among the road test data, they had examples of tests where the result differed with the speed, eg by direction. These tests drop out as a basis because they were not comparable with the principle of the testing, constant Q, V and BPP. One wonders if such non constancy by direction in a test was not the reason for the error.

The equations in Fig 16 of L116 do not demonstrate a basis for altering Q, simply playing with the concepts “left over”, not used so far in trying to explain the anomalies. The Derby test officers had observed some peculiar effects of different speeds, which is perhaps why they thought speed was playing a part in explaining the determining the influence of BPP on the Q passing the Blast Pipe. They did not think that through. See my tests below. Note also that where they claimed that the system worked, that a correct LR, or correctly shaped LR, results, there is no case where a correct LR comparator exists. Nor the basis for declaring how a LR would be established from first principles. No prospects for science there.

The officers considered that there must be more to it, however. They considered that the reason for the error was that their assumption held over the whole seven years of testing that Q varied only with BPP was wrong, that the relationship between Q and BPP was affected by V. They therefore sought a relationship among Q, BNP and V. they also believed that the error in procedures and/or measurement were in the EDBTE, which was measured by Derby. But that also depended on ITE registered on the road.

Although L116 was partially accepted and some adjustments made with it, there are memoranda within it from D R Carling, Supervising Engineer of the Rugby plant, and E S Cox, Chairman of the Locomotive Testing Committee. Both have considerable reservations about the report. Both note that no explanation is offered for the supposed effect of V on the relationship between BNP and Q, Cox saying as much as that the variation with V was not established scientifically. Cox believed that the range of experimental data was to a large degree the range of experimental error.

Carling said that on the whole the data examined until then could only be regarded as supporting the method proposed in the report as a workable method for use where necessary, without any pretension to confirming it as a fundamentally correct method.

Neither of these gentlemen called in aid S O Ell or his staff, who were in charge of testing at Swindon. The CRTs conducted at Swindon depended on duplication of the results of boiler and efficiency tests conducted on the Test Plant at Swindon. Ell claimed that the road tests confirmed the plant tests. Ell was surely the person most likely to discover the defect in the Derby practice.

There are more and better reasons for not accepting the correction method. The authors of L116, presumably Rugby officers, were not content with the conclusions and intentions of L116. On p 8, under (2), Joint Analysis of Results, they say “It is desirable that test results should be pooled, so that Indicated and Drawbar Characteristics can be constructed together. Hitherto, the curves have been drawn up entirely independently, and small differences in the methods of construction have added to the difficulties of reconciliation.”

In similar vein, they go on “(3) Elimination of Differences in Test Procedure. Testing methods have been developed at Rugby and Derby separately, and the results of tests at both centres are valid for the respective conditions under which the tests were made. It is desirable however, if agreement is to be achieved with joint tests, for local differences to be eliminated as far as practicable . In this connection, it must be mentioned that the mean blast pipe pressure curve established at Rugby cannot be reproduced when a locomotive is subsequently subject to tests on the line. A re-calibration of the orifice meter was therefore necessary , and this work was to be undertaken while the main tests are proceeding. It is considered that anomalies of this nature could be readily eliminated by close co-operation with regard to choice and siting of instruments”.

These comments are indicative that the joint tests did not agree for seven years because the procedures were sloppy, and did not lead to automatic reconciliation of results.

Experimental Data on 9F

In L116 the experimental data on Q, V and BPP used in formulating the correction process are presented in Figure 11, ten observations at 15 mph, five at 30 and five at 50 mph. I have transformed these data into Table 2. In Fig 13 of L116, appears another set of BNP against Q for 92050 with differing figures. To increase the number of observations, especially at 30 and 50 mph, the data in Tables 2 and 3 below have been combined into one series, to give 18 observations at 15 mph, ten at 30 and nine at 50 mph, a total of 37. The results are very little different, both in actual answers and goodness of fit. (the comparison was with the 20 observations of Table 2 and the 37 of Tables 2 and 3).

Table 2 Data in Fig 11 of L116 on Blast Pipe Pressure, V in mph, and Q, 9F 92050

BPP gauge lbs/sq in	Q lbs	V mph
1.6	11900	15
1.95	13200	15
2.15	14000	15
3.2	16100	15
3.4	16700	15
4.75	19000	15
4.83	19800	15
5.5	20200	15
6.6	21600	15
7.1	22400	15
2.8	15600	30
4.55	19000	30
6.6	22300	30

7.1	23400	30
8.5	24800	30
2.5	15000	50
3.55	17400	50
4.6	19600	50
5.8	21400	50
7.1	23300	50

Table 3 Data in Fig 13 of L116 on Blast Pipe Pressure, V in mph, and Q. 9F 92050

BPP gauge lbs/sq in	Q lbs	V mph
1.972	13122	15
2.018	13900	15
3.236	16144	15
3.388	16749	15
4.786	18281	15
5.623	20277	15
6.607	21135	15
7.08	22491	15
2.818	15596	30
4.571	19055	30
6.025	21528	30
6.607	22284	30
8.414	24717	30
2.4547	15066	50
3.3884	17378	50
4.5709	21478	50
7.0795	23227	50

In the same Figure 11 of L116 are freehand lines which are meant to represent the relationships among these items, judged to be:
At 15 mph Q = 9900 $P^{0.415}$
At 30 mph Q = 10,200 $P^{0.415}$
At 50 mph Q = 10,400 $P^{0.415}$

Where P is blast pipe gauge pressure. It is argued in L116 that as these lines are parallel in non-logarithmic form, the index or power can be made the same for each line. The lines in non logarithmic form are not straight, and are therefore cannot be parallel. Nor is the slope of each the same in non-logarithmic form (change in BPP divided by change in Q). (This was a rich claim in any case with only five observations in Fig 11 at each of 30 and 50 mph). They are *in part* the same distance apart, in log form because the centre of the curves of each at that point has been moved a certain distance. A mathematically correct analysis of the data of Both Figs13 and 15 together gives:

At 15 mph, Q = 55$BPPabs^{1.964}$

At 30 mph, $Q = 121.5BPPabs^{1.705}$
At 50 mph, $Q = 95BPPabs^{1.798}$,

Which are mathematically and statistically respectable, whereas the L116 figures are not.

Analyses of 9F Data

There are three important defects in this work. First BPP is measured at atmospheric or gauge pressure, whereas it should be in pressure absolute, as even an apprentice scientist should have known. Second, the three curves in Fig 11 from which Table 2 was drawn above were fitted by freehand, with the initial pressure for each speed picked by eye. More importantly, the data are fitted to lines for the speed at which the tests were made, 15, 30 and 50 mph, and the curves for each speed drawn by eye. That means that the relationship with V is assumed to be that drawn in Fig 11.

Regression of this very same data both with and without its relationship to a speed being assumed finds the effect of V on the relationship between Q and BPP to be in effect nil. Regression also avoids guesswork and has the enormous advantage of giving as a test statistic whether there is any significant (statistically significant) difference in curvature or constant by speed. There is not (see below), which means that eyeing up the gradient and constant introduced a serious bias. Fortunately, it is possible to do this analysis properly, at least in principle.

Third, there are insufficient observations at each of 30 and 50 mph (ten each) to analyse the effects at those speeds properly. It is also desirable to analyse the data in such a way to see whether the implied assumption on the part of the testing officers that the speed effect differed by speed, an assumption for which no reasons are given.

Regressions obtained from these data follow. The physical act of passing a given quantity of steam through a restricted nozzle should have BPPabs on the vertical axis as the result, and Q as the cause, on the horizontal axis. As however, the system is used as a meter, the reverse arrangement of the data is used, ie Q on the vertical, BPP on the horizontal.

The regression results which follow are all in terms of BPPabs, ie in absolute pressure, and speed in RPM.

The following are the results of regressing the useful permutations of BPPabs, Q and RPM, the figures or values in Tables 2 and 3:

1 $BPP\ abs = 0.126Q^{0.52}.RPM^{-0.025}$

Effects and comparisons: A 10% increase in Q, RPM constant, leads to a 5% increase in BPPabs.
A 10% increase in RPM, Q constant, leads to a 0.25% increase in BPPabs (ie a quarter of one per cent) .
If the RPM term and data are eliminated, ie the regression is of Q on BPPabs, the best fit equation of BPPabs on Q scarcely changes. It becomes $BPPabs = 0.133Q^{0.51}$
2 $Q = 65.BPP^{1.83}.RPM^{0.05}$
Effects and comparisons: at a Q of 16,000lbs, if there is a 10% increase in BPPabs, RPM constant, Q rises 19%.
A 10% increase in RPM, BPPabs constant, leads to a 0.78% increase in Q (ie four-fifths of one per cent)

If the RPM term and data are eliminated, the best fit equation to Q on BPPabs changes only a little from the above with RPM included, to Q = 74. $BPPabs^{1.87}$. A 10% increase in BPP abs at a Q of 20,100 lbs, leads to a 19.5% increase in Q

This equation without the RPM term (ie Q = 61. $BPPabs^{1.9}$) is that used successfully in Swindon testing with BPP as the meter of Q. It was also used by Derby, but not successfully). The equation with the RPM as an extra term, shows how the Q/BPPabs relationship is unaffected by V, ie by RPM) (ie 65.$BPP^{1.83}$.$RPM^{0.05}$).

3 Speeds considered separately as in L116 (as above)

At 15mph, 18 observations, Q = 43$BNPabs^{2.06}$
At 30 mph, 10 observations Q = 137$BNPabs^{1.66}$
At 50 mph, 9 observations Q = 85$BNPabs^{1.33}$
(Compare all speeds together, as in 2 above, Q= 65.$BPP^{1.83}$, or Q = 74. $BPPabs^{1.87}$.)
These equations differ vastly from those in Fig 11 in L116.
L116, because they are not based on freehand or by eye curve fitting, and because they employ BPPabs and not BPPgauge, represent the best statistical (scientific) fit to the data. The ratios involved with any change in BPPabs are much smaller than those used in gauge or atmospheric pressure, as in Fig 11 of L116. As before, five observations at each of 30 and 50 mph are totally insufficient for investigation, and ten only on the verge of sufficiency.
The coefficients on RPM are always very small. Q is always a large number in thousands, and RPM always a small number in comparison (50 mph, 280rpm, or less). The effect of V is very small indeed, as in the notes above about the effect of 10% increases in determining variables.
This data from Figs 11 and 13 of L116 does not contain and cannot reveal a speed effect on the relationship between Q and BPP, because there is none. (Statistical tests revealing probabilities available).
There is similar data on a Royal Scot, source now forgotten or lost. The Scot data have been analysed similarly to those of the 9F.
Table 4 Data on Blast Nozzle Pressure, Q and V in mph, Royal Scot

V mph	rpm	BNP, abs pressure, lbs/sq in	Q lbs	ln BNPabs	ln RPM	lnQ
20	83	15.6	11340	2.747277	4.418394	9.336092
35	145.25	15.6	11510	2.747277	4.97801	9.350972
50	207.5	15.6	11590	2.747277	5.334685	9.357898
20	83	16.6	14980	2.809403	4.418394	9.614471
35	145.25	16.6	15260	2.809403	4.97801	9.63299
50	207.5	16.6	15510	2.809403	5.334685	9.64924
20	83	17.6	17640	2.867899	4.418394	9.777924
35	145.25	17.6	18000	2.867899	4.97801	9.798127
50	207.5	17.6	18390	2.867899	5.334685	9.819562
20	83	18.6	19800	2.923162	4.418394	9.893437
35	145.25	18.6	20230	2.923162	4.97801	9.914922
50	207.5	18.6	20750	2.923162	5.334685	9.940302
20	83	19.6	21660	2.97553	4.418394	9.983223

35	145.25	19.6	22160	2.97553	4.97801	10.00604
50	207.5	19.6	22790	2.97553	5.334685	10.03408
20	83	20.6	23310	3.025291	4.418394	10.05664
35	145.25	20.6	23870	3.025291	4.97801	10.08038
50	207.5	20.6	24600	3.025291	5.334685	10.1105
20	83	21.6	24790	3.072693	4.418394	10.1182
35	145.25	21.6	25420	3.072693	4.97801	10.14329
50	207.5	21.6	26240	3.072693	5.334685	10.17504
20	83	22.6	26160	3.11795	4.418394	10.17199
35	145.25	22.6	26830	3.11795	4.97801	10.19728
50	207.5	22.6	27740	3.11795	5.334685	10.23063
20	83	23.6	27440	3.161247	4.418394	10.21976
35	145.25	23.6	28150	3.161247	4.97801	10.2453
50	207.5	23.6	29170	3.161247	5.334685	10.2809
20	83	24.6	28620	3.202746	4.418394	10.26186
35	145.25	24.6	28150	3.202746	4.97801	10.2453
50	207.5	24.6	29170	3.202746	5.334685	10.2809

1 Regressing Q on BPabs and RPM:
$Q = 62.BNPabs^{1.87}.RPM^{0.047}$
If BNPabs rises 10%, RPM constant, Q rises 19.5%; if BNPabs rises 10%, RPM constant, Q rises .044% (ie less than one twentieth of one percent).
2 Regressing BNPabs on RPM and Q
$BNPabs = .16RPM^{-0.234}.Q^{0.50}$
If RPM rises 10%, Q constant, BNPabs falls 2.36%; if Q rises 10%, RPM constant, BNPabs rises 7.4%.
As with the 9F above, the **Royal Scot data shows no relationship between Q and RPM, and that speed has no effect on the relationship between Q and BPPabs,** in complete contrast to the hypothesis of the testing officers.
3 Speeds considered separately as in L116 (as above)

At 15mph, ten observations, $Q = 43BPPabs^{2.06}$
At 30 mph, five observations $Q = 137BPPabs^{1.66}$
At 50 mph, five observations $Q = 85BPP^{1.33}$

These equations differ considerably from those in Fig 11 in L116, because they are not freehand curve fitting, and because they employ BPPabs and not BPPgauge, and because they represent the best statistical (scientific) fit to the data. The ratios involved with any change in BPPabs are much smaller than those used in gauge or atmospheric pressure, as in Fig 11 of L116.

Use of gauge pressure instead of the correct absolute pressure will have considerably distorted any relationships including BPP, including the idea that V is needed in determining a relationship between Q and BPP.

The coefficients on V are always very small. Q is always a large number in thousands, and V always a small number (50 mph, 280rpm, or less). The effect of V is very small indeed, as in the notes above about the effect of 10% increases in determining variables above. This Royal Scot data does not contain and cannot reveal a speed effect on the relationship between Q and BPP.

Analysed via Perform

There is no effect of V on the relationship between Q and BPPabs in the data, data gathered in the testing of 92050 and Royal Scot. Further, despite the claims in L116, the three lines, one for each speed, in L116 are not straight, nor parallel (not that those charactristics matter provided a good fit is obtainable). Fig 15 does not connect C to V; while one can be graphed against the other, a constant remains a constant. Something else must have been in mind. That can be observed in test applications of the Perform program, as could be done by considering indicator diagrams. What can be learnt from Table 5 is that at low RPM, the exhaust from each stroke is separate, but as rpm rises, the exhausts merge, to give a continuous flow of the Q. This can be observed through trials of Perform.

Table 5

Perform Exercise to Show Effect of V on Relationship between Q and BNP abs, Standard 9F

V	Q	CO	BNP Gauge*	Inlet Steam Temper-ature, ℃	Release Pressure	ITE Perform	ITE, Table 10 of Bulletin 13	Ratio ITEs, Perform to Bulletin 13
20	12000	19.2	1.9	316	40	13,700	14,400	0.95
40	12000	13	1.71	316	52	8040	7200	1.12
60	12000	11	1.7	316	52	5420	5000	1.08
20	18000	33.9	4.75	366	91	22,800	22,000	1.036
40	18000	21.7	4.45	366	52	12,900	12,000	1.08
60	18000	18.2	4.36	366	39	8,850	8,100	1.09
20	24000	44.9	9.2	377	110	27,600	28,000	0.99
40	24000	29.1	8.67	377	78	16,400	16,000	1.03
60	24000	24.5	8.45	377	52	11,300	11,000	1.03
20	30000	57.1	16.3	393	143	31,800	32,600	0.98
40	30000	36.6	15.1	393	84	19,200	19,000	1.01
60	30000	30.8	14.8	393	72	13,300	13,000	1.02

*Perform works in pressure absolute and automatically converts gauge pressure to absolute. Absolute pressure is simply gauge pressure plus 14.6 lbs/sq in.

In each set of 20, 40 and 60 mph at a certain Q, BPP gauge is close, but falls a little from 20 to 60 mph. In each set of three, BPP gauge is highest at 20 mph, because at 20 mph, the BPP discharges are more distinct than at higher speeds, but decline from 20 to 60 mph. The Perform ITE is close to the Bulletin 13 figure. There is no evidence here for a speed effect on BPP at each speed at each Q.

This table can be rearranged to have sets of three for Q and for CO.

Speed in Normal Test Results

In the Test Bulletins, ITE and DBTE are shown in Figures or Graphs against Speed, with the following shown across the Figures: Q, CO, fuel and efficiency. These performance maps clearly show that ITE varies at a given Q with speed, that as V increases, the ITE curve declines with speed.

That has to be. As speed increases less steam is available per stroke, and ITE from a given Q falls. The issue in L116 is different. The issue in L116 is whether during a given test, changing V affects the relationship between Q and BPP. Fig 11 in L116 is drawn to imply that it does. The data collected to test that point for the 9F and Royal Scot, and the simulations with Perform show that V does not make the slightest difference to the relationship between Q and BPP, that it is BPP which affects Q, unaffected by V.

The officers were not clear about this distinction. It is said on p 5 of L116 that it was first observed with B1 61353 that during the course of a day's test running from Carlisle to Skipton and return, the Q produced by a particular BPP during the outward run could not be accurately repeated on the return. The only difference of any significance between the two test runs was that the overall average speed was lower on the return, owing to the nature of the test route.
They refer to average speed. The whole aim of the MTUs was to keep speed constant for a given Q. The aim would have been to select the speed to be run for a given Q on the ruling gradient of 1 in 100, and to add to the resistance to maintain the speed where the gradient eased. The Q and V were to be maintained for the whole test section, downhill as well as up. The average speed was of no significance. Nothing is said about how much downhill running was converted to 1 in 100 (or other desired gradient) uphill, but it would have been useful for testing for significant periods at the higher speeds. Average speed is not of interest for either CRTs or the effect of V on the relationship between BPP and Q.

There is also confusion on pp 5 and 6 of L116, and in Figures 7, 8, 9 and 10 of L116. There is mention of constant BPP, mean Q for a test applying only to power developed at the mean speed for the test, adjustments being needed at all other speeds. There is obviously mistaken thinking and measurement here, because the data are those which appear in the EDBTE diagram in Test Bulletins, Q and EDBTE on the vertical axis, V on the horizontal axis, with a series of lines showing the relationships among those items (usually with CO and efficiency superposed). There is nothing to do with the effect of speed on the relationship between Q and BPP here. Nothing is discovered through the idea that effects differ at high and low speed tests, so called.

Experimental Data on Duchess 46225
The Last Attempt to Get it Right, Rugby and Derby?

This was the testing of Duchess 46225 in 1958, which involved use of the steam flow meter to ensure constancy of Q on the road, and the MTUs. There was an extensive time gap between the Rugby tests of this engine and those on the road. Further the valve heads were set back for the road tests to even the tractive effort produced from the front and rear ends of all cylinders; that led in turn to a given CO giving a higher ITE at any speed, ie plant and road ITEs could be expected to have differed for a given Q. On the other hand, the road tests were separately indicated at all cylinder ends (the same applied to other of the joint tests, but Rugby results were preferred for indicated results in those cases. L116 had just been published, identifying the defect, and proposing a solution. Efforts were made to confirm the discrepancy between ITE on the plant and on the road, but the results were not conclusive. For all that, report R13 says that the L116 method of adjusting CRT results was used and brought agreement between the two types of test. In addition, LR was measured directly on line as the difference between ITE and EDBTE (but very poorly presented – in specific terms with exact weight indeterminant, wind effects unknown, the statement "average service conditions" undefined; even broad values or conditions still leave an unusually low rate of increase with speed. Although water consumed (Q) was measured incrementally on a time basis, the BPP was used to measure the instantaneous Q. For all the care taken, plant and road ITEs differed by speed, as in Table 6 above.

Neither R13 and L109 give any detail on the application of the L116 method of reconciliation. No reasons are adduced for the ability of the method to reconcile the results from the two types of test; there was none.

No progress was therefore made in comprehending the problem of wrong EDBTE values encountered by the Derby testing people, explaining it, and finding a solution, in these last tests. No correct cure.

L116 admits the error and the period over which it prevailed. It is singularly deficient in not saying what went wrong and why. There is the idea that the error was the result of failing to take into account the effect of speed on the relationship between Q and BPP, but the above analyses show that idea to be fanciful. In particular, allocating observations into speed bands vastly exaggerates the effect of V in the results of the analyses. For all that, it is obvious what was going wrong. The method was not connecting ITE at Rugby with ITE on the road for, so far as we readers sixty years later can tell, a given BPP advised to the driver in a CRT. Even if the ITE on the road for a given Q and V was equal to the Rugby ITE, the EDBTE for that ITE (hence Q and V) was not measured or calculated properly.

ITE on the road and on the Rugby plant ought to differ for reasons already given, to do with draft on the fire and exhaust effects on the road compared with the plant, and the road figures ought to be preferred. That does not really answer the question of what happened in Derby controlled CRTs. L116 does not give the road ITEs, except in a very indirect way for the Crostis. (The fig 11 data in L116 is wrongly presented and analysed, as discussed above). The exception is in yet another internal report L109, in Fig 20. This shows ITE recorded by Rugby and Derby for various Qs from 16,000 to 38,000 for a Duchess at speeds from 20 to 80 mph. At a Q of 28,000 (one of many Qs available), Derby ITE differs from Rugby ITE as follows:

Table 6 ITE Recorded by Rugby and Derby at Q of 28,000 lbs/hour Duchess 46225, 1956

mph	Rugby ITE	Derby ITE	Rugby ITE/Derby ITE
25	25,500	24,000	1.0625
30	22,400	21,700	1.032
40	17,700	17,400	1.017
50	14,600	14,600	1
60	12,200	12,500	0.976
70	10,400	11,000	0.945
80	9,000	9,800	0.918

Source, Table 20, Internal Report L109. The road tests (Derby figures) were conducted March to May 1956.
Here remerges the pattern of Table 1. The Derby figure is the lower from 20 mph to 50, and the higher from 50 to 80 mph, with results equal at 50 mph (39 mph for the Crosti 9F). Note above that the indicators were compared. So were the Qs (ie water consumption) and not found to be the source of error or explanation. Even if there was error in measurement of Q, it would be expected to be a constant quantity or proportion, not one operating in one direction below 50 mph and the other above 50 mph, and to different extents. Nor would it be expected that the ITEs would be equal at 50 mph. There is no measurement of EDBTE in this data, but if EDBTE were properly measured relative to Derby ITE, it would follow a similar pattern of ratios.
This data does not appear in R13, reporting the same tests of the same engine. But R13 says:
When the two sets of test results were first compared there appeared to be an even larger discrepancy between them as regards power output than there was between similar tests on

the plant and on the line in the case of the (Class 9 locomotives). The extent of the disagreement was shown in Fig 20 of L109 (and in part in Table 6 just above).

Application of the methods (in L116) has, it is claimed, however, brought agreement of the two sets of tests within the normal limits of experimental error, having regards to the circumstances of the tests mentioned above (ie the time gap). This does not apply, however, because the correcting equation is wrong in principle.

The Duchess data on the Rugby plant and on the road are definitely not comparable. Between the tests at each place, the valves were set back to increase the work done at the rear end of each cylinder. Nevertheless, the pattern and extent of the ratio of Rugby to Derby ITEs, as in Table 6, could still not be explained. That is of course if any consideration had been given to why it could exist.

However, Report L109 states:

An attempt was made to determine whether the same blast pipe pressure produced different rates of evaporation under constant and variable conditions of speed respectively. The constant speed tests were carried out during the first two weeks, and difficulties encountered during the early stages of the tests (Effect not given) … prevented them being strictly comparable with the remainder of the tests. The results were therefore not conclusive. Despite which:

As regards the degree of reconciliation with the results obtained during the Stationary Plant tests, …..as on previous occasions, however, there is some discrepancy between the ITE characteristics established on the Stationary Testing Plant and the road. Results were of the type appearing in Table 6 above.

It then goes on ……"Tests will be carried out in the near future at Rugby to investigate this discrepancy." So only after testing had ended was the error to be investigated, and then only on the test plant.

So no progress was made in understanding the difference between road and plant ITEs from a given Q, even at the very end of steam testing.

Unscientific Presentation of the Results of the Derby tests and the Supposed Correction Procedure

As the commission of the error was so long lived, its effect was so unusual and gross, and the correction procedure was of such doubtful validity, a lot more explanation should have been given than is present in L116. The following would be expected:

1 Showing the Error – about 30 examples of what were meant to be corresponding Rugby and Derby results, the Q, BPP, V of the test, ITE, EDBTE and any Vs which might have affected the BPP/Q relationship. In particular, additional characteristics of the Derby ITE and EDBTE results, especially such as Derby and Rugby ITEs which are the same at some central speed but which are different at other speeds, and to increasing or decreasing extents from some central value.

2 Application of the Intended Correction, in particular the application of Figure 16 of L116. What adjustments are made to the Derby Q for road ITE and DBTE tests. Then, for a given recorded erroneous road ITE and EDBTE, the source of the corrected ITE and EDBTE (what is their source without running special tests; were the corrected values interpolated from

other data, and if so, what? Are there examples of whether during a given test, changing V affects the relationship between Q and BPP. Even more basically, it cannot be expected that variations in Q on the basis of the correcting equation can be correct . how is it supposed to produce what it is said to do. The ITE and EDBTE developed on the road should be derived from accurate measurement, not an invalid formula.

3 Results of the Correction Made – the different Q, and the associated road ITE and EDBTE; where did they come from, how do they fit into a continuity of ITE and EDBTE, ie the results of the corrected Q and associated ITE and EDBTE, for both Rugby and Derby.

4 The LR of the loco for which these adjustments were made and what was the comparison locomotive, and how its LR was obtained. That and any easier and more accurate tests, such as road tests run at a constant speed and CO for ITE, EDBTE and LR.

The Correction Equation

This is of the form $Q = CP^n$. Its derivation is not explained, either what it is intended to do, nor its origin. There is ready comparison with the equations derived above from the research data for the 9F. The conclusions reached, however, are very different. P is BPP, which is probably in gauge pressure, whereas it should be in pressure absolute. C varies with V, according to Figure 15, from 99 at 15 mph to 104 at 50 mph, or by a ratio of 1 to 1.05. That is the ratio of the constants in Figure 11, remarked upon above as a bias towards a speed effect. In my regressions, across all speeds together, the value of this constant is 57 with a speed term present, or 61 with no speed term present (as above).

In the L116 correction equation, the index on BPP is 0.415 in all circumstances. By the regression of the test data on which it is based, the index on BPPabs is 1.9, whether a term for RPM is included or not, a vastly greater influence of BPPabs than the index on BPP in the freehand L116 equations.

The correcting equation is therefore Q= (99 to 104, depending on speed)$BPP^{0.415}$. As the regressions of the same data show there is no dependence on speed, a conclusion confirmed by the Perform analysis, and no explanations or instructions are given in L116 (despite Fig 16) on the circumstances in which the correction equation is to be used and how, it should not be used to correct any data. And it cannot correct the old Derby data. In L116 not only is the correcting equation based on wrong thinking, it is based on wrong data and relationships.

The correct equation relating Q to BPPabs is $Q = 61BPPabs^{1.9}$, at all speeds and BPPabss. That is based on the test data collected for 9F, and applies to that class. See the analyses and results of the data above. Subject to the reliability of that data, it gives correct Q for any BPPabs for a 9F.

These two equations (99 to 104, 61 etc) are not correcting equations, but relationships between Q and BPPabs. The 99 to 104 equation is wrong, for reasons already given, and the 61 equation is the best fit to the data collected to research the V effect on the relationship between Q and BPP. L116 gives no rules for declaring that a Q is incorrect, although an LR might be judged to be the wrong shape. Even if a Q can be said to be incorrect, where does the correct BPP to obtain a correct Q come from, and from that the correct ITE and EDBTE. As Derby had made so many mistaken estimates of road ITE and EDBTE, it is not satisfactory to suggest that it will have a large notebook of observations for each engine tested, certainly not correct ones, because it had no way of saying which if any were correct. Nor should any further tests a Rugby be expected to solve the problem

Conclusions

The conclusions are not favourable to the Derby team. First, the results being anomalous over the whole testing period, it follows that the Derby team did not know how to achieve satisfactory road ITE and EDBTE results for a given Q despite years of practice. They wasted time in developing a supposed speed effect on the relationship between Q and BPPabs and V. The same applies to the supposed correction equation and procedure.

Different and more scientific expertise (including statistical) should have been called in early in the testing programme (before the end of the first year say) rather than tolerate anomalous results for years on end, ie better technical expertise on the generation and detection of correct data on the road of ITE and EDBTE, the function of the Derby Testing Section.

This paper first considered the large number of wrong results, admitted in internal report L116. It then considered how incorrect results could have arisen, and the modest research conducted to allow correct the incorrect results to be corrected, research which was extremely poorly applied. The officers concerned considered that their results were wrong because they had not taken into account the effect of speed on the use of the Blast Pipe Pressure on the metering of steam. In that they were mistaken, for there was no such speed effect. The correcting mechanism and equation they devised did not fit the data available, which led to wrong conclusions. They believed that they could conduct desktop corrections of results, but in that they were mistaken also, and no corrections of results proved possible. Nor did they perfect the testing and measurement, and to the end the Derby measurements of ITE proved defective, including that of a Duchess. Although Derby thought it had a system which could correct LR, it never explained where the comaparator locomotive came from. Checks were made of the apparatus and procedure, but the Derby errors were never corrected. This failure by Derby is surprising because testing procedure with similar intentions took place at Swindon and seemed to operate satisfactorily – it was Derby which did not succeed in measuring properly, and which devised correcting mechanism which was not a logical explanation for the mismeasurement which occurred.

The data available has been analysed much more soundly here than was done for L116.

Derby did not run its side of the joint Rugby – Derby testing soundly.

Some conclusions are drawn in the text on the peculiarities of some of the testing.

The conclusions of L116 should be forgotten, such as they are. That includes the supposed LR of a 9F.

Locomotive Resistance - Doug Landau Dec 2019

A response to John Knowles letter 4 July 2017, is somewhat overdue. In the interim since my letter March 17 2017 I have undertaken further examination and analysis of the available test data from the Rugby test plant together with material from internal reports, technical papers, correspondence, and the various test bulletins. This has involved two further trips to the NRM archive at York, the latest in March 18 2019. My response, I'm afraid, covers over 26.000 words, of which only part is directly dealing with John Knowles letter. Additional analysis of the available data takes up much of the text. Three examples of the "simple proof" promised in my letter 12th October 2017, are included. The predominant approach remains presentation of the empirical evidence, avoiding the need for estimates as far a possible. Some call on the latter in some circumstances is unavoidable. Estimates can be a bit fluid at times, such as estimating aerodynamic effects subject to natural variation, for example.

The paper trail is currently by no means complete, and further visits to the NRM are required to establish an acceptably complete chronology and record of the various, trials, tribulations encountered, solutions and improvements achieved, during the operating life of the test plant. One thing that emerges from the archive is that the approach of the test staff was meticulous; every aspect of test plant instrumentation was subject to calibration on a fairly regular basis. On occasion outside organisations such as the National Physics Laboratory or manufactures such as Kent Instruments carried out independent calibration tests. Plant tests were preceded by calculations on the theoretical critical speeds for the various Belleville washer options. Calculations were also made of the mediating gear correction required for shifts from top dead centre on the rollers. These also allowed for shifts from TDC of the bogie and trailing truck wheels resting on stationary rollers. Where results appeared suspect, calibration tests, investigations and experiments were undertaken ad hoc.

When tested with the troublesome hydraulic dashpot emptied of oil, of 11 drawbar pulls recorded with 45318 on variable speed test run 156, 19 January 1950, no mediating gear corrections were required When the mediating gear did indicate such a need, the corrections were often as little as 10 lb, sometimes even less; the highest noted from a very limited sample is -54 lb at 20 mph (3 HP) for 45218 on test run 148/2 on 12 January 1950. Corrections recorded were both positive and negative, so the shift was not always forward as might be expected from a locomotive trying to break free from its tethers. By this time, whenever the dashpot was operating with oil, the test sheets also record a 'differential pressure' correction recorded by a manometer. This first appeared in the record for test run 128 on 9th November 1949 with WD 2-10-0 73788. This provision did not appear on the test sheet for run 126 five days earlier (no oil). The manometer, apparently appearing in the interim in an attempt to correct for the wayward behaviour of the dashpot damper when operating with oil. The damper was not given up readily, not only was it seen as potentially of operational benefit, it had become an intellectual challenge. Various combinations of by-pass and pump pressures up to 15 psi were tested or with the pump not running. This produced a variety of outcomes with both positive and negative corrections indicated; the highest discovered was – 1,587 lb at 45.7 mph (-193 HP) on test run 130, 10th December 1949. The day before at a similar speed the correction was +779 lb (95 hp). In both instances no mediating gear correction was required.   When not filled with oil there was a fixed drawbar pull correction of +60 lb, to allow for the non buoyancy of the dashpot pistons.

The apparently satisfactory situation with the dashpot emptied of oil notwithstanding, intermittent dashpot tests occurred for some time, as new ideas, tweaks and different types of oil of were tested to no avail. In the end a satisfactory solution appears to have defeated the best brains at Rugby, the Derby research department and the manufacturers Heenen & Froude.

The visit to the NRM archive in September 2018 produced some interesting material, and significant dates. .

Dashpot Removal

A test sheet for Black 5 44862 12th December 1950 was revealing. The significant point being that the items recorded no longer included any corrective adjustments for dashpot "differential pressure", as when the dashpot was still in use following experimental modifications, or compensation for "buoyancy" when operated filled with air; such adjustments being as included in the test sheets earlier that year. The absence of these tabulations is taken as evidence the dashpot was no longer in operation, confirming Jim Jarvis's recollection that he "thought it was eventually removed" A letter to the Railway Executive dated 15th January 1951 headed Damping Dashpot Investigation confirms this, it begins: "In connection with the experiments in hand to establish streamline flow of the oil, it has been decided to transfer the experimental equipment, rigged at Rugby, to Derby, where greater resources are available and more continual attention can be given."

44862 Test Run No. 422 12 December 1950 15% Cut-Off - Part Regulator

MPH	Pull from Work Lb	Med Gear Correction	Corrected Pull Lb	WRHP	SC PSIG (Approx)	Superheat (Approx)
73.5	1200	-20	1180	231	133	550
67	1450	0	1450	258	132	540
62	1700	0	1700	282	133	540
57	1900	0	1900	289	133	525
52	1980	0	1980	275	132	510
46	2140	0	2140	263	132	505
42.6	2340	0	2340	266	132	505
36.6	2860	0	2860	279	134	510
31.5	3200	0	3200	269	137	515
27	3615	0	3615	260	141	515
22	4195	0	4195	245	148	515
16.8	4820	0	4820	216	153	510

At this stage of development the test reports omitted details of steam rate, making the outcome impossible to cross-check for specific steam consumption and other comparisons. The results of this low power test are nevertheless not without interest when plotted as below.

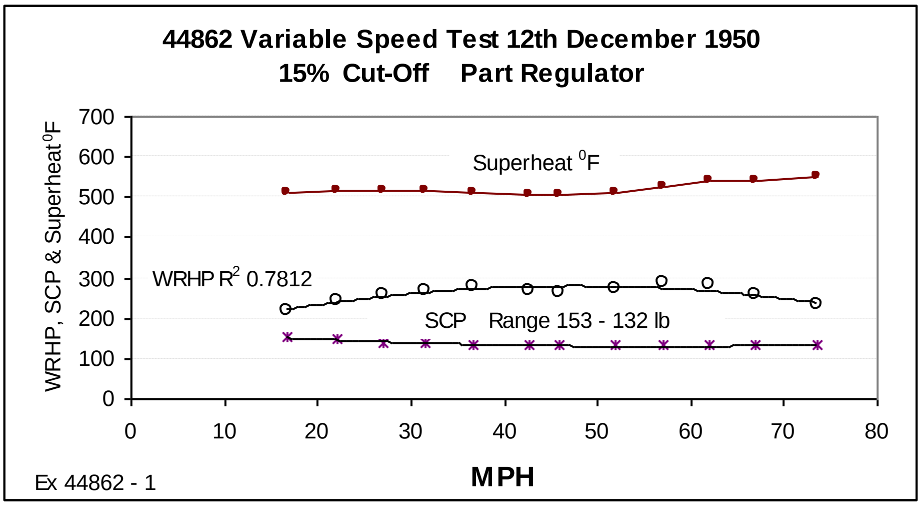


Figure 1 A power sensitivity to superheat appears apparent across the middle speed range. Note the sixth and seventh WRHP plots. The plot progression appears well behaved, free from any deviant changes.

Theoretical Critical Speed Calculations.

A calculation sheet dated 16th April 1951 examines the theoretical critical speeds for impending tests with the *Britannia*. The scope of damping considered ranged from no damping whatever, up to 10 pairs of Bellville Washers. It is evident that the critical speeds occur at the bottom end of the speed range, that speed decreasing as additional washers are brought into play. I have plotted the results in Figure 2 below. The Amsler dynamometer could function over 3 ranges of force; up to 12,000 lb, 36,000 lb and 96,000 lb. Only the two lower scales were considered for this exercise, and it seems likely the highest scale was seldom deployed. It emerges that critical speeds over the speed range encountered on the plant (to over 100 mph on the Duchess tests) was primarily a function of the uneven traction forces, most notably for 2 outside cylinders, and not as the result of dynamic imbalance at speed. The critical speed could be arranged to occur well below the planned test range and would be quickly passed as a locomotive got into its stride under low power at the start of a test. This contradicts John Knowles numerous suppositions and assertions as to how the damping must have malfunctioned, had not been adjusted to suit circumstances and so on. The dynamometer was not existing under constant risk of damage or even destruction, the damping arrangements did not screw up the test results (more on this below). Obviously commissioning and operating a complex test plant was to some degree beyond the experience of the engineers, and they would be treading a capricious learning curve along the way, but the problems were tackled with due diligence and they were not making the supposed oversights and basic mistakes that have been inferred. Please note I am not saying the plant and its operation achieved a state of perfection. How could it, given the inevitability of the metrological limitations, the extensive and varied instrumentation, and the mischief of small remainders.

