THE YOUNG ACHIEVER AWARD 2019

This award is open to all involved in rail infrastructure engineering projects; engineers, project & maintenance staff, graduates, apprentices etc.
For details of the 2020 award see page 56.

RADEAN: Use of data to analyse rail wear and identify trends

AUTHOR:

Timothy Atkinson
Network Rail

FIRST PRIZE

ABSTRACT

The role of data on the railway is increasing in value. RADEAN, or Rail Degradation Analysis, is a new tool built into an Excel workbook and is designed to sort Ultrasonic Testing Unit data and arrange it into an easier-to-read format and provide a user with options to analyse various rail characteristics and graphically display the deterioration of any site within a route.

This report covers the initial use of formulas and their subsequent removal due to high processing duration and the development of Visual Basic for Applications code to allow for more data processing at higher speeds. It discusses the uses RADEAN can have for identifying trends on sites and the prediction of wear-related failures and non-compliances, and its possibility for use in the wider rail community in the future.

INTRODUCTION

By the very nature of wheel-rail interface, rails will always wear. This can be extremely problematic. Regular measurements using NR4 step gauges and Network Rail's Ultrasonic Testing Unit (UTU) supply data to Track Maintenance Engineers (TMEs) and notifies them of non-compliances. These non-compliances are then corrected within a given timeframe, with potential cost implications for line blocks, possessions and staffing. However, in an age where data drives key business decisions, (Figueres-Esteban, Hughes, & Van Gulijk, 2015), is this reactive approach really the best way? Certainly, within Network Rail very little research has been conducted into predictive maintenance and cascaded to depots and TMEs.

This report covers the development of a project assigned to a graduate by a TME on the London North Western (LNW) route. The remit of this project was to:

- Make the UTU data easier to read and provide a list of the highest deteriorating sites.
- Provide assurance that the frequencies of inspections for sites within the Maintenance Delivery Unit's (MDUs) patch were compliant with those specified in standard NR/L2/TRK/001/mod02.

The project developed into a fully coded inbuilt software in Excel, with potential for predicting rail wear and the report has been structured logically, presenting the original formula-reliant Mk.I RADEAN, its drawbacks and the issues with the UTU data. The 27th and most recent version of RADEAN is then presented. The code itself is referred to but is not discussed as it is not relevant to railway infrastructure, but its outputs are relevant, and they are discussed. For completeness, relevant parts of the code are included in the appendix with some annotation explaining them.

DEVELOPMENT

RADEAN MK.I

UTU data arrives in the form seen in Figure 1. The first issue with the data is apparent; the worksheet(s) is crowded, with various colours present. There is no Network Rail standard for the presentation of UTU data meaning that it may not be fully clear to an engineer unfamiliar with it. There are titles in the top row, but these do not explain the colours or show what the optimum values would be to allow a comparison. It is not explained whether measurements such as head loss are indicating the head loss since the last measurement or the head loss relative to a new rail. In addition to this, the measurements are split into left and right rail; there are multiple track IDs; some ELRs have measurements taken on different dates and some UTU runs do not take measurements at the same yardage each time, resulting in gaps.

The UTU folders are structured as in Figure 2. With each workbook containing multiple ELRs, locating required data can take some time. With that in mind, a directory was created. Shown in Figure 3, the directory contains a hyperlink to the relevant workbook and the dates and ELR covered in it. The data can then be found and filtered using the Filter function so only the desired ELR is selected, and then copied into RADEAN.

With a directory complete, the next step was making the data easier to read. The 2 most recent measurements from an ELR were copied into a blank workbook. The data was split into a Left Rail and Right Rail worksheet, sorted as in Figure 4, showing two different UTU runs together allowing an instant comparison to be made between measurements.

Calculations were conducted in additional columns to show the rate of deterioration at each site as seen in Figure 5. Each of the characteristics seen in a green column in Figure 1 were to be analysed so an initial formula was written involving the subtraction of one value from the one preceding it. A second column was added for calculating the wear rate, dividing the change in wear by the number of days in between measurements. Figure 6 shows the formula for change in wear and Figure 7 shows the formula for the rate of change of wear.