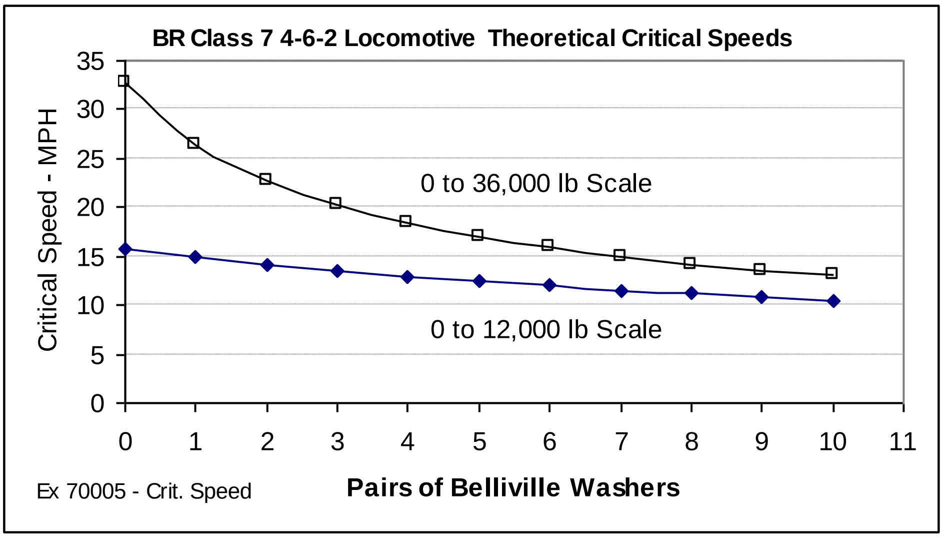


Figure 2 Plot of Rugby calculation sheet 16th April 1951.

Amsler Calibration Tests

Later that year on 28th November 1951; "The work done integrator was checked by pumping up a predetermined load on a National Physics Laboratory (NPL) standardising box and winding through a set distance on the recording table.

The recorded drawbar pull showed negative deviations at a pull of 2 or 3 tons and positive upwards of 8 tons, exceeding 1% positive over 20 tons, which was outside the tractive powers of any locomotive tested on the plant. It was noted that 1679 revolutions of the Amsler speedometer drive disc equalled 5277.37 feet travelled and 1680 equalled 5280.52 ft. In other words, over a mile (1680 revs) the distance error was 1 in 10,000. Below an abstracted data summary from the calibration test excluding data for pulls of over 20 tons (1.157% high at 40 tons). The work-done integrator was checked by pumping a pre-determined load and winding through a set distance on the recording table. This showed the recorded work done 1% high compared with the figures obtained from the standardising box.

This last observation passed without further comment, perhaps because 1% was within the Amsler guarantee. If systematic it would represent +10 HP per 1000 WRHP; 188 lb at 20 mph falling to 54 lb at 70.

Dead Weight Calibration of Amsler Dynamometer Table against NPL Standardisation Box
28 November 1951

Load Tons	2	3	4	5	6	8	10	15	20
Error %	-1.41%	-1.16%	0.0021%	-0.117%	-0.117%	0.021%	0.546%	0.205%	0.021%
Error Lb	-63	-78	0	-13	-16	4	122	69	9
MPH	70	60	50	45	35	30	20	20	15
HP Error	-12	-12	0	-2	-2	0	8	4	.0.4

Only the first two lines are as documented, I have added some notional speeds on the basis that the lower the drawbar pull the higher the speed, in order to give some inkling of the WRHP error magnitudes that would occur given the percentage errors indicated.

There were further calibration tests in 1953, 1955 and 1957. Remedial maintenance and refurbishment work to the Amsler integrator mechanism and mediating gear resulting from wear and tear was carried out from time to time.

1953 & 1955 Amsler Dynamometer Calibrations

	Work Done	Correction 1953	Correction 1955
12,000 lb Scale	6,000 lb	N/A	-0.1%
	12,000 lb	N/A	-0.75%
36,000 lb Scale	12,000 lb	N/A	-0.23%
	18,000 lb	N/A	-0.75%
	Scaled Pull	**Correction 1953**	**Correction 1955**
12,000 lb Scale	6,000 lb	+1.87%	-0.57%
	12,000 lb	+0.125%	-0.06%
36,000 lb Scale	12,000 lb	+0.71%	-0.4%
	18,000 lb	0	-0.1%

May-June 1967 Amsler Dynamometer Calibration

The report summary took a different form to the earlier reports. The calibration of the Dead Weight Tester indicated the actual pull was 285/286 of the calculated pull, a correction of - 0.35%.The Work Done integrator error was 361/360, a correction of +0.27%

Indicating Developments

The early commissioning phase gave little attention to cylinder indication, though ultimately of importance, such measurements were not integral with the functioning of the plant test bed and dynamometer. During the various interregna when the commissioning of the plant dynamometer was halted for one reason or another, the opportunity was taken indicate D49 62764 with Reidinger poppet valve gear and Capprotti Black 5 44752 in 1949. I have no experimental data for these tests. Perhaps, with an eye to the forthcoming BR Standards, it was done to discover if poppet valve gear potentially offered a better way forward. The first locomotive on the plant after the first commissioning phase was 45218, undergoing 137 test runs between 3rd January and 19th May 1950. This early post commissioning phase in the history of the test plant could be dubbed the “working up phase” which lasted about another two years. 45218 only appears to have been indicated during its last few days on the plant, notwith- standing that the tests were investigating the effects of changes in lead. Such determinations were evaluated by the changes in the recorded WRHP. As the official report notes: *“Unfortunately, no consistently reliable indicator cards were obtained either from the Farnboro indicator which is still in the process of adaption to work on a* *steam locomotive, or from a borrowed Crosby indicator, so that no assistance could be obtained in this way to explain the somewhat irregular sequence in the rates of consumption for the various leads. As all the above mentioned curves are intended only for comparison with one another they have been left on a basis of horsepower at the wheel rim."*

The tests with 44765 comparing the efficacy of single and double chimneys and the steaming tests with B1 61353 have handed down WRHP and boiler performance only, though a note in the correspondence mentions that the B1 was indicated at the end of the final test series, recording very low or negative machinery friction (no data available). The data base boiler performance for 44765 and 61353 is poor in regard to specific evaporation rates (lb/steam per lb coal). It is concluded that the steam rates given in the data base are in fact the feed water rates only, and that the exhaust steam injector was in use. The steam temperatures reached support this view. This is known to be the case in regard to 61353; it says so in the test bulletin, but only in passing. The true steam rates were therefore about 6 to 6.5% higher than shown in the data base up to the ESI limit around 20,000 lb/hr.

Indicator shortcomings notwithstanding; 45218 was indicated for its last few days on the plant. The data base I am working from has no data on this, an internal report (20 May 1950) gives some details: "*In order to attempt to isolate the apparent error in the Farnboro attention focussed on the LH cylinder exclusively (to which the Crosby was fitted) and a number of diagrams taken with a Farnboro element while indicating by the Crosby."* The initial results with the Crosby showed a mechanical efficiency of 0.95, - with some lapses to 1.02." Some experiments concluded that the Crosby indicator was subject to a phasing error caused by the length of pipe between indicator and cylinder. Reducing the pipe length in stages. Eventually the Crosby MEP results were *"sensibly the same as the Farnbro element".* Both were *"less than the measured Amsler drawbar figures and therefore the latter also are in error to the extent of about 12%. The Rugby (Farnbro) indicator appears to be correct. Action. Indicate the Amsler cylinder as originally suggested many weeks ago."* The actual report the previous day put the probable error between 7 % and 10%.

It took over a year to organise such tests. A letter dated 8th August 1951 refers to "Dynamic Calibration Of Amsler Dynamometer" involving 61353, The last B1 test was a week earlier on 1st August. On what appears to have been an adaption of the Farnboro indicator, the peak and minimum hydraulic pressures of the dynamometer were monitored and compared to the recorded WRTE test value. There was no attempt to integrate the monitored readings into WRHP on a work done basis. More details of these tests on page 93 below.

Comparison of the WRHPs recorded at this stage with later periods, *when positive MFs were being routinely returned*, does not support the idea of WRHP errors as high as 12%, since the overlapping WRHP Willans Lines were closer or similar across time.

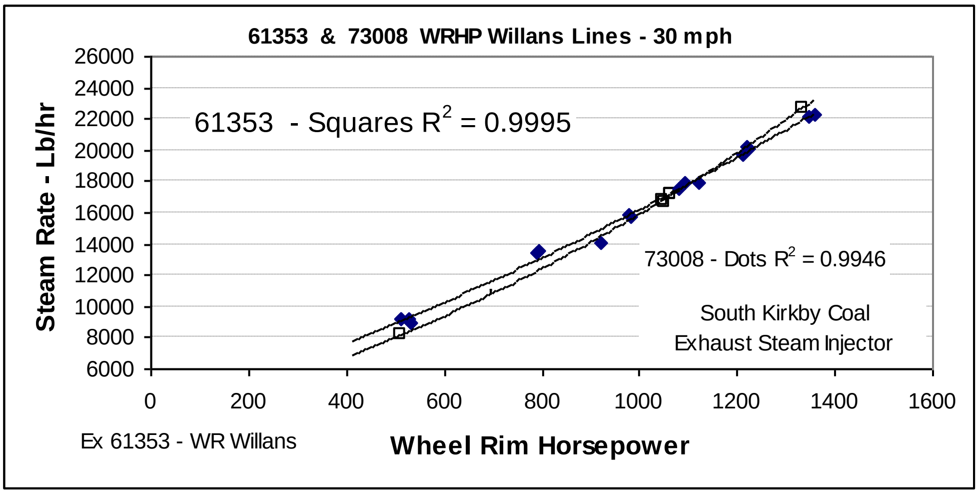


Figure 3. Diverging overlap with mid-range agreement. ESI contribution assumed at 6%.

Some lfurther comparative indicator tests with 70005 in December 1951 returned results for the Crosby (LH cylinder front only) averaging 2.8% below the Farnboro' (16 plots). Presumably the Crosby pipe set-up was along the lines developed for 45218. The conclusion in May 1950 that the Farnboro' indicator "appears to be correct" is put at odds to some extent by later IHP Willans Line outcomes for the Britannia which improved over time. In example the 40 mph IHP Willans Lines from the Rugby data and Test Bulletin at a steam rate of 20,000 lb/hr yield the following results.

	IHP	Index
70005 1951	1374	100
70025 1952/53	1420	103
Bulletin No.5 - April 1953	1445	105

It would be misleading however to conclude that this level of increase applied uniformly across the full speed and power range portrayed in the test bulletin. In contrast to John Knowles claim that the Rugby IHP data was "consistent", detail scrutiny of the IHP data for 70005 and 70025 reveals disparities at times verging on the chaotic, a situation applicable to some of the IHP data generally. The second test series for 9F 92050 showed a measurable decline in cylinder efficiency compared to the first; the WRHP reduced accordingly. In his case the change was real enough, attributable to steam leakage as traceable by exhaust steam temperature and pressure changes.

Correspondence from Ron Pocklington, the test engineer intimately involved with the operation and development of the Farnboro' "balanced pressure" indicating equipment reveals shortcomings in regard to reliability and performance in its first years of operation: "We used to get semi or complete snowstorms before an improved spark generator was obtained (1954). I endeavoured to sort it out to become reliable and precise, including an accurate assessment of the dead centre as a reference and the compilation of the stroke diagram and its IHP determination. If this is not carefully done then a direct fattening up, or down of the stroke based diagram appears." This level of reliability and performance was not the situation as he first found it when he started work at the plant at sometime in 1952.

The case made for correcting the Crosby result in 1950 was straightforward and persuasive. However; *"….the Farnboro' element had in effect been used as a stop watch to time the delay of the pipe line and as such had measured a delay of the time lag as about 4 milisecs.*" This effect fattened the Crosby indicator diagram. This assumes the Farnbro was accurately plotting stroke dead centre at the then stage of development. Commenting on the indicator diagram in the test plant brochure (70005
Test Run 665, 1.12.51), Ron Pocklington observed: *"If you look at the slide bar contact marks you will see some wobble due to slackness in the universal coupling to the indicator drum*." Written communication.

The Farnbro. "balanced pressure" indicator encounters some intrinsic "lag" in another way. It operates on the principle of those coloured tinplate clicking novelties popular in Christmas crackers. A shallow dish pressed into the tinplate makes a click when the dish is reversed by pressing on the convex side. The so called "balanced pressure" Farnbro indicator requires a finite pressure differential to operate. This is defined as the "lag", and ideally should be of very low magnitude. The contact with the diaphragm as originally set up at Rugby was spring loaded, this will have introduced a slight increase in the degree of "lag" when breaking contact. The final improvement of the Farnboro' indicator was achieved by the simple expedient of substituting a fixed electrical contact for a spring loaded one. *"One element was fitted with a new arrangement of centre contact and it was soon found this produced the standard of diagram so long sought after. No scatter was apparent even at the highest speeds."* This was early on in the Duchess tests starting at the end of January 1955. Quite late in the day, in the history of the plant. This outcome makes sense; a spring loaded contact would slightly delay circuit interruption and the spark generated pin holes that formed the diagram. The spring loaded contact was effectively minutely increasing the system lag by delaying contact separation and spark generation.

Progress achieving positive IHP-WRHP relationships is mapped out below in Figure 4.

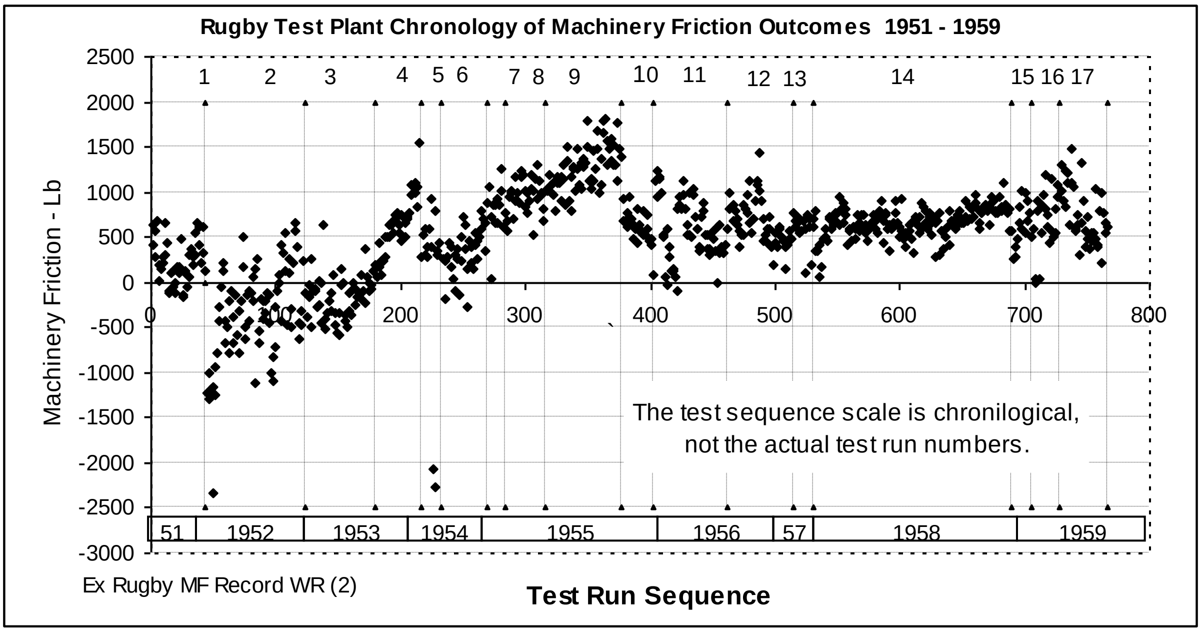


Figure 4 Earlier WRHP data available for 45218, 44765, 61353 and 70005 lacked any correspond- ing IHP data. The numbered data sets are indentified in the table below. 1953 was something of a watershed year since from that point, negative MF outcomes only rarely occurred, at a rate predicted by random number experiments. There were a number of developments and improvements in 1953 of which more later.

Absent through lack of data are further tests for 35022 with a single chimney following on from 70025 in March 53 (26 test runs), and again later that year after 73030, and 70025 (5 demonstration runs) for tests without thermic siphons (36 test runs), Also absent is data for two test series with Crab 42824 fitted with Reidinger poppet valve gear, following on from 70025 at the end of 1953, and later after 46165 in June 1956; 47 & 56 tests respectively. EE GT3 tests occupied much of 1957.

Key to Figure 3

Ref	Locomotive	Ref	Locomotive	Ref	Locomotive
1	73008	7	42725	13	92050
2	35022	8	46225	14	73131
3	70025	9	92023	15	92166 Stoker
4	73030	10	92050	16	92250 D/C
5	42725	11	46165	17	92250 Giesel
6	92013	12	45722		

It seemed that the tests starting with 73008 in April 1951, imperfect though they were, with mixed MF outcomes, represented the dawning of some light. It was to be a brief victory of sorts, the tests that immediately followed with 70025 represented a serious relapse, which only became worse when with the turn of 35020, which proved to be something of a law unto itself. Somehow, when 73030 put in an appearance in July 1953, things seemed to be on track.

During this period the Farnboro' indicator equipment underwent many modifications as recorded in official correspondence and private communications from Ron Pocklington. This included several modifications to the spark generating circuitry, the diaphragm material, and the spring contact set-up prior the adoption of a fixed contact. The changes were driven by frequent failures of the spark circuit, cracked diaphragms and an ambition to reduce chronic scatter. In its final form the diaphragm could be operated “with a breath”. At operating temperatures this sensitivity may have been slightly reduced. Some of the changes along the way may have had a retrograde

outcome. This could explain some of the set-backs as evidenced by the see-saw nature of both the early MF outcomes and apparent IHP variations. Figure 5 below, though representing some progress, is not without its obvious imperfections.

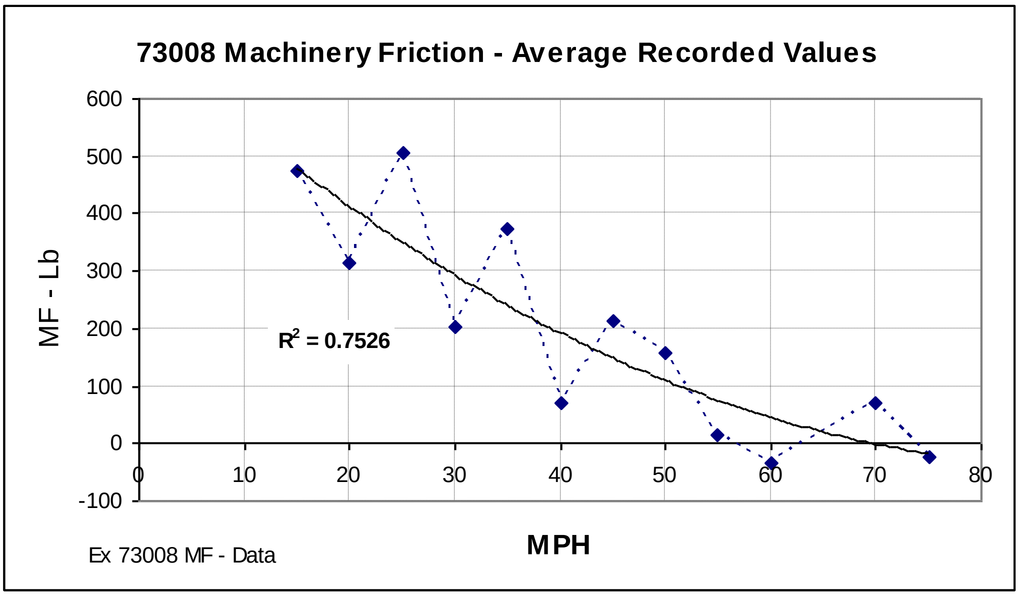


Figure 5 Here the scattered MF outcomes for the speed sets have been averaged and plotted against speed. The overall trend, clearly and illogically, is saying that MF is an inverse function of speed. However, when the plots were joined together, note how the resulting zig-zag trace follows the overall falling trend. As randomised number experiments have shown, speed data sets may cluster to form high and low biases as evidenced here.

Some degree of the scatter is ‘true’ in the sense that small variations in steam pressure and temperature will influence the result

When the 73008 MF outcomes are examined in order of sequence a different picture emerges. MF data was late to emerge in the test programme, since the Rugby test team had little confidence in mechanical indicators, and post commissioning, cylinder indication was largely absent from the early test programme as tabled below.

Rugby Test Plant Programme & Data Record 1951-53						
Engine	**Test Runs**	**Dates**	**IHP**	**WRHP**	**MF**	**Notes**
61353	449-508	15.1.51-30.3.51	-	25	-	
70005	509-543	17.4.51-28.5.51	37	-	-	1st Application Farnobro' Indicator
61353	544-589	7.6.51-1.8.51	-	26	-	
73008	590-657	13.8.51-5.11.51	-	50	-	
Amsler Calibration 28th November 1951						
70005	658-691	3.12.51-3.12.51	41	9	-	
73008	692-714	30.1.52-21.2,52	35	65 #	35	
35022	715-821	19.3.52-2.10.52	75	133	74	
70025	822-895	31.10.52-20.2.53	67	63	47	
35022	896- 923	10.3.53-7.5.53	-	-	-	Single Chimney Tests
73030	924-1022	22.7.53-3.11.53	35	94	35	$5^{1}/_{8}$", 5", & $4^{7}/_{8}$" Blast Pipe Caps
70025	1023-1027	25.11.53-27.11.53	-	-	-	Demonstration Runs
35022	1028-1063	5.12.53-25.1.54	-	-	-	Without Thermic Syphons Tests
		Total	290	465	191	
# A few test runs at miscellaneous speeds omitted						

It was not until April 1951 the Farnboro' indicator was available for testing with the initial trials of 70005. Following these tests, there was a 6 month interlude before indicating was tried again, presumably to deal with development problems that had emerged regarding the electrical circuitry and diaphragm durability. As a consequence the first test series with 73008 was not indicated. Cylinder indication for the second test series starting in January 1952, was confined to 35 test runs. When sequenced, the MF outcomes fell into two distinct groups: the 1st group comprising 21 test runs included 7 negative MF outcomes with an overall average of 95 lb; the 2nd series of

14 runs was free of negative outcomes, with an overall average of 411 lb. The specific IHP steam consumptions for the seven negative MF outcomes were all significantly high when plotted against the BR5 test bulletin IHP SSC Willans Lines as indicated in Figure 6. The implication being the IHP was under-recorded.

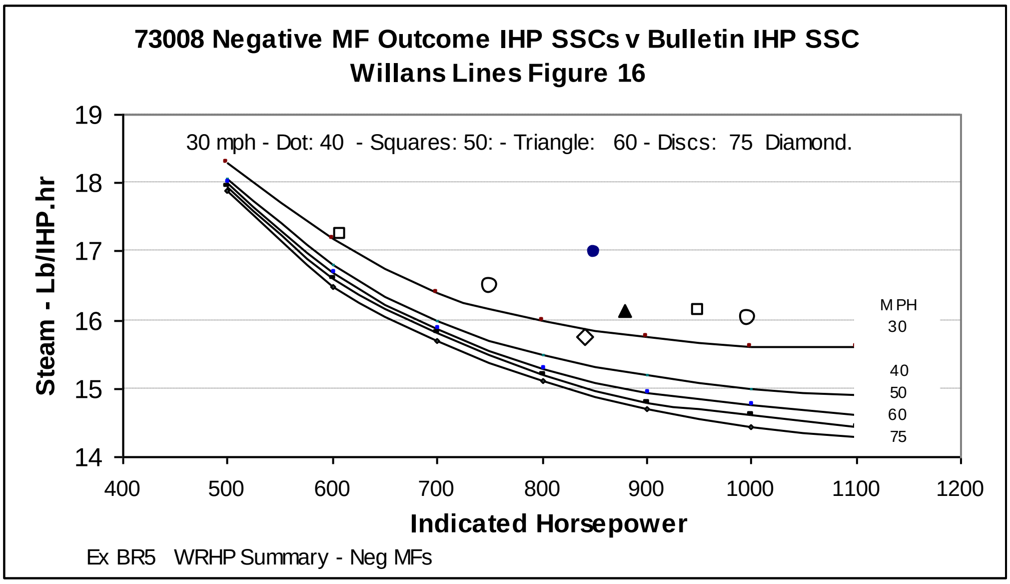


Figure 6. All the IHP SSC plots, as associated with negative MG outcomes, fall significantly above the related speed IHP SSC Willans Lines.

Overlapping test data for the 73008 and 73030 test series when both were fitted with 5.125” blast pipe caps is limited to WRHP data at 35 mph with 12 and 15 plots respectively, as plotted in Figure 7. The available overlapping IHP data is minimal.

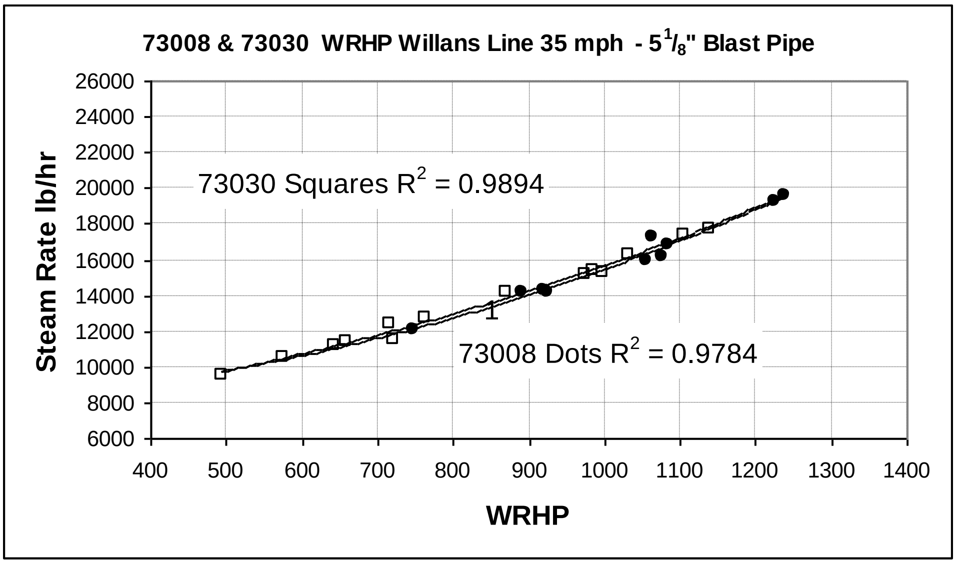


Figure 7 The 73008 plots include examples from the initial test series in the latter part of 1951 and the later tests early in 1953. The 73030 tests were in the second half of 1953. The slight Willans lines separation falls within the guaranteed dynamometer accuracy. Combining the plots returns an R^2 value of 0.9905.

In late July 1951, some 15 months after the 45218 tests, when it was proposed to
“Indicate the Amsler cylinder as originally suggested many weeks ago”: the decision was enacted upon for the last few tests with B1 61353 (report dated 8th August 1951).

“*The discrepancies between the WRHP and the IHP obtained from the ER B1 Class Engine No.61353 has caused further investigation into the accuracy or otherwise of the Amsler measuring equipment. A differential pressure element has been made at Rugby, and after a very limited attempt to calibrate same inserted into the Amsler dynamometer cylinder”.*

The report included a note of caution. “*As stated earlier, calibration of the element was found very difficult in view of the limited facilities available for pressure calibration at Rugby Testing Station. And the result obtained should be treated with the utmost* caution. since an error of 1 lb in the gauge used in the air side will cause a resulting error of 114 lb on the pull." A diagram of the apparatus has not been found.

61353 Amsler Indicator Calibration Test - 25% Cut-Off - August 1951					
MPH	Recorded Pull - lb	Indicated Maximum Pull		Indicated Minimum Pull	
		Maximum	Minimum	Maximum	Minimum
20	11,300	10,600	10,070	10,200	9,660
20.25	11,930	10,600	10,070	10,370	9,870
29.7	9,810	9,360	8,840		
40.5	8,850	8,420	7,910		
60.9	7,495	7,100	6,580		
60.9	7,505	7,100	6,580		

The “peak” calibration indications averaged only 95% of the recorded pull of the Amsler. The peak value should have been higher since the recorded pull was the average value. On an average of the maximum and minimum pulls, the indicated results were only 90% of the Amsler. No explanation is given for the absence of “Indicated Minimum” pulls above 20 mph. It may be that the differences were insignificant at the higher speeds. As Lomonossoff pointed out*, the flywheel effects of the coupled wheels and motion smooth out the fluctuations in turning moments such that they “*cannot perceptibly vary its speed*”. It is therefore, difficult to model the drawbar pull profile per revolution directly from the simultaneous MEP pressure record of the four cylinder ends as recorded in these tests.

Obviously the results of these tests are problematical, at face value supporting the suspicion that the Amsler dynamometer was at fault. The problem remains, that later results, when positive MF outcomes were being returned, no change in the measured
WRHP obtaining when negative MF values were endemic is obvious: vide Figure 7.

It is perhaps not without interest that among the improvements listed in 1953, were improvements to the Farnboro’ Indicator diagram converter. *“A new crank and connecting rod with ball bearings were fitted and the base board stiffened up. Following the successful improvised drive by a meccano electric motor, a permanent Hillman motor was obtained and a gearbox assembled at the plant.”* .

Pocklington was not impressed with the situation as he found it when he arrived on the scene in 1952, citing among other things, the difficulty in establishing the true ‘dead center’ for the Farnboro’ radial indicator diagrams.

A situation further complicated since the dead centres for the cylinder front and rear power strokes occur at different, crank angles, having to accommodate for cylinder thickness.

Notwithstanding the apparent indications of dynamometer malfunction as manifest in the Crosby/Farnboro' tests with 45216 in 1950, and the calibration experiment with 61353 in 1951, the WRHP outcomes seem little changed over time, notwithstanding that MF outcomes had become positive in the interim, as exampled in Figure 7.

I have looked into the effects of dead centre error, converting a sample Rugby indicator diagram for one cylinder front end to a stroke base, then repeating the exercise, first with 'dead centre' moved $^{1}/_{32}$" to the left, then $^{1}/_{32}$" to the right ($^{1}/_{896}$ of the stroke).

70005 40% Cut-Off - 20.28 mph Potential IHP 'Dead Centre' Error Effects			
Item	As Diagram	1/32" 'Early' Admission	1/32" 'Late' Admission
MEP	144.9	146.84	142.0
MEP Index	100	101.4	98.0
IHP	1149	1165	1126

…………………………………………………………………………………………………

* Introduction to Railway Mechanics , G Lomonossoff, Oxford University Press. 1933; page 105.

The calculated 1149 IHP assumes equal MEP for the four cylinder ends which is of course contrary to the actual case (1125 IHP). The tests at Rugby routinely followed a lwarming up period to stabilise any thermal effects on valve setting and dead centres.

The IHP test data from 1951 to early 1953 involving 70005,73008, 35022 and 70025 falls someway short on consistency, at times, things seem to have been going backwards. Starting with the BR7, the tests with 70005 and 70025 thread different paths when plotting Steam Rate v Speed & Cut-Off. In relative terms the two paths shown, Figure 8, are likely real enough, the difference are probably attributable to the subtleties of valve setting. Valve setting, long held as something of a black art, often with secretive ideas as how to best do it, provides scope for different outcomes. Some careful thought and experiments on thermal expansion allowances are said to have reduced Britannia water consumption on the Great Eastern section by about 12%.*

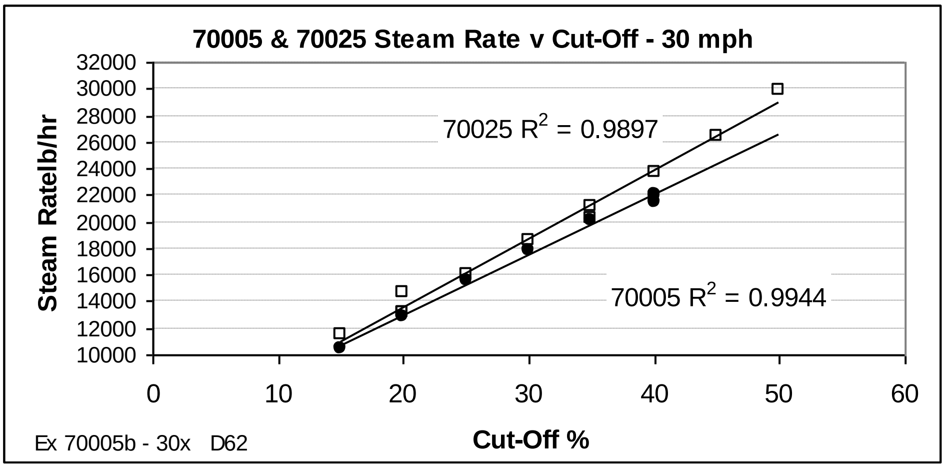


Figure 8 The trend for 70025 is the basis of the test bulletin cut-off curves; Figure15.

The recorded WRHP data for 70005 was not simultaneous with any IHP data, so there is no direct MF record. The comparative WRHP Willans Lines for 70005 & 70025 at 40 mph are plotted below. The 70005 XL extrapolation beyond 1400 WRHP is unreliable.

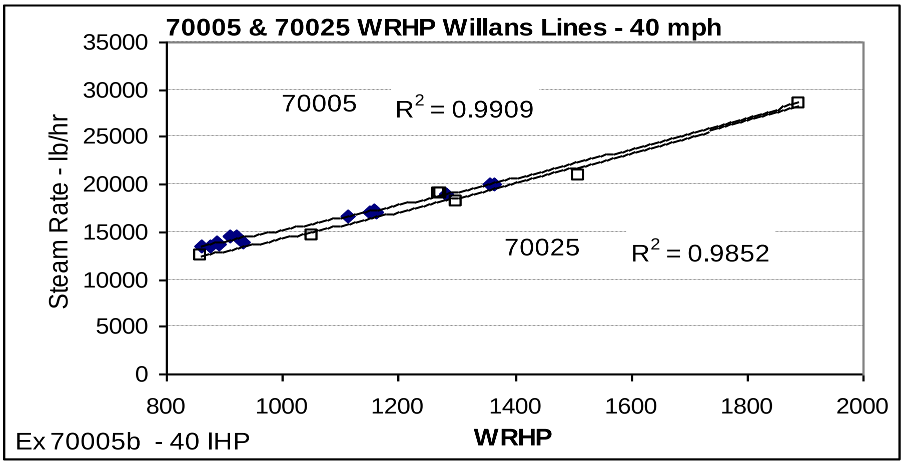


Figure 9 Unlike the WRHP data above, the 70025 IHP data features wide scatter when plotted on a specific steam consumption basis; R^2 0.2964. The data base at 40 mph lacks any coal rates and is endorsed "LSI assumed" (Live Steam Injector). In the absence of firing rates it's not possible to cross check this by calculating the specific evaporation rates Assuming the ESI was applicable to the outlying plots brings them into line. It is not possible to verify such changes

Merchant Navy 35020 treated the Rugby test team to a harvest of negative MF outcomes and one or two idiosyncrasies. One example was the dip in indicated horsepower at 24 mph as speed increased at cut-offs between 10 and 20%. A
similar eccentricity was evident when 35005 was road tested with a mechanical stoker in 1950. In this instance the dip was at 20 mph between 15 and 30% cut-off,

…………………………………………………………………………………………
……..

* Bill Harvey's 60 Years In Steam, D W Harvey, David & Charles, 1986; page 202.

The one uncertainty 35022 did avoid was the use of an exhaust steam injector, since none were fitted. In that regard, at least the data base steam rates are unequivocal. Some of the IHP data is clearly aberrant in character, with no potential explanation on the grounds of exhaust steam injector participation or lack of it. Said aberrations are best seen when the data is examined in enlarged form; that is IHP and WRHP specific steam consumption, as Figure 8.below. Following on is an orderly set of WRHP Willans lines for 15, 20, 30 & 40mph - Figure 10.

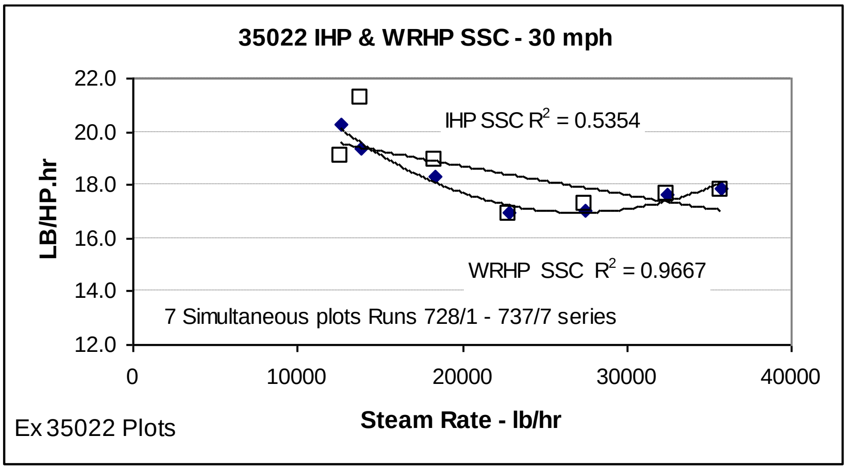


Figure 10   The IHP & WRHP plots are clearly in collision, as was endemic at this stage of development, but, unlike the IHP trend line, at least the WRHP curve is the right shape, and returns a respectable R^2 value. A similar exercise for 40 mph delivered a similar result. Removing the low LH IHP SSC plot, clearly an outlier, delivers a concave trend line,

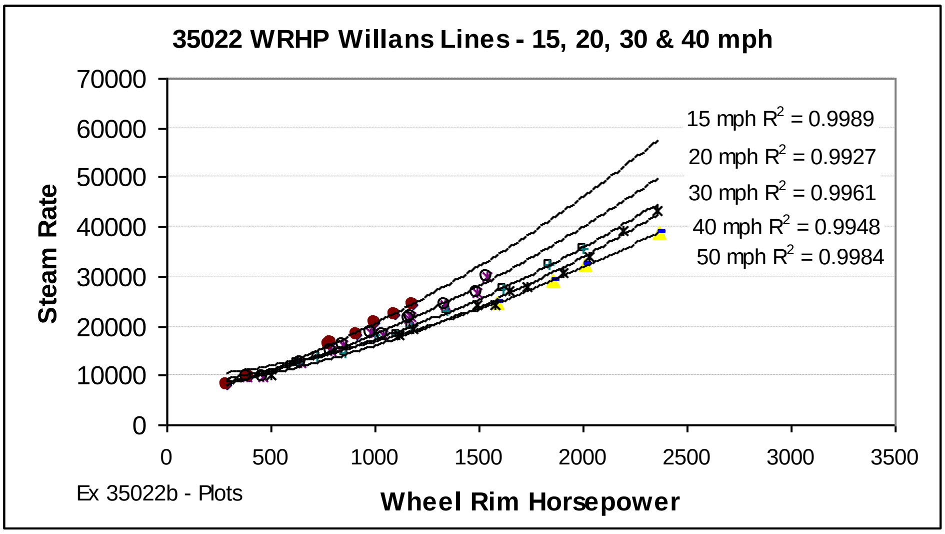


Figure 10   The orderly pattern as a function of speed and power follows the intrinsic characteristics of reciprocating steam. The equivalent diagram for the indicted horsepower is equally orderly at this level of magnification. The problem was the IHP/WRHP data at this stage of development was mostly in collision, with over 80% of the MF outcomes returning negative values. The recorded cylinder efficiency was about 12% low compared to a Duchess.