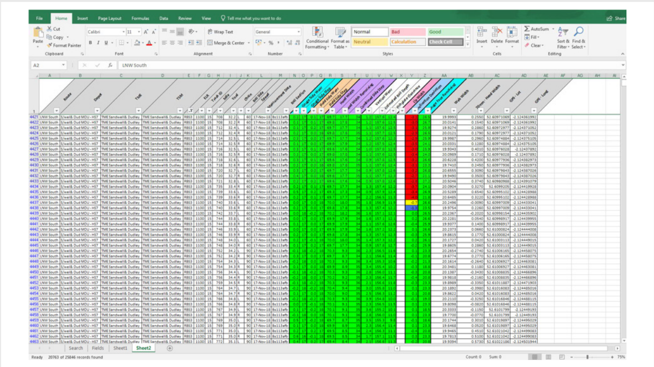


Figure 1: Raw data as captured by the UTU

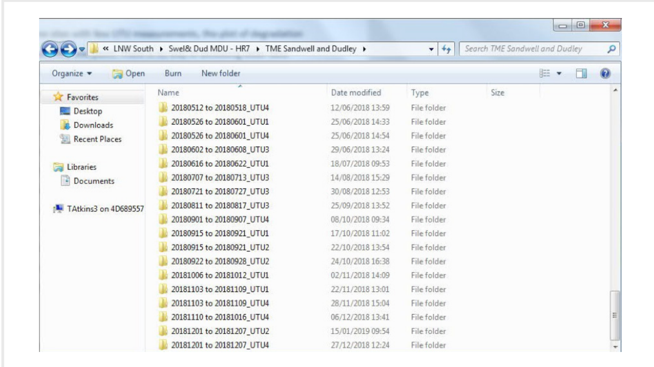
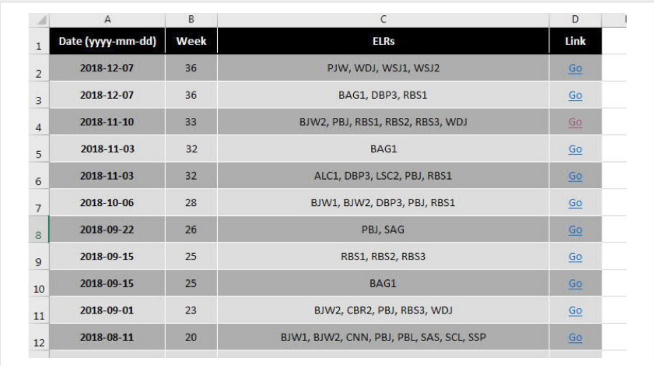


Figure 2: The folder containing UTU data for Sandwell and Dudley MDU, with the dates in yyyy-mm-dd format.



	A	B	C	D
1	Date (yyyy-mm-dd)	Week	ELRs	Link
2	2018-12-07	36	PJW, WDJ, WSJ1, WSJ2	Go
3	2018-12-07	36	BAG1, DBP3, RBS1	Go
4	2018-11-10	33	BJW2, PBJ, RBS1, RBS2, RBS3, WDJ	Go
5	2018-11-03	32	BAG1	Go
6	2018-11-03	32	ALC1, DBP3, LSC2, PBJ, RBS1	Go
7	2018-10-06	28	BJW1, BJW2, DBP3, PBJ, RBS1	Go
8	2018-09-22	26	PBJ, SAG	Go
9	2018-09-15	25	RBS1, RBS2, RBS3	Go
10	2018-09-15	25	BAG1	Go
11	2018-09-01	23	BJW2, CBR2, PBJ, RBS3, WDJ	Go
12	2018-08-11	20	BJW1, BJW2, CNN, PBJ, PBL, SAS, SCL, SSP	Go

Figure 3: The created UTU directory.