Mechanical Efficiency

Mechanical Efficiency is a simple relationship: Mη = WRHP/IHP or WRTE/ITE

Firstly, a look at the combigned raw MF data for stoker fitted 9F 92166 and 92250 in double chimney and Giesel ejector guise reveals wide scatter, a 'high' bias at 40 mph and a vestigial R^2 value, as evident in Figure 12 below. Some of said scatter is real in the sense that it reflects variations in effort. When re-plotted in mechanical efficiency form as Figure 13, the scatter is much attenuated, the 40 mph bias reverses, falling generally in line with the overall trend against speed, and the R^2 value, though remaining medioccre, is significantly improved.

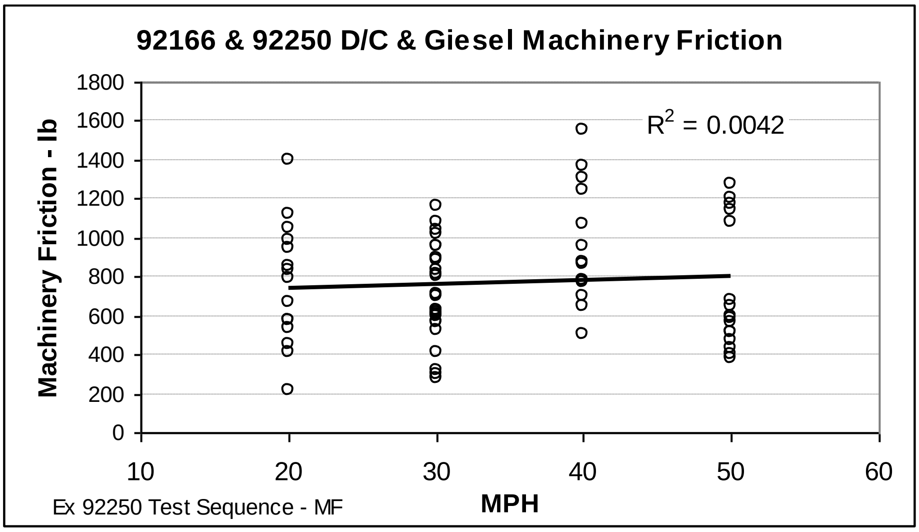


Figure 12. Wide scatter and some random bias as seen here is an inherent characteristic of small remainder data sets.

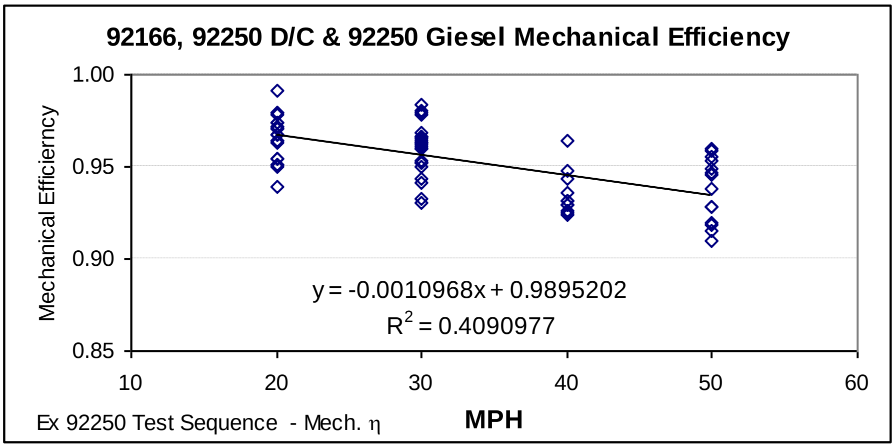


Figure 13. Expressed in Mech.η form, the Figure 12 data assumes a more orderly outcome with an unequivocal overall trend.

A similar exercise for the two 92050 test series produced a similar result – Figure 14

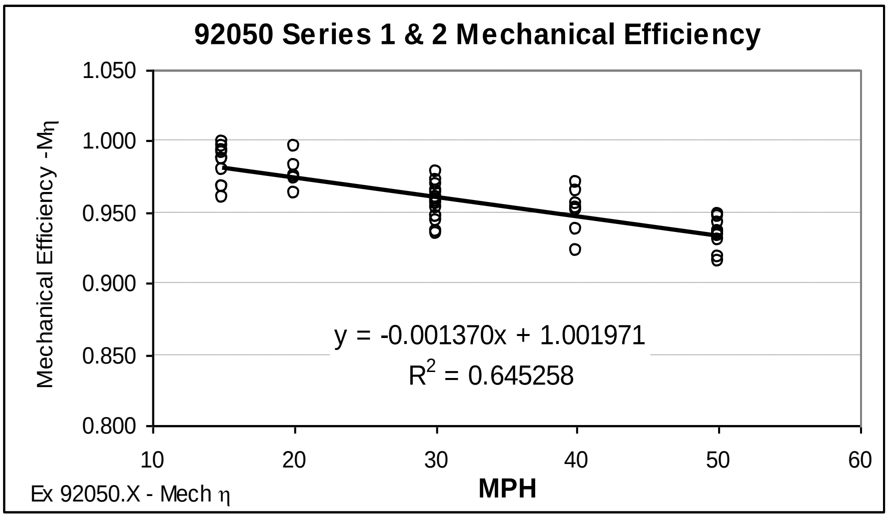


Figure 14. The overall trend and Mech.η values are similar to Figure 13.

The mechanical efficiencies for 92050 and 92166 & 92250 derived from Figs 13 & 14 are tabled below, they fall within +/-1%.

92050 & 92250 Mech. η			
92050 $y = -0.00137x + 1.001971$			$R^2 = 0.6453$
92250 $y = -0.0010968x + 0.98952$			$R^2 = 0.4091$
92050 & 92250 Mechanical Efficiency			
MPH	92050	92250 *	Δ Mech.η. 050 v 250
15	0.9814	0.9731	0.9%
20	0.9746	0.9676	0.7%
30	0.9609	0.9566	0.4%
40	0.9472	0.9456	0.2%
50	0.9335	0.9347	-0.1%
60	0.9198	0.9237	-0.4%
* Includes 92166 runs at 30 mph + 1 at 40.			

At face value the mechanical efficiency formulae as derived in Figures 13 and 14 provide a simple way of plotting WRHP across the speed range as a function of IHP, as exampled in Figure 15 below.

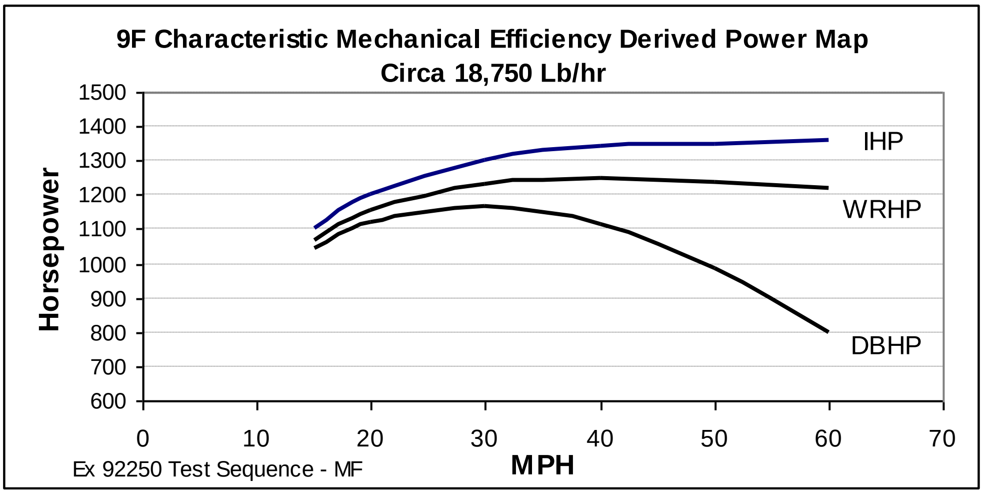


Figure 15. The average steam rates for Figures 13 & 14 data varied slightly for each speed set, the IHP values plotted here have been pitched to the mean rate. The DBHP curve assumes Report L116 Figure 3 locomotive resistance curve. Unfortunately, the Mech.η formulae are only a snapshot representative of the average steam rates obtaining for the available data sets, and cannot be used across the full working range, since the mechanical efficiency improves slightly with the level of effort - Figure 16.

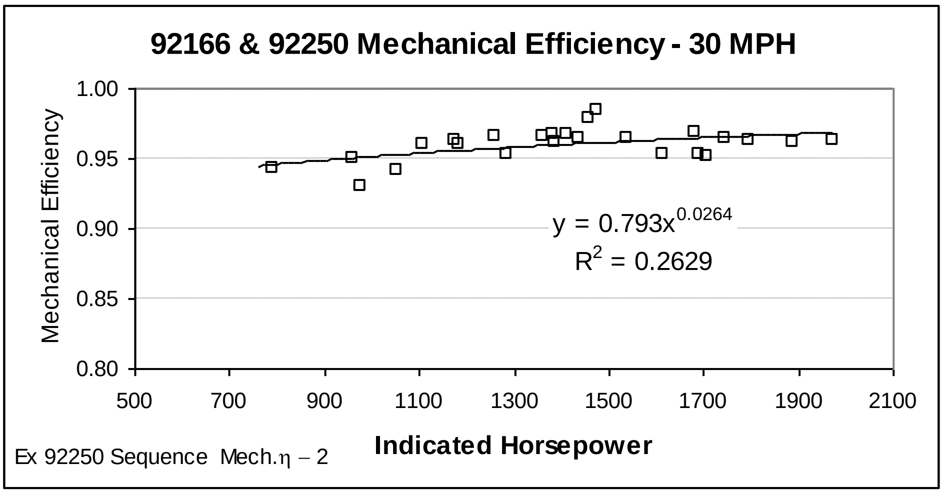


Figure 16 The somewhat scattered outcome and low R^2 value is characteristic of small differences and low rates of change. In this instance the spread is +/- 2.7%.

The small differences in mechanical efficiency for 92050 and 92250 tabled above notwithstanding, they are sufficient to generate significant differences in machinery friction outcomes at a given IHP power output, as tabled below.

92050 & 92250 MF Outcomes v IHP & Speed IHP

MPH	IHP	WRHP		MF LB		Δ MF HP 050 v 250
		92050	92250	92050	92250	
15	1275	1251	1241	592	858	-10
20	1400	1364	1355	668	851	-9
30	1510	1451	1444	739	819	-7
40	1560	1478	1475	773	795	-3
50	1590	1484	1486	793	779	2
60	1600	1472	1478	802	763	6

While in horsepower terms the discrepancies of up to 10 HP appear quite modest,
differences of over 250 lb at15 mph seems less impressive. So here we have equipment performing within the specified uncertainty, while the two WRHP sets at a given IHP and speed within 0.8% deliver measurably divergent MF outcomes.

Such differences fall within the expected range of experimental error, small wonder then, that Carling thought it difficult to confidently plot WRHP and likewise locomotive resistance. It is unlikely that such small differences are entirely down to experimental error alone. Given manufacturing limits and fits and such matters as machinery alignment and lubrication integrity, it does not seem remarkable to suggest that machinery friction for individual locomotives might vary by +/- half a percent, possibly more. Such small differences are more than enough to challenge the test engineer endeavouring to reconcile the divergent data of small differences. In WWII the performance of military aircraft as delivered was found to vary up to 2.5%. This was attributable to power unit variations and airframe quality, the latter having a long list of potential flaws. Obviously the scope for variation with a locomotive running indoors on a test plant is much reduced compared to aeroplanes, and anything serious will quickly manifest itself in the guise of hot boxes and so on. However, as already touched on, test outcomes will be sensitive to valve setting, other things being equal.

WRTE v ITE is Linear

That this relationship is linear is one of few certainties that emerges from the test data. Beyond that, when plotted, the outcome is not always reliable. For given types it appears unaffected by single or double chimneys, the Giesel ejector and blast pipe changes notwithstanding; ITE rules. The fundamental characteristic of the linear relationship is that as ITE increases WRTE increases at some slightly reducing overall rate (Figure 16). Such plots are confined to speeds sets, and if they provide only a few plots covering a limited range of power and steam rate, they sometimes deliver a trend line sloping the wrong way - falling from left to right. Such an outcome implies WRTE still available at zero steam rate. An outcome attributable to the vagaries of scatter.

The linear relationship is simple: $Y = fx – C$.

On occasion, notwithstanding a seemingly adequate number of plots and wide working range, the constant sign turns out to be positive. This again implies power at the wheel rim at zero ITE. This contradicts John Knowles assertion that more data axiomatically provides more accuracy. The reality is that some measurements are more accurate than others, and the sequence of delivery is entirely random. The nth plot might readily bring confusion where relative order otherwise prevailed. A good example is to be found in the data for 9F 92166 – Figure 17. In terms of WRTE v ITE, the outcome was in close accord with the data for 9F 92250, but the trend line constant for 14 tests at 30 mph delivered the wrong sign; WRTE cannot be positive when X is zero.

It took some weeding on a trial and error basis to eliminate the positive sign, the removed plots were randomly distributed – Figure 17B .

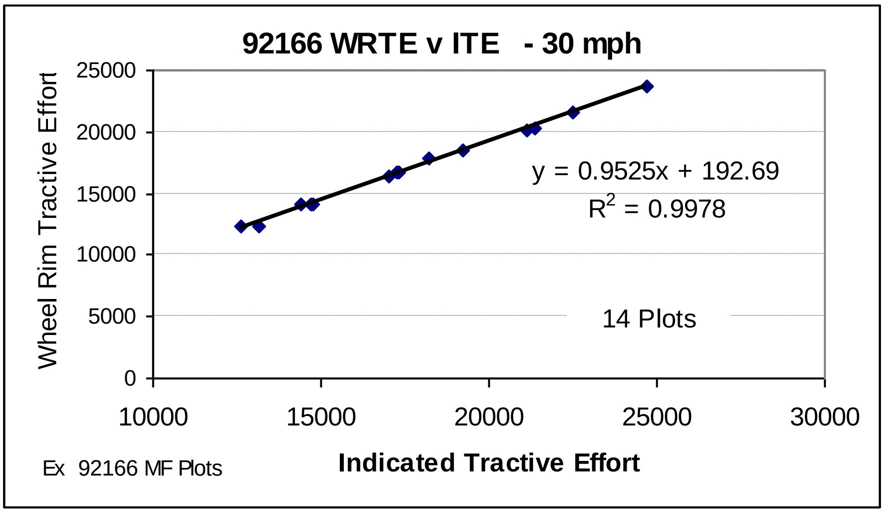


Figure 17 Visibly the scatter is low, as corroborated by the high R^2 value. However, delivering what would be 15 WRHP with the regulator closed is not to be countenanced (positive constant).

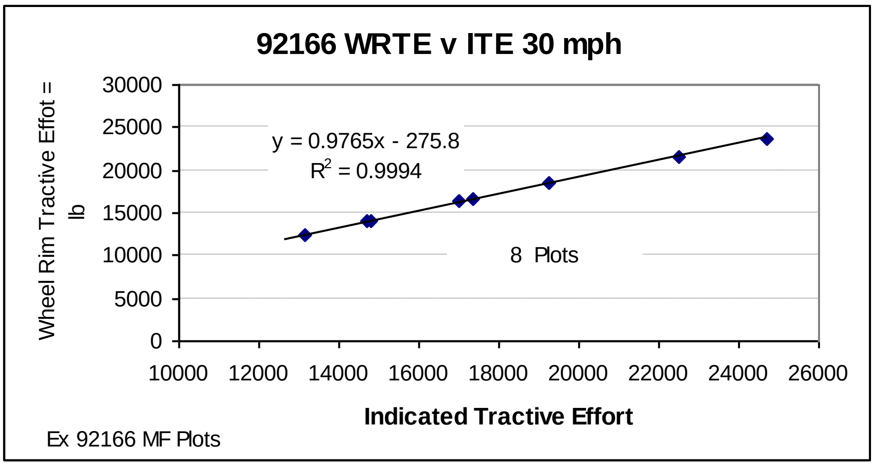


Figure 17B 40% fewer plots delivers a negative constant. Visible scatter reduced,
R^2 outcome improved.

Given sufficient range of output (more important than the amount of data), most WRTE v ITE plots are not troubling in the way of 92166 exampled above. An 'untroubled' example is shown below for 92250 – Figure 18

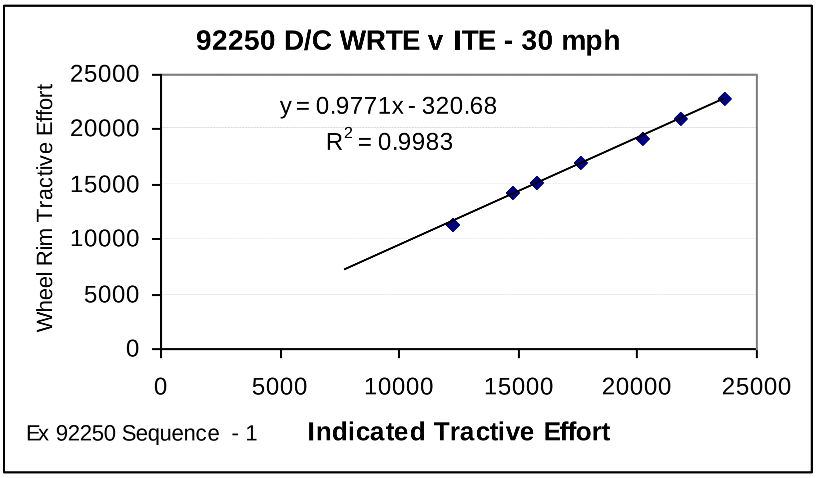


Figure 18 This straightforward relationship notwithstanding, note the slight differences in the x variable compared to Figure17B. This affects the slope of the trend line and thereby the derivation of the constant, which inevitably, will also differ. These small differences are the product of the random scatter, or may reflect slight differences resulting from manufacturing tolerances,. .

Looked at on an indices basis, the differences in the WRTE outcomes for 92166 and 92250 across the power range are negligible, under $^{1}/_{2}$%.

92166 v 92250 WRTE - 30 MPH				
ITE	WRTE		WRTE Index	
	92166	92250	92166	92250
10000	9489	9450	100	99.59
15000	14372	14336	100	99.75
20000	19254	19221	100	99.83
25000	24137	24107	100	99.88

However, when the small remainder problem raises its head, the MF outcomes are inevitably more tangible than a mere half a percent difference would seem to suggest.

92166 v 92250 Machinery Friction - 30 MPH				
ITE	Machinery Friction - LB		MF Index	
	92166	92250	92166	92250
10000	511	550	100	107.61
15000	628	664	100	105.71
20000	746	779	100	104.41
25000	863	893	100	103.46

It is all too apparent that small remainders (SRMs) can make mischief with trivial deviations in the cylinder ITE and WRTE data, *even within the supposed accuracy of measurement limitations.* Figure19 below plots the potential MF deviation ranges resulting from no more than 1.5% SRM compounded error.

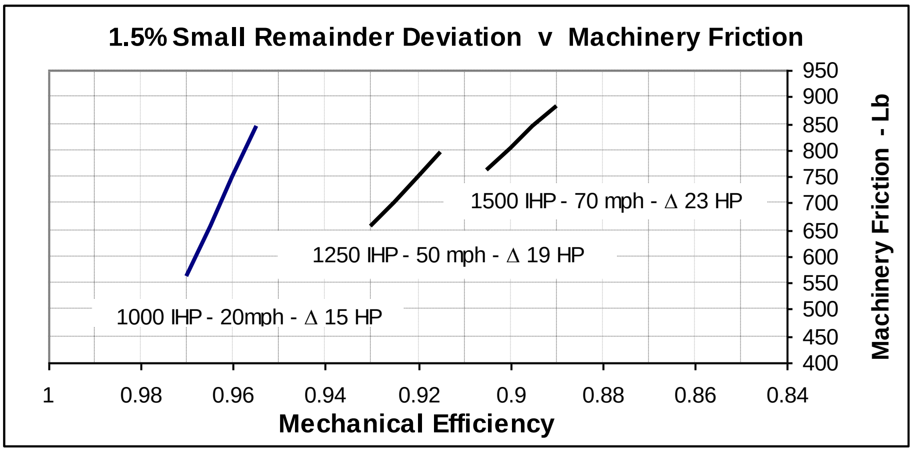


Figure 19 Given that Carling* put the accuracy of the Amsler dynamometer work
done measurement at 1½% and the Farnboro' indicator as "probably within 2% or less.", the scope for uncertainty is over 3%, and that's without things going wrong as they sometimes did. Carling* thought individual locomotives might vary by up to1%.

John Knowles call for around a dozen plots carries more weight in regard to small remainders. The random number experiments tabled below clearly support this point. The Rugby data sets are often limited to only a few plots at given speeds.

Randomised MF Outcomes @ 800 lb +/- 2% # 10 Data Sets of 10 Plots x 6 (20 to 70 mph)		
Average 600 Plots	782	98%
Set Minimum - 6 x 10 Plots	723	90%
Set Maximum - 6 x 10 Plots	847	106%
Average 10 x 5 Plot Sequences	682	85%
Minimum 5 Plot Sequence	379	47%
Maximum 5 Plot Sequence	1125	141%
# Randomised variation limit for ITE & WRTE entries		

....
....
* Model Engineer 17 October and 7 November 1980

Uncoupled Locomotive Vehicle Resistance VRU – A Key Constant

Here we look at the "simple proof" alluded to earlier in this correspondence.

WRHP minus DBHP = VRU = a constant

The uncoupled vehicle resistance component of locomotive resistance, VRU, can be discovered by deducting the drawbar horsepower (DBHP) as derived from road tests, from the wheel rim horsepower (WRHP) as recorded on the test plant. If the test WRHP and DBHP data is accurate, this exercise *should return a constant VRU value for any given speed irrespective of power output and steam rate.* Such an outcome assumes the DBHP data has been regularised to a uniform situation in regard to wind and track conditions. The plausibility of this result, can be verified as within credible limits or otherwise by comparison with estimated values of VRU (VRUe) based on a body of empirical evidence in regard to the available experimental and technical data. The VRUe values calculated therefore represent a band of possibility within which the experimental VRUx values should fall. Where wind conditions pertaining for the road tests are known, as in the case to be exampled, the 'band of possibility' can be narrowed down to some extent. VRUx indicates as derived by experiment from the test plant WRHP in association with the road test data. For an examination of LR, MF and VRU, the following relationships obtain:

LRHP = IHP - DBHP (1)

WRHP = IHP – MFHP (2)

MFHP = IHP –WRHP (3)

VRU HP = LRHP – MFHP (4) & WRHP – DBHP (5)

LRHP = MFHP + VRU HP (6)

DBHP = IHP – LRHP (7) & WRHP – VRU HP (8)

These same relationships apply where using force, i.e.; ITE, WRTE, DBTE.

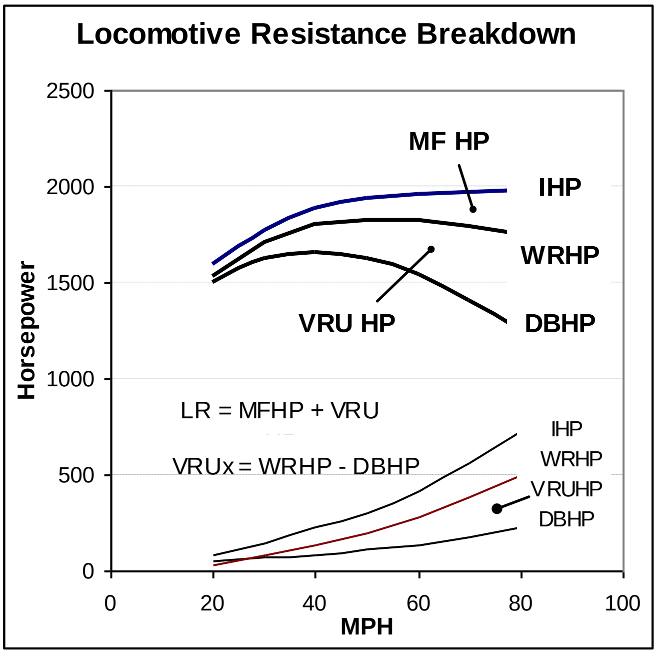


Figure 20 Plotted curves are notional values,

VRU Comprises 3 Elements

1, The rolling resistance of the locomotive and tender carrying wheels. This element is absent for tank locomotives without carrying wheels such as 0-6-0Ts etc.

2. Vehicle resistance is usually expressed in the form: $R = A + V/B + V^2/C$ Lb/ton, where the 1st term A represents rolling resistance as 1 above, and is assumed, as a convenience, to be a fixed value independent of speed. The 2nd term is attributed to the track and ride losses resulting from the behaviour of the vehicle and its interaction with the track. This term is usually derived as the remainder after the rolling resistance and aerodynamic drag (3rd term), has been deducted from the total resistance as established by experiment. The extent to which the 2nd term losses are replicated at the coupled wheels of a locomotive working on the test plant rollers is uncertain. These losses running on the spot will be reduced to some extent The absence of percussive rail joint losses on the rollers is estimated to save 0.015V pounds per ton.* Since the rollers are mounted on more solid foundations, further reductions are probable given the behaviour on the more flexible permanent way and track bed. In reality the 2nd term would also include an element of coupled wheel rolling resistance since this gradually increases with speed (ZN/P); this occurs on both plant and track.

3. The 3rd term, an intrinsically squared function, is exclusively ascribed as aerodynamic drag in regard to rolling stock. Where locomotive resistance as determined by experiment is concerned, the 3rd term will also include an element attributable to the dynamic losses of the motion and coupled wheel windage, which will occur as part of the power transmission losses (MF), and not as part of the uncoupled vehicle resistance losses, VRU, as considered here.

Aerodynamic drag is problematical since it is a variable subject to the moods and direction of wind, which potentially, may have a significant impact. Although aerodynamic drag can be estimated for an assumed set of conditions in regard to speed and direction, it will always remain an estimate of some uncertainty. Wind conditions tend to vary by the hour if not the minute, and are constantly affected by the shifting local topography. Some of the Swindon derived test bulletins declared wind conditions: a $7^1/_2$ mph, 45^0 headwind, and later 10 mph un-vectored; such specific information was absent from Rugby/Derby derived test bulletins and reports.

Test Bulletin Locomotive Resistance.

The test bulletins mostly return constant locomotive resistance at given speeds across the full working range. In some instances, including the Duchess, Report R13, deducting DBHP from IHP returns increasing LR with the level of effort; likewise the 9F bulletin. Assuming the data is regularised for a constant wind condition, then the VRU value at a given speed is a constant. This obtains whether it is VRUx as determined from deducting DBHP from the experimental WRHP, or using a VRUe estimate to crosscheck VRUx. Accurate WRHP data (assuming reliable DBHP values) theoretically returns constant VRUx values at a given speed across the working range. Such is the case for 46225 as below.

Scope of Experimental DBHP Data.

To determine cross checks on a VRUx based analysis it is necessary to have reliable DBHP data, so this potentially limits the types available for examination to the Duchess,. The Derby derived DBHP data for the Britannia, BR5, and the 9F is unreliable – Report L116. A Crosti locomotive resistance curve is included in L116, also for a standard 9F, and for the Duchess in Report R13.

----

* *How Long-Welded Track Aids the Rolling Stock Engineer*, J K Koffman, *Modern Railways* May 1965. *Traction Supplement*, D H Landau 1998.

The Duchess, 46225 (Report R13), incorporates DBHP data across the speed range, as determined by Report L109 and the L109 Supplement. The road tests for the 70005, 73008, 92050 and Crosti 9F 93023 were carried out under the "controlled road test procedure", as pioneered and developed by Sam Ell at Swindon in the early post war years, by the Derby road test team. The nub of this concept was maintaining a constant steam rate throughout the test period irrespective of changes in speed. It was claimed such control could be maintained by working at a constant blast pipe pressure. Given this assumption it was concluded by the Derby test department that this rendered indicating on road tests redundant, since, if the steam rate was so controlled at a known steam rate using the blast pipe pressure as a meter, backed-up by Sam Ell's 'summation of increments' procedure, the IHP data as determined at Rugby would be automatically replicated on the road tests. As things turned out this proved not to be the case. At a given steam rate, blast pipe steam temperature falls as speed increases. Since cylinder efficiency increases with rising speed, increasing the heat drop resulting in falling exhaust temperature and increased steam density, steam flow variations with speed at a given blast pipe pressure will occur. A problem was first suspected on the B1 road tests in 1951; action was long delayed.

Realisation of the problem eventually heralded the reinstatement of cylinder indication on road testing and periods of constant speed testing were also reintroduced, as applied for the Duchess road tests. As a consequence of this problem, the road test DBHP data for the B1, Britannia, BR5, 9F and Crosti 9F was compromised; the actual working steam rate tending to be lower than assumed at the lowest speeds and higher at the highest, and only coincident somewhere in the middle speed range. Consequently DBHP tended to be under recorded relative to what the supposed steam rate would have produced at the low end of the speed range and over recorded at the upper end. The resulting locomotive resistance curves were of strange form and improbably flat when extracted from the test bulletins. This problem gave fruit to Reports L109 (Duchess road tests), and L116 (9F & Crosti 9F), which investigated the roots of the problem and developed a procedure for correcting the road test data in line with the true steam rates obtaining. The report included before and after locomotive resistance curves for the Crosti 9F and an LR curve for the standard 9F. When the latter is plotted against the LR curve as derived from the test bulletin, these lines cross at about 39.5 mph;

and likewise for the Crosti as first determined from the road tests, and as the corrected LR curve.

On the assumption the equivalent null point for the BR5 and BR7 would be at the same piston speeds as the 9F, it would occur at about 48 mph. The relative blast pipe areas differed however, on an index basis: BR7 = 100, 9F = 95 and BR5 = 91. This may have influenced the outcome beyond piston speed alone. Notwithstanding the many test runs conducted on the test plant, the data available for individual locomotives is sometimes quite limited in scope. In the case of the Duchess for example, adequate IHP and WRHP data is only available at 50 mph. Comprehensive IHP and DBHP data plus a locomotive resistance curve is available from report R13 based on report L109 and the "L109 Supplement". It is fortunate that at 50 mph the road test steam rates were in accord with the theoretical Rugby values throughout the working range, so the Rugby IHP determinations could reasonably be assumed as having been replicated. Report L109 investigated departures from steam rate over the working speed range, and determined the actual steam rates obtaining in regard to the recorded DBHP. "Corrected" DBHP curves were produced accordingly and these were incorporated in the final report. Oddly, the drawbar figures in the 9F report were as uncorrected, notwithstanding that report L116 was issued a year before the 9F test bulletin was published. Internal correspondence reveals E S Cox was unwilling to accept the idea of steam rate deviations; as being without a theoretical basis, and likely simply a case experimental error. At this point a departmental impasse is apparent. Exhaust steam temperature and specific volume at a given pressure falls with rising cylinder efficiency (density increases) as a function of speed and heat drop. Road test steam rates could deviate from the assumed value by over 1000 lb/hr.

46225 - A VRU Test Case

The available test plant ITE, WRTE and MF data at 50 mph for the Duchess, 22 plots,
is set out in Figure 21.

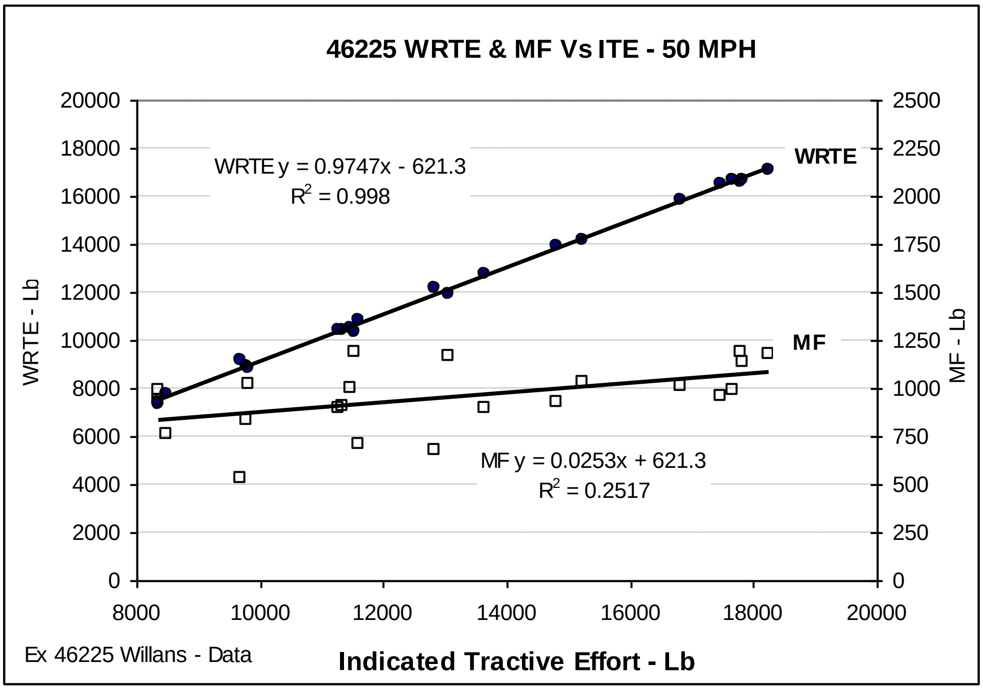


Figure 21 A similar chart using only15 of the available plots appeared in my letter 17 March 2017.This yielded the formula WRTE = 0.9708 – 545 lb.

The differences in the MF outcomes are slight. .

46225 MF Outcomes - 50 mph.				
IHP	1000	1500	2000	2500
15 Plots	764	874	983	1093
22 Plots	811	906	1001	1095
Δ MF Lb	47	32	18	3
Δ MF HP	6	4	2	0

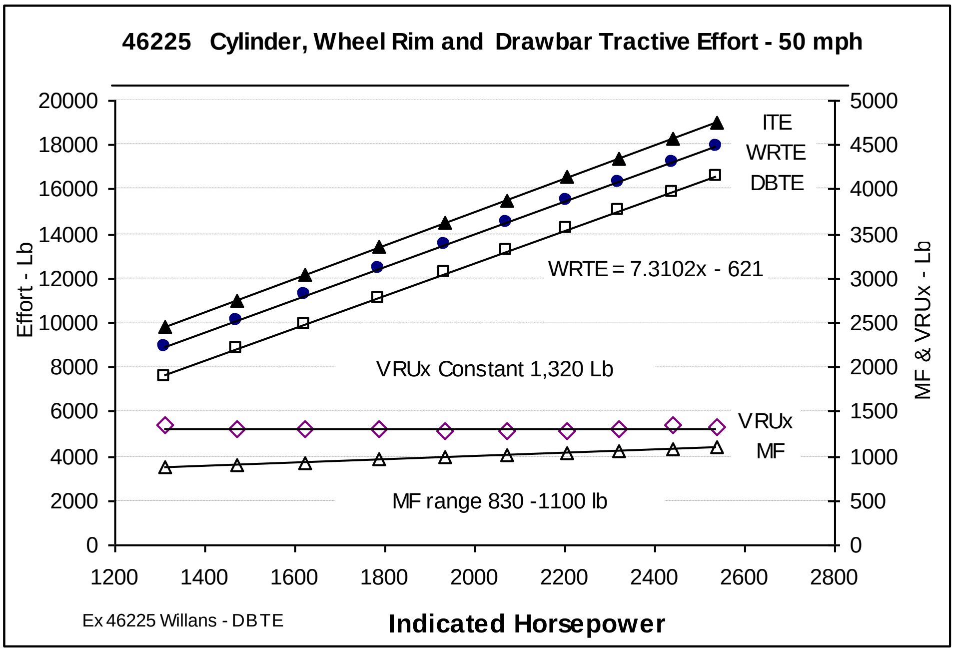


Figure 22 The WRTE & MF plots are 'smoothed' as derived from Figure 21. The VRUx scatter is within the range + 21 - 9 lb. The bulletin graphs are not drawn with

tool room accuracy, likewise recovering said data by scaling off is short of high precision. The DBHP Willans Lines so derived from L109 return high R^2 values, sometimes achieving unity, but this is no guarantee of spot-on determinations.

The Report R13 locomotive resistance curve is in lb/ton (Figure 18). At 50 mph the LR is given as 14.4lb/ ton; 2327 Lb in total. This is coincident with a steam rate of 30,000 lb/hr, a coal rate of 4,110 lb/hr, IHP 2072. The smoothed experimental data for ITE, WRTE, MF and Report R13/L109 DBTE, and the derived VRUx values are plotted above in Figure 22. Since LR = MF + VRU (5), then:

ITE @ 2,072 IHP = 15,540 Lb; WRTE 14,526 Lb; MF 1,014 Lb + VRUx 1,320 Lb
= LR 2,335 lb. Report LR at 2,327 lb is effectively identical..

Tabled below a VRUe estimate for the Duchess. It is assumed the 2nd term losses for the coupled wheels will be reduced to some extent when running on the test plant relative to the losses that occur working out on the line. This reduction occurs on two counts. Firstly the percussive losses at rail joints will be absent, and secondly, given the more solid foundations of the plant, the degree to which the adhesion weight LR 2nd term ride and track losses are encountered on the test plant. It seems likely that these losses will be reduced running on of the test plant. In this example the plant losses appear reduced to around 60% relative to what is normally encountered on the more flexible track and track bed of the permanent way. Obviously, given the estimated make-up of VRUe, this determination is tentative.

Most of the limited WRHP data available for 46225 is at 50 mph, this was coincident with the speed at which *the assumed steam rate was accurately replicated on the road tests.* The Derby Farnboro' indicator was deployed throughout the road tests. The comparative Rugby plant and Derby road test indicated horsepower results were in agreement at 50 mph: no revision of road test IHP and DBHP data applicable.

46225 Estimated VRUe 50 mph *			
Uncoupled Wheels 1st Term			R Lb
Bogie	2 x10.75 tons	4.45 lb/t	96
Truck	1 x 16.8 tons	3.75 lb/t	63
Tender	3 x 18.8 tons	2.8 lb/t	188
Uncoupled 2nd Term 94.65 tons		3.125 lb/t	296
Aero $3^1/_2$ mph 45^0 Headwind			645
Coupled Wheel Percussion Losses		0.53 lb/t	50
Coupled Wheel Track & Ride Losses **		0.5 lb/t	34
Total VRUe (= VRUx + 4% = 52 lb, 7 HP)			**1372**

The wind conditions for the road tests over the S & C are on record and were atypically moderate. The VRUx and VRUe outcomes in this instance are tolerably close. On the basis of these figures about 40% of the 2nd term coupled wheel LR losses are avoided when running on the test plant. The remaining 60% will primarily relate to the journal ZN/P losses and the coupled wheel windage as part of the overall machinery friction. The modest track ride losses are based on a relatively recent paper on train performance hailing from the USA. **

***** 1. The 1st term as tabulated is based on bearing loadings, mechanical advantage, and friction coefficients derived from Ell's wagon resistance data in his 1958 I. Loc. E paper; *The Mechanics of the Train in the Service of Railway Operation.* It's purely a mathematical fit to the data, effectively a rolling resistance constant, excluding the ZN/P frictional speed increment.

2. The 2nd term assessment assumes some of the normal coupled wheel adhesion weight track and ride losses will be absent when running on the test plant. Namely the percussive losses at the rail joints and some of the losses involving the ride interaction with the track and track bed. The rail joint losses were determined some years ago from an article by J L Koffman: *How Long-Welded Rail aids the Rolling Stock Engineer*, Modern Railways, May 1965. $R_p = 0.015V$ lb/ton.

3.. The aero term assumes a drag coefficient of 0.77 as LMS wind tunnel tests, a net frontal area 101.5 sq.ft and a $3^1/_2$ mph headwind. The latter value is the average of the road test wind record.

** *Train Performance: AREMA Manual for Railway Engineering-* American Railway Engineering and Maintenance-of-Way Association,1999. It elegantly described these losses as attributable to the "wave action of the rail".

Drawbar Horsepower Derived Locomotive Resistance

Back in 2013 I investigated the veracity of the Duchess resistance curve included in the Report R13. The resistance curve was regarded by many as being too low. The examination subjected the data to four tests which were satisfied (DHL R13 Audit). The 4th test was the derivation of locomotive resistance from the DBHP data.

This method of approximating LR is derived from the zero root point of DBHP Vs Steam rate linear trend lines at given speeds, the root point (negative value**)** being representative of LR (Figure 23). The proximity of these results to the R13 LR HP curve is striking – (Figure 24). The underlying theoretical point is that no horsepower appears at the drawbar until the locomotive resistance has been overcome. The linear projections represent the tangential mean of the recorded data. Having explored this method extensively, the outcomes are very sensitive, notably at low speeds, to the steam rate range selected to find a tenable data set. There is some scope for geometric mean solutions; in the case of the R13 data, this proved unnecessary, no weeding required.

This method was inspired by reading Stanley Hooker's autobiography *Not Much Of An Engineer*, Hooker was an engineer at Rolls Royce, initially specialising in super chargers. Backwards projection was used to determine aero engine frictional losses.

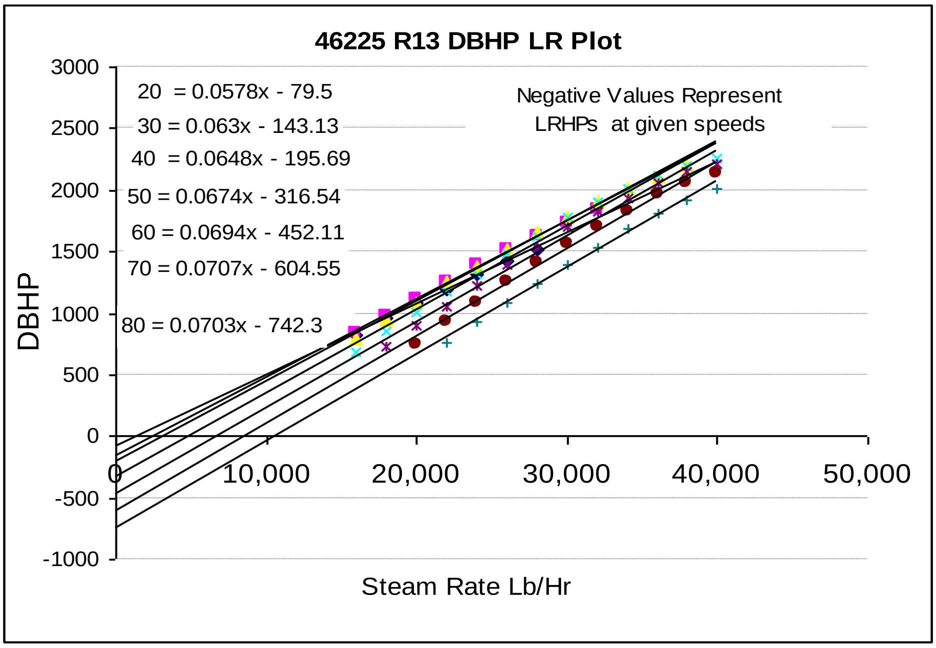


Figure 23 The plotted data covers the full test bulletin power envelope. The outcomes theoretically approximate to mean steam rate LR.

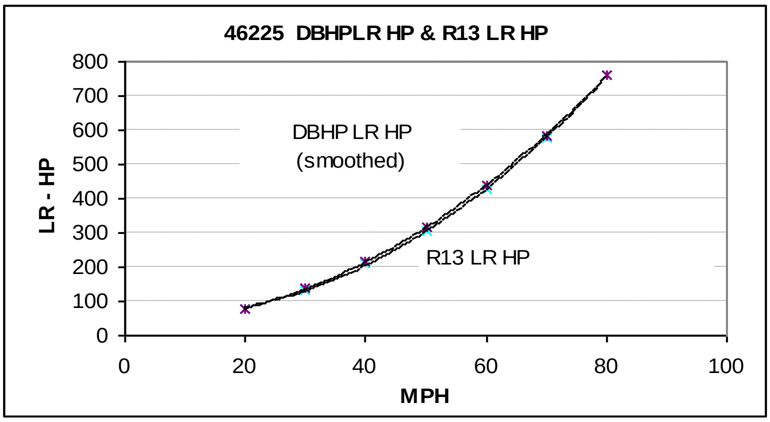


Figure 24 The smoothed DBHP derived LR HP is barely distinguishable from the Report R13 Figure 18 derived LR HP.

Road Test Steam Rate Anomalies

Report L116 treating the steam rate anomalies in regard to the Crosti and Standard 9Fs showed, as with 46225 (Report L109), the same trait of deviation in steam rate at given speeds across most of the working range. The machinery friction for Crosti 9F 92023 as tested at Rugby was significantly higher than as recorded for the standard 9Fs tested on the plant. This difference was confirmed in road tests as below.

LMR No.3 Dynamometer Car and Mobile Test Unit *

Steam Rate 16,000 lb/hr

Speed MPH	Drawbar Horsepower (DBHP)	
	Crosti 92023	Standard 92050
20	862	917
30	900	960
40	875	939
50	827	903
Average	866	930

The Crosti drawbar deficiency was 55, 60, 64 and 76 HP for the speeds shown. This was attributable to reduced indicated horsepower of the Crosti resulting from higher back pressure (offset to some extent by higher superheat), and increased machinery friction as evidenced on the test plant. Subsequently, 92050 underwent further tests at Rugby eighteen months later to “resolve perceived differences between results obtained on the stationary test plant and the road tests.” No indicating was carried out on the standard 9F and Crosti road tests.

The nominal road test steam rates were not held constant across the speed range, tending to increase with speed, the test plant indicated horsepower/steam rate only being replicated on the road tests at about 39.5 mph. The steam rate deviations as determined in report L116 were significant.

Post the road tests, some satisfactory comparative tests between the Rugby and Derby versions of the Farnboro indicators were conducted at Rugby in 1957: 92050 Series 2 tests . These tests post-dated the significant improvements to this equipment reported by Ron Pocklington.

92050 Comparative Indicator tests IHP Indices 1957						
Steam Rate	IHP - Rugby-Derby Mean Value Indices					
	15 MPH		30 MPH		50 MPH	
	Rugby	Derby	Rugby	Derby	Rugby	Derby
12,300			100.6	99.4		
13,100	99.9	100.1				
14,900			99.8	100.3		
15,500			100.4	99.6	99.1	100.9
16,150	99.3	100.7				
17,400					98.4	101.6
18,500			98.8	101.2		
18,900	98.4	101.6				
19,100			99.3	100.7		
19,500					101.1	99.0
19,750	99.8	100.2				
21,400					100.1	99.9
22,400	100.4	99.6				
23,400			100.4	99.6	100.4	99.6
Averages	99.6	100.5	99.9	100.1	99.8	100.2
Averages	All Rugby		99.75	All Derby		100.26

----

*A Detailed History of British Railways Standard Locomotives, Vol. 4: The 9F 2-10-0 Class, page 217. RCTS, 2008

The 92050 Series 2 tests at Rugby in 1957 returned reduced IHP and WRHP outcomes relative to the 1955 Series 1 tests. The Series 2 tests recorded higher exhaust steam temperatures for given steam rates at 30 and 50 mph. (Comparative data at other speeds unavailable). Such an outcome is symptomatic of steam leakage, The Series 2 tests also showed an increased steam consumption of around 2 percent at a given cut-off. 92050 was in traffic for 18 months between the Series 1 and Series 2 tests 92050 and will have clocked up around 35,000 miles in the interim. The BR Standards with the 3 bar crosshead slidebar arrangement were notorious for high piston valve ring and piston ring wear.

92050 Test Series 1 & 2 IHP & WRHP Comparison - 50 mph						
Steam Rate	IHP Willans 50 mph			WRHP Willans 50 mph		
	16,000	20,000	24,000	16,000	20,000	24,000
Series 1	1,170	1,500	1,770	1,090	1,415	1,680
Series 2	1,100	1,415	1,670	1,010	1,315	1,562
S2 Δ HP	-70	-85	-100	-80	-100	-118
S2 Δ HP %	-6.0%	-5.7%	-5.6%	-7.3%	-7.1%	-7.0%
The Series 1 tests 1955, and the Series 2 1957 tests post dated the final improvements to the Farnboro Indicator early in 1955.						

The comparative exhaust temperatures are consistent with increased leakage for the Series 2 tests – Figure 25. Curiously the 9F test bulletin IHP appears to have combigned and thereby averaged the Series 1 and 2 IHP data. Possibly this was a deliberate decision to reflect typical operating conditions.

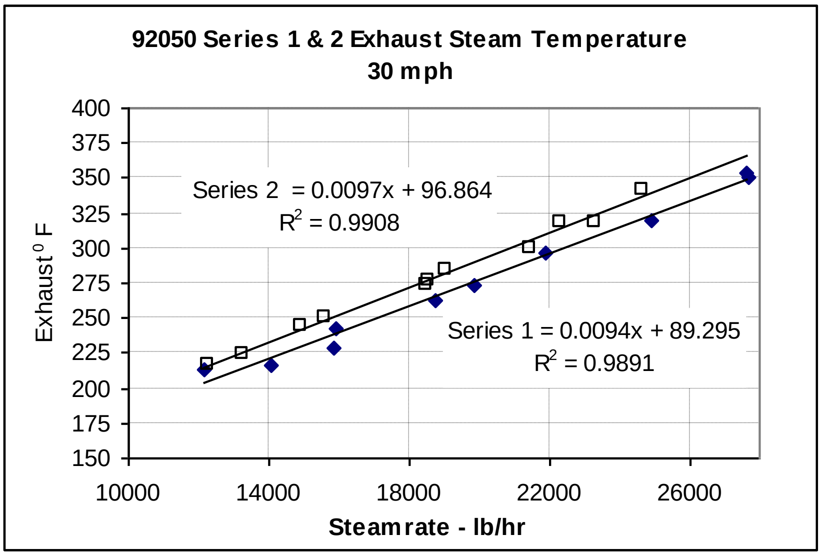


Figure 25 The higher exhaust temperatures of the Series 2 tests are indicative of increased steam leakage. This may occur as both a constant loss to atmosphere from the steam chest, and a cyclic loss via the cylinder during compression, admission and expansion.

The apparent and eccentric road test locomotive resistances of Crosti 9F 92023 and 9F 92050 were subject to correction in Report L116, after adjustment for significant steam rate departures from the assumed constant rates. These deviations from the nominal test rate could be over 1000 lb/hr, positive and negative, crossing over from negative at some point roughly two thirds through the speed range.

Report L116 gives 'before and after' LR curves for the Crosti, and an LR curve for the standard 9F. The degree of adjustment for the Crosti was striking (Figure 26). The standard 9F Report L116 LR curve was of similar form and crossover point relative to the
9F LR curve as derived from the test bulletin.

The outcome of the steam rate deviations, aside from the crossover point, was that the recorded DBHP related to other than the supposed steam rate and related Rugby IHP data, hence the eccentric L116 LR curves as initially derived from the road tests.

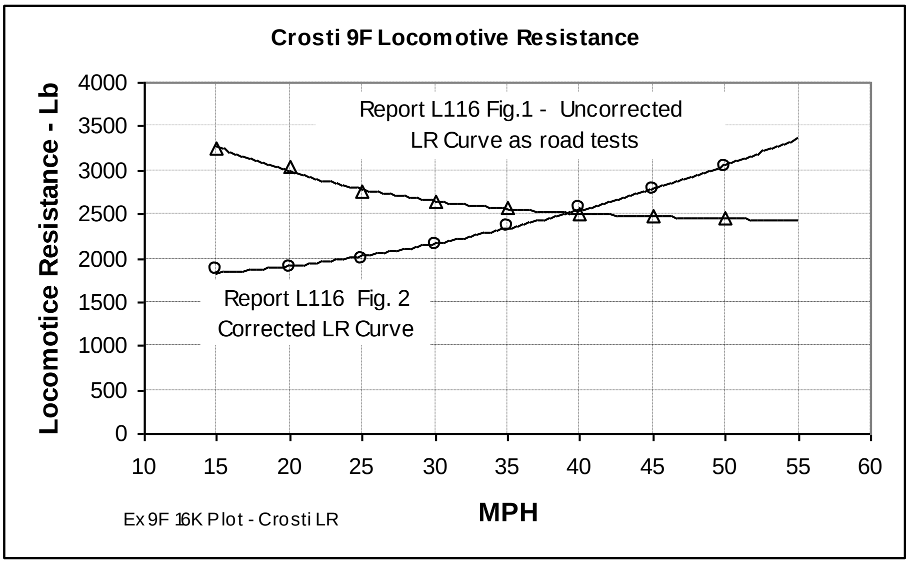


Figure 26 The uncorrected curve reflects a trend for the steam rate initially to fall below the nominal test rate as an inverse function of speed, an error diminishing to zero at the crossover with the corrected curve, and increasing as a function of speed thereafter. A similar pattern is apparent for the standard 9F L116 Fig. 3 LR curve when plotted against the LR curve derived from the test bulletin. Both the Crosti and standard 9F share a common crossover point of 39.5 mph. The steam rate anomalies for Duchess 46225 as evaluated in Report L109 follow a similar pattern; crossover point 50 mph. The BR5 crossover relative to the estimated LR (dashed lines) is less distinct.

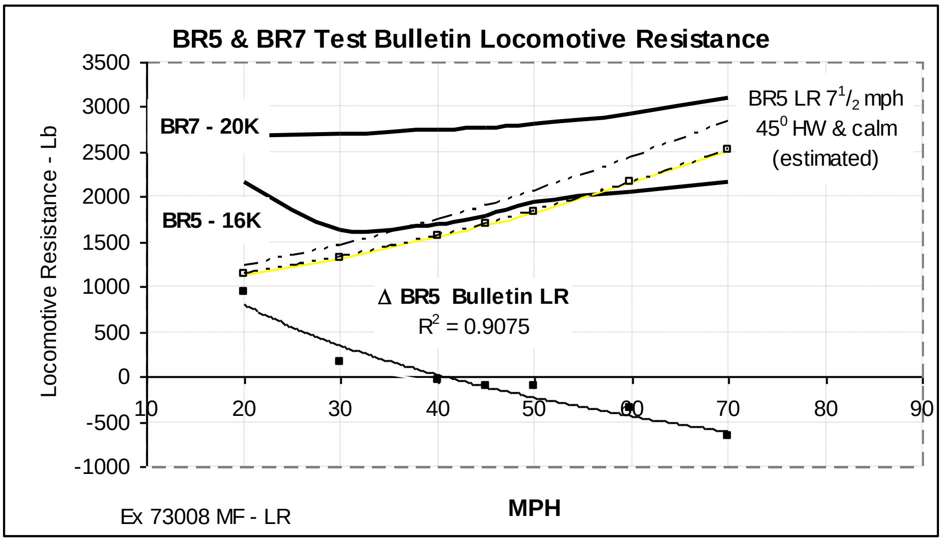


Figure 27. The high flat lining LR curve for the BR7 is an extreme example of how things could go wrong. The BR5 appears somewhat undecided, with a plausible outcome

somewhere in the middle steam range. The falling error curve shown is for the test bulletin derived curve difference relative to the estimated LR curve for 7 1/2 mph headwind.

The key change increasing steam rate with speed at a given blast pipe pressure is the fall in exhaust steam temperature and density that accompanies increasing cylinder efficiency and heat drop as exampled below for the BR5. An characteristic example of along the lines of Report L116 Figure 11 is portrayed in Figure 28.

On the basis of piston speed relative to the 9F, it has been calculated that the point of zero steam rate error on the road tests would occur at 48.7mph, this is considered sufficiently close for the test bulletin DBHP curves for 50 mph to be suitable for the analysis, as set out in Figure 29, as derived from the procedure set out for Figure 22.

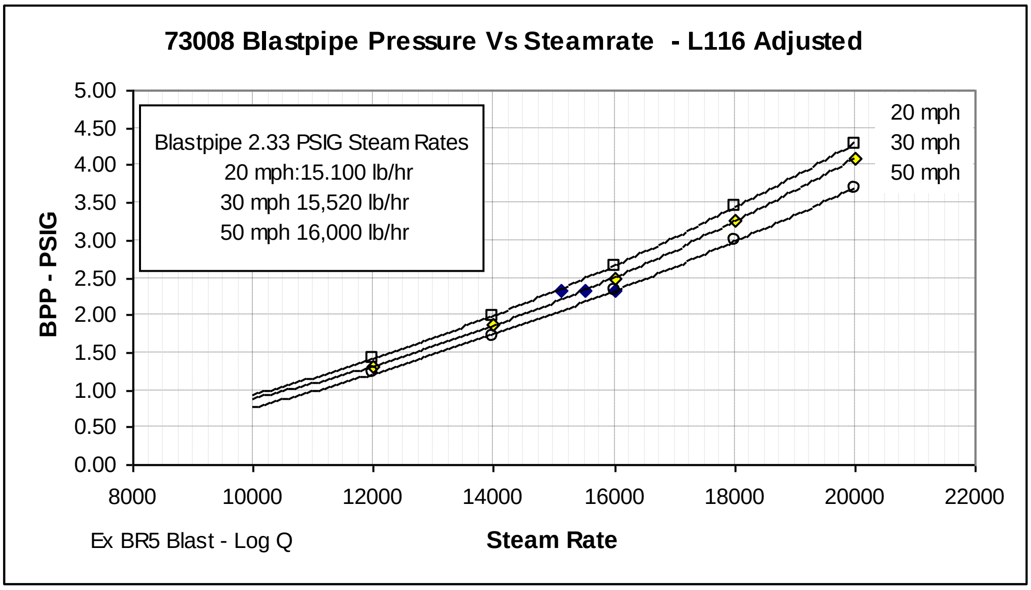


Figure 28 This is of equivalent form to Figure 11 for 92050 in Report L116, as determined from Rugby test plant experimental data using the Log Q = Log C + n Log P relationship.

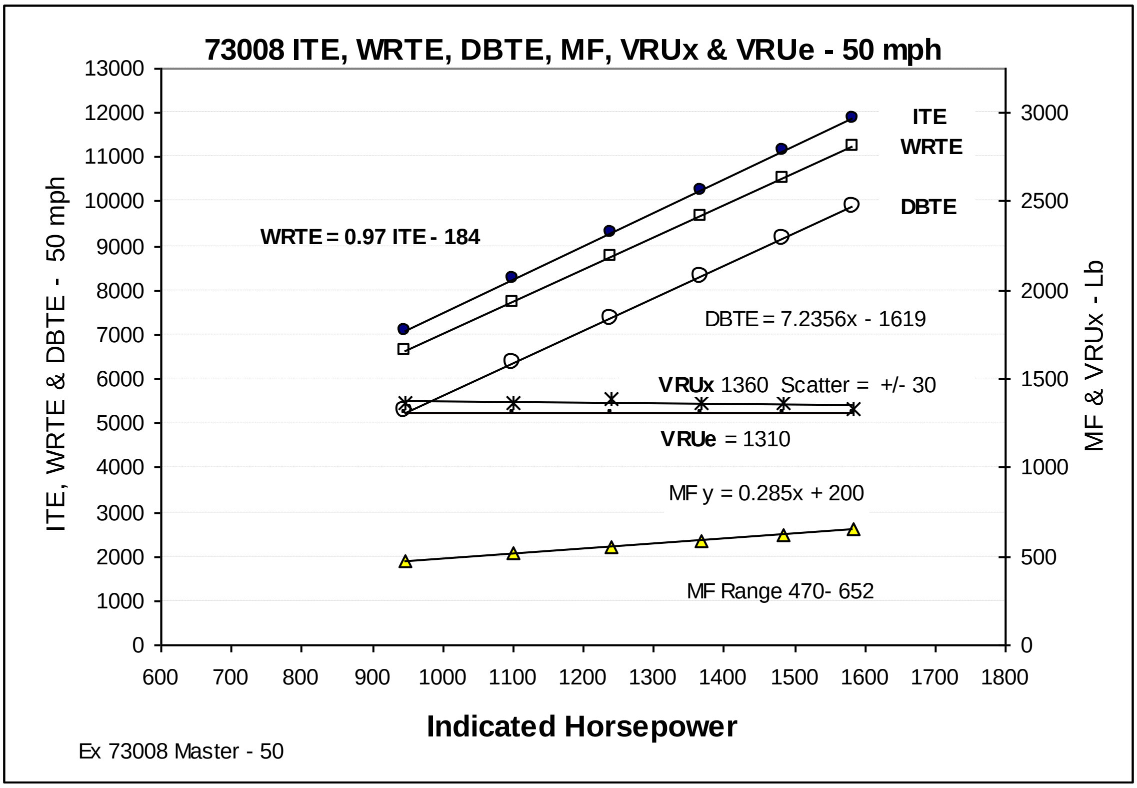


Figure 29 The IHP/ITE data used is as test bulletin, WRTE as Rugby Willans Lines; 7/2 mph 45⁰ headwind assumed. In the event, the BR5 road tests were subject to unusually high wind speeds averaging 14 mph south westerly – 270⁰; as derived from Beaufort Scale median values. Line headings Carlisle – Appleby SE (135⁰); Appleby – Settle Jcn SE (170⁰).

BR5 73008 Figure 29 LR Derivations 50 mph			**73008 Estimated VRUe - 50 mph** *			
Steam Rate	18,000 lb/hr	24,000 lb/hr	Uncoupled Wheels 1st Term			R Lb
IHP	1238	1580	Bogie	2 x 8.95 t	5.27 lb/t	94
ITE	9,285	11,850	Tender	3 x 16.4 t	3.94 lb/t	194
DBTE	7,353	9,813	Uncoupled 2nd Term 67.1 t		3.125 lb/t	210
LR	**1,932**	**2,037**	Aero $7^1/_2$ mph 45^0 Headwind			739
MF	553	650	Coupled Wheel Percussion Losses		0.75 lb/t	44
VRUx	1360	1360	Coupled Track & Ride Losses **		0.5 lb/t	29
LR	**1,913**	**2,010**	Total VRUe			**1310**
Figure 27 Estimated LR 50 mph - 2054 lb			∆ VRUx v VRUe = 50 lb, 7 HP			

A "Simple Proof" along the lines of the Duchess procedure Figures 21 & 22 has also returned constant VRUx of 1190 lb for the 9F at 40 mph. The speed was selected on the grounds that there was minimal departure from the supposed steam rate, corrections unnecessary, the bulletin DBHP curves at 40 mph were assumed satisfactory. At 1190 lb the VRUx plotted scatter was +/- 35 lb, +/- 4 HP.

92050 16,000 lb/hr - 40 mph	
IHP Bulletin Figure 11	1115
DBHP Bulletin Figure 2	899
Fig. 11 - Fig. 2 = LR - Lb	2025
MF - Lb	796
VRU = LR - MF Lb	1229
VRUx (Δ VRUx v VRU = - 4 HP)	1190
L116 Figure 3 LR - Lb	2062
Δ Fig. 3 LR v Fig. 11 - Fig.2 LR	37 Lb, 4 HP

A Simple Proof?

While the simple proof described appears satisfied within tolerable limits, SRMs are not a simple case for verification, as compound errors they are beyond simple calibration, and therefore best avoided where alternatives exist.. Many of the measurements on a locomotive testing station involve complex instrumentation subject to finite degrees of potential error, which though small, is sufficient to play havoc in the small remainder situation. Such outcomes are the inevitable result of randomised scatter, a problem considered further in the addendum. Absolute proof is elusive. As far as is practicable, the constant VRUx outcome "simple proof" has been demonstrated for 46225, the BR5 and the 9F. Given all this, some prerequisites must be satisfied:

1. Repeatability.

Though combigned WRHP Willans Lines for locomotives of the same type have returned high R^2 values and generally low scatter with few 'strays', this is not proof in itself. Systematic errors may occur. Willans lines do however return relative order whereas the small remainder MF outcomes deliver confusion; hence the low R^2 values. Repeatability nevertheless remains a prerequisite of proof, but SRMs are unlikely to be of any use in this regard. Plots of WRTE against ITE are generally even better behaved than Willans Lines, but even when returning visually near identical trend lines as plotted immediately below, the curve fitting formulae may return little agreement regarding the coefficients and constants involved as exampled in Figure 30.

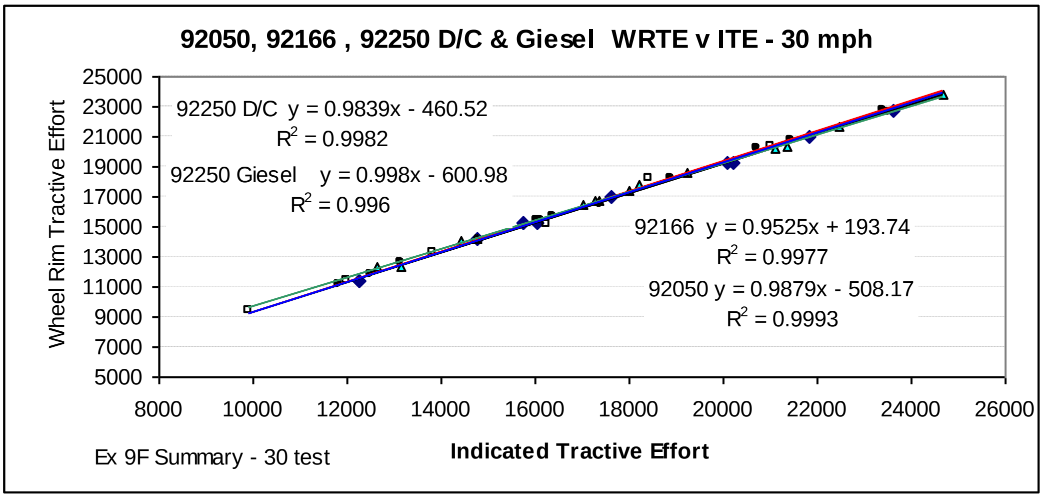


Figure 30 The four trend lines bundled together here are indistinguishable over the middle range. Of the four constants, three are of the same sign and general order of

magnitude. Perversely, such are the joys of random scatter, 92166 contrives to change both sign and magnitude. (This was corrected above - Figure 18).

An assumption that for a given indicated tractive effort and speed, machinery friction will be the same, irrespective of the back pressure and superheat obtaining resulting from changes in blast pipe area, appears to be bourn out by the pooled data, as for the 9Fs plotted in Figure 30 The 92250 Giesel data, comprising only 6 MF plots, has been combigned with the 11 plots available in double chimney guise yielding outcomes, along with those for 92050 and 92166, as tabulated below.

9F Collective WRTE v ITE Machinery Friction Outcomes @ 1600 IHP, 20,000 lb ITE - 30 mph					
Engine	Plots	R^2	Formula	20K ITE MF	20K ITE MF HP
All	44	0.9978	y = 0.9779x - 308.16	710	57
92050	12	0.9993	y = 0.9879x - 508.17	750	60
92166	15	0.9977	y = 0.9525x + 193.74	756	60.5
92250	17	0.9974	y = 0.9865x - 476.3	746	60
Averages		0.9981	y = 0.9820x - 390	740	59

The MF returns, representative of an effort of around 24,000 lb/hr steam rate, fall within +/- 2 HP, 25 lb of the mean value. While not proof of accuracy in itself, it does satisfy the repeatability criteria, and even then, only up to a point. As will be seen the various formulae fitted show differences in the x coefficient, representing the work sensitive friction coefficient (1-Function x), and more markedly for the constants, including the anomalous positive constant for 92166 (as examined above- page 16). The x term outcome is very sensitive to the tilt generated by the random scatter of the data set. It is noted that 92166 returns the highest implied frictional coefficient, approaching 5%, and that a false compensating positive constant is returned in order to fit the recorded values.

The 92166 IHP and WRHP SSC curves return mediocre R^2 values, 92166 involved a mechanical stoker, and allowing for the steam consumption involved may on occasion have led to some miscalculation of the steam reaching the cylinders. Given this possible potential for error, or for whatever reason, the ringed SSC plots below possibly relate to steam rates other than shown. The R^2 values are accordingly compromised.

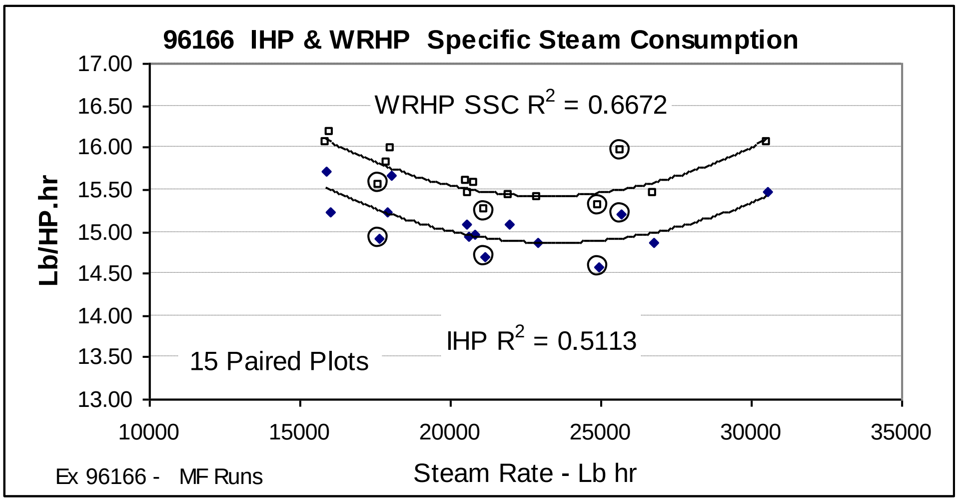


Figure 31a The master/slave relationship of the IHP./WRHP vertical paired coupling displacements are clearly in evidence here. I have ringed four pairings, and have likened this in the past to a dog following on a lead, with the slack or tension in the lead being analogous to the potential small remainder experimental error when determining the distance between man and dog.

John Knowles has disputed the existence of this relationship in his letter 12 July 2017 and elsewhere, Like it or not, WRHP is ever the child of IHP. Given the matching vertical shifts of the IHP-WRHP pairings shown here, it is apparent the IHP deviations from trend are in most cases are the outcome of real shifts rather than measurement errors. The usual 'elasticity' of small differences of course remains.

.

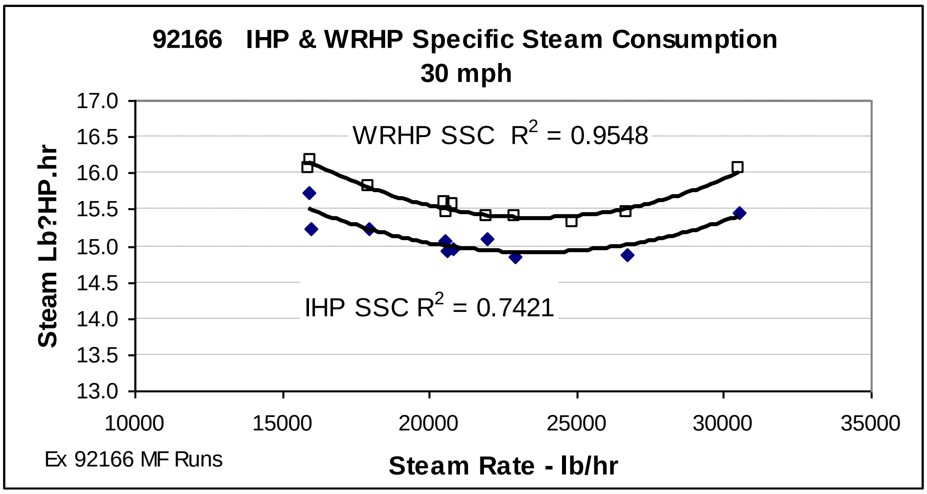


Figure 31b Removing the 4 vertically displaced pairings improves the SSC curves R^2 values;

The data set for 92166 includes 49 WRHP readings against steam rate. The associated Willans Line gives an R^2 value of 0.9946. Reducing the data set to 42 by removing randomly distributed plots not in contact with the trend line marginally increases R^2 to 0.9974. Another example that more data does not necessarily lead to more accurate outcomes. A poor plot or plots can occur at any point in the testing cycle. Simultaneous IHP and WRHP plots are limited to 15 for 92166, and as explained (page16), the positive remainder it returns for the WRTE v ITE formula is unsatisfactory. Such an outcome can only be eliminated by reducing the data set to 8 pairings, as determined by experiment. The revised outcomes, along with 92050 and 92250 are tabled below.

9F Modified* Collective Machinery Friction Outcomes @ 1600 IHP, 20.000 lb ITE - 30 mph					
Engine	Plots	R^2	Formula	20K ITE MF	20K MF HP
All	37	0.9984	y = 0.98530x - 449.48	743	59
92050	12	0.9993	y = 0.9879x - 508.17	750	60
92166	8	0.9994	y = 0.9765x – 275.8	746	60
92250	17	0.9974	y = 0.9865x - 476.3	746	60
Averages		0.9986	y = 0.9840x – 427.4	747	60

At 10,000 ITE, the MF outcomes average 588 lb, 47 HP, spread 40 – 50; at the highest output, ITE 24.000, MF averages 812 lb, 65 HP. Spread 64 – 67.

2. Sensitivity.

This is observable in the linkage of IHP - WRHP master-slave coupled plots. In the main, the IHP/WRHP scatter pattern pairings move in the same direction, up or down in elastic harness. It is that elasticity of small errors born of large numbers that generates the small remainder scatter. Outliers exceeding +/- 100% of the mean experimental value and the occasional negative outcomes may occur, as demonstrated in random number experiments,

While the above describes the responsiveness of the dynamometer to changes in drawbar pull, the collective sensitivity of WRTE v ITE data sets is very sensitive in regard to the tilt of the simple $Y = C_f\, x – R$ relationship as generated by the random scatter pattern of the data sets as exampled for the 9F in Figure 30 and the associated tabulations above. Since the trend line constant notionally represents the resistance of the of the power transmission machinery (including of course the coupled wheels) *when not under power*, some relationship of the contsant as a function of speed is to be expected. In practice the random scatter is often sufficient to frustrate clear outcomes in this regard. As demonstrated for 92166, the constant outcome was not even the right sign. Other examples can be found in the Rugby data generally. The hostage to scatter is heightened when the ITE – WRTE relationship only covers a limited range of steam rate and power. The tilt outcomes do not necessarily improve as a function of the plot numbers available, a trend wrecking plot or plots can occur at any point in a test series.

Some plots are obviously more accurate than others, and in some instances so wayward as to be beyond the definition of 'outliers'. In this situation, something has obviously gone wrong

3 Veracity**.**

This is something of a judgement call: does it all make sense? The determination of VRU, an idea of fundamental logic, has satisfied the theoretical outcome of returning constant values, and perhaps is the nearest thing to a “simple proof”. Said VRUx values however must be considered close approximations at best. In reality, that caveat applies to the test bulletin data generally, whether it originates from Rugby/Derby or Swindon. It was sometimes more wanting from both camps. Understandably high cylinder efficiency will be welcome, but if accompanied by unusually high locomotive resistance should it be believed? The ultimate comparator of locomotive performance at a given steam rate and speed is the DBHP, but even that measure has sometimes proved unreliable due to assumed steam rate errors. This applies to both the Rugby/Derby and Swindon bulletins.

4 Uncertainty

Even if the test plant performed perfectly to the design specification in all respects throughout its operating life, the small remainder problem would not disappear. The delivery of empirical data that falls into place with the precision of a perfect jig-saw is inevitably beyond reach given the metrological limitations. While Chapelon opined that the Rugby data was the most accurate he had seen, this was against the notably chequered history of locomotive testing generally. I think Carling was right to be equally circumspect about the determination of both locomotive resistance and machinery friction. This he attributed as intrinsic to the small remainder problem. If anything, locomotive resistance is more problematical since it is determined in uncontrolled, and typically, unstable atmospheric conditions. One certainty is that WRTE will fall somewhere between ITE and DBTE, the problem is exactly where? It can tentatively be approximated by adding VRUe to DBHP where the latter is thought reliable. At best such estimates can only produce a plausible band within which the WTRE, and the MF thus implied, could fall.. Unfortunately most of the DBHP data in the Rugby/Derby derived test bulletins is wrong (Report L116). Report R13 for the Duchess is the only example where the DBHP data was fully reconciled with the Rugby IHP data (Report L109 and L109 Supplement). The available WRHP data for 46225 is only sufficient at 50 mph. The WRHP data for the BR7, BR5 and 9F is more comprehensive; but the DBHP data is deficient. The bulletin derived LR for the BR7 even appears to elude a ‘no error’ crossover point - Figure 27. Locomotive resistance determinations, given the small remainder problem can be no better than as for WRHP, and are additionally subject to climatic variation. At least WRHP, along with IHP and DBHP can be measured and scrutinised as a quantity; MF and LR and are forever a small remainders.

Addendum

First and foremost, the data base drawn upon must be credited to an XL spread-sheet put together by David Pawson in 2009, following an epic stint of research at the NRM. Comprising over 2,200 rows with up to 50 data entries per row chronicling boiler, cylinder and dynamometer performance, temperatures, pressures and gas analysis, it must comprise between 50 and 60,000 entries . It is a truly monumental piece
of research. Additional to the Rugby data, there is some Swindon plant and road test
data for 6001 and 71000. The Rugby data covers 10 locomotive types and 22 allowing for sub types. Additional to this, various reports and correspondence came to light.