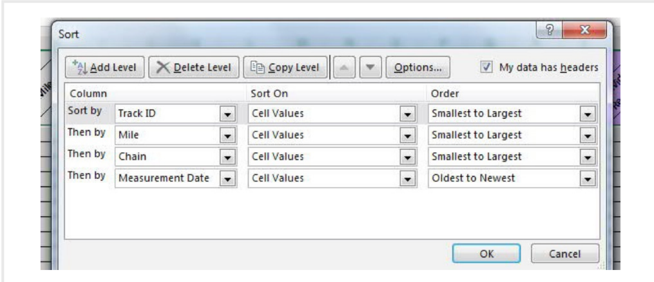


Figure 4: The custom sorting for RADEAN Mk.I

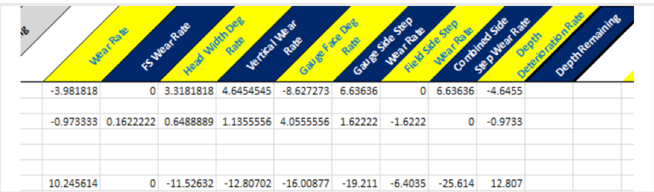


Figure 5: New columns containing calculated values

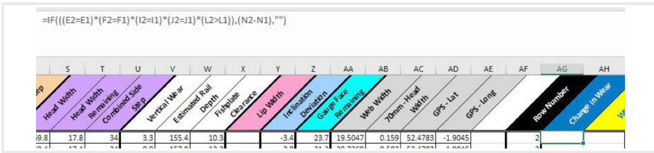
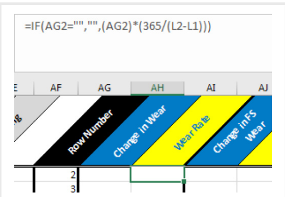
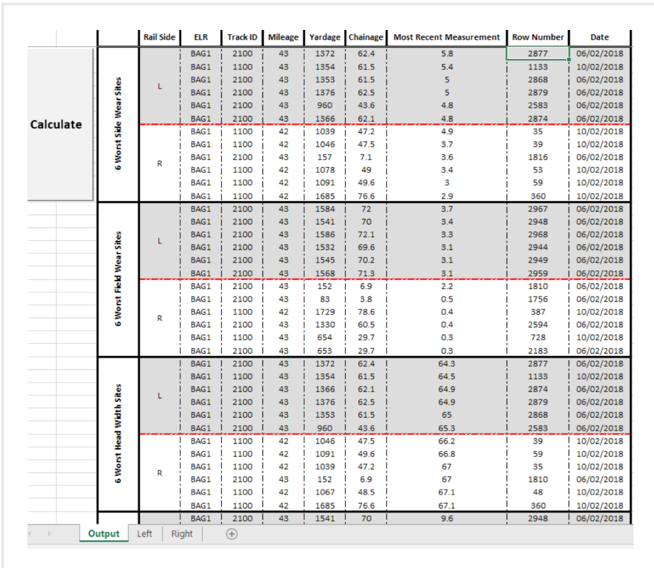


Figure 6: Formula for calculating the change in wear

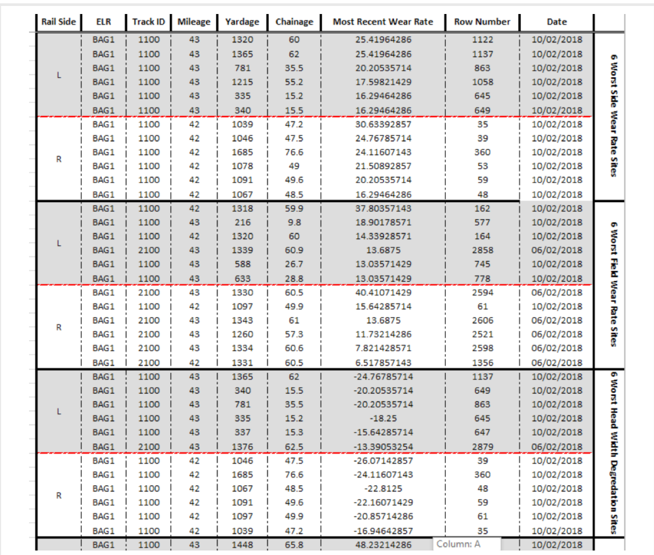
Figure 7: Formula for calculating the rate of wear