As alluded to earlier in this correspondence, Dennis Carling is on record as thinking the
determination of locomotive resistance and machinery friction as troublesome. Having been privy to what at first sight is a vast body test data, my impression is that putting together a test bulletin was not exactly easy either; it was inevitably something of a

black art. It was akin to working with a shoddily manufactured jig saw with a large number of missing pieces, both randomly distributed and whole missing sections. When the data is broken down for particular speeds, it is often sketchy or absent altogether. A significant amount of interpolation, extrapolation and tweaking will have been unavoidable.

"*When a sufficient number of values of indicated pull or power had been obtained over the necessary range of speeds and rates of steaming, the values of each speed were plotted to obtain the relevant Willans Line: these are compared to those of adjacent speeds and slight adjustments are made to obtain a regular family of curves fitting as nearly as possible to <u>all</u> the points. No two draughtsman will draw exactly the same curve through the points as to what fits best, and indeed, they may be influenced to some extent by the set of French curves available in the drawing office!" **

This may sound unscientific, but it is very much the practical reality, moreover, the XL curve fitting programme is not necessarily better at it, and can be notably poor at extrapolating much beyond the maximum and minimum recorded values. The randomness of the experimental data sets and the formula thus generated is nothing less than a lottery. Wide variations of coefficients and constants are evident as demonstrated. The most reliable first steps for analysis is plotting Willans Lines, steam rate against IHP. WRHP and DBHP, or ITE, WRTE and DBTE. The drawbar data is only available by scaling off the test bulletins. Steam rate, particularly when working with the live steam injector, was thought the most accurate determination of the Rugby test data, with experimental error "*probably well under 1%" **

*"Amsler of Switzerland, guaranteed an accuracy of 1% of the scale (dynamometer pull) used, and $1^1/_2$% for the work done. ** "A calibrating device, itself checked at the National Physical Laboratory, showed this value was in fact substantially improved upon, tending to fall from close to 1% at quarter scale to 0.75-0.5% at three quarters scale, in which range most of the work would be done."* See page 91 for an NPL test record.

While IHP and WRHP Willans lines at particular speeds uniformly returned R^2 values approaching unity (not in itself is not proof of veracity), they do not extrapolate reliably much beyond the minimum and maximum plotted values, and are influenced by the particular random scatter pattern obtaining in a data set. Plots of WRHP v IHP or
WRTE v ITE provide a direct relationship where scatter is typically low as a percentage of the quantities measured, but as already demonstrate , the linear

trend lines are sensitive to the scatter in regard to 'tilt'. Some of the data base steam rates are unclear in regard to the use or otherwise of the exhaust steam injector. These uncertainties can be sometimes be resolved by examining specific evaporation rates (if coal rates available) and the steam rate v cut-off relationship. Adjustments can then be made accordingly where necessary.

Below, demonstrating the sensitivity to scatter, 3 doctored outcomes of an 8 plot MF data set, as derived from WRTE v ITE for 92050 at 40 mph when a single WRTE plot is removed. Note the varied outcome of the constant. The MF outcome at a steam rate of 20,000 lb/hr, roughly midway of the range examined; ranges from 608 to 671 lb: +/- 5% of mean. The range of uncertainty, maximum v minimum, is +/- 0.46% of ITE

92050 WRTE v ITE MF Plot Variation Outcomes 40 mph				
Plots	R^2	Formula ITE v WRTE	20K MF*	MF Index
8	0.988	Y = 0.9889x-465.5	618	98
- Minimum	0.9972	Y = 0.9798x-330.3	608	97
- Maximum	0.9974	Y = 0.9679x-229.68	671	107
- Middle	0.998	Y = 0.9892x-463.3	612	98
* Q - Willans IHP 1465		Average	627	100

----

* Dennis Carling: *An Outline of Locomotive Testing on British Railways*, * Model Engineer, 7 November 1980. Page1331. ** Ibid 17 October 1980, Page 1253.

Work done was the basis for calculating the WRHP, and for the most part it probably achieved the +/-1.5% standard. At 15 HP per 1000, up to 1.5% seems to be a realistic assessment regarding the range of uncertainty that accompanies the Willans lines. There are occasional plots where this standard of accuracy was obviously not achieved. The scatter problem is further complicated beyond experimental error in that some of the scatter is real, given the small variations in steam chest pressure and superheat. The Willans lines for IHP & WRHP routinely deliver R^2 values approaching unity, which accords with low measurement deviations from trend in percentage terms. When the difference between theses two large numbers is examined, the MF, then the data set R^2 values approach zero due to compounded error; the randomised "high" or "low" bias of speed related data sets relative to the overall trend of all the MF data independent of speed are frequently in evidence. Random number experiments have shown that such MF data set biases may not imply a real shift in measurement accuracy since exactly the same ITE & WRTE values are always entered. The resulting experimental outcomes showing clear "off-trend" bias are entirely the result of random variation within the set measurement accuracy parameters. High R^2 squared values are not axiomatically an indication of accuracy. Consistent error would also score high.

The limited scope of the experimental data, routinely fails to cover the full range of power and speed portrayed in the test bulletins. The published data for the lowest and highest working rates is evidently often based on extrapolations, and as such is sensitive to the French Curve syndrome described by Carling. As explained above, extrapolations using the XL curve fitting formulae cannot be relied upon either. This problem was apparent when looking at the VRUx determinations, when it was found constant values did not obtain over the full working range, though they did for the bulk of it. The outcome for BR5 73008 in Figure 29 for example; covered a range of 12,000 to 24,000 lb/hr as against 8,000 to a little over 26,000 in the test bulletin. This degree of cover, around 70% of the working range, was typical.

Finally, returning to the constant steam rate deviations encountered on the Derby road tests, it should not be thought the Swindon road tests were immune from this problem. The locomotive resistances evident from the Swindon derived test bulletins, though at least satisfactory in regard to the general shape of the LR curves, are far from anomaly free. Below the LR curves as derived from Test Bulletins Nos. 3 & 4.

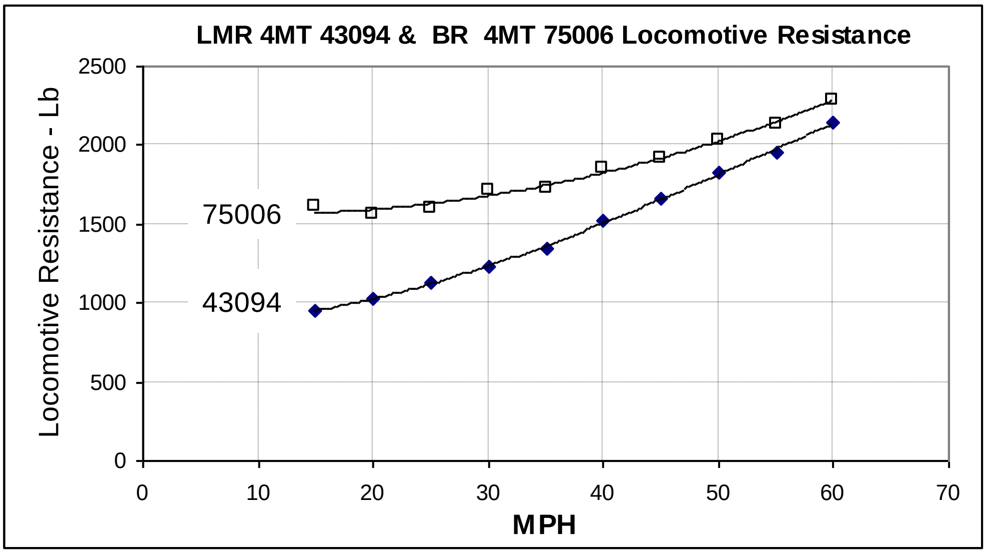


Figure 32 Note the marked LR separation at low speed

The LM4 weighs in at 99.4 tons and the BR4 at 110.05 tons. At 20 mph the respective resistances are 55 and 83 HP, a difference of 50%.

The Swindon test team had the advantage of a test route featuring fewer and less severe gradient changes, enabling longer periods of relatively steady pace. This will likely have simplified controlling the steam rate, though nevertheless, the diversity in LR outcomes as shown above, and in other cases, was at least in part, contributed to by steam rate uncertainties.

On the evidence of the Swindon road test data for 75006 and 71000, significant steam rate deviation tended to occur at the lower end of the speed range when speed was changing more rapidly, acceleration forces, and steam rate increments potentially rising quickly.

The mean steam rate of 23 spot readings based on speed and cut-off for 75006 works out at 15,214 lb.* This is not representative of the overall average for the test, since it is based on instantaneous values rather than a summation of all 48 cut-off changes of varying duration shown in a series of steps, and the associated speed changes. The overall test average was probably closer to the nominal rate. The point that emerges here is that significant departures from the nominal test steam rate could pass undetected; the summation of increments procedure with a metered water supply notwithstanding. Unseen short term boiler water level changes and shifting gradients and inertia effects provided a cushion of uncertainty. From MPs 103 -106, for example, on a constant gradient, cut-off is *shown* held at 24% for approximately 2.8 minutes as speed rose from 60 to 68 mph. Steam rate will have increased about 12% over this section. The bulletin of course, working with the visible metering summations, showed only minimal drifts from the nominal steam rate at any point, as published in the bulletin.

It was perhaps inevitable that cut-off adjustment of steam rate and the available instrumentation had its limitations as a means of controlling Q. The increasing heat drop and reducing exhaust steam specific volume with rising speed and cylinder efficiency for given steam rates was challenging on road tests, even when the density effect was understood. It maybe, the cut-off changes were more gradual than shown. This pretty well concludes my investigations for now, at least I think it can be agreed that the determination of locomotives resistance and machinery friction was no easy matter, or for that matter, the production of test bulletins more generally.

John Knowles Submissions 4 July 2017 and 2 April 2018

As previously, points raised will not necessarily be taken in chronological order, words in quotation marks and emboldened for clarity are his own. The underlined subheadings
are mine. Quotations by others are in italics. There may be some repetition here and there involving points raised above or in the earlier correspondence. This occurs because the same points keep re-emerging, often in mutated form, calling for further comment.

Some General Points.

"Doug seems to believe the data are sacrosanct, apparently perfect, or if not perfect (a real world situation?) they are as good as can be obtained in the real world, and are not to be questioned."

This is far from the case, contradicting my many writings on the subject down the years, of which he is aware. Were it so, I would not have spent years tying to make sense of locomotive experimental test data generally and the Rugby and Swindon record in particular. I have posed many questions and identified numerous anomalies over the years and extensive correspondence since 1970 testify. Even within the contractual measurement limits, the randomised scatter in the small remainder situation is fundamentally problematical. Some disparity is a statistical inevitability. Obviously a satisfactory standard within the understood limitations was not always achieved, some highly aberrant outcomes affecting various aspects of the data is evident; systems can malfunction. A key point here is 'measurements' as opposed to the lottery of small remainders. On a direct measurement basis the WRHP data (Willans Lines) returns higher consistency over time than the IHP data in the early years. Overall, the latter was more erratic in this regard (higher scatter- lower R^2) and inconsistent with later outcomes. More on this below.

- Test Bulletin No.4. Road Test No.1   14,200 Lb/hr steam rate.. Cut-offs shown as a series of steps. Steam rates calculated from steam Rate v cut-off and speed – Figure 15.

My very first writings on this topic in 1970 began:*

"The steam locomotive is not an animal the test engineer would fondly regard, for as the discrepancies in the BR Test Bulletins bear witness, it does not readily give an accurate result. And later -*.These results (LRs) can thus be taken to show constant losses. We thus have nine sets of results, seven of which suggest that locomotive resistance at any given speed is a constant independent of power output, and this has been taken to be the case. In stating the above however, it should be noted that this runs contrary to engineering experience and logic, and some rise in losses with effort should occur."*

"Doug uses Carling's belief that because the ITE results for the same test circumstances fall in a narrow band, the ITE data are acceptable, even accurate."

I don't know where this idea comes from. On the contrary, the opposite is true of IHP and ITE over the history of the plant. Perhaps he meant to say WRTE. The performance of Farnboro'' indicator took some years to reach a satisfactory level of performance and was not free from some setbacks along the way. It is the WRTE Willans lines that I have generally found consistent for different test series of the same locomotive type. In contrast to the claim of "consistent" IHP data early in this correspondence, it is often poor. This emerges most clearly when the IHP data is examined in the basis of specific steam consumption. The outcomes often verge on the erratic, with evident 'strays' and poor R sq'd values.

It took some years for the indicating equipment and process to reach a satisfactory standard of performance, and progress was not without some setbacks along the way.
Even then, the occasional episode of wayward performance was not unknown in later years such occurred as late as 1959 with 92250 in Giesel ejector guise. The IHP SSC data for 50 mph produced a medley of strays: Figure 34.

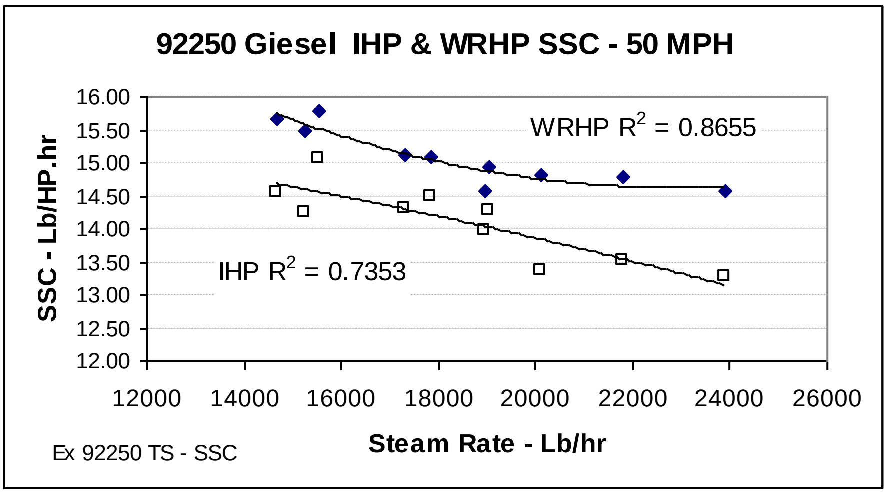


Figure 34 Most of the IHP 'strays' from trend evident here are likely of spurious value since for the most part, the corresponding WRHP plots remain un-persuaded and stick close to trend. The IHP's slightly convex IHP trend line is the wrong shape.

<u>Indicator Calibration Tests</u>

There were three episodes of comparative indicator tests. The first series compared the Rugby Farnboro' in dicator with Maihak and Dobbie mechanical indicators supplied and operated by visiting Swindon engineers in January 1953, The Rugby v Derby Farnbro indicators were matched later that year, and again in March/April 1957. Only this last test series achieved, for the most part, close agreement, with average results within +/- 0.5%.

…………………………………………………………………………………………

* Test Result Anomalies – An Interim Study; D. H. Landau; Stephenson Locomotive Society Journal, December 1970.

Initially the 1953 mechanical indicator MEPs were up to 10% higher than the Rugby Farnboro' outcomes. Subsequent calibration checks reduced the discrepancies to +2% for the Maihak, with the D & M still 7% high at low steam rates, then falling to about ½% at 23,300 lb/hr. On this showing the D & M indicator was an unsatisfactory piece of kit. The Maihak indicator re-calibrated results were consistently 2% higher than the Farnboro'. The differences here perhaps represent a margin of uncertainty.

The intermediate 1953 tests deemed the Derby Farnboro' to be indicator erratic, with mixed results overall. The Derby variance with Rugby was up to +13% - 3.4%. Full data sets are available for Rugby tests 872 to 882 immediately preceding these tests. Each test involved averaging up to 10 indicator diagrams. Maximum scatter was +/- 2.9%, averaging +/- 1.5%. Speeds covered 30, 50 and 70 mph. The final Rugby/Derby Farnboro' indicator results were as tabled for the 92050 Series 2 tests - page 24.

“**It would be wrong to regress DP against Q. Q has already influenced ITE, at a rate varying with Q per se and V, and as seen in the Specific Steam Consumption**.”

This objection is without any rational basis. The relationship rejected is as would be derived from WRHP Willans lines. It removes the obvious way to compare WRHP outcomes of other test series with the same type at given speeds. Steam rate (Q), is the most accurate baseline of available from the Rugby data, (perhaps not quite so secure when the exhaust injector was (rarely) in use). The WRHP relationship with Q is unaffected by whatever the IHP *measurements* turn out to be. The determination of WRHP is an independent function. There were several episodes where cylinder indicating was omitted and the measurement of WRHP continued. Presumably the indicating equipment was undergoing repair or modification. The WRHP Willans lines were then the adopted basis of comparison, as for example the 92015 regulator experiments. The plotting of WRTE against ITE gives a direct measure of mechanical efficiency. Such plots for given speed sets have established one of the few certainties to emerge from within the Rugby data: WRTE v ITE at a given speed is a linear relationship.

“**Doug should not be concerned about a proper regression line (rather than an EXCEL trend line) not passing through the actual data. A best fit will often not pass directly through any of the data. No method of analysis can make up for poorly measured/inaccurate/inconsistent data or improper specification of the equation to be fitted**.**”** (JK letter 25 October 2016)

“A best fit not passing through the actual data” sounds like a mathematical aberration rather than a revelation of a supposed statistical reality. Something akin to walking on water or flotation without getting wet. It is absurd. A good example emerges in his letter 4 July 19 2017 (page 37) where he cites a Graph that I gave him some years ago that has not appeared in this correspondence – Figure 35a.

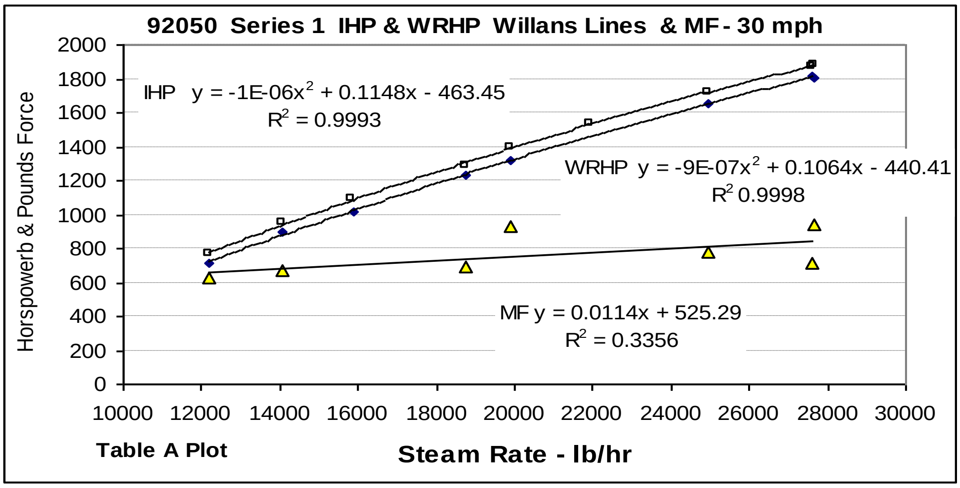


Figure 35a The 9F returns a positive ITE – WRTE separation. The MF values average 763 lb, the smoothed outcome ranges from 706 to 840 lb.

He comments;

“This exercise was supposed to show that TSR was constant at 30 mph (like a dog following its master on a lead he claimed – see *Backtrack*, April 2014, p 253). It does the exact opposite. It shows TSR supposedly varying with Q, but not as fast, and at a declining rate, to high levels.”

It was most certainly not originally presented to show “constant TSR”, from a long correspondence John should know that is not a view I hold. What he actually said at the time was that seven plots was too few, rendering the positive MF outcomes worthless

John goes on to calculate the smoothed MF outcomes derived from the formulae shown in Figure 35a. While this exercise is mathematically correct, the outcome from the smoothed results significantly raises the MF from an average of 763 to 1270 lb. A comparison of the “before and after” IHP and WRHP Willans Line proved revealing as Figures 35b & 35c below.

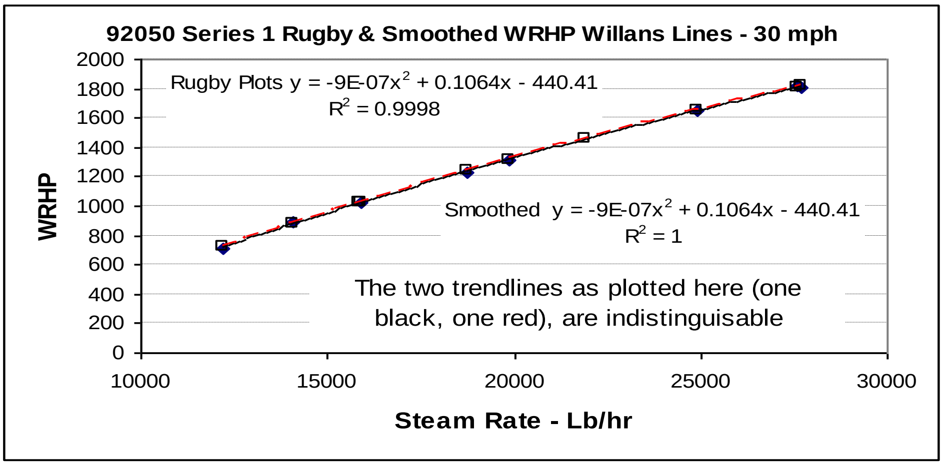


Figure 35b There was little adjustment to the Rugby WRHP plots. They fell within 0.6% to – 1.7% of the smoothed values; the average deviation was 0.7%.

The smoothed IHP plot, Figure 35c is unsatisfactory, inflating the IHP outcomes.

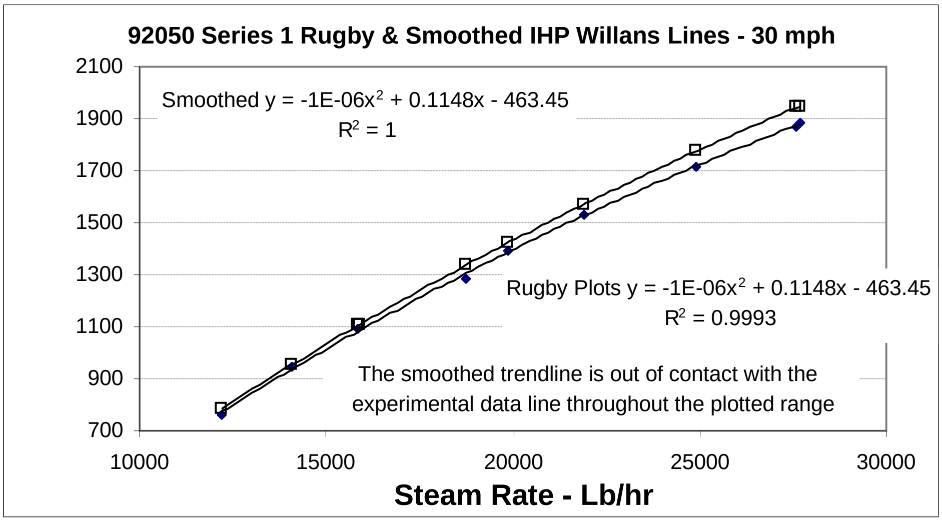


Figure 35c The upper "smoothed" IHP trend line makes no contact at any point
with the Rugby plotted data. This is clearly a mathematical aberration, hence the erroneous uplifting of the MF outcomes in which the smoothing of the WRHP trend line plays no part.

The smoothed IHP values are clearly an aberration and are seriously in error. The answer has proved quite simple; the XL curve fitting programme *defaults to four decimal places.* An override option increasing the decimal places is available: RH click on the trend line equation, and then choose 'Format Trend line label', select 'number' then choose 'decimal places'. In this instance 9 was selected, the aberration disappeared, refer Figure 35d.

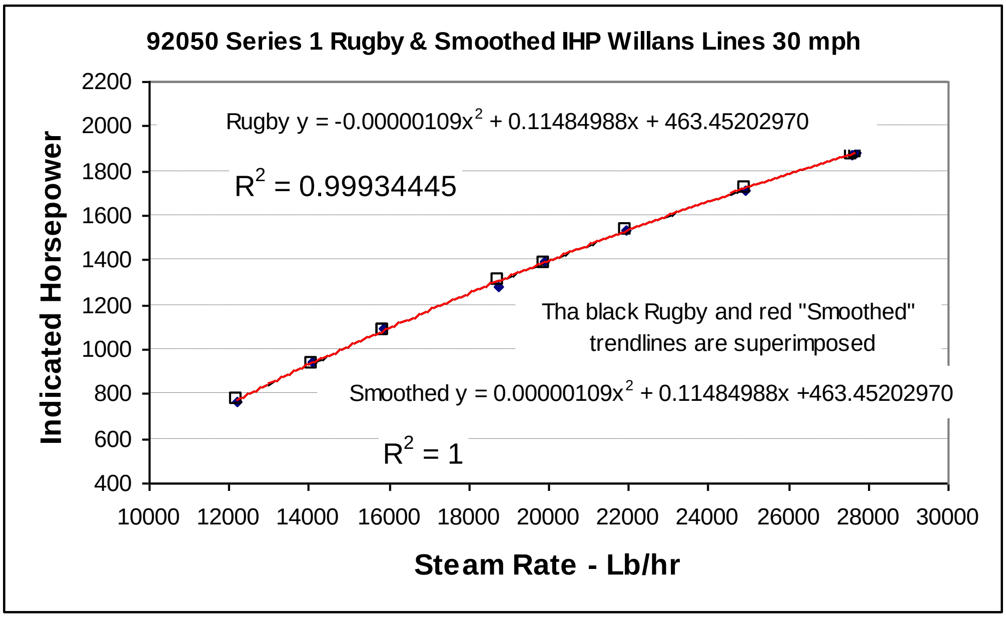


Figure 35d The enhanced decimal place formula and Rugby trend lines are indistinguishable. The average "smoothed" IHP correction was 0.1%

“The Rugby indicator results are highly consistent for a given engine when regressed against Q and V.” “In addition he calls on repeatability as a criterion for acceptability or accuracy of data, when all the repeated data can all be wrong.”

We don’t disagree on this basic point. While repeatability is a prerequisite, it not in itself an axiomatic proof of accuracy, as I have written elsewhere. The same limitations apply to high R^2 values as also pointed out, obviously fixed calibration or systematic errors might be in play. I note that early in this correspondence John was content to cite the indicated horsepower data as “consistent” in an attempt to infer WRHP data shortcomings implied by negative MF outcomes fell entirely on to the shoulders of the Amsler Dynamometer. This supposed “consistency” was inaccurate; the said data appears to have been taken on trust without due scrutiny. The chequered history of indicator development described in the Ron Pocklington correspondence receives no mention. The recorded IHP for the BR7 increased with time, as I have shown. Indicator performance was not deemed satisfactory from both the reliability and diagram quality standpoints until early 1955. The differences between the 92050 test Series I & 2 IHP results were overlooked. (The difference in this case proved to be steam leakage, not IHP measurement,)

“Only late in the testing was it discovered by simple consideration of the data, that for LR in this case, that such was not correct.”

The Rugby/Derby test staff certainly seem to have been slow to take action; this was likely down to the test plant work-load, but they could easily have re-introduced indicating for the road tests at an earlier stage. However, contrary to the above assertion, Report L116 indicates the LR problem was recognised early on, as indicated in its opening sentence: “*In all cases where locomotive trials at Rugby have been followed by road tests carried out with the LMR Mobile Test Plant there has been a lack of reconciliation of the results to the extent that values of locomotive resistance obtained by subtracting road T.E. from Rugby cylinder T.E. have not been acceptable.”*

It later continued: *“It was first observed with the E.R. B.1 Class 4-6-0 Engine No. 61353 during the course of a day’s running from Carlisle to Skipton and return, the steam rate produced by a particular setting of the blast pipe pressure during the outward run could not be accurately be repeated on the return. The only difference of any significance between the two test runs was that the overall average speed was lower on the return, owing to the nature of the test route.”* The road tests were in 1951.

<u>“Perform”</u>

“…the Perform program gives results a little higher than those from Rugby. Perform is by far the best way of approximating cylinder outputs.”

This is an optimistic view of the Perform programme. For those unfamiliar with the late Professor Hall’s “Perform” programme, herewith some brief notes.

Hall, a nuclear power engineer, did some ground breaking research using a live steam model, demonstrating that even with superheat, under some circumstances condensation could occur in the course of a power cycle. In summary he developed a programme embracing the many complexities of thermodynamics, fluid dynamics, valve events and the various dimensionless coefficients involved to compute IHP. He then compared his theoretical results against the published data.

He was not privy to the actual experimental Rugby and Swindon test data that has later become available. His matrixes for comparison were confined to the data available in the *Britannia* Test Bulletin (N0.5) and S Ell's 1953 I.Loc.E paper *Developments in Locomotive Testing*; essentially a test report for high superheat *King* 6001.

Hall was unaware of the notoriety that surrounded the test data for 6001, distinguished by high LR with a distinctly high sensitivity to the level of effort, when he commented ; *'However it has been possible to infer enough information for a start (comparison) to be made using an excellent paper by Ell which describes controlled road tests made in 1953 on the former G.W.R 4-6-0 4-cylinder "King" class locomotive No. 6001'.*

As things turned out the computed results for IHP v speed at constant cut-off traced a similar parallel path to the report data but were over 10% higher at 40 and 50% cut-off. Hall was unaware of the disparate outside/inside cylinder performance of the King; the inside delivering only around 70% HP relative to the outside, and the high pressure drop from boiler to steam chest; about 10PSIG more than a Duchess at the same steam rate, and more still compared to the Scot. Had Hall had access to this data he would likely have been less encouraged. The IHP Willans Line R^2 returns for 6001 covering 14 road tests were mediocre, averaging 0.7933; the range 0.6451 to 0.9002.

The later comparison by Hall for the *Britannia* was generally close to the bulletin values at given speeds and cut-offs. There was however some difference in regard to the actual steam rate at 15% cut-off, and to a lesser extent at 25% up to 40 mph. Hall also converted a few bulletin indicator diagrams in radial form to the conventional stroke base, with an overall trend for the computed admission PSIG values to be a little higher than the actual. Of the indicator diagram conversion for 25% cut-off at 40 mph, Hall concludes that the *'result appears to somewhat out of line with the others, and leads me to wonder whether the location of top dead centre has been correctly defined on the indicator record'*. Shades here of Ron Pocklington's concerns when he forst arrive at Rugby in 1952.

David Pawson, is an expert in using 'Perform'. His recent (MP 38) *How Powerful are UK Steam Locomotives?,* with its *Perform* computed IHP results are tabled below.

Perform Power & Steam Rate Estimates at 25% Cut-Off, 60 mph v Test Bulletin record							
Loco	Perform Estimate		Test Record			Perform indices v Test Record	
	Q Lb/hr	IHP	Source	Q Lb/hr	IHP	Q lb/hr	IHP
Duchess	33,600	2440	R13	31,500	2195	107	111
Reb Scot	23,400	1720	Rugby	27,930	1945	84	88
BR5	18,700	1410	Bulletin 2	17,750	1230	105	115
BR7	22,000	1740	Bulletin 5	21,500	1610	102	108
BR9	23,600	1880	Bulletin 13	24,500	1770	96	106
V2	20,300	1610	Bulletin 8	24,180	1665	84	97
King 4RS	28,700	1730	S O Ell	27,800	1910	103	91
Mod Hall	17,500	1230	Bulletin 1	24,250	1630	72	75

The test record data shown is as interpolated by measurement.

It is apparent, that Perform is unable to replicate the test record steam rates at a given cut-off with both under and over estimates returned. The test plant derived water rates are the most accurate data available, Carling reckoned steam rate experimental error to be "well under 1%". Given the nuances of valve setting, cut-off introduces some uncertainty, but the deviations from nominal values inherent from the crank angularity effect tend to cancel out front and back, and seem insufficient to explain the differences tabled. The valve settings were checked by the Rugby test staff and the practice at Swindon was probably the same. In exception, quite what the true cut-offs were for the V2 middle cylinder is difficult to determine from the bulletin indicator diagrams. That the *Perform* estimated steam rates fall both above and below the test plant values suggests that uniform assumptions for steam port friction coefficients and other design details affecting steam flow are more nuanced than supposed. The measured test plant IHP data is also of course subject to uncertainty, notably the early Rugby data and the Swindon data generally. Had life given Bill Hall more time, and he'd had more access to the experimental record, his ground breaking work may well have acquired a few more tweaks.

.At an estimated 5% accuracy, Perform may well have outperformed many mechanical indicators, but with uncertainty up to 50 HP per 1000; it would play havoc in small remainder situations.

All of the above on the *Perform* programme is a bit of a diversion, and not really relevant to the discussion in hand; but John Knowles having referred to it, it seemed an outline of would be helpful to those unfamiliar with Hall's work.

"It is therefore extraordinary that Doug Landau, after all these years, claims to be able to judge the Rugby data better than Carling, and to want to do so without explaining how. That is the same as setting his face against regression results – nothing declaring against the Rugby regression results, specially by me, is to be tolerated I suspect too, that he believes that scatter is evenly distributed and that the true answer lies in some sort of average of all the data. I fear not. The testing and consideration of the data requires consideration of the scatter, its

extent and an examination for biases. Simply declaring that the Rugby data are fit for providing TSR values avoids crucial steps in showing that it is fit. Declarations are empty if the steps have not been taken. Doug Landau has never shown that he has considered the data, so it follows his declarations are empty. "

These imaginative assertions are travesty of my thinking and methodology. *Pure rubbish* would be a fair description. Not content with putting words and thoughts into the mouths of the dead, he now seeks to do the same for the living; desperate stuff. I have not challenged the powers of regression. What is being challenged is flawed thinking and misapplication, reducing the exercise to the status of reading tea leaves. What I am supposed to explain? Essentially, all I have done is present the recorded test data in clear unequivocal form. What could be more straightforward for example, than the linear WRTE v ITE relationship?; a simple representation of the recorded data; likewise Willans lines. In that form scatter is generally of low magnitude as a percentage of the values directly measured. Estimates have been avoided as far as possible, are few in number, and when deployed, their basis is explained and open for challenge if thought at fault. If my experiments removing one or two plots from data sets is deemed 'playing with the data' so be it; I am simply doing so to demonstrate the random uncertainties and sensitivities of the data sets exampled. Some plots are inevitably more accurate than others

"I suspect too, that he believes that scatter is evenly distributed and that the true answer lies in some sort of average of all the data."

Why would anyone think anything so silly? My randomised small remainder experiments show the complete opposite.

The actuality is that the scatter falls into two camps. Though random, scatter is small when referenced to direct relationships such as ITE and WRTE Willans Lines or WRTE v ITE, where the scatter generally falls within the understood metrological limitations. The second category is the chaotic statistical joint venture of small remainders where scatter can readily exceed +/-100% and random clusters of bias and the occasional negative outcome may occur.

As to **"doing better than Carling**", I agree with Carling that it was not possible to determine internal friction *within fine limits free of some uncertainty*. He attributed this to the small remainder problem and thought the same in regard to locomotive resistance notwithstanding a larger remainder. In regard to the direct measurement of WRHP, he said *"We got the results right".*

The advantages I have had over Carling is considerably more time, a comprehensive overview of perhaps 80% or so of the Rugby test programmes data, and the time saving powers of the Excel programme when it comes to plotting graphs, fitting trend lines and calculations. A considerable degree of the mental labouring aspect is eliminated. That is not to say that Excel is free of limitations and potential pitfalls.

I feel compelled to repeat and elaborate: the last thing I think is that "scatter is evenly distributed". Indeed the random distribution of speed specific data set groupings on occasion show clear signs of positive or negative bias relative to the overall trend for locomotives data sets. The idea that more plots axiomatically deliver sounder outcomes is not bourn out by examination. The last plot in a data set may well be a wild card disturbing what would otherwise have been a plausible relationship. "Unbalanced" outliers may occur. The best way to minimise this sensitivity is to plot WRTE against ITE. This relationship follows a straight line law in the form WRTE = Ax – B, where x is a coefficient sensitive to A, the ITE, and B a negative constant notionally representing the resistance of the power transmission machinery including the coupled wheels when coasting without any application of power. Such outcomes should deliver a negative constant. In other words as long as the locomotive is moving the power transmission machinery including the coupled wheels will encounter some machinery friction with steam shut off. Compression effects in the cylinders when coasting may of course add to the friction losses, but theoretically this should not effect the constant as derived under power. Some experimental error, will however be attached to said constant, given the sensitivity of the linear trend line tilt sensitivity to the distribution of the scatter.

The 'constant' outcomes as tabled for four 9Fs on pages 124-125 above examples these uncertainties. As things turn out, the constant may sometimes be falsely positive as cited for 92166; an unequivocal example of random scatter mischief..

A reproduction of my chart plotting the recorded MF data for Jubilee 45722 is criticised as below.

"These trendlines are not regressions. As immediately above, there is no discipline to them – Doug Landau has used them here to obtain relationships which do not exist in physics or mechanics. They can be done without any of the tests possible with regressions." (Reference to 45722 chart of Machinery Friction v Speed – plotted Rugby data.)

The chart is simply a plot of the recorded test data using the Excel curve fitting programme.. Contrary to his assertion that the relationships shown "do not exist in physics or mechanics", there are very sound theoretical reasons why the MF v speed relationship *may* take the dished form as represented by the trend line. At low speed the traction force piston thrusts are at their highest, initially falling rapidly with speed; in parallel, the rotational and sliding friction is increasing as a function of speed, and
the dynamic forces are increasing as a square of the speed. In such circumstances a dished MF trend line is entirely possible from the theoretical standpoint.

I accept that the outcome shown for 45722 might equally be simply down to the randomised bias of error within the scatter pattern of the overall data set . In contrast some of the Rugby data sets seem to flat line across the speed range. Such outcomes, based on the small remainder data could equally be the product of randomisation. The iterations of force, friction, dynamics and inertia within the span of each revolution are complex. While the shape of the MF v speed relationship may remain an open question, a flat-lining outcome is theoretically difficult, but cannot be ruled out. It is no wonder Dennis Carling thought the determination of MF (and likewise LR) to be problematical. A situation he attributed to the small remainder problem.

"Doug Landau appears to be unaware of the convention applying to the term static axle or bearing load. He thinks it means without the wheels turning. It applies to both circumstances. There are plenty of examples of the term static in the sense in which I have used it – see for example the paper by Cox on locomotive axleboxes, which he quoted, with the flavour that Cox's paper proves I am wrong in some way. If this still offends him, he can ignore the word static."

I don't know where John gets this idea of my objection to 'static' load comes from, I think nothing of the kind. Obviously the 'static' load is a constant that never goes away, whether stationary or in motion. In motion, dynamic effects, track behaviour, and imperfect balance will augment said vertical load both positively and negatively within the course of a revolution. In citing Cox's diagram of the forces acting on the coupled wheel axleboxes when under power, I was making a point he seems unable to understand. (my letter 7th March 2017). (He has also not revealed the **"other analysts"** that, apparently, do not consider the resultant (journal) loading part of. MR.) The point being that the sum of piston thrusts, dynamic forces and the vertical (static) load on the coupled axlebox journals is *less* than the mathematical sum of these forces. In other words, there is a degree of opposing forces and vectors cancelling out. It is a shared mitigation.

His idea of dismissing coupled journal friction and V squared losses as part of MF, in order to determine the notional values of 'Pure Machinery Friction (PMF), overlooks this mitigation. (His 9F statistical analyses pages 45 - 48). Deducting a questionable friction estimate for the coupled wheel journals when notionally behaving as a passive unpowered vehicle, in order to discover the delusion of PMF, is an exercise without any conceivably useful purpose. This corruption of the measured evidence by interference is further compromised by subtracting a doubtless dodgy estimate of losses attributable to dynamic effects. The PMF idea as an analytical approach can only be described as utterly clueless. Whose "playing with the data" now.

I put his idea to Adrian Tester, he replied: "*As you correctly point out, WRTE and the pull recorded on the Amsler dynamometer were one and the same. Also, WRTE has to be net of all the machine friction inherent in driving the locomotive. Axlebox friction forms part of MR. it does not appear in WRTE; it represents part of the difference between indicated power and WRHP. It cannot somehow escape to be part of the WRTE, only to be absorbed later. I don't see the logic of that."*

Note also, that this exercise creates a smaller remainder to be tested against the previously existing levels of scatter and uncertainty. Such an exercise is implicitly inferring that the recorded IHP data was perfect - blameless. The dynamometer was entirely at fault; strident confirmation bias on the march.

Such unnecessary meddling is wholly avoidable by simply treating the coupled wheels as part of the power transmission system which is exactly how they function. That is what is actually measured, it constitutes the overall *mechanical efficiency*; as referred to as such by the Rugby test staff. It's interesting to imagine, how, in the absence of any adhesion weight, power would be transmitted. The statistical analyses on pages 45 – 46 of his letter 7 July 2017 are worthless: Pure guff.

The estimates of journal resistance (CWBR) are, according to his earlier citation, based on a misunderstanding of what Ell's paper on rolling stock resistances reveals. The paper was of much interest since it concerned 520 ton freight trains of varying length and vehicle type in both the empty and loaded condition. The related resistance formulae fitting the data in all eleven cases took the form $R = A + BV + V^2/C$ lb/ton. While it is true that notional frictional rolling resistance relationships as a function of axlebox loadings were determined from the constant A term values fitted to the experimental data, they cannot be construed as actually representing the journal friction *across the speed range*. It has been shown (DHL R13 Audit) that the individual values of the ABC resistance formulae will accommodate some permutation of the coefficient values while still delivering a satisfactory fit to a given curve. In other words any seemingly causal relationships of A, B & C are tenuous.

Even the simplified form, $R = A + V^2/B$ can sometimes do the job. In summary, the three elements of the classic resistance formula, may at best only approximate to some causal functional relationship; obviously the squared function will have lot to do with aerodynamic drag, but close representations of the causal realities cannot be assumed. This does not matter of course if overall, the curve fitted is considered sound and the formula fits the purpose of estimating *total* vehicle resistance.

It is another matter when trying to determine the true journal friction across the speed range; it is not a constant. JK's Tables 1 to 4 for 92250 data for erroneously show constant a CWBR value of 228 lb across the speed range; given established bearing theory, this is wrong.

Coefficient of friction $\mu = ZN/P$ where Z = Viscosity, N = RPM, P = Bearing pressure

It is apparent that μ is a function of speed and an inverse function of P. The rising ZN/P relationship only obtains for values upwards of around 25 once *hydrodynamic lubrication* has been established. Starting from rest the *boundary film lubrication* zone (otherwise known as stiction) is encountered, then falling rapidly, the intermediate *mixed film lubrication* zone being reached from about ZN/P = 5, then falling to a minimum on reaching the hydrodynamic state; from this point μ increases with speed.

These uncertainties are of course wholly avoidable; PMF = Pure Mechanical Fallacy..

“He gives no reference for the claimed confirmation of TSR by road tests for the Crosti and standard 9Fs, nor explained how he reconciled what are essentially different measurements – TSR given on the test plant and LR on the road. Given the lack of repeatability in the Rugby data, he does not say which 9F data among the non-repeating 9F data he picked for his own use as the resistance of the 9Fs”.

A curious statement; obviously, if two locomotives of similar vehicle architecture, size weight and shape display significantly disparate machinery friction, the locomotive resistances will be inevitably be similarly disparate. The Crosti 9F 92023 returned notably higher MF on the plant than 92050 and all the other 9Fs tested. The confirming details of these road tests were covered on page 117 above. The other 9Fs were consistent in regard to MF when examined as WTRE v ITE.

The more significant comment here is **“which 9F data among the non-repeating 9F data he picked for his own use as the resistance of the 9Fs”.**

This supposed “non repeatability” is based on small remainder data sets, when such disparate outcomes are highly probable. It’s about as meaningful as comparing the results on a Bingo night. His claim of **“non-repeating”** is thus entirely erroneous.

His statistical analysis of truncated small remainder data (PMFs) for 92250, vide Figures 1 to 4 for 92250, pages 43-44 is flawed at every level from conceptual to execution. Small remainder outcomes are the results of joint enterprise, not direct

measurements, and are no place to start with statistical dissection in the first place. Other, more direct relationships are available. His first step is to corrupt what is already inherently troublesome data (SRMs) by deducting highly questionable and wholly avoidable estimates. Thus the SRM gets even smaller whilst retaining the same degree of scatter. All apparently a the outcome of single handed of dynamometer malfunction. This approach can justly be deemed clueless. The scatter originally displayed, generally falls within the SRM potential scatter given the known accuracy limitations of the indicating gear and dynamometer, as has been demonstrated by the randomised small reminder experiments.

Notwithstanding that the same notionally perfect measurements were entered on dozens of test runs, it was all too apparent that individual speed data sets and whole data sets could sometimes display upward and downward bias relative to the perfect fixed remainder value entered. Off target clusters may occur. The spread of individual remainders could exceed +/-100% of the actual fixed true value adopted for the exercise, with the occasional negative result. Notionally the averaged data sets should loosely approximate to something approaching the true mean value across the working range tested, but discrepancies for individual data sets may average significant diffrences in small remainder form, or worse with only limited SRM plots available. That's about all such data sets are good for at best, a rough approximation. Some outcomes may also be hostage to the average work rate of the individual test series, which may differ sufficiently in magnitude to skew average outcomes.

When, as in Fig. 4, page 43, July 2017, 3 plots are cited **"as good as it gets**" but a fourth as demonstrating the **"lack of consistency or repeatability.",** one can only ask **;** Whatever ever happened to the call for plots in double figures as essential to providing meaningful samples for analysis? Back to reading tea leaves it seems.

"Even knowing these ranges (measurement limitations), the effects of the small difference between two large numbers problem could well prevent satisfactory data and analyses emerging." *(My italics).*

Exactly: A moment of sanity? The moral here would seem to be, where possible, leave small remainders alone.

In forming his **"consistent IHP"** conclusion, this was presumably by plotting the IHP Willans Lines at given speeds The relationship with speed and cut-off perhaps providing scope for secondary analysis. High R^2 values alone are not proof of accuracy, merely consistency as mutually agreed. The WRHP data Willans Lines return equally high, and often superior R^2 values. It can sometimes be more revealing to examine the IHP and WRHP data in specific steam consumption form (SSC). This is a form of amplification, and aberrations sometimes emerge. Some examples are below.

"but scatter is lack of repeatability," Some experimental instrumentation error is inevitable , normally falling within known limitations for direct measurements. In the small remainder situation, the margin of potential experimental error is intrinsically magnified, and are a troublesome basis for statistical examination since two uncertainties of unknown deviation contribute to every outcome. The line seems to have been taken here that poor statistical outcomes are solely indicative of dynamometer malfunction; an unlikely scenario. The WRTE and work done was recorded and summated mechanically over the course of test period. The IHP and ITE was determined on a sampling basis during the test period, the average of around half a dozen readings being taken as the test value. The indicator diagram determination was literally a case of "joining the dots", not that easy when faced with joining a "snowstorm" of dots in the early years of the test plant, as attested by Ron Pocklington. It was not until sometime in 1952 that RP took up the reins at Rugby, when indicator diagram "snowstorms" were evidently a problem.

The IHP data was less than "**consistent"** in the early years of the plant. The tests with Merchant Navy pacific 35022 were notorious for delivering negative MF values. Onwards from test run 744/1 to the last of the 1952 test series, the recording of IHP discontinued; WRHP recording continued. These tests involved variable speed at constant steam rate and cut-off, a test scenario inevitably involving part regulator working, with steam chest pressure reduced as speed increased. A procedure possibly adopted to replicate the way the Bullied pacifics were often worked in traffic. The WRHP curves recorded in these tests, as plotted in Figure 36 were of consistent form collectively for the four steam rates shown. Yes, of course such consistency is not in itself proof of accuracy, but it is a long way from the 'tea leaf' chaos delivered by the small remainder data, and is therefore the proper subject for regression analysis or any other means suited to testing the data's veracity. It eliminates the problem of apportioning the random dual contributions to joint error as delivered by small remainders.

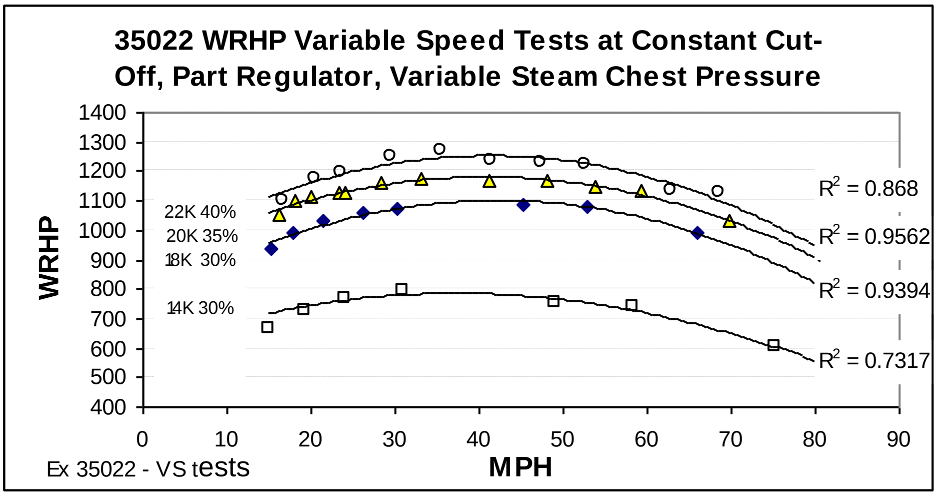


Figure 36. An orderly distribution of plots. The speed steps were initiated at 5 minute intervals.

This discovery prompted a comparison of the simultaneous IHP and WRHP SSC data where available. On this basis the IHP R^2 values were poor relative to the WRHP data, as exampled in Figure 37 below. Similarly erratic results obtained at 30 mph.

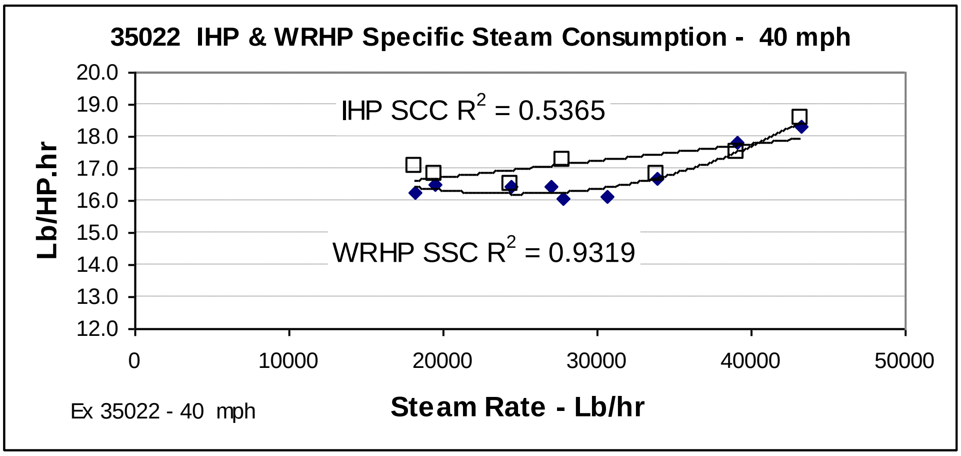


Figure 37. Both trend lines are polynomials, the IHP plots were unable return the characteristic shallow dished SSC curve as returned for WRHP. All the IHP

plots should of course fall below the WRHP plots. A similar outcome was found at 30 mph with the IHP SSC poly trend line flat-lining; R^2 0.534; WRHP; R^2 0.9667 (Figure 10 page 12). The recorded WRHP SSC values are unexceptional in regard to implied thermal efficiency. These were full regulator tests, boiler pressure averaged 272lb and steam chest pressure 261 lb.

Noteworthy here that is that the nine WRHP SSC plots do not all slavishly follow the usually observable 'dog on a lead' response to the linked IHP pairings, but stick close to the trend line. The inference here is that the IHP plot at circa 28,000 lbs/hr is erroneous.

As possibly, in a different way, are all the other IHP plots. The WRHP SSCs at over 16 lb per WRHP hour are unremarkable.

"Fig.4. "…….. the three observations in the far top left of Fig 4 are as good as could be expected, but the fourth observation at 16,800 lbs demonstrates the lack of consistency, or repeatability*."*

I could not see the "far top left" plots (I think top right was intended). The alleged rogue fourth 16800lb plot lower down falls within normal small remainder scatter. As far as I can see, all the plots shown fall within the predictable scatter range. Potentially, the trend for such small samples of small remainders over a short abscissa range could point anywhere. The approach portrayed is about as meaningful as reading tea leaves. Applying regression to random small remainders rather than the direct measured relationships generally returning high R^2 values is beyond logic.

"Heat from any effect (the Belleville washers and dashpot) will be lost from measurement, so that measured DP will have been too low and measured TSR too high. "

Any heat generated by the Belleville washers was minimal, resulting from the slight hysteresis effects. The force at the drawbar and Amsler dynamometer were exactly the same, simultaneous, equal and opposite. The dashpot, being in parallel, rather than in series was another matter. As Carling pointed out;* "*Being wise after the event he considered that, had the whole of the system been suspended on the drawbar, not fixed to the foundations, and acted as an inertia damper, there could have been no falsification of mean pull. It would have involved a major engineering modification and was not justified."* The dashpot falsification was plain to see; under steady state running conditions the recorded drawbar pull steadily increased. However, as now established, by the end of 1950 the dashpot had been decommissioned and is irrelevant. .

"Adrian Tester has informed me (personal communication) that Carling, superintendent of the plant, noted that the Amsler could record to +/- 1% for pull, and provided data within a +/- 1½% range for work done and +/- 2½% range for power (these are presumably at its own recording table, as might be expected from what these terms represent and the accuracy of the components. Only the pull, however, was recorded."

Did Adrian Tester realy write the the last sentence? Writing in Backtrack* he explains that *"speed was recorded in miles per hour via the Selsyn drive* *thereby enabling work done in ft lb to be integrated by means of an Amsler spherical integrator to give rail power."* If only pull was recorded, how would the "work done" (HPHRS), as clearly referred to, be determined? John has been given a sample Rugby test sheet: Drawbar Pull, Work Done, Speed,. Distance (miles), and the Mediating Gear Inch Seconds are among the items recorded. The mediating gear inch seconds recorded the net deviation of the coupled wheels from top dead centre on the rollers over the course of the test. If the recorded value was the same at the start and finish of the test no deviation had occurred. There was provision on the test sheet to record corrections as necessary. The rollers were manually rotated during calibration tests to determine the accuracy of the work done function.

Pure Machinery Friction

Some further points. It was about 16 years ago John Knowles conceived the notion of *Pure Machinery Friction* (PMF). The idea was to describe the machinery friction of the locomotive pistons and motion, free from the friction arising from the locomotive's vehicular aspects – the coupled wheel journals machinery *friction, and windage losses.* What was to be gained from such a concept remains a mystery. It thus might

---

- Locomotive Testing Stations (Part II), D R Carling. Proceedings of the Newcomen Society Volume 45 .1972, p. 173.
- *Stationary Locomotive Testing Part 3 – Adrian Tester; Backtrack; October 2013*

be construed as a corollary, that the vehicle resistance was somehow. *Impure Machinery Friction.* Did this involve different mechanical laws? If the idea of PMF seemed to be simplifying any analytical approach it could hardly do so.

Machinery friction is a complex iteration of ever shifting simple, dynamic and inertia force vectors in the course of each revolution. It was the resolution of all these forces *that was measured on the Rugby test plant. The PMF idea inevitably interferes with the recorded evidence to no conceivable purpose.* The small remainders involved are trouble enough without making them smaller and inevitably subject to flawed estimates in order to extract some supposed item of purity. How is the missing quantity to be apportioned between the PMF and the subtracted VR element? Since the manifold forces and resultant outcome (MF = ITE – WTRE) is less than the mathematical sum of the forces involved, how is this mitigation to be determined and divided when breaking down the measured outcome into two separate quantities? 'Pure' MF and 'impure' MF. Even if the notional VR element of the coupled wheel journal friction estimate was accurate it would, lacking a mitigation allowance, deduct too much, rendering the PMF element dubiously low as the default outcome. It is also noted, that any apparent improbabilities resulting from the supposed scientific analyses of these truncated remainders are axiomatically presented as proof of the Amsler Dynamometer deficiency: the Farnbro indicating equipment being assumed satisfactory. A travesty of supposed objective analysis..

46165 Tests Analysis

These tests are examined on an SRM basis by John Knowles in detail at 40 and 50 mph, tests at other speeds having insufficient plots. The conclusion that the 40 mph data is sufficient for close analysis is ill judged, 13 observations notwithstanding (actually only 12 returning an MF plot). The range of power and steam rates covered is very narrow, both increasing only 5% from the lowest to highest values:1220 IHP for test run 1492 to 1287 IHP for test run 1504. Under these circumstances the hazards of small remainder random scatter outcomes can potentially tilt the overall trend in numerous directions with both positive or negative constants of wide extremities possible. This is exactly what happens in this instance. The “so poor” data falls within the understood metrology limitations when examined over such a narrow range.

Some of the IHP data for 46165 is erratic when amplified to the SSC format, vide Figure 38 below. Whilst the WRHP plots the characteristic shallow dish shape, the IHP poly trend line is linear. Shades here of the problems with the 35022 IHP SSC data. There is an intriguing note in a list of modifications to the Farnboro’ Indicator set-up dated 31st December 1955: “Improvements in 1955”.

“3. Further developments were made in the mid-stroke devices, including one for the inside slidebar of Engine No. 46165.which could be adjusted from the outside whilst the engine is running.” Had a problem come to light?

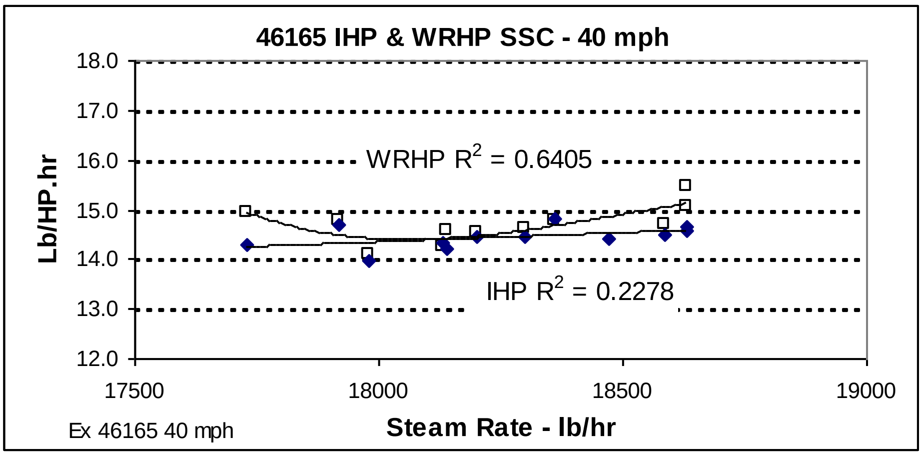


Figure 38. The WRHP plots (squares) generates a characteristic dished trend Line. The R^2 value is mediocre. The IHP trend (dots) flat-lines, R^2 poor.

As already referred to, the determination of dead centres was critical to the accurate determination of indicator diagrams. Some of indicator diagrams for 35022’s middle cylinder in the test bulletin feature compression loops and rats tails, which given the 9.8% clearance volume seems unlikely if dead centre had been properly established.

The other cylinders were not exactly anomaly free. The left cylinder front delivering a very skinny outline at 15% cut-off and low speed, and the right hand only achieving about half boiler pressure on admission back and front. All this was on full regulator, the diagrams only achieving a degree of even work back and front for all three cylinders at speeds of 45 mph and over. The Bulleid gear was clearly a law unto itself.

Returning now to the analysis of 46165 plots given in John Knowles Fig. 6, page 52, It is difficult to see where the numbers come from.

46165 Rugby Power & Tractive Effort Test Range at 50 mph						
Status	Test Run	JK Fig.6 PTTE	!HP	WRHP	ITE	WRTE
Minimum	1564	C,13,750	1130	1076	8,475	8,070
Maximum	1544	C.16,400	1957	1909	14,678	14,318

The mysterious PTTE on the Figure 6 x axis is described in the glossary of abbreviations as the Piston Thrust Tractive Effort, it being defined on page 58 as the net sum of the PTTES and the PTTEVsq'd; these being defined as "Piston Thrust Tractive Effort propulsive and compressive.", and "Piston Tractive Effort forces from unbalanced reciprocating masses dependent on speed squared". Note that the outcomes shown and tabled above exceed the minimum and maximum recorded ITE
outcomes for 46155 at 50 mph. The point at which force PTTE impinges itself on 46165's anatomy is not explained, no force diagrams, sample calculations etc.

The outcomes are hard to follow. The ITE and WRTE working range recorded at Rugby increases by over 70%, in contrast the PTTE increases by only 11%, and at the lowest output contrives to exceed ITE by over 60%. What do the numerical values given for PTTE actually represent? On what parts of the Scot's anatomy is PTTE supposed to impinge? This is quite aside from the fact that the whole exercise is a conceptual misadventure.

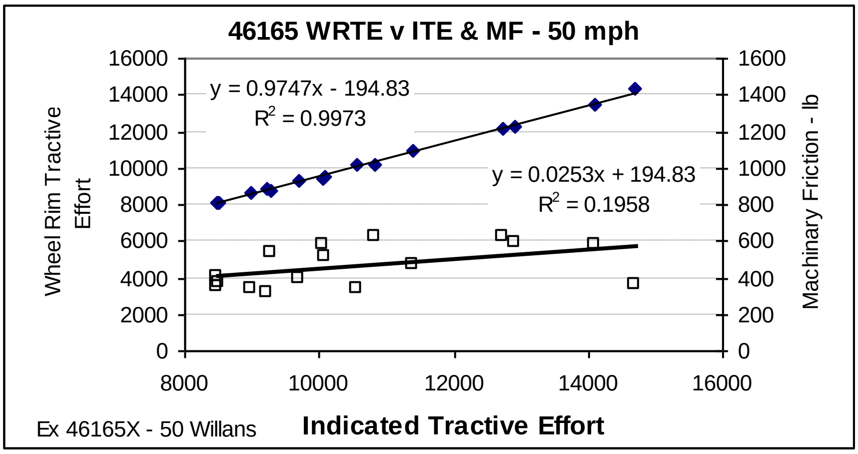


Figure 39. Aside from 1" smaller cylinders, the architecture of the Scot's and Jubilee's power transmissions are essentially identical. Combining the 50 mph test data for 46165 & 45722 (33 plots) returns y = 0.976 – 250.79, R^2 0.9943.
The average mechanical efficiency for the combigned outcome is 95.4%
In both instances the variable is around 2.5% of ITE. Both constants look low.

At one point John suggests a peer review. Confused thinking aside, his presentations fall a long way short of adequate explanation and clarity. Such things as force diagrams, assumed friction coefficients and basis for same, shifting force iterations, sample calculations, explanation of statistical dissection method and theory, etc are notably absent. The prime weakness is the lack of any convincing argument as to why the measured machinery friction, an intrinsically troublesome small remainder, is unnecessarily corrupted in pursuit of notional imaginary quantity – Pure Machinery Friction.

Among a long period of correspondence with John, I recall the following. "**I make no apologies for treating the coupled wheels as part of vehicle resistance, it is after all a vehicle."**

The locomotive is an active traction unit, not a passive vehicle.

I'm reminded by this of the civil servant at the Ministry of Agriculture and Fisheries, who wanted the welfare conditions of captive live crayfish to be the same as for aquatic vertebrates on the grounds it was called a fish. Shakespeare's *Merchant of Venice* also comes to mind, when, paraphrasing a little, Portia says; *You can have your pound of flesh, but do spill one drop of blood.*

This concludes my comments on John Knowles' July 2017 letter at this point; more will follow in my final summary. I now turn to his letter 2nd April 2018:

"A defective approach in UK to UK Loco testing."

This is largely focussed on Report L116 and its implications regarding locomotive testing in the UK generally. While it broadly covers the scope and substance of the report,
there are one or two critical omissions that would undermine the arguments he develops. Before dealing with this however, I will first make a few general points of clarification regarding Report L116 and the related report L109.

Scientists

My mention of "scientists" was with the Amsler design and commissioning staff in mind, not the Rugby staff. As manufacturers of international renown in the field of scientific instruments, the Amsler team may have included one or two scientists; but perhaps they were all engineers. Any distinction between the two professions in the context to the tasks in hand will be of little significance. Engineers such as Dennis Carling and Jim Jarvis will have shared a common understanding in the fields of applied mechanics and mathematics.

Report L116

Report L116 was focussed on the road test results for 9F 92050 and Crosti 9F 92023. Both locomotives had been tested at Rugby prior to the road tests. These locomotive were only indicated on the test plant The anomalous road test results prompted a second series of tests at Rugby with 92050. This second series included comparative tests between the Rugby and Derby versions of the Fanboro' indicators. These tests proved satisfactory (page 25 above). The fundamental problem was that when the recorded road test DBHP data was subtracted from the Rugby IHP data at given speed speeds and steam rates, the locomotive resistance curve was the wrong shape. The resistance curves for the BR5 and BR7 as derived from the test bulletins were similarly anomalous, but the LR curves were of varying form. See Figures 25 & 26 on page 25, and figure 37 below.

Report L109 and the “Supplement to Report L109” concerned the road test anomalies with Duchess 46225. Report R13 essentially took the form of the BR Test Bulletins, and incorporated the corrections in report L109, namely corrected DBHP curves (Drawing DTG .976). Unlike the 9Fs, 46225 was indicated on both plant and road tests. These tests too were anomalous, only coincident with the road tests at 50 mph.. The 9F test bulletin as published retained the anomalous DBHP data. Some unresolved departmental politics were perhaps in play here. E S Cox was reluctant to accept that in practice, the *Controlled Road Test* procedure (constant steam rate), was flawed in principle; the theory of constant blast pipe pressure for a given steam rate independent of speed having proved not quite so straightforward as originally thought.

Below, Figure 40 illustrates the extent to which the locomotive resistance curve as initially derived from the road tests for 92050, was “the wrong shape”.

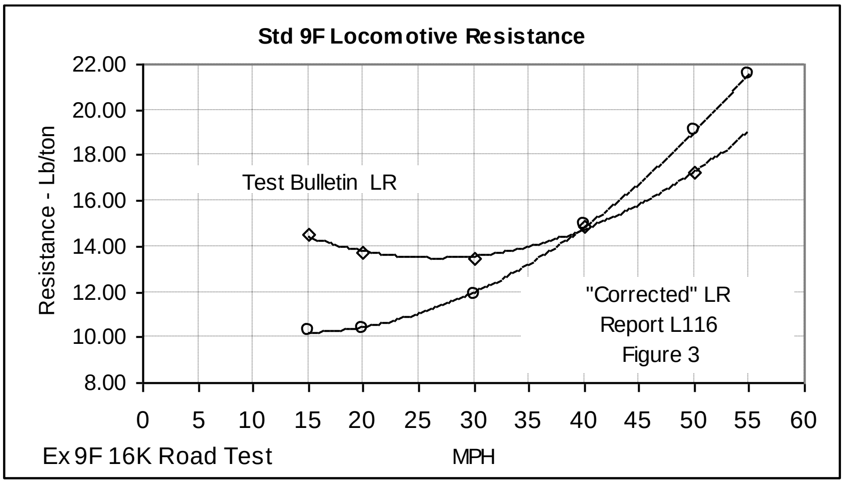


Figure 40 The bulletin curves takes on a slightly different form to Figure 26 page 26 above for Crosti 9F 92023. while having a similar crossover point. The bulletin locomotive resistance curve derives from Figure 11 - Figure 2 as for 16,000 lb/hr steam rate.

“They could not find any thermodynamic reason, which probably meant there was none, and picked, in speed effect, something which did not exist, as I show below. It is true that among the road test data, they had examples of tests where the result differed with the speed, eg by direction. These tests drop out as a basis because they were not comparable with the principle of the testing, constant Q, V and BPP. One wonders if such non constancy by direction in a test was not the reason for the error.” (my underlining)

This is with reference to road test anomalies involving steam rate variations under constant blast pipe pressure irrespective of speed. It is an inaccurate representation of what report L116 actually says. (The idea that direction may have changed the thermodynamics is most amusing.)

Report L116 Page 2 *“It is possible to correct the steam rate resulting from constant blast pipe pressure testing by two alterative methods, i.e.*

(a) Variable heat drop in exhaust steam according to temperature.
(b) Variable Density of Exhaust steam (Swindon Method).

Neither of these methods will entirely eliminate the discrepancy between Derby and Rugby.”

Note the word “entirely*”,* this is in deference to experimental error uncertainties (of which there are several mentions in the report), to which all aspects of measurement are subject. Note the reference to the *Swindon Method* regarding variable steam density. On L116 page 7 it states:

“It has been stated elsewhere that the steam rate variation which occurs when at fixed blast pipe pressure and variable speed is familiar at Swindon. A condition due to uncompensated change in density. Assuming that the flow rate is proportional to the product of the square roots of the differential pressure and density, it was suggested that the constancy of the steaming could be maintained by suitably varying the nominal blast pipe pressure to compensate for any observed change in density.”

This was considered impractical for variable speed road tests where speed was frequently changing, and the exhaust temperature responses lagged. It was seen however as a suitable basis to amend the test data.

“………**and picked, in speed effect, something which did not exist.”**
“Their analysis of the data was defective and biased the results of their thinking towards the idea that there was a speed effect.”

At no point does John Knowles mention report L116 Figs. 5 and 6 showing variation in steam temperatures with speed. He appears to be unfamiliar with Charles Law con-
erning the temperature/volume relationship of gases, or to have ever looked at a Molliere Diagram. He describes these variations as **“peculiar effects”**.

Unfortunately the blast pipe pressure data is missing from the 92050 Series 2 Rugby tests data base. It does however include exhaust steam temperatures against steam rate. When plotted as T against Q in speed sets, the temperature separation, and by implication density variation that emerges, is plain to see.

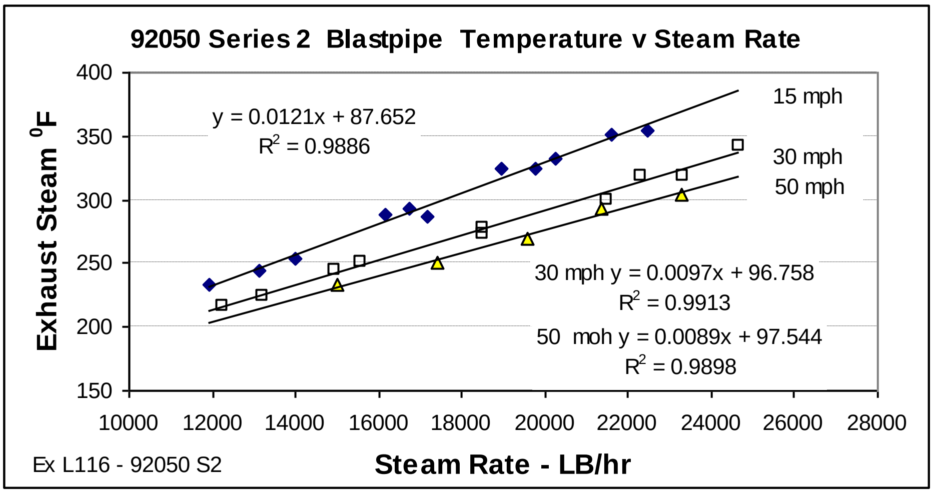


Figure 41 The reducing separation with speed accords with the trend indicated in L116 Figure 6 and is co-incident with the characteristic cylinder efficiency curve as a function of speed.

Examples of the temperature effect from test plant data are given in the internal "Comments on Test Report L116" document (Rugby June 1958) to which he has access.

Blast pipe pressure is a difficult measurement on account of the changing pressure during the exhaust cycle between exhaust release and compression. Some experiments comparing steady and pulsating gas flow through an orifice found that while the recorded manometer pressure in the pulsating situation was the mean of the maximum and minimum pressure per cycle, the quantity differed to that obtaining for steady flow at the same pressure. The effect varied with the frequency of pulsations, up to 200 per minute. The experiments were not entirely free of some uncertainty. *"…..the result did not indicate any improvement in the scatter of the final results, suggesting that the complexity of the problem is more fundamental than has been thought up to now."* *****
Also**,** close control of inlet steam temperatures was not possible.

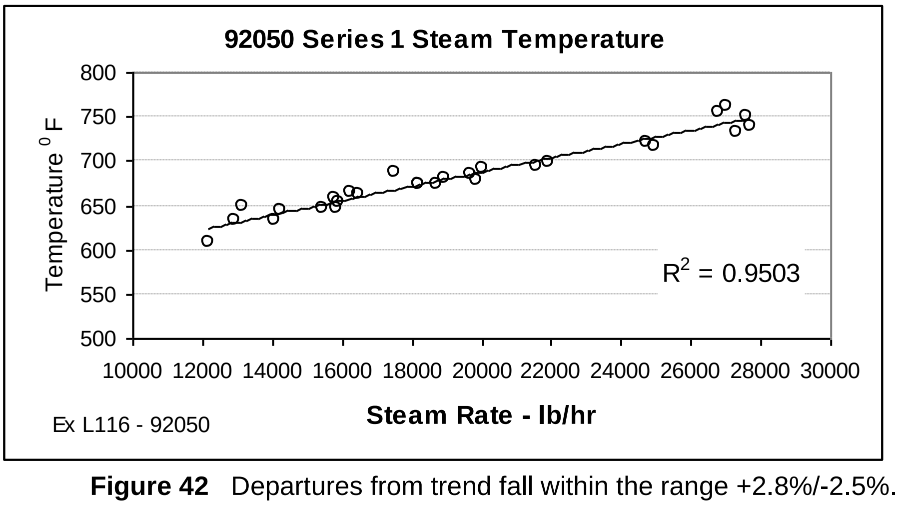


Figure 42 Departures from trend fall within the range +2.8%/-2.5%. Another potential source of scatter is variations in steam chest pressure. This ranged from 234 to 241 PSIG against the average value 239.1: +0.75/-2%.

- *Metering Pulsating Flow – Coefficients For Sharp-Edge Orifices;* J M Zarek, The Engineer, January 7 1955.

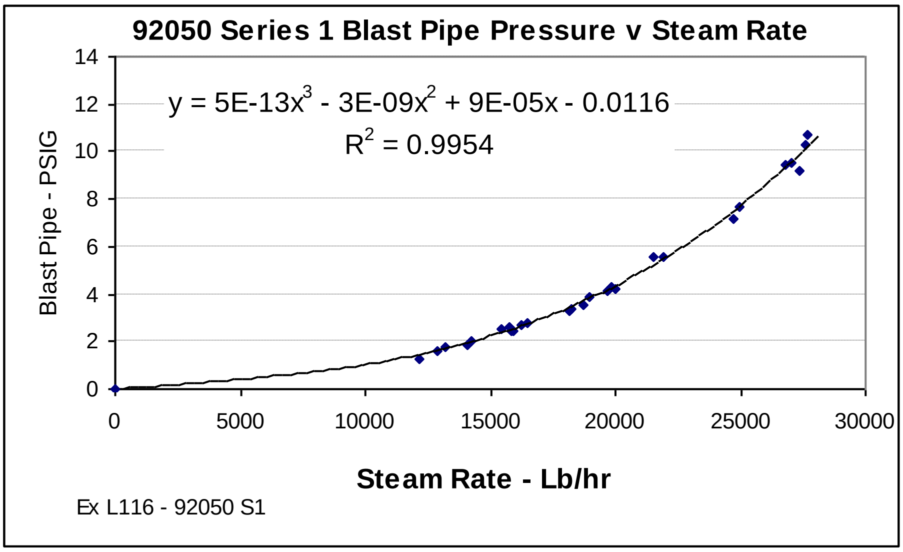


Figure 43 In the absence of any blast pipe pressure data for the Series 2 tests with 92050, the Series 1 tests must suffice. The low scatter here with only one or two visible strays from trend, and the high R^2 value, is typical of such data generally. The plots shown cover four speeds at 20, 30. 40 and 50 mph. An additional anchor point has been added to the plotted data, that being that when at rest, steam rate and blast pipe pressure will be zero: a simple matter. The constant shown should of course be zero, not -0.0116 lb.

At face value, Figure 43 supports the impression that blast pipe pressure is constant at any given steam rate independent of speed. Analysis of Q v BPP in separate speed sets reveals otherwise, as Figure 44 below.

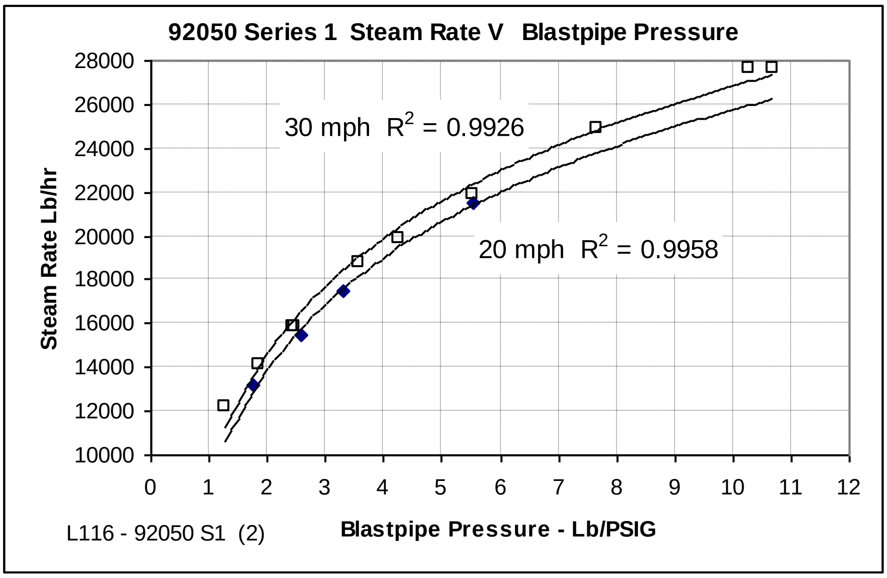


Figure 44, A clear increase in steam rate with speed is evident.