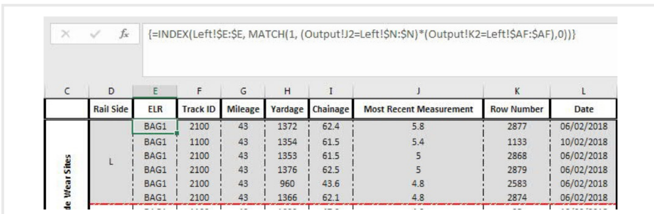
Rail Side	ELR	Track ID	Mileage	Yardage	Chainage	Most Recent Measurement	Row Number	Date	
L	BAG1	2100	43	1372	62.4	5.8	2877	06/02/2018	6 Worst Side Wear Sites
	BAG1	1100	43	1354	61.5	5.4	1133	10/02/2018	
	BAG1	2100	43	1353	61.5	5	2868	06/02/2018	
	BAG1	2100	43	1376	62.5	5	2879	06/02/2018	
	BAG1	2100	43	960	43.6	4.8	2583	06/02/2018	
	BAG1	2100	43	1366	62.1	4.8	2874	06/02/2018	
R	BAG1	1100	42	1039	47.2	4.9	35	10/02/2018	
	BAG1	1100	42	1046	47.5	3.7	39	10/02/2018	
	BAG1	2100	43	157	7.1	3.6	1816	06/02/2018	
	BAG1	1100	42	1078	49	3.4	53	10/02/2018	
	BAG1	1100	42	1091	49.6	3	59	10/02/2018	
	BAG1	1100	42	1685	76.6	2.9	360	10/02/2018	
L	BAG1	2100	43	1584	72	3.7	2967	06/02/2018	6 Worst Field Wear Sites
	BAG1	2100	43	1541	70	3.4	2948	06/02/2018	
	BAG1	2100	43	1586	72.1	3.3	2968	06/02/2018	
	BAG1	2100	43	1532	69.6	3.1	2944	06/02/2018	
	BAG1	2100	43	1545	70.2	3.1	2949	06/02/2018	
	BAG1	2100	43	1568	71.3	3.1	2959	06/02/2018	
R	BAG1	2100	43	152	6.9	2.2	1810	06/02/2018	
	BAG1	2100	43	83	3.8	0.5	1756	06/02/2018	
	BAG1	1100	42	1729	78.6	0.4	387	10/02/2018	
	BAG1	2100	43	1330	60.5	0.4	2594	06/02/2018	
	BAG1	1100	43	654	29.7	0.3	728	10/02/2018	
	BAG1	2100	43	653	29.7	0.3	2183	06/02/2018	
L	BAG1	2100	43	1372	62.4	64.3	2877	06/02/2018	6 Worst Head Width Sites
	BAG1	1100	43	1354	61.5	64.5	1133	10/02/2018	
	BAG1	2100	43	1366	62.1	64.9	2874	06/02/2018	
	BAG1	2100	43	1376	62.5	64.9	2879	06/02/2018	
	BAG1	2100	43	1353	61.5	65	2868	06/02/2018	
	BAG1	2100	43	960	43.6	65.3	2583	06/02/2018	
R	BAG1	1100	42	1046	47.5	66.2	39	10/02/2018	
	BAG1	1100	42	1091	49.6	66.8	59	10/02/2018	
	BAG1	1100	42	1039	47.2	67	35	10/02/2018	
	BAG1	2100	43	152	6.9	67	1810	06/02/2018	
	BAG1	1100	42	1067	48.5	67.1	48	10/02/2018	
	BAG1	1100	42	1685	76.6	67.1	360	10/02/2018	
	BAG1	2100	43	1541	70	9.6	2948	06/02/2018	

Figure 8: The Output worksheet for RADEAN Mk.I



Rail Side	ELR	Track ID	Mileage	Yardage	Chainage	Most Recent Wear Rate	Row Number	Date	
L	BAG1	1100	43	1320	60	25.41964286	1122	10/02/2018	6 Worst Side Wear Rate Sites
	BAG1	1100	43	1365	62	25.41964286	1137	10/02/2018	
	BAG1	1100	43	781	35.5	20.20535714	863	10/02/2018	
	BAG1	1100	43	1215	55.2	17.59821429	1058	10/02/2018	
	BAG1	1100	43	335	15.2	16.29464286	645	10/02/2018	
	BAG1	1100	43	340	15.5	16.29464286	649	10/02/2018	
R	BAG1	1100	42	1039	47.2	30.63392857	35	10/02/2018	
	BAG1	1100	42	1046	47.5	24.76785714	39	10/02/2018	
	BAG1	1100	42	1685	76.6	24.11607143	360	10/02/2018	
	BAG1	1100	42	1078	49	21.50892857	53	10/02/2018	
	BAG1	1100	42	1091	49.6	20.20535714	59	10/02/2018	
	BAG1	1100	42	1067	48.5	16.29464286	48	10/02/2018	
L	BAG1	1100	42	1318	59.9	37.80357143	162	10/02/2018	6 Worst Field Wear Rate Sites
	BAG1	1100	43	216	9.8	18.90178571	577	10/02/2018	
	BAG1	1100	42	1320	60	14.33928571	164	10/02/2018	
	BAG1	2100	43	1339	60.9	13.6875	2858	06/02/2018	
	BAG1	1100	43	588	26.7	13.03571429	745	10/02/2018	
	BAG1	1100	43	633	28.8	13.03571429	778	10/02/2018	
R	BAG1	2100	43	1330	60.5	40.41071429	2594	06/02/2018	
	BAG1	1100	42	1097	49.9	15.64285714	61	10/02/2018	
	BAG1	2100	43	1343	61	13.6875	2606	06/02/2018	
	BAG1	2100	43	1260	57.3	11.73214286	2521	06/02/2018	
	BAG1	2100	43	1334	60.6	7.821428571	2598	06/02/2018	
	BAG1	2100	42	1331	60.5	6.517857143	1356	06/02/2018	
L	BAG1	1100	43	1365	62	-24.76785714	1137	10/02/2018	6 Worst Head Width Degradation Sites
	BAG1	1100	43	340	15.5	-20.20535714	649	10/02/2018	
	BAG1	1100	43	781	35.5	-20.20535714	863	10/02/2018	
	BAG1	1100	43	335	15.2	-18.25	645	10/02/2018	
	BAG1	1100	43	337	15.3	-15.64285714	647	10/02/2018	
	BAG1	2100	43	1376	62.5	-13.39053254	2879	06/02/2018	
R	BAG1	1100	42	1046	47.5	-26.07142857	39	10/02/2018	
	BAG1	1100	42	1685	76.6	-24.11607143	360	10/02/2018	
	BAG1	1100	42	1067	48.5	-22.8125	48	10/02/2018	
	BAG1	1100	42	1091	49.6	-22.16071429	59	10/02/2018	
	BAG1	1100	42	1097	49.9	-20.85714286	61	10/02/2018	
	BAG1	1100	42	1039	47.2	-16.94642857	35	10/02/2018	
	BAG1	1100	43	1448	65.8	48.23214286	Column: A	10/02/2018	