"They could not find any thermodynamic reason which probably meant there was none, and picked, in speed effect, something that did not exist,"

Really? The *Comments On Test Report L116* states: "*Variation in steam density is accepted by L116 as a condition which properly requires compensation."*

Page 8 of the "Comments" cites the conclusions of looking at other test series when they were carried out *"on the assumption the effect did not exist."* Some of the earlier test series were handicapped by the manometers then in use. Nevertheless some evidence was found for 45218, 44765, the BR7, 35022 and 46165.

The mean steam rate was 20,467 Lb/hr at a mean speed of 40.4 mph. On these figures the potential drift from the assumed Rugby steam rate on a road test at 20 mph would be about -700 lb/hr increasing the apparent LR based on the *supposed*

replication of the Rugby plant IHP data by about 42 HP, 790lb. It is apparent from the disparate locomotive resistances resulting from the Swindon controlled road tests that similar problems sometimes obtained. Hence the low speed LR anomaly identified for the two 4MT locomotives in Figure 32, page 33 above.

46165 Steam Rate Variation at 4Lb Blast Pipe Pressure.

MPH	20	35	50	65	80
Steam Lb/hr	19,772	20,325	20,588	20.936	21,142
% Mean	96.6%	99.3%	100.06%	102.2%	103,3%

The effects of changing temperature on steam density, and thereby the discharge rate through an orifice at a given pressure had been well understood long before the Rugby Test Plant was up und running. An undated booklet, probably dating from the 1930s, gives the following formula;*

Q = C Sq.rt P x W Lb/hr

Where C is the orifice constant as from tables, P is the pressure head across the orifice, and W is the steam density in Lb/cu.ft.

"There are three important defects in this work. First BPP is measured in atmospheric pressure or gauge pressure, whereas it should be in pressure absolute, as even an apprentice scientist knows."

This is incorrect. As Kent's formula shows, the discharge from an orifice is a function of the pressure head, steam density and the orifice discharge co-efficient. For "pressure head" read pressure differential, so if you adopt absolute pressure you have to set it against atmospheric pressure. So what differential do you end up with? Gauge Pressure!

A Swindon road test diagram with King 6013 in 1955 traces steaming rate as .a function of sq.root P, defined as "Orifice Differential Pressure in PSIG.

"Second, the three curves in Fig.11 from which Table 2 (JK's page 73) **was drawn above were fitted by free hand, with the initial pressure for each speed picked by eye."**

The curves appear in accord with the formulae derived from the Rugby test data plots.

"Thirdly, there are insufficient observations at each of 30 and 50 mph (ten each) to analyse the effects of those speeds properly."

This is unsubstantiated dogma.

In summary John Knowles assertion that there was "no thermodynamic reason to be found (in L116) why steam rate at a given blast pipe pressure varied with speed" is in defiance of the thermodynamic reality. Likewise his belief that for the purposes of analysis, blast pipe pressure should have been expressed as absolute pressure. It all amounts to another travesty of confused thought, and supposed science.

<u>A few more general points</u>.

"The higher (Crosti) **LR accords well with the back pressure, as shown by the Perform program. The frequently quoted idea that the resistance of the Crostis was high because they had weak frames is unsubstantiated; those quoting it as the reason for the high LR need to consider where the effects of the higher back pressure were felt,"**

………

***** *Flow Measurement Memoranda,* George Kent Ltd, undated. The firm later became Kent Instruments Ltd, and provided instrumentation for the test plant.

The point regarding frame flexure as unproven is fair enough, there was however a significant reduction in the inherent stiffness of the Crosti arrangements. Back pressure affects the mean effective pressure as determined by 'Perform' or an indicator diagram. and thus the Indicated Horsepower. There is no evidence of MF sensitivity to back pressure within the Rugby data. 9F 92250 returned the same mechanical efficiency for a given effort (ITE) in both guises; double chimney or Geisel ejector. The Giesel back pressure reduction was significant.

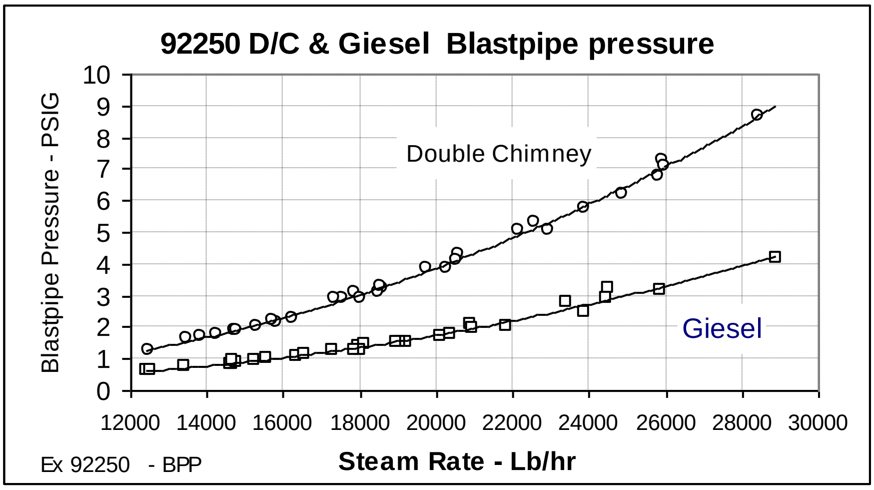


Figure 45 The significantly reduced back pressure with the Giesel ejector is evident.

The improved cylinder efficiency and reduced back pressure brought no measurable changes in mechanical efficiency, ref Figure 46. The back pressure reduction is implicit in the lower specific steam consumption and the increased blast pipe area with the Giesel injector fitted: the total nozzle area ratio was 302 sq.in. v 25.1 sq.in.

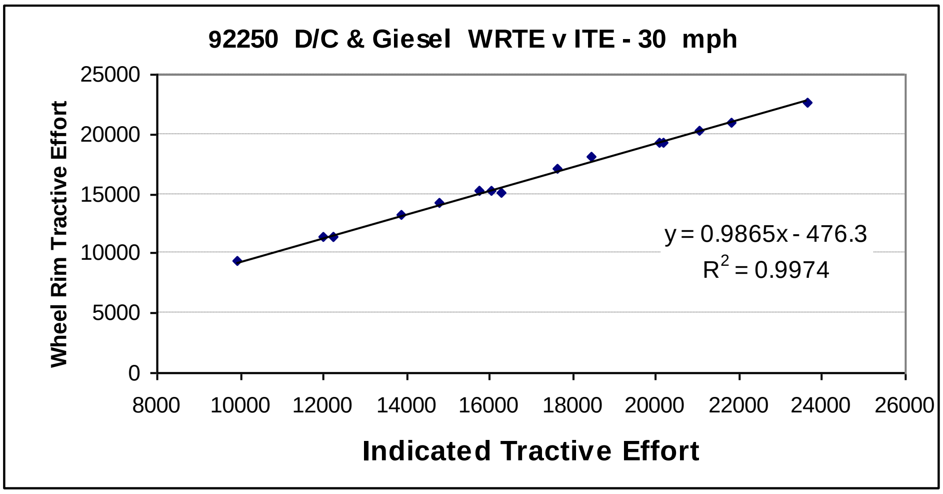


Figure 46 No discernable evidence here of a back pressure effect on mechanical6
efficiency. If such an effect exists, it must be very small.

“The conclusions of L116 should be forgotten, such as they are. That includes the supposed LR of a 9F.”

Report L116 may not have been without some questionable facets, but it’s general scientific thrust was sound, unlike John Knowles’ tendentious ideas as exposed above. No locomotive resistance curve can be declared as perfect simply because it is a variable: modestly with the level of effort, and potentially more significantly, according to environmental circumstance. The latter itself can only be roughly determined, and can vary from minute to minute. Beyond that, as amply evidenced by this long running debate, small remainders inevitably render such determinations at best approximate in outcome. Whether on test plant or road test, possible error bars of ten or horsepower seem realistic. The Crosti and standard 9F LR resistance formulae given in Report L116 closely reflect the differences in machinery friction established on the test plant

and manifest on road tests - Figure 47 It has been assumed the LR values are for a steam rate of 16,000 Lb/hr, as on the comparative road tests.

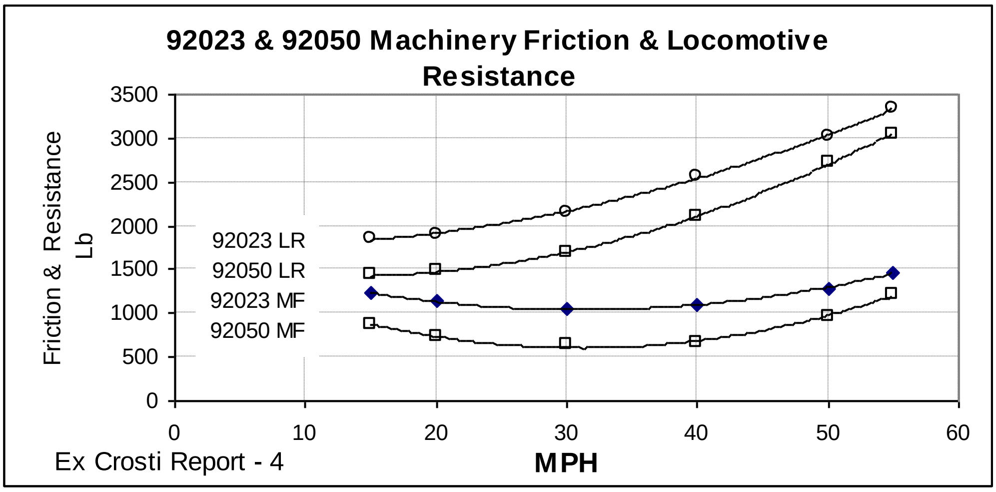


Figure 47. The plotted MF data at given speeds is as determined by Willans lines at a 16,000 Lb/hr steam rate. The Crosti 92023 v 9F 92050 differences in MF and LR are similar for both conditions in accord with the trends and magnitudes recorded on the test plant and the L116 LR formulae as Figures 2 & 3.

The 9F test bulletin includes a resistance curve for 16 ton mineral wagons as for a 7.5 mph 45 degree headwind. Presumably similar conditions apply to the L116 LR curves.

“It was Doug Landau who changed the subject to Steam Locomotive Resistance. Why did he do that? In my view he has not advanced the subject of steam locomotives one jot.”

I will simply reply by asking if he thinks that such poor work, untenable concepts, statistical misadventures and false attributions put into the public domain should be beyond challenge?. Well over 90% of what I have presented is simply setting out the empirical evidence as recorded at Rugby in various ways and the difficulties and uncertainties associated with it. It is ironic to be accused of “playing with the data” given his corruption of the recorded data in the futile pursuit of dissecting ESRMs (even smaller remainders!). Far from playing with the data, I have highlighted its limitations and uncertainties, how, even within the contractual measurement limits, exact fits falling neatly across the full data range remain elusive. Ultimately therefore, stitching test data and bulletins together was inevitably something of a black art.

Overall, the Rugby test data was far from perfect, but it was also by far the best and most informative locomotive test data to become available. The simple linear relationship between WRTE and ITE comes through loud and clear in principle, uncertainties as to exact magnitudes notwithstanding.

Intrinsically, road testing, away from the ‘steady state’ conditions of the test plant, proved to be a more difficult proposition. Anomalies in both the Derby and Swindon test data reflect this. Derby road tests in particular, were compromised by the assumption that the Rugby cylinder characteristics, in the absence of indication, would be safely replicated by the supposed control of steam rate alone.

Readers will have to make up their own minds. My own view is that aside from one or two statements of the obvious, John Knowles has been wasting everybody’s time, including his own. Likewise his website on L*ocomotive Resistance*: another charade of confused thought and superficial scrutiny. He needs to have a serious rethink.

Doug Landau

30 December 2019

Reply to Doug Landau of 30th December 2019 from John Knowles on Steam Locomotive Resistance

A few introductory remarks:

1 Starting the above submission to the RPS as he did shows what was on Doug Landau’s mind, ie replying to me, but also putting me in my place in his opinion, rather than putting forward the things he says he discovered at the NRM.

2 His submission of 26,000 words (some 84 A4 pages, a modest book) was far too long for a Word document. It should have been in several parts, and those parts put on the site as separate PDF documents. One of the parts could have been about me, if he cannot help himself, another on what was really new in what he saw at the NRM. The supposed Spreadsheet where the material is located in the RPS website needs tidying, so that each submission can be opened separately, preferably with each submission a separate PDF with date. Some of the diagrams in DHL’s December 2019 submission did not open in a download. I hope nothing longer than about 7500 words is ever accepted for this site again.

3 I have had the Rugby test data since 1988, and gave copies to DHL plus some information from files before 2000. I really wonder what is new.

4 Is the RPS Journal (Milepost) or website the right place for his voluminous ideas on Steam Locomotive Resistance? Members are interested in well founded resistance information for steam locomotives to evaluate exceptional runs by such locomotives, but I judge from

Milepost that such is a minority interest. They would normally look for such information in refereed technical journals. As steam is itself a minority subject in such journals now, I suggest that the Stephenson Locomotive Society Journal is the proper place for this subject, and then after the material has been distilled and much reduced in length – no journal will publish 26,000 words – and refereed. I wonder how many RPS members read the whole of DHL's submission? There was no need for such verbosity and multitude of graphs. Members who are interested but deterred by the length and debate might well say: let these people debate person to person; if they agree on something, then publish a summary in Milepost. Doug was a frequent contributor to the SLS Journal, but does not now contribute to it.

5 His submission (p 84) is preceded on the website by a piece of mine (p 64) on a defective approach adopted in BR steam locomotive testing. This appeared in the SLSJ for November/December 2019. I have no objection to that appearing on this site, although it would have been usual to ask my permission first. I certainly see no need for it to be published again almost immediately on the RPS website. As it happens, Doug Landau wrote a piece in reply to that, which was also published in the SLSJ in January/February 2020, in which he claimed that the method of Report L116 was scientifically sound. I disagree with that claim, but my reply awaits publication. I see that despite his having put forward his point of view, he wants to put it forward again on this website. Again that seems to be a vendetta against me.

6 I accept DHL's remarks on my use of the terms pressure absolute and pressure gauge.

Critical Speeds

When reading Doug Landau's preliminary remarks, especially were he says that the plant tests were preceded by calculations of the critical speeds for setting the Belleville washers, and presumably earlier the oil and air dashpot which were present in the connection to the dynamometer which gave the readings **o**f drawbar pull (DP), I looked for discussion on what they were intended to do, how that intention was fulfilled, what was the effect of the damping on pull, and how much there was of it. I was disappointed. The intention was to dampen the to and fro forces in the drawbar pull (item 1 below), but it was considered that damping falsified the readings, presumably of DP. Doug seems to think that it was the result of high traction forces, by which I think he means high piston thrusts at low speeds which caused the falsification. The lowest speed usually found in Rugby tests was 20 mph. The forces which are more likely to need damping because they can vary tremendously and have high values are the dynamic forces listed below, especially the first, because only the first passes through the drawgear.

Dynamic Forces in the Machinery Resistance

1 An alternating force along the locomotive and train through the drawbar, the result of unbalanced reciprocating masses in the mechanism. Work is necessary to create this force, which work is a source of MR. If any of that work is returned to the mechanism in moving the locomotive and train, that returned work would represent a reduction in the force created, and work performed. That would be possible only if the forward aspects of the alternating force consistently added to the tractive pull along the drawbar, and there was no other effect. The problem with that possibility is that the negative aspect of the alternating force reduces that pull by an equal negative amount. In addition, the forces in both directions, while of high maximum values at all but the lowest speeds, move only a tiny amount before reversing, rendering it impossible to obtain any sustained pull. Rather the alternating force is observed as a vibration. By observation, the alternating forces are gradually absorbed in the drawbar

springs and in the inertia mass (unwillingness to move or change movement or direction) of tender and train. If not absorbed along the train by or at any particular vehicle, this force is the to and fro motion or rake felt by the passengers in that vehicle, causing them to nod slightly in unison and to hear the drawbar springs moving in and out. Even with the absorption in the inertia mass, the work required to create the forces remains as MR.

2 An alternating couple about the crank axle, again needing work for its creation, causing the locomotive proper to sway from side to side, resisted by the inertia mass of the locomotive proper, and in the side operating springs (or other devices such as inclined slides) of the leading or trailing trucks or bogies, so sprung or inclined to control curving. This sway has to be resisted because excessive sway, which rises with speed squared, can be dangerous.

3 A couple at the origin of the drawgear in the locomotive frame on account of any vertical difference in the height above rail of the coupled wheel centres and the drawgear, really a leverage, usually small.

4 An oscillation in the vertical plane at the line of reciprocation leading to up and down motion on each side of the balancing masses in the coupled wheels on to the rails and bridges, the well known hammer blow, resulting from the balancing masses in the coupled wheels rising and falling as those wheels revolve. Hammer blow worries the bridge engineers, and leads to the percentage of the reciprocating masses balanced being less than 100, usually 40 to 70%. The hammer blow is itself not part of MR, but the work involved in making the oscillation is. (Revolving parts are usually completely balanced, but the balancing masses in the coupled wheels are not in the same plane as the parts concerned, leading to modest couples as well.) The 90 degree separation of the cylinders leads to uneven occurrence of the maximum hammer blows, requiring work to oscillate these uneven masses, and to operate a couple of the same kind as 2.

The operation of these four effects all require work, which is part of MR, and cannot be avoided with the engine moving, even when not under steam. I have noted the alternating force in trains hauled by many locomotives, from 15 inches gauge upwards, from four to ten coupled wheels, even with three cylinders.

The only way to eliminate the alternating forces is to balance the reciprocating masses completely or to oppose the forces directly. At Rugby, neither was done. Instead resistance was incorporated into the connection from the drawbar to the dynamometer to dampen (not eliminate) those forces, as Carling said. He did not explain how it was set up or operated, the extent of damping for any test, nor the resistance (DR). He considered that the testing apparatus needed protection from effects of resonance. Doug Landau has not found out or explained the matter either (as above). In any case, he does not give a file or library reference or title for these calculations, for those interested in discovering the meaning or value of the calculations. On a Testing Station, the distance between the coupled wheel centres and the dynamometer was not allowed to change, to avoid the coupled wheels moving back or forward and the locomotive to be subject to slight up or down gradients on the rollers, a source of defective readings compared with being on the level. The distance was controlled by a hydraulic mechanism, which recorded the effect on the work done, as Doug has previously explained.

The history of the damping was: having oil in a dashpot, air in the same dashpot, no damping, and Belleville washers. Jim Jarvis (a deceased engineer at Rugby) explained that there was a valve controlled by-pass to the dashpot, used to vary the extent of the damping, down to nil. I know of no record of the damping employed in particular tests. That means that the presence of DR affected the DP and consequently the value of (ITE – DP) (ie the TSAMR, or Testing Station Apparent MR) in any test to an unknown extent. Without knowledge of the DR employed in any test, the TSAMR cannot be interpreted properly (the constant of the CWBR, also present in the (ITE – DP) can be reasonably calculated).

The value of all the dynamic forces, the work involved in operating them, at every point in coupled wheel revolutions (or in the piston strokes, or at every crank angle) and the average for every stroke or revolution can be calculated from the masses and distances involved. They vary considerably during a stroke, and the maxima can be very high, even exceeding the tractive effort (here I quote from Professor Dalby in his book "Balancing Engines"). A method of exactly opposing these forces in the rod between the drawhook and the dynamometer at all speeds and DPs would require a much more complicated device, I suggest, than the dashpot or a stack of Belleville washers employed. Even if the damping could absorb say 75% and above of the alternating forces at a certain speed, the remainder are still considerable, and need to be known for interpretation, as do those absorbed. Any friction in the mechanism arising from these forces (say at the coupled wheel axleboxes) is also part of MR.

Readers are told that the calculations do not support my suppositions about how the damping functioned. So DHL interpreted them to that extent. If critical speeds are of interest, however, that can leave many other matters necessary to evaluating the Ruby DPs and their usefulness if any. If DHL was able to check and decry my so-called suppositions he must have been able to decipher more than he has declared about what was considered in setting up the Belleville washers and why, especially the resistance built into the Belleville washers for a particular test.

I have previously calculated MR including 1 to 4 for the LMS 5 and associated LR. These are much higher than anything DHL has ever deduced or calculated, and than the figures for the very similar BR 5 from Rugby LTS. That seems to have been the inspiration for all the thousands of words from him criticizing me ever since.

Analysis of Data

Rsqd has a special meaning when numbers are presented. It is called the coefficient of determination, and gives the extent to which the data concerned has been explained by some relationship, one inherent in the science of the subject, or a hypothesis being tested for such. Doug uses it for all data he presents in graphs. Because he is limited in his ability to form and test hypotheses and conduct multiple regressions, using two or more explanatory variables, indeed any regressions, the Rsqd he uses in many graphs is misleading or wrong, as in purely demonstrative graphs (figs 7, 8 and 9 are not causal), while WRHP against speed on a graph does not explain anything. It can of course exemplify, but then it is necessary to explain whether it comes from the pure science or some experiment; in the latter case, if a relationship has been fitted, giving Rsqd would be correct. I notice that Doug always presents the Willans lines as near straight lines, and says he doubts my claim that the relationship between ITE and Q and V together in Rugby ITE data is well founded. Try ln ITE = a ln Q/b lnV. Just to make sure I am understood, however, I am not claiming that regressions of experimental data make the data correct. Anything but. If ITE or DP is poorly measured (see below), nothing can rectify that. Indeed trial regressions are often the best way of showing the hypothesis to be completely wrong, or that the experiment to test it has been poorly done. It is odd that Doug does not know of multiple regression – it is the basis of the quadratic equation used to give the VR of railway vehicles, which he uses in the calculation of locomotive power.

What happened in May – June 1967? The LTS had long closed. Am I accused of 12% error in WRHP? Reference needed.

Figure 5 has MR falling from 15 mph to almost zero at 75. A little thought will show that this is rubbish. As speed rises, all the dynamic forces rise with speed squared, the frictions in the mechanism mostly rise with speed, while no sources of friction fall to zero as speed rises.

I know that the DP reading for any test was averaged and I have some ranges. Why were recorded DP readings for a test greater than any observed, whatever the date? Doug expresses satisfaction that observed MR (as ITE – DP) turns positive (with Fig 7). That MR includes CWBR and DR discussed above. When these are deducted, the TSAMR is not necessarily positive.

As usual, Doug uses his own version of how to analyse data and how to reach and present conclusions on the value of what he does. He has learnt nothing of statistics as a science (he seems to think it has to do with anything which can be shown on a graph, which limits him to a single explanatory variable, as in every relationship he puts forward). It is also why he puts forward relationships between ITE and DP to explain MR, and discusses Mechanical Efficiency in the context of MR (I realise Mechanical Efficiency on a LTS can be defined as DP/ITE), but the subject is MR, and the mechanical efficiency figures he gives are improbably high, reflecting the same old problem, the TSAMR figures are improbably low. And he regards linearity as of great value, but then notes that the sign on the constant implies impossibility with the regulator shut. That is correct as a point, but the lowest x variable he shows is 14,000 lbs ITE. The linear form does not allow for the behaviour to change at lower ITE values; anything can happen at lower ITEs; the form of the relationship he chose does not allow whatever that might be to emerge.

He makes a big thing out of eliminating observations to improve Rsqd, already 0.99, although no other statistical tests are given, Rsqd being the only test he uses, which he says runs counter to an edict of mine that the more observations the better. That interpretation is part of the indulgent approach he adopts to science and statistics, and his search for negative things to say about me. It all depends on what observations might be removed and the criteria for their removal, as a good book on statistics will tell him. See also the end of this note, where I remove some observations, and reduce the value of all remaining observations by the same amount in an attempt to extract some slightly better value from Rugby TSAMR data than is available from a "first pass" regression. The point is that all the data from an experiment which might influence the explanation of the subject under investigation should be used (experiments which turn out badly should not have been recorded), all speeds together, because speed can be expected to be a major determinant of MR, and should therefore be an explanatory variable among others. Apparent outliers should not be excluded in a first pass. Apparently imperfect data should be tested to see if some influence not initially considered by the analyst should be tried, and tested for whether the outliers show any influence on the subject, especially if only single explanatory variables are being used, as in the Landau method.

At the end of his enormous document, DHL claims that one of the reasons he changed the subject to steam locomotive resistance was my poor dealing in statistics. What a hide! What an accusation from he who knows so little about the subject! I have some education, and considerable experience in the subject and can say with confidence that he knows so little about the subject that his declarations are waste of space. The things to use to try to explain TSAMR are at least and at first, piston thrusts and speed squared, because from first principles they are the major variable influences on it. A constant should also be allowed to emerge. CWBR (calculable from bearing dimensions and masses borne by them with high accuracy) should be deducted, as should DR. Then all test statistics should be examined. Despite my attempt to show DHL in a previous post what these are and what they show, he shows no evidence of knowing about them, as he should before he tells the world about my qualifications and abilities in statistics. If the fitted equation proves little, I would change its form from added terms to multiplied, or to have an index on speed to be given by the regression. See at the end my reasons for finding that the Rugby data does not explain MR.

I suggest DHL forget his ideas on explaining the data, or finding "proof", and obtain a good book on statistics and study it.

What is a differential pressure element in a hydraulic dynamometer? What is the Mi-o index, what does it show and how is it obtained?

Fig 14a – why is data from 92166 included with that from another locomotive? Are the delta mechanical efficiency figures statistically different from zero? In the context of MR, why such attention to mechanical efficiency? (I know the connection between them, and would say that such very high ME is an indicator of inadequate registering of MR).

With Fig 17, the reader is informed that WRTE is linear, with ITE, but that is not said in the headline, and is not shown to be universally so, being exemplified only for 30 mph. As ITE at a given steam rate is not linear with speed, and as MR is composed of elements which are not linear with speed, such linearity would be surprising, indeed wrong. I conclude that it is the result of a single variable equation as DHL uses with an Excel line to a set of data. See above for what he should have done. He should also calculate MR from first principles based on its causes or sources. I have done that. I find it much higher than anything derivable from the Rugby data. It is all very well to claim that people at Rugby were highly qualified, the devices the best available, and properly checked, but that does not give the results special credence. Those results must satisfy explanatory equations which accord with the physics.

I consider (based on the evidence of all he says and does in these posts, especially the claims he makes) that Doug Landau is not competent in technical and statistical terms to derive MR from the Rugby data, or to express any judgement on the value of that data for such a purpose. He seems to believe that MR is a function of ITE. That is simplistic and wrong, but his view has not been refereed and will mislead others (that is not to say, that MR might be expressed in certain ranges by lbs mep as a shortcut, as I have done, but that is a stage down the line), hence the following. The MR of a two cylinder steam locomotive with the usual arrangement of the pistons on the two sides separated by 90 degrees and results from items which reciprocate (move to and fro) and which revolve. The parts which reciprocate are the pistons with their rings and rods, the crosshead and its slipper, the leading end of the connecting rod, and similar parts of the valve gear. Those which revolve are the remaining mass of the connecting rod, the big ends, the coupling rods, the bushes of these rods on their pins, the journals of the axles in the CW axle boxes, the balancing masses in the wheels, while the links in the valve gear oscillate. Friction occurs on the cylinder walls, the glands, the crosshead pin, the crosshead slipper on its guides, the big end, the coupling rod pins, and at various points in the valve gear. MR also arises from the work done in the operation of the dynamic forces, and in the oscillation from lack of balance of various masses, including the reciprocating parts and the balancing masses themselves, plus any friction from the action of those forces.

On a Testing Station like Rugby, there are problems in measuring the WRTE, which is in principle DP deducted from the ITE to give the MR. At Rugby, this was measured not at the CW rims, but by a dynamometer at the end of the drawbar, on the assumption that the pull there (DP) and the WRTE are equal. This assumption was obviously not true. The CWBR, usually considered part of the Vehicle Resistance, occurs between the ITE and the DP, and its constant value can be calculated, from the bearing dimensions, the mass borne by them, the resulting pressure, the appropriate coefficient of friction (Cf) for that pressure, and the ratio of the bearing diameter to the CW diameter, and should be deducted. So long as that is understood, it matters little whether the CWBR constant is deducted from the raw data or from the regression result. In addition, at Rugby, an unknown DR was incorporated between ITE and DP. And of course no tender was attached; on the road, it was a source of VR before the drawbar. It had no MR.

The Rugby TS also had water dynamometers, through which the load was put on the locomotive at the CWR by churning water and raising its temperature, one per CW. Each gave a reading of WRTE directly, but Rugby preferred to obtain what it considered to be the WRTE for the whole locomotive at the drawbar hydraulic dynamometer, because it was considered that was easier. In his piece, Doug Landau thought the change in water temperature not very accurate; he is poorly informed. Many powers and efforts are accurately obtained through change in water temperature, of motor vehicles especially. Furthermore, it is amazing that no check was ever made at Rugby using the two sources of information on effort at the CWRs, especially as the braking dynamometers were used in monitoring the braking, and the braking had to be accurate, and the inability to find sense in the readings at the drawbar.

No tests were conducted at Rugby with all the reciprocating masses balanced solely for the tests, and the DR eliminated. That would have revealed proper WRTE as the DP of the dynamometer which was used, subject of course to the ITE being correctly measured, and the values of the WRTE as a thick line of small variations being capable of interpretation.

I consider after all the above, that the Rugby plant does not reveal steam locomotive MR for four major reasons:

1 The TSAMR is so low in itself, and does not accord in any respect with the factors which should, from first principles, explain it. For the latter point, I have taken the LTS data for tests on 9F 92250, the last steam engine tested on the plant, from April to September 1959, by when the testing procedures should have been as perfected as they were to be. There were 62 observations. (ITE – DP) (TSAMR) is low by observation, from -37.5 to 1472, with an average of 719.6. Four are less than the expected constant value of the CWBR of 228 lbs, eight less than 400 lbs and 22 less than 600 lbs. The PTWR can be expected to be a little less in tests where the engine had a double chimney or a Giesl Oblong Ejector, but the effect of those fitments on MR should be comparatively modest in relation to all the sources of MR together. I have done some new regressions, five with one explanatory variable, of the kind Doug Landau favours, and three with two. In the following, Q is the steam rate in lbs/hr, PTWR the piston thrusts at the CW rims, Rsqd the coefficient of determination, SEE the Standard Error of the Estimate, t a test statistic, with 1 for the constant, 2 for the first or only variable, and 3 for the second variable. All the data from the Rugby LTS records at the

NRM. The low value of most coefficients is the result of the high values of the variables, Q, ITE and PTWR occur in tens of thousands, and Vsqd at and above 40 mph in thousands.

Equations for TSAMR in lbs

(1) -73.3 + .0417 Q; Rsqd .285, SEE 279, t1 -0.43, t2 4.81
(2) 498 + .0145 ITE; Rsqd .07, SEE 318, t1 4.33, t2 2.06
(3) 433 + .015 PTWR; R sqd .04, SEE 324, t1 2.22, t2 1.51
(4) 629 + 2.59 V; Rsqd .008, SEE 328, t1 4.39, t2 0.69
(5) 1775 - .05Vsqd; Rsqd .02, SEE 2867, t1 1.47, t2 -1.1
(6) 666 + .046 PTWR - .0339 Vsqd; R sqd .02, SEE 2888, t1 0.23, t2 .04, t3 -0.57
(7) 1695 +.024 PTWR -33.2 V; Rsqd .027, SEE 2883, t1 .041, t2 .024, t3 -0.73
(8) 16.9 +.047 Q -.01 PTWR; Rsqd .30, SEE 278.7, t1 1.09, t2 4.59, t3 -1.02

These equations do not reveal any close dependence of TSAMR on the variables which, from first principles, were chosen to explain it. I shall leave it to Doug to find out the usefulness of t and the SEE, both of which have to do with the range in which the results overall and for particular coefficients can be accepted with a certain level of confidence. In this case, however, the data are such that it would not be expected that they would display anything much of value on MR. It is however possible for the analyst to use experience and observation to perform some slightly different regressions. One is to exclude observations of TSAR which are clearly not sensible. As the calculated constant of the CWBR of a 9F is 228 lbs, and the DR might be considered as 172 lbs (my estimate after considering the maximum values of the alternating force at the higher speeds) total 400 lbs, any values of TSAR of 400 lbs or less are clearly wrong, and have been excluded for the next two regressions. That reduces the observations to 52. A regression of Vsqd and PTWR on the remainder yields TSAMR of

(9) 514 + .04Vsqd + .0115 PTWR; Rsqd .02, t1 1.85, t2 0.7, t3 1.05, SEE 269

Apart from giving a reasonable blessing to the constant, this is of greater help. If 400 is deducted from each value, and the remainder regressed on V sqd and PTWR with no constant, and the 400 added back to the result, a reasonable equation is obtained

(10) (400) + .605 Vsqd + .0106 PTWR; (numbers apply only to the second and third terms) R sqd .70, SEE 269, t2 1.62, t3 5.4.

This shows that the estimation of (9) is taken over by estimation of the constant, leaving little data for the other terms to explain the rest of TSAMR. In (10) with the 400 being added back, it is not possible to say how good it is, the t2 for Vsqd has a wide range of uncertainty, but the PTWR term is well established. The procedure is repeated for the sum of the CWBR constant and the DR to be 500. The number of observations is reduced to 46. The results are:

(11) 36 + .08Vsqd + .01PTWR; Rsqd 0.04, SEE 253, t1 .135, t2 1.32, t3 0.96
(12) (500) + .09Vsqd + .011 PTWR; (second and third terms) Rsqd 0.66, t2 2.23, t3 3.95

In (12), if only the reliability of the added back constant in this case were available (and obviously it should be high), this would be regarded as a useful equation. Hence, if the true constant of TSAR is 500 lbs or a little below, this is a reasonable equation for TSAMR from the data for 92250. It is important, however, to say that the cofficients on Vsqd and PTWR are decidedly low compared with values calculated from first principles, the result of the generally low values of the data.

3 Values of MR calculated from first principles are much higher than observed at Rugby, even more so if CWBR and plausible values of DR are deducted from TSAMR as above.

4 D R Carling, Superintending Engineer of the Rugby plant, considered that what the plant registered was DBTE and not ITE. He considered that the ITE recorded there was suspect. That does not cause me to withdraw my views on the problems with the DP and associated MR recorded there – they are too well founded, but if ITE was suspect, it can be the source of low MR, and aspects of the apparent behaviour of MR; indeed, tho' Carling does not say that a consequence of suspect ITE is suspect MR, it can explain the difficulties with interpreting the low values of MR. It is worth saying that a low ITE leads to low values of PTWR.

The Perform Programme

I explained that I used this programme to obtain indicator diagrams, from which to derive piston thrusts, leading to another series of negative remarks from Doug Landau. He criticises the programme, derived by the late Prof W B Hall. He has not so far as I know made these comments public previously, nor has anyone else, in the 21 years since the programme was released in the SLS Journal for May/June and July/August 1999. Prof Hall was very much aware of literature going back a very long way, especially efforts of many people to explain the "missing quantity" in steam consumption. The coefficients he uses and suggests are used in science for steam expanding under load, and uses a system of continuous differentiation and integration to explain the behaviour of valves, and pressure difference to explain steam flow and action. The programme has been used to explain as never before many aspects of the operation of steam locomotives, as he should be aware. I have used it myself to explain how and why certain locomotives I have known well have behaved as they did, and how improved versions of certain locomotives improved efficiency compared with the unimproved. How else would Doug Landau obtain the pressure along a stroke to calculate piston thrusts, positive and negative? It must be better than any mechanical indicator. Given what is known of BR testing, it is perhaps unfortunate that Prof Hall used some BR test results to exemplify the programme. How does Doug Landau, stout defender of everything done at LTS Rugby, criticise BR testing as implied by his remarks?

Until be died, Prof Hall made alterations to the programme which satisfied various criticisms, leading to Perform 2, which I use. David Pawson used that version to test a large number of the Rugby ITE results, and found, with slight tweakings of certain items that Perform reproduced the BR results very well (all published in the SLSJ in the early years of the century). Doug Landau's concerns are therefore not valid, nor Mr Carling's concerns about the ITE readings there (4 immediately above).

The Crosti 9F

It is almost standard for those writing about these engines to say that they had higher TSAR than the standard 9Fs, on account of reduced depth of frame stretchers, to fit in the preheater drum below the boiler barrel. That is not why I think they had higher TSAR. It is because they had very restricted blast nozzles in the chimneys, which created very high back pressure in the exhaust strokes of the cylinders, as can be seen using Perform, and calculating the piston thrusts. They were designed with numerous U shaped frame stays under the preheater drum, of some strength in the U itself, intended to reproduce the frame stiffness of the standard 9Fs, and were of course well stayed at the cylinders, firebox and dragbox. It is not clear that shallow frame stays would in any case have led to high friction in the axleboxes; if the frame was less stiff, the friction might well have been reduced.

Conclusion

I shall return to comments on the rest of DHL's massive piece of December 2019, although I expect to do that in less detail. I doubt that the LTS as set up and operated can reveal MR. I am perfectly willing for the content of this piece to be refereed. Indeed, I would like it to be.

6th November 2020.

Reply to John Knowles of 6th November 2020 from Doug Landau on Steam Locomotive Resistance

As previously, emboldened words in inverted commas are John Knowles' own. The many points raised may not follow the same chronological order as his letter. Some of his themes keep reappearing. In particular a seeming obsession with unbalanced masses, which in reality were only of minor significance.

Critical Speeds

"Doug seems to think that it was the result of high traction forces, by which I think he means high piston thrusts at low speeds which caused falsification." "…it was considered that damping falsified the readings"

Not so. The falsification was entirely down to malfunction of the dashpots, not damping per se. A problem not confined to particular speeds. The drawbar pull was observed to increase with the engine working in steady state conditions. At one time the idea was mooted by the test staff that the problem arose because the drawbar pull did not follow the expected sine wave. Numerous dashpot modifications and experiments reduced the falsification, but failed to eliminate the problem and the dashpots were decommissioned by the end of 1950. Experience had proved the dashpots were an unnecessary over precaution. A dysfunctional belt to the braces provided by the Belleville Washers.

The 'Critical speed' problem was attributed to high and asymmetric traction forces at low speed when the frequency coincided with the natural frequency of the test plant. The very high unbalanced reciprocating mass forces at speed were regarded of trivial consequence, notwithstanding these forces could substantially exceed tractive effort, as enlarged upon below.

On Test Run 126, 7th November 1949, WD 2-10-0 73788 (no reciprocating balance), was run up to 45mph with the dashpot drained of oil. Slipping curtailed tests below 30 mph. The peak transient accelerating force vector on the locomotive mass per cylinder was 0.2G. No problems encountered.

Dynamic Forces in the Machinery Resistance

"An alternating force along the locomotive and train through the drawbar, the result of unbalanced reciprocating masses in the mechanism. Work is necessary to create this force."

The first point is that on the test plant, the train and tender are absent and the locomotive is stationary. He has previously suggested the vibrations generated would disturb the accuracy of, and possibly damage the Amsler dynamometer. No such damage is on record. The frictional losses resulting from the centrifugal force loadings of the reciprocating masses on the various motion crankpins, are the same whether balanced or unbalanced. Obviously, these frictional losses would diminish the available drawbar pull accordingly, as do any other frictional losses in the power transmission machinery along the way. The accelerating forces applied to these masses cancel out, both hindering and helping in the course of each revolution.

An alternating shuffling force however, will obtain as suggested, and this can indeed substantially exceed the tractive effort, but the maximum horizontal reversing force vectors arising from imbalance only occur fleetingly in the course of a revolution. High though these momentary forces are, the available acceleration G forces are low relative to the locomotive mass. The imbalance shuffling effect was of little practical concern under all normal operating conditions. In the unlikely event that these reversing forces did not exactly cancel out, the slight net effect on tractive effort, positive or negative, would be real, not some kind of falsification as appears to be suggested.

Unbalanced Reciprocating Mass Force per Cylinder									
Class	Coupled Wheels "	Stroke " d	Unbalanced Mass Lb	Percent Balance	MPH	RPM	Force lb	Loco Mass lb	Loco Mass/ Force Ratio
9F	60	28	519	40%	15 30 60	84 168 336	1,448 5,791 23,166	198,240	136.9 34.2 8.6
WD 2-10-0	56.5	28	1298	Nil	15 30 45	89 178 268	4,089 16,358 37,081	175,392	42.9 10.7 4.7
Black 5	72	28	467	50%	30 60 104	15 29 52	3,641 14,564 44,968	161,504	44.4 11.1 3.6

$F = (W\varpi^2 r/g)\cos\theta$ ϖ = Angular velocity radians per second. r = Crank radius in feet.

The maximum transient resultant force vector for the 2 cylinders is □F x 0.8325

In 1939 the LMSR conducted some stationary slipping tests involving three Black Fives with 66.6, 50 (as per standard) and 30 percent reciprocating balance. At 104 mph the oscillations (shuffling) of the standard engine were described as ‘Moderate’, ‘Excessive’ for the 30% engine at 99 mph, and ‘Nothing abnormal’; for 66.6% at 103 mph. Vertically the situation was more serious, with the coupled wheels of the 66.6% and 50% engines momentarily leaving the track once per revolution, and actually justifying for once that horribly over-egged term “hammer blow”. In the USA it was more sensibly described as “axle-load augment”, a smoothly approaching maximum value in all normal circumstances. No sudden shocks or sledge hammers involved.

All this is covered in a long paper to the by E S Cox to the Institution of Civil Engineers in 1941: *Balancing of Locomotive Reciprocating Parts.*

Among the general conclusions was “… the additional variation in drawbar pull (traction) is very small and can be ignored at anything but very slow speeds.”

Unbalanced reciprocating mass dynamic forces were transient and adequately suppressed by the locomotive mass under all normal operating conditions. At the highest speeds the accelerating G forces remain fractional. Reciprocating mass as a function of locomotive mass tended to reduce as the latter increased.

Length was also a factor. The Stanier 2-6-4 tanks, weighed in at 87.85 tons, with 5’9” coupled wheels and 66.6% reciprocating balance as standard. The wheelbase at 37’ 1”, was one inch longer than a Duchess. One example had been running for eight years without any reciprocating balance; “No adverse report has ever been made on its riding”.

Throughout this long correspondence John Knowles has been much exercised by his ideas on unbalanced reciprocating masses, with a hand wringing innuendo that only he had properly considered these matters, and the Rugby test plant was by implication a design and operational fiasco. The reality was otherwise; his understanding is based on little more than flawed opinion rather than thorough research. His long explanations of imbalance affects, give insufficient significance to their transient reversing nature and the inertial mass of the locomotive. In summary, his concerns are yet another red herring.

Why he continues to dwell on the performance of the dashpots is a puzzle. They proved wholly unfit for the job, notwithstanding considerable attempts to rectify, and in the end proved unnecessary. They played no part in the later test results that have been subject to challenge.

At the end of this letter is an appendix showing that the design and operating circumstances of the Rugby Test Plant in regard to traction and reciprocating forces were thoroughly understood and accommodated with due diligence.

“Even if the damping devices could absorb 75% and above at a certain speed, the remainder are still considerable.”

Not so, for reason of the inertial reluctance as explained in detail above. Ironically, he does recognise this effect: with his own words; **“resisted by the inertia mass of the locomotive”,** but brushes its significance aside with all the authority of ill informed opinion. The dynamic G force vectors are both fractional and transient!

“I know of no record of the damping employed in particular tests”

Not much information is available to me, more probably exists at NRM. It is however apparent that the number of Belleville washers deployed varied from test to test depending on the locomotive type involved and the planned speed and work rate. Test run 156, 19th January 1950.with Black 5 45218 for example, involved 3 pairs of Belleville washers, dashpot drained of oil.

The Belleville washer hysteresis was very low. The theoretical critical speeds plotted for the BR7 in my last letter (Figure 2) was based on tabulated results. The formula given as: Critical Speed (mph) = V = 4.26 Sq.Rt K. For K, “$ee report 1949/276.” The natural frequency of the plant was low. The disturbing forces of unbalanced masses were insufficient at low speeds given the locomotive’s inertial mass, to set up any resonance, at higher speeds the frequencies were out of step. The reversing acceleration forces of the unbalanced reciprocating masses cancel out to zero.

Any parasitic motion of the locomotive resulting from imbalance effects cannot falsify the *actual drawbar pull*. At any given moment, that pull is the reality, net of all the manifold frictional and dynamic work done losses of the power transmission system, by whatever cause. The force encountered by the dynamometer will be equal and simultaneous.

Water Brake Dynamometers

“The Rugby TS also had water dynamometers, …… but Rugby preferred to obtain what it considered to be the WRTE for the whole locomotive at the drawbar hydraulic

dynamometer, because it was easier….. Many powers and efforts are accurately obtained through change in water temperature.”

The prime function of the Heenan and Froude water brakes was to provide a load. Each roller was metered so they could be adjusted to provide a similar load for each coupled wheel. Exact similarity of load was not considered critical. The method of measurement and control was by metering the torque (measured between the water brake stator and rotor paddles), and adjusting a water control valve. The accuracy was nominally +/- 5% minus an unknown missing quantity. As configured the frictional losses of the $10^1/_4$” roller bearing journals of the rollers were by-passed and not picked up by the dynamometers. This being so the water brakes were recording losses relative to a quasi-twelve-coupled engine when a six coupled was on the rollers and no less than a quasi-twenty-coupled when it was a ten coupled. The 4’ 9” roller mechanical advantage (MA) to the bearing rollers pitch circle was low at 4.23. This compares to the MA of 8.7 for a BR5 and 8.1 for a Duchess (plain bearings). That is not the end of it, the extension shaft to the dynamometer has two further bearings with losses not picked up by the dynamometer. They were however under lower vertical loads and of lighter construction.

Water brake dynamometers can indeed function by the determination of water flow and temperature rise as suggested, but for the Rugby Test Plant set-up this would have in involved lagging the dynamometers. Not a very practicable proposition. They would however be second hand measurements based on the conservation of energy principle. The losses of the roller journals, as described above, would still have escaped measurement. It would have been quite an involved experiment to set up a rig in the manufacturer’s works to determine these losses under simulated working conditions. Such an operation apparently was not thought worth the trouble, since the function of WRTE determination was entrusted to the Amsler dynamometer. The likelihood that the Froude system could be more accurate would remain very low. The Amsler, along with its servo mediating mechanism, remained essential to hold the coupled wheels over top dead centre. So, all in all; not a sensible proposition.

The application of water brake dynamometers to IC engine testing is considerably more straightforward than a loco test plant, with a simple direct coupling of the prime mover and dynamometer drive shafts. Froude water brake dynamometers are still going strong, and still function by torque measurement and not water temperature rise.

“…..the Rsqd he uses in many graphs is misleading or wrong, as in purely demonstrative (Figs. 7, 8, and 9 are not causal), “

The determination of R^2 relationships is the statistical “least squares” mathematical outcome of the data provided, so how is that ‘wrong’ in the pure mathematical sense? High values are indicative of low scatter, a desirable characteristic of test data more likely to reveal relationships than chaotic scatter such as routinely occur in small remainder outcomes. Obviously high R^2 values are not proof of accuracy, a poorly calibrated instrument may perform with respectable consistency. Another potential polynomial curve fitting hazard is using insufficient decimal places as previously exampled in my December 2019 letter; significant mathematical errors may occur.

To say the relationships displayed in the graphs cited is not ‘causal’ seems something of a moot point. It is apparent that ‘No steam’ = ‘No horsepower’, but since the relationship is clearly not linear, some underlying factors beyond Q must also obtain. It is a 2nd hand relationship. The shape of the Willans lines (Fig.7), indicate a trend of initially rising then falling efficiency over the working range.

“**WRHP against speed on a graph does not explain anything.”**

Really? Figure 11 for example, shows a very tidy set a WRHP Willans Lines at various speeds fully reflecting the expected pattern given the inherent characteristics of IHP v speed efficiency trends and Willans Lines. Such order was something the related data for the IHP Willans Lines seems to have been unable to replicate at the period of the test plant history.

Some General Points

The prolific use of 9F machinery friction data at 30 mph in my last letter was simply because most data was available at that speed, and involved 3 individual 9Fs and 5 test series. Close agreement of the WRTE v ITE outcomes was evident, indicating degrees of uncertainty falling within the Amsler contractual guarantee. In addition, the 26 test runs for Crosti 9F 92023 at 30 mph were also predominant. The irrational 9F WRTE positive constant only occurred in the instance of 92166, hence the weeding of some data as explained above. Data below 14,000 lb/hr was not available for 92166.

The linear relationship of WRTE v ITE unequivocally obtained for all other 9Fs at other speeds examined along with negative constants, as it did other locomotive types tested at Rugby. A similar relationship is evident from the limited test data available from the Vitry test plant for EST 241-004 at 37, 50 and 75 mph, which returned 3 parallel lines.

As demonstrated in my last letter, the MF outcomes for 3 9Fs and 5 test series were within 1% in the middle power range, notwithstanding variations in the formulae coefficients obtaining (Figure 30). The latter are highly sensitive to the random variations in the scatter of the individual data sets.

The variable coefficient representing effort sensitivity generally falls within the 2 to 3% range. When the variable is lower, a higher negative constant follows, as the number crunching juggles with the scatter, and vice versa. The sensitivity of the linear slope and therefore the variable to the plots at the extremities of the plotted range is high. In his *Tribology and Lubrication,* L D Porta put the frictional losses sensitivity in the 2 to 3% range.

It is notable that such a mechanically determined iteration of the complex shifting force vectors, cross couples, dynamic effects, windage and the shifting resultant frictional loadings in the course of each revolution should resolve into such a simple relationship as a function of ITE.

$$WRTE = ITEx - C$$

Where x is analogous to the overall frictional sensitivity as a function of effort, and C is a variable constant as a function of speed or RPM and among other things. The R^2 values for these plots uniformly approach unity. By comparison the small remainder derived machinery friction outcomes are a randomised confusion of reality about as useful as bingo results. To flatter them with statistical analyses as sources of scientific revalation is ridiculous.

“He makes a big thing out of elimination observations to improve the Rsqd. already 0.99, although no other statistical tests are given. Rsqd. being the only test he uses”

No so, the improvement in R^2 values was purely coincidental, not the aim. The statistical test was the elimination of a positive constant, which was a technical impossibility, thus changing it to a negative sign. It was a process of irrefutable logic unhampered by mindless dogma.

The rest of the paragraph that precedes and follows the quotation above is a misrepresentation of what I have said and shown. The positive constant was a solitary aberration attributable to the hazards of random scatter among numerous satisfactory date sets. Linear trend line outcomes with a single variable are not within my control, it's simply the way the plots resolve. The various Willans Lines plotted are quadratics.

"Figure 5 has MR falling from 15 mph to almost zero at 75. A little thought will show that this is rubbish."

The explanatory caption below Figure 5 notes: "The overall trend, clearly and illogically, is saying that MF is an inverse function of speed." It was presented as an example of the troublesome experimental outcomes that were evident at that stage of the test plant history. JK's comment adds nothing.

This, as are the bulk of my presentations, was a demonstration of what the empirical data was showing, warts and all.

"The things to use to try to explain TSAMR are at least and at first, piston thrusts and speed squared, because from first principles they are the major influences on it."

The iteration and resolution of forces: static, traction, dynamic, inertial, frictional and various vector shifts of the locomotive power transmission are far more complex than John Knowles assumes above, and cannot be reliably dissected. In the absence of the complex force diagrams and numerous complex iterations involved, he falls a long way short of the standards of presentation, diagrams and analysis that would pass muster in any design office. To work from first principles, it is first advisable to understand them, and most important, be aware of what is unknown or uncertain. The net significance of piston thrust is far less than he seems to suggest.

The abstract below is a remarkable example yet another fallacious concept, presumably derived from JK's supposed "First Principles."

"On a Testing Station like Rugby, there are problems in measuring the WRTE, which is in principle, DP deducted from the ITE to give MR. At Rugby, this was measured not at the CW rims, but by a dynamometer at the end of the drawbar, on the assumption that the pull on there (DP) ant the WRTE are equal. This assumption was obviously not true/ The CWBR, usually considered part of the Vehicle Resistance, occurs between the ITE and the DP, and its constant value can be calculated, from the bearing dimensions, the mass borne by them, the resulting pressure, the appropriate coefficient of friction (Cf), for that pressure, and the ratio of bearing diameter to the CW diameter, and should be deducted. So long as this is understood, it matters little whether the CWBR constant is deducted from the raw data or from the regression result. In addition, at Rugby, an unknown DR was incorporated between the ITE and DP."

"The unknown DR between ITE and DP" did not and could not exist. On the test plant, under steady state conditions, that being running at constant speed, the wheel rim tractive effort and the drawbar tractive effort will always be exactly the same. It cannot be otherwise. It there was a mismatch constant speed stability would no longer obtain. The DP is the passive slave of the WRTE, and vanishes should a slipping incident occur. The forces encountered by the dynamometer and the drawbar and coupled wheels are inevitably equal and opposite at all times under constant speed conditions. John Knowles' "unknown DR" affecting MF is just another of his several fantasies

“No tests were conducted at Rugby with all the reciprocating masses balanced solely for the tests. and the DR eliminated.”

DR, Drawbar Resistance, a supposedly interposing disturbance of the drawbar pull when encountered at dynamometer cannot exist as explained above. What does not exist cannot be eliminated. Reciprocating balance brings with it its own problems which is why engineers sought to keep the percentage balance as low as practicable. The tests with full balance suggested would be pointless

The next part of his submission turns to his supposed scientific analysis. It begins with: **“1. The TSAMR is so low in itself, and does not accord in any respect with the factors which should, from first principles explain it.”**

He goes on to adopt the test data for 9F 92250 as his choice for statistical analysis on the grounds that as the last locomotive tested on the plant **“the testing procedures should have been as perfected as they were to be”.** The data presented comprises four small remainder MF sets at 20, 30, 40 & 50 mph.

He goes on, inter alia: **“Four** (of the MF small remainders) **are less than the expected constant value of CWBR of 228lb.”**

As my randomised small remainder experiments have clearly demonstrated, the range of small remainder scatter evident in the 9F test data falls within the predicted range of possibility given the understood metrological limitations, and that includes the two negative MF outcomes. Less than 228 lb is of no statistical significance within this realm of possibility

More significant, the constant CWBR referred to is fundamentally flawed on two counts. Firstly, it is not constant across the speed range, the friction coefficient will increase as function of speed as I explained in my last letter: □□= ZN/P. Secondly the axle load and piston thrust force vectors are in directional misalignment, with the latter constantly changing in magnitude and angularity, constantly changing the resultant force resolution. Thus, the resultant bearing stress and friction will be less than the mathematical sum of the forces involved.

In other words, there will be significant mitigation of the losses involved. Lomonosof, who was no slouch when it came to complex analysis, hesitated when it came resolving this complex iteration mathematically.

“…..experiments that turn out badly should not have been recorded.”

How can an experiment be deemed to have turned out badly until its results are determined? Was “rejected”, rather than “not recorded” the intended meaning?

Surely negative machinery friction small remainders are failed experiments, but they remain in the data he adopts for statistical analysis. They would certainly pass the Grubbs Test justifying outlier elimination (not available in the 1950s), but simple logic alone is sufficient in such circumstances. Given the vestigial R^2 values of MF small remainder data sets, subjecting such data to regressions can be guaranteed to return unfavourable statistical outcomes. In the past John Knowles has opined to me personally that for a data set to worth looking at, the R^2 value should be at least 0.4 and the higher the better; the degree of scatter and inconsistency being inversely proportional the R^2 value. In this regard the options available for analysis stack up as below as exampled by the Graphs as below that appeared in my last letter.

Figure12. 9F MF Small Reminder MF Outcomes – All Speeds R^2 0.0042

Figure 13 9F Mechanical Efficiency - All Speeds R^2 0.4090

Figure 46 9F WRTE v ITE - 30 mph R^2 0.9974

The three outcomes above involve the same data set. The second option increases the R^2 value almost one hundredfold and the third well over two hundredfold. The sensible matrix for scrutiny seems rather obvious.

"Equations for TSAMR in lbs"

"(1) -73 + .0417; Rsqd , 285, SEE 279, tl -043, t 24.81
(2) 498 +.0145 ITE; ……." Etc, etc.

The above runs to 12 formulae, each of differing make-up, The explanatory notes are poor, I have no idea what the first number in each set represents for example. No matter, it is apparent from the preamble that the data under examination is the small remainder data set; about as useful as the results on a Bingo night. Again, the spurious constant, CWBR 228lb is referred to, and, lo and behold, the non-existent DR from the world of fantasy appears in the text rated at 172lb! The whole exercise is worthless. It is notable that he ascribes any apparent MR anomalies or improbabilities that fall out of his flawed analytical approach as solely attributable to dynamometer error. As the difference between two large numbers both subject to randomised scatter, it is witless. This implies the ITE data was perfect and totally devoid of scatter. The division of supposed error cannot possibly be sensibly determined by such a fundamentally flawed procedure.

"D R Carling, Superintending Engineer of the Rugby Plant, considered that what the plant registered was DBTE and not ITE."

What is this is supposed to mean? Both values were determined at Rugby, assuming DBTE refers to Plant DBTE, in other words, WTRE. ITE is determined from the measured IHP at the cylinders, so inevitably the effort at the drawbar and rollers is reduced. If as stated, Carling was simply stating fact.

"He considered that the ITE recorded there was suspect."

A citation is needed here. I cannot recall or trace such a statement from Carling. Interesting if true, he would have effectively been calling the measured cylinder performance into question. He did however say the indicator "spread of values was in the order of 3%", and "the mean values for each test was probably within 2%.". In relation to comparative tests with the Derby version of the Farnboro' indicator and the two mechanical indicator types used by Swindon, he did express regret that the test were only conducted up to 60 mph, and that systematic errors "might have become apparent at 80 or 90 mph." Something he thought more likely to afflict mechanical indicators.

If "ITE" was intended to mean the WRTE Carling simply said "We got the results right."

"What happened in May – June 1967?" Obviously 1957 was intended.

The Perform Programme

I am criticised for questioning the use and veracity of *Perform* as a tool for dissecting the Rugby experimental data, apparently in the guise of piston thrust determination. For a start this was perfectly possible before the advent of *Perform*, and simply done using the recorded test plant data. The substantive problem is that the *Perform* and test plant outcomes are routinely disparate, they cannot both be correct. Take your pick. John Knowles has consistently maintained the Rugby IHP data as sound (contrary to some evident conflictions), almost it seems, to the point of being unimpeachable. The published estimates for *Perform* as cited do not replicate the test plant in regard to both steam rate and IHP at a given speed and cut-off, with the results falling above and below the test data. The steam rate estimates are all out of step and by too much to be explained by the minutiae of valve setting. Carling considered the water and steam rate data as better than 1%.

The steam rate outcomes are at variance ranging from -27.9 to 6.7% with an overall negative trend averaging -5.7%. In these circumstances the test plant IHP data is unlikely to be replicated. The closest, within 2.3%, 500 Lb/hr, was the BR7. On the basis of IHP specific steam consumption, that would be worth nearly 40 HP. The *Perform* programme somehow contrives to make it 130 HP. The average SSC for 8 types on the plant is 14.27 and 13.85 lb/IHP.hr using *Perforrn*; a difference of -4.4%. Not bad perhaps, but too large in the context of machinery friction determination. None of the *Perform* estimates come closer than 4%. Plant and *Perform* cannot both be right.

The Crosti 9F

"It is almost standard for those writing about these engines to say that they had higher TSAR then the standard 9Fs, on account of reduced depth of frame stretchers, to fit in the preheater drum below the boiler barrel. That is not why I think they had higher TSAR, it is because they had very restricted blast nozzles in the chimneys, which create very high back pressure in the cylinders, as can be seen using Perform, and calculating the piston thrusts."

The evidence that clearly contradicts his conclusion, has been ignored. As demonstrated by Figures 45 and 46 in my last letter showing that, notwithstanding the significantly higher back pressure of 9F 92250 in double chimney guise compared to the Giesel ejector application, no discernable difference in WRTE at a given Indicated Tractive Effort was apparent. (The area of the Giesel blast arrangement by the way is 30.2 sq.in. not 302 as was shown). For the record, while the difference in back pressure at an ITE of 28,000lb for 92250 with double chimney and Giesel ejector was about 5lb, there was only a 2lb difference between single chimney 92050 and Crosti 92023 at this work rate. Any increase in back pressure would *reduce* the net piston thrust and frictional implications, not increase it.

It is evident from the numerous examples in the Rugby data that WRTE resolves into the simple relationship WRTE = ITEx – C, the constant C varying with speed. Coefficient x, the factor responding to effort sensitivity mostly falls within the 2 to 3% range and the formula routinely resolves into a linear trend line. The limited Vitry data displays similar characteristics. This variable is a summation of all the frictional sensitivities subject to the forces of effort, weight and dynamic effects obtaining. These are subject to the complex resultant mitigations of opposing forces such as the constant vertical coupled wheel axle load, piston thrusts (at moments opposing) and the shifting force vectors of the dynamic masses and cross couples. It's all a rather complicated resolution of forces which the power transmission system, without any resort to mathematics or computation, resolves and delivers a single and equal force at the drawbar and wheel rims. It cannot do anything else.

It is worth noting in regard to coupled wheel bearing stresses, that they are similar to those for passive wheels or even slightly higher; for example, 255lb/sq.in. for the Duchess coupled axleboxes as against 221lb/sq.in., for the trailing truck. The expected augmentation of bearing stresses when under power, are evidently relatively insignificant. An exception is bogie and pony wheels which routinely involve lower bearing stresses, 150lb/sq.in being typical. These axles are subject to significant lateral forces.

Small Remainder Regressions

A statistical fiasco, spurious outcomes guaranteed.

"With Fig, 17 the reader is informed that WRTE is linear, but that is not said in the headline , and is not shown to be universally so, being exemplified only at 30 mph. As ITE at a given steam rate is not linear with speed, and as MR is composed of elements which are not linear with speed, such linearity would be surprising, indeed wrong."

The main substance of the second sentence is correct, but that is the very reason Figure 17 was at constant speed, so it is not wrong or misleading. The linear relationship holds at any given speed, with the constant increasing with speed. Numerous examples are available from across the Rugby data. This relationship, in a different setting, appears in Figures 21, 22, & 29 of my December 2019 letter. Several components of machinery friction at a given speed are constant independent of effort, hence the constant coefficient.

"He should also calculate MR from first principles based on its causes or sources, I have done that."

I did that back in 2005, of which he is fully aware, having been informed of my analysis as it evolved. Such exercises can only be estimates amounting to "the likely order of magnitude". My process involved the summation of nine factors, each individually sensitive to the nature the of forces and losses involved in various ways. While the determination of forces, bearing stresses, dynamic loads and so on can be readily obtained, the selection of friction coefficients had to fall on the available empirical evidence and technical data sheets. The idea was to err on the pessimistic; I had no idea on how the numbers would stack up. The first runs came out in hundreds of pounds, not thousands, in other words similar to the Rugby data. A number of engine types were examined. In the light of more detailed analysis of the Rugby I would now approach such an exercise a bit differently.

I have no knowledge of the MR estimates claimed by John; given his serial mechanical misconceptions they seem unlikely to reflect reality.

"I consider (based on all the evidence he says and does in these posts, especially the claims he makes) that Doug Landau is not competent in technical and statistical terms to derive MR from the Rugby data or to express any judgment on the value of that data for such a purpose. He seems to believe that MR is a function of ITE. …. his view has not been refereed and will mislead others.

MR = ITE – WRTE; nothing more nothing less. I think I can manage the maths.
Of course, in essence, MR is a function of ITE: When ITE = 0, then MR - 0.
I think what he means is that it is not the sole cause of MR which is perfectly correct, but that is not what I say or the data says. The sensitivity to ITE is solely a function of the first term of a very simple equation; The negative constant second term (variable with speed), represents

the manifold losses independent of effort. This is just another episode of misspeak and misrepresentation. On the refereeing point I'll leave that for the moment.

Conclusions

I'm afraid an appraisal of John Knowles latest contribution to this correspondence can be nothing but harsh. It is not a situation that would be willingly chosen. He just digs deeper into a charade of serial misconceptions, cursory scrutiny, and a misguided and erroneous statistical approach posturing as science. His request to be refereed could not survive informed scrutiny. He might even consider it to have been already done, likewise the suggestion that I have not been refereed. When I gave for comment my last contribution, December 2019, and this one, along with copies of John Knowles' submissions including the most recent to Fred Rich C.Eng., M.I.Mech,E, he opined: "Well that demolishes John Knowles." Fred's career started out an engineering apprentice at Brighton Works, he worked at the Rugby Test Plant from June 1957 to October 1958. His summary view of John Knowles' submissions was dismissive.

Perhaps Andre Chapelon could have the last word when he said of the Rugby test data for the 9F and Crosti 9F; "The most consistent and accurate in his experience": *Riddles and the 9Fs*, Colonel H B C Rogers, Ian Alan, 1982. That is not to say the Rugby test data constitutes an impeccable anomaly free record, it was far from that. Much time was expended in the earliest days trying to achieve dashpot functionality, an endeavour that ultimately proved futile. Fortunately, its abandonment was of no operational consequence. The evolution of the Farnboro' indicator to a satisfactory level of performance and reliability was a protracted process with setbacks along the way. Ultimately the expected standards of all round accuracy were not consistently achieved until nearly half way through the plants operational history, excluding the early commissioning period. Even then the occasional hiccup could occur. Accuracies within 1% are impressive, about as good as then possible, but that's still enough potentially to significantly alter outcomes, especially small remainders, if it's a systemic error of constant sign. Thus, that some uncertainty remains is in the very nature of such tests.

Doug Landau

18 February 2021

Appendix 1

Introductory Note
Considerations of potential plant resonance were evaluated in the early stages of the test plant project. An early report on this topic is as below (my copy incomplete, date and authorship missing, some clipping).

THE REPORT OF THE OSCILLATING FORCES APPLIED TO A STATIONARY TEST PLANT BY A STEAM LOCOMOTIVE IN MOTION

Due to variation in traction force in the course of a revolution.
Due to the forces set up by the unbalanced reciprocating parts.

Introduction

This memorandum contains an investigation of the vibrations that may be set up in various parts of the dynamometric equipment of the Locomotive Testing Station about to be built by the L.M.S.R. and L.N.E.R. companies at Rugby.

In such a plant the locomotive with its wheels running on rollers is attached by a drawbar to a hydraulic dynamometer, the latter measuring and recording the tractive effort exerted. The plant forms a composite elastic system and is, therefore, capable of being set into vibration by any disturbing force. By reason of the unbalanced reciprocating parts, a periodic disturbing force operates while the locomotive is in motion, and, therefore, the whole plant will execute forced vibrations.

Apart from the general desirability of investigating the possible magnitude of these vibrations, another reason was provided by the experience of the French Railway Companies' plant at Vitry. When this plant was built, no means of damping the vibrations was incorporated, nor apparently was the possibility considered that the disturbing force might at some speed be of the same frequency as the natural oscillation of the plant.

During the early tests these conditions of resonance were actually attained with the result the whole plant was suddenly thrown into violent oscillation, the trouble was later removed by the introduction of a damping system (Belleville Washers).

Summary

The general nature and magnitude of the alternating forces in action upon a steam locomotive in motion are briefly considered. They consist of two principal components:

The first varies above and below a steady mean value, but always in one, direction whereas the second varies from maximum positive to maximum negative. once in each revolution of the coupled wheels. It is well known that the second component can often reach a value many times in excess of the steady mean tractive effort.

It is matter of common experience that these large variations are not to a train or recorded in the dynamometer car, and it can be proved that they have no significance in the steady motion of the train. The explanation lies in the fact that the frequency of the disturbing forces is many times greater than the natural frequency of the elastic system as a whole. Complete mathematical analyses of the problem as they affect a locomotive and train in service and conditions on a stationary test plant are presented (These are missing from my incomplete copy of the report as provided by a third party).

The conditions under which resonance can be set up in a locomotive and train by the action of periodic forces are investigated and it is found such resonances only occur at very low speeds. Resonance due to irregularity turning moment occurs at a lower speed than in the case of forces due to unbalanced reciprocating parts, and while the former causes a large amplitude of vibration it is shown the irregularity of turning moment becomes of no importance compared with the effect of the unbalanced reciprocating parts above a speed of about four of five miles per hour.

Irregularity of the turning moment can, therefore, be neglected when considering the conditions to be met on a stationary test plant,

Although is has been shown that in ordinary service on the road resonance will not cause trouble, the natural period of oscillation of a stationary test plant is very different from that of a locomotive and train. It is, therefore, most important that consideration should be given to the conditions under which resonance might occur on a testing plant and the steps necessary to ensure its suppression.

This matter is fully considered by means of mathematical analysis. At the same time attention is directed to the stresses which may expected in various parts of the system and investigation is made as to the requirement s necessary to ensure a smooth record of the mean draw bar pull og the locomotive under test.

The top of the third and last page of the Summary is cropped.

…..value of a particular locomotive, assumed to have considerable unbalanced forces at high speeds have been calculated together with the amount of damping required in the system for its complete suppression*.* It is shown that there will be no difficulty in so choosing the value of the damping coefficient that resonance will not occur under any practicable working conditions.

Further, by a suitable choice of the damping factor, stresses in the drawbar and the dynamometer system will be kept can be kept within satisfactory limits, and, on the assumptions made, it is calculated that the first section of the drawbar between the locomotive and the dashpots should not exceed approximately 12-ft in length with a diameter of $4^1/_2$".

The characteristics of the auxiliary springs introduced for the purpose of ensuring a smooth record of the drawbar pull have been calculated and it is demonstrated that they should have as low a modulus as possible, the limit being determined by the maximum deflection per ton of steady tractive effort which can the dealt with by the compensating mechanism for maintaining the locomotive in its initial position over the vertical centres of the supporting rollers.

Tables and graphs of amplitudes of vibration and stress in various parts of the drawbar and dynamometer are included (not in my possession). Values are tabulated and plotted for representative combinations of damping coefficients and stiffness of auxiliary springs. They cover the three sensitivity ranges of the hydraulic dynamometer and all speeds up to 70 radians per second.

It should be borne in mind that the calculated values of vibration and stress have been determined on the basis that extreme conditions exist. In the majority of cases the disturbing forces acting on the plant will be much less than assumed in the memorandum as high-speed tests will undoubtedly usually involve multi-cylinder engines, the disturbing forces in which are very much less than in the case of the two-cylinder locomotive assumed for the purpose of this investigation.

Remarks on Doug Landau comments to the RPS of February 2021

Doug Landau's submissions on Steam Locomotive Resistance are mostly related to data collected on the Locomotive Testing Station at Rugby, in much more detail than is needed to draw conclusions on steam locomotives running on the road hauling trains, which is the subject of interest to RPS members. I don't think the RPS website is the place for these discussions about the Rugby LTS and the scientific derivation of relationships for the Resistance of SLs. No members contribute on the subjects. It would be usual for a non-technical Society like the RPS to draw data on SLR from a technical journal, which employs a system of peer review of submissions on the subject, and encourages articles on SLR.

Above this submission are earlier pieces by Doug and myself. One includes several dozen pages with Doug's analyses of mostly Rugby data, which are of limited usefulness/value in my opinion because they are single variable equations for complex relationships, often fitted after Doug has weeded the data. They really require multiple variable regression equations, without the weeding.

I confine my remarks on what he has written this time to:

1 His claim that MR at coupled wheel rims and the pull on the dynamometer (DP) are precisely the same cannot be right, certainly on the road. Between the two are the resistance of the coupled wheel bearings (CWBR), a resolved sum of those bearings carrying much of the weight of the locomotive and the near horizontal tractive forces. If the mechanical resistance is expressed as (ITE – DP), it also includes the work done rotating and oscillating wheels and rods in the drive.

2 The 228-lbs is the minimum value or constant in a CWBR of a medium sized locomotive. It increases to a maximum per unit area with increased loads.

3 He claims that the MR of a Crosti 9F is reduced because the ITE is reduced. The ITE is reduced, however, because the back pressure in the cylinders is so much higher than that of a standard 9F on account of the very restricted exhaust nozzles on the Crosti, rendering the ITE so much lower. I enquired of officers who observed the Crosti 9F at Rugby, and they were unaware of weak frames leading to more parts incurring friction or higher friction. See reference to my 1988 notebook in 6 below, this time p 40. The engine resistance of a Crosti is said to overcome the gain from the boiler. More flexible frames were blamed. Chapelon was asked to report and accepted lesser rigidity of frames as reason for higher loss ITE to DP.

4 The assertion that Damping Resistance (DR) did not and could not exist. Belleville Washers were in the connection to the dynamometer right to the end at Rugby, and required work to operate them. That increased the value of ITE-DP and the TSAR.

5 He has never before published his criticisms of the Perform program. It is such a path breaking and valuable program, with the variables all known, that he could rewrite it if he so chose, using values he considered appropriate.

6 Carling's remarks on the low value of ITE recorded on the LTS appear in the file on reconciliation of results from Mobile Tests and those from the LTS (the subject of Report L116). These are: it is difficult to obtain ITE with great accuracy; error in the indicator, especially at high speeds; determination of area difficult and not very accurate. (Notes taken at Rugby 1988, p 44 of my notebook, copy of which given to Doug in the 1990s).

No doubt these few remarks will incur the wrath of DHL. I hope he sends them to somewhere more appropriate than the RPS website. I shall be happy to respond wherever they appear.

Locomotive Resistance - Reply to John Knowles letter 22 May 2021

It is commendably short, but it contains misrepresentations, and some plain mistakes calling for comment I will deal with his six points in the order presented.

The preamble and final paragraph suggests that the RPS is no place for a discussion of this nature and it should continue elsewhere. It's rather late to make this suggestion at such a late stage, and the suitability or otherwise of the RPS is something of a moot point. It does after all contain material of a technical nature from time to time.

The second paragraph bewails the fact that WRTE v ITE at a given speed resolves into a simple equation of the form WRTE = ITEx - C. on the grounds of the "limited usefulness" of "single variable equations for complex relationships." Complex indeed, but this simple single variable relationship is what consistently falls out of the test data time after time. It was only necessary to weed a single data set to eliminate an erroneous negative constant, An aberration entirely explicable, given the sensitivity to the potential hazards of random scatter. There's no guarantee multiple variables would indicate a reliable causal relationship, quite the reverse for such complex iterations is more likely, rendering them unfit for purpose.. A "bumblebees can't fly" outcome probable.

1. "His claim that MR at the coupled wheels and the pull on the dynamometer (DP) are precisely the same cannot be right, certainly on the road."

   I said nothing of the kind, I said the plant Amsler DBTE and WRTE were the same. Otherwise constant speed would not obtain.

2 The 228lb is the minimum value or a constant in a CWBR ...."

   It is more complex than that. CWBR, is not constant as a function of speed/RPM as previously explained. The WRTE outcome is a very complex mechanical iteration of manifold shifting forces that resolve into a linear relationship of single variable, negative constant form. .

3 .''He claims that the MR of a Crosti 9Fis reduced because the ITE is reduced."

   Really? Reduced! Surely this cannot be what was intended. The Crosti MR recorded on the test plant was substantially <u>increased</u> compared to a standard 9F, a characteristic confirmed by inferior DBHP on road tests. Obviously the reduced cylinder efficiency attributable to increased back pressure contributed to this, but said back pressure had no influence whatever on the increased MF. This was clearly demonstrated by the 92250 tests when fitted with Double Chimney and Giesel Ejector, where the differences in back pressure were significantly higher than was the case for Crosti v 92050. No difference in MF was discernable for 92250 in both Guises. See Figure 46 , my submission November 2019.

4 "The assertion that Damping Resistance (DR) did not and could not exist" "Bellville Washers required work to operate."

   DR is an irrelevance. If I stretch an elastic band the applied force and restraining force will be exactly the same. As Carling pointed out, the abandoned dashpot would have worked had it been installed in series rather than in parallel.

5. "He has never before published his criticisms of the Perform programme."

   It is a matter of fact that published Perform estimates do not tally with the Rugby test data in regard to steam rate and IHP for given working conditions It follows that they cannot both be right, perhaps both are wrong. Perform may have some utility, but has insufficient authority to substitute for MR evaluation. My reservations are therefore primarily empirical. It is interesting that Perform contradicts the Rugby data in which JK has previously expressed much confide“Carling’s remarks on the low value of ITE recorded on the LTS ”

6. “Carling’s remarks on the low value of ITE recorded on the LTS”0

I’m not sure what point is being made here. Carling did express the wish that the comparative tests with mechanical indicators in 1953 had been carried out at higher speeds than 50 mph. As things turned out the two mechanical types had to be returned to Swindon for recalibration. It is interesting that the corrected results for the Maihak indicator remained 2% higher than Rugby’s Farnboro indicator results. The recalibrated Dobbie & McInnes indicator was up to 7% high on Rugby at low steam rates. Report L116 identified steam rate anomalies as the prime source of the IHP disparities between that arose between plant and road tests.

It is disappointing that John Knowles appears unable to change his mind when the experimental evidence unequivocally contradicts his views, as exampled by his supposed effects of back-pressure on machinery friction; ref point 3 above. The chances he will concede his several conceptual errors seems remote. I have no wish to continue these discussions elsewhere