Figure 9: The calculated degradation rates on the Output worksheet for RADEAN Mk.I



{=INDEX(Left!$E:$E, MATCH(1, (Output!J2=Left!$N:$N)*(Output!K2=Left!$AF:$AF),0))}

C	D	E	F	G	H	I	J	K	L
	Rail Side	ELR	Track ID	Mileage	Yardage	Chainage	Most Recent Measurement	Row Number	Date
de Wear Sites	L	BAG1	2100	43	1372	62.4	5.8	2877	06/02/2018
		BAG1	1100	43	1354	61.5	5.4	1133	10/02/2018
		BAG1	2100	43	1353	61.5	5	2868	06/02/2018
		BAG1	2100	43	1376	62.5	5	2879	06/02/2018
		BAG1	2100	43	960	43.6	4.8	2583	06/02/2018
		BAG1	2100	43	1366	62.1	4.8	2874	06/02/2018

Figure 10: Formula for extracting the ELR of the site with the highest wear

Both formulas involve IF statements; Excel will only calculate the result of the formulas provided that: the ELR, mileage, chainage, rail side, track ID and rail section type are the same, and the dates of the two measurements are different. Using these IF statements meant that only relevant data was processed, hence why some rows are blank.

Figure 8 and Figure 9 show the Output worksheet for RADEAN. The left half of the work-sheet contains the 6 worst values of each characteristic and the right half presents the cor-responding rates of degradation. The grey rows show the left rail, and the white ones show the right. This output sheet searches through both the left and right worksheets, finds the 6 worst sites for each characteristic and then extracts them to Output. After researching vari-ous formulas available in Excel, it was found that the use of the LARGE function was the most appropriate.
The function is used with an IF statement to ensure that only the most recent values were selected; the formula can be found in Figure 11. It states, "give the highest value for gauge side wear IF the date of the measurement is the most recent one for that site". The LARGE function can give the nth largest value by changing the number at the end of the formula.

The 6 worst sites are extracted by the row number seen in column K. Excel pulls the relevant location details from that row number. By doing this the formulas for the location cells are much simpler, bring the value from the required column and the provided row number, using INDEX and MATCH as in Figure 12.

Other ELRs were inputted to RADEAN and the workbook was evaluated. The main issue was found when a larger ELR was placed in it. The amount of time taken to process each of the many formulas in Output alone was reaching hours, and ultimately would discourage any serious use. In addition to this, there was a risk that the formulas could be edited in error and the workbook would fail to operate at all.

Methods of increasing processing speed in Excel were researched; almost all articles suggested reducing the amount of formulas which was an inappropriate solution given the nature of RADEAN. The solution lay in coding; after a brief foray into using MATLAB (rejected due to the licence cost of £1400 per user), Excel's own Visual Basic for Applications was discovered. It is based on Microsoft's Visual Basic (VB), discontinued in 2008. VBA is nested into Microsoft Office applications. It can be extremely useful in scenarios where a process is repetitive, or an Excel workbook is formula-heavy.

RADEAN MK.XXVII

Through 27 iterations, RADEAN progressed from using formulas to full coding. Due to no experience with computer programming, coding with VBA had to be self-taught; it is commonly used by computer programmers so multiple resources exist to aid the budding programmer.

VBA operates using modules, sub-modules and user forms. Each separate section of code lies in a sub-module, and a collection of sub-modules make up a module. User forms are necessary when a user is required to input a value; in the case of RADEAN, the user must choose which characteristic to analyse. Figure 13 shows the VBA menu of RADEAN; there are 7 worksheets within the workbook which will be explained later in this report, 3 user forms and 8 modules.

Three command buttons were placed into the RADEAN home screen in Figure 14; a specific process occurs when a user clicks them. Clicking Initialise RADEAN will display the user form of Figure 15. This is how the user selects the rail characteristic for analysis.

The example used here will be the RBS3. The data is imported from the UTU folder, Initialise RADEAN is clicked and the Gauge Side Wear is selected as the characteristic. Once OK is clicked, the user form ufProgress appears as in Figure 16. As the amount of data increased with more UTU measurements being taken, the length of processing time was increasing, and it was decided to provide the user with a progress bar to indicate how long the analysis would take.

The coding behind this was very simple; VBA counts the number of rows it must calculate and perform processes on, calculates the percentage completed and then displays that percentage as a portion of the blue bar.

VBA can create pop-up message boxes informing users of developments or issues within RADEAN. Figure 17 shows one such message box; the Excel icon on the taskbar flashes orange so a user may leave RADEAN to run in the background and be notified when results are available. Upon clicking "OK", the screen will be filled with the results as seen in Figure 18.

The output column has been conditionally formatted; the colour and point at which cells are coloured is dependent on the characteristic being analysed; in the case of gauge side wear, any wear above 3mm is highlighted in yellow.

Having completed the first deliverable of the project in making the UTU data more user friendly, attention turned to the second deliverable; Provide assurance that the frequencies of inspections for sites within the Maintenance Delivery Unit's (MDUs) patch were compliant with those specified in standard NR/L2/TRK/001/mod02.

The track categories were obtained for the route and inputted to VBA so once the inspection frequency monitor ran it would check the data for its ELR, TID and mileage and from that identify the correct inspection frequency as per module 002 of TRK/001. This figure was then compared to the actual time difference between the measurements and any non-compliances were extracted to the worksheet "Overdue Sites", as in Figure 19.

Using the top row as an example, the required inspection frequency is 8 weeks. The previous two measurements were 33 days apart, therefore compliant. However, for that site the last measurement was taken 300 days ago meaning it is overdue an inspection. A similar conditional formatting sub-module is applied to Overdue Sites as in Output, where historical non-compliances are coloured orange and current non-compliances are coloured red.

EXTRA FEATURES

After completing the code for monitoring the inspection frequency, further additions to RADEAN were explored.

The first target was to create a button to automatically extract a user-defined site and plot a graph showing its degradation over time; this would be an incredibly useful visual tool to aid in monitoring how rapidly a site deteriorates. Figure 20 and Figure 21 show the graph of gauge side wear for a site. This is where the worksheet "Graphit" is used; a user selects a cell in Output and selects "Analyse the degradation trend for this site" from the Options menu in Figure 22. RADEAN will find that site in the raw data, collect all historic measurements and plot a chart of the measurements.

The theory behind the Graphit worksheet is that as more data is added, more patterns will appear for sites and cases can be made to Route Asset Managers (RAMs) for strategic renewals before the rail reaches the point of non-compliance. First though, a check would be suitable to ensure a site isn't already due for renewal. Therefore, the CP6 workbank was imported into RADEAN. When a cell is selected and "Search for this site in the CP6 workbank" is selected in the Options menu, the workbank is checked for any renewals that will cover the site selected. If one is present as in Figure 23, RADEAN takes the user to that site in the workbank for closer inspection. If no job is found, an error message appears informing the user that no renewal is scheduled to take place for that site. This means that if a TME were to find a site that had a pattern of degrading every 3 years for example and found out it had been 2 years since the previous renewal, they would be able to search for the site in the workbank, and if there was no renewal planned within the next year they would be able to put a case to the RAM for renewal.

The final addition to RADEAN was a conditional formatting module to highlight any sites near to a user-defined one, as in Figure 24. This came following a discussion with a RAM about renewals, and how large a site needed to be before a full renewal would be

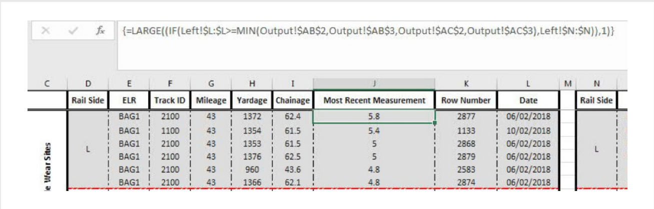


Figure 11: Formula showing for extraction of the highest wear value

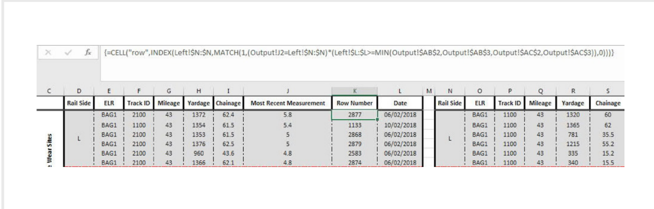


Figure 12: Formula looking up the row number of the highest wear value

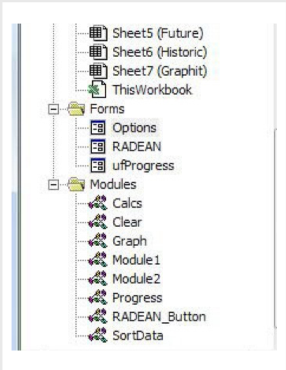


Figure 13: VBA Menu showing the worksheets, user forms and modules

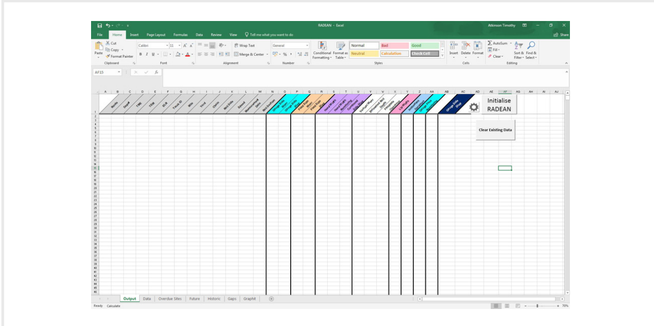


Figure 14: The RADEAN home screen

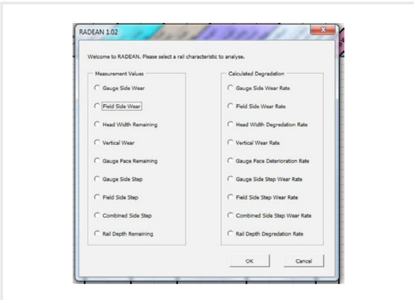


Figure 15: RADEAN user form

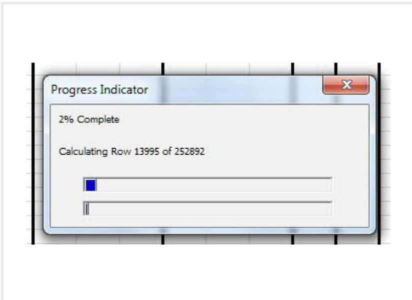


Figure 16: RADEAN progress bar

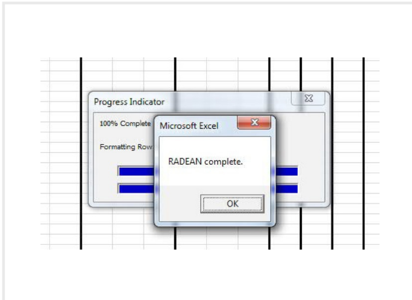


Figure 17: Message box informing user of RADEAN's completion

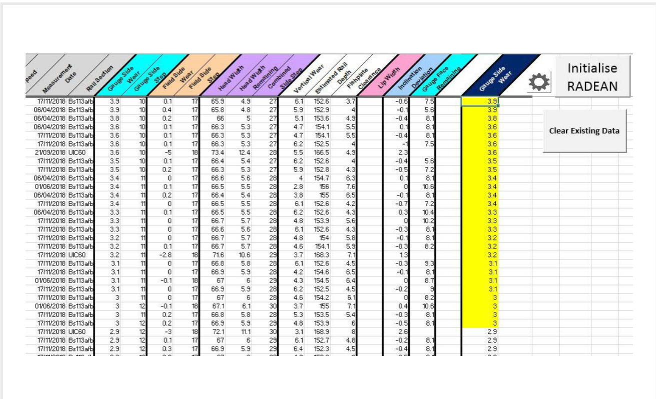


Figure 18: Completed RADEAN for gauge side wear

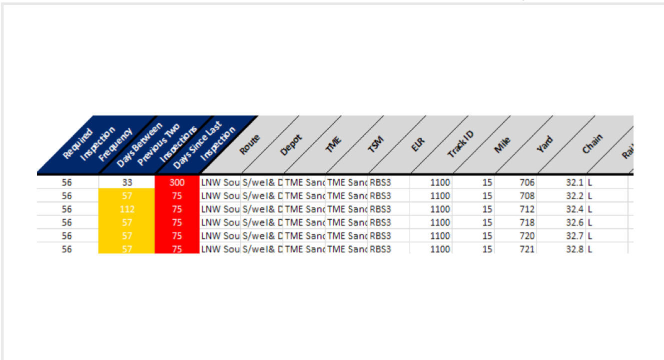


Figure 19: Inspection frequency non-compliances. Dates have been altered to provide non-compliances of varying degrees to highlight conditional formatting

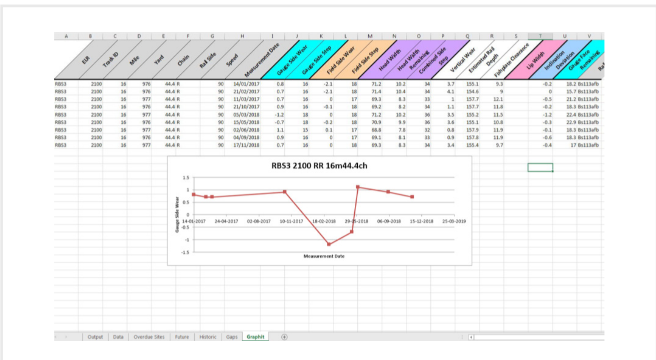


Figure 20: Gauge side wear trend for a specific site

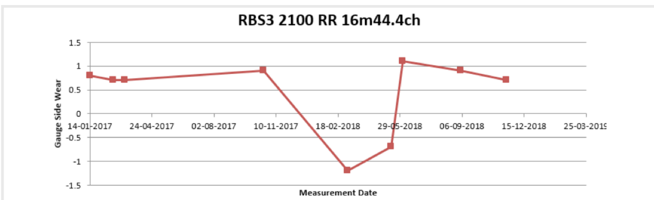


Figure 21: A closer look at the graph of degradation. The wear becomes negative around February 2018 due to maintenance being conducted on that site, causing an apparent decrease in wear compared with the previous UTU measurement

considered. With that considered, RADEAN will highlight any site within 400 yards either side of the user-defined site so a TME can see if a high-wearing site is part of a larger issue.

CONCLUSION

What started as a project for a graduate developed into a piece of software which could potentially prove vital as the role of data in the Railway increases. Continuous improvement has been in practice, constantly adapting formulas and increasing processing speed to allow RADEAN to work on any PC at a good level. The ability to display graphically the deterioration of a site can aid the understanding of vast amounts of data which engineers encounter daily and inspection frequency can be monitored to ensure compliance. Going forward, if RADEAN were to continue to be developed and expanded it could help to predict wear-related failures before they happen and keep MDUs compliant. RADEAN is currently being used to monitor ELRs within one MDU and identify previously unidentified patterns in wear and could be involved as evidence for a renewal in the future.

NEXT STEPS

Issues were identified with the data, both in terms of errors and format as more data was imported from various ELRs and updated with more measurements. It is likely that there are more issues that could be caused by data that have yet to be encountered, so since RADEAN has been trialled with the data from one MDU, the next logical step would be to pass it to another MDU with different data.

Writing the track categories does take a substantial amount of time and it would be unwise to attempt to write the track categories for every section of line in the country. Therefore, going forward, options are being explored in which a variable is assigned to a track category and that category is inputted by the user, rather than all line categories being assigned. A failsafe could be incorporated that if no line category is defined then it is assumed for all sites to be category 1A.

There is also work underway to begin importing data into RADEAN directly from UTU folders, removing the manual copy-and-paste element that is still present in the current version.

This would be the final step in automating the process as far as possible and speeding up the process even further.

REFERENCES & BIBLIOGRAPHY

Attoh-Okine, N. (2014). Big Data Challenges in Railway Engineering. Newark: Department of Civil and Environmental Engineering, University of Delaware.
Figueres-Esteban, M., Hughes, P., & Van Gulijk, C. (2015). The role of data visualization in Railway Big Data Risk Analysis. Huddersfield: Institute of Railway Research, University of Huddersfield.
The Permanent Way Institution. (2014). Understanding Track Engineering. Great Britain: The Lavenham Press.

TECHNICAL ARTICLE

AS PUBLISHED IN

The Journal July 2019 Volume 137 Part 3

Permanent Way Institution

If you would like to reproduce this article, please contact:

Kerrie Illsley
JOURNAL PRODUCTION EDITOR
Permanent Way Institution
kerrie.illsley@thepwi.org