

Enhancing Railway Pantograph Carbon Strip Prognostics with Data Blending through a Time-Delay Neural Network Ensemble

Alexandre Trilla¹, Verónica Fernández², and Xavier Cabré³

^{1,2,3} *Alstom R&D Services, Santa Perpètua de la Mogoda, Barcelona, 08130, Spain*

alexandre.trilla@alstomgroup.com

veronica.fernandez@alstomgroup.com

francesc-xavier.cabre@alstomgroup.com

ABSTRACT

Energy supply for high-speed trains is mainly attained with a high-voltage catenary (i.e., the source on the infrastructure) in contact with a sliding pantograph (i.e., the drain on the rolling-stock vehicle). The friction between these two elements is minimised with a carbon strip that the pantograph equips. In addition to erosion, this carbon strip is also subject to abrasion due to the high current that flows from the catenary to the train. Therefore, it is of utmost importance to keep the degradation of the carbon material under control to guarantee the reliability of the railway service. To attain this goal, this article explores an accurate (i.e., uncertainty bounded) predictive method based on a robust online non-linear multivariate regression technique, considering some factors that may have an impact on the degradation on the carbon strip, such as the seasonal condition of the contact wire, which may develop an especially critical ice build-up in the winter. The proposed approach uses a neural ensemble to integrate all these sources of potential utility with the carbon strip data, which is convoluted in time with a set of spreading filters to increase the overall robustness. Finally, the article evaluates the effectiveness of this prognosis approach with a dataset of pantograph carbon thickness measurements over a year at the fleet level. The results of the analysis prove that it is definitely possible to deploy a fine prediction, and thus yield a new avenue for business improvement through the application of the predictive maintenance approach to pantograph carbon strips.

1. INTRODUCTION

The railway environment in general, and the maintenance of rolling-stock in particular, are recently experiencing great benefits with the deployment of data-driven Prognostics and Health Management (PHM) technology (Atamuradov, V., Medjaher, K., Dersin, P., Lamoureux, B., and Zerhouni, N.,

Alexandre Trilla et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



Figure 1. Alstom TrainScanner deployment at the Manchester Traincare Centre.

2017; Tsui, K. L., Chen, N., Zhou, Q., Hai, Y., and Wang, W., 2015). In line with this source of innovation, Alstom has developed the TrainScanner, which is a track-side train monitoring system that is aimed at optimising the maintenance of brake pads (Trilla, A., Dersin, P., and Cabré, X., 2018), pantograph carbon strips, and wheelsets (Trilla, A., and Cabré, X., 2018), see Figure 1. This product is based on a set of computer vision technologies with lasers and 3D cameras that capture the degradation-related measures for each component as the trains traverse its portal. Then, it automatically triggers the analysis of the collected data, and advises the maintenance team with data-informed prescriptions. This work is particularly focused on the pantograph prognostic enhancement that may be attained with the carbon strip thickness measurements over time.

The British Rail Class 390 rolling stock is an electric high-speed passenger train that conducts the current collection through a pantograph. Therefore, the pantograph is an essential element of the traction chain because it provides access to the power to drive the traction motors, among other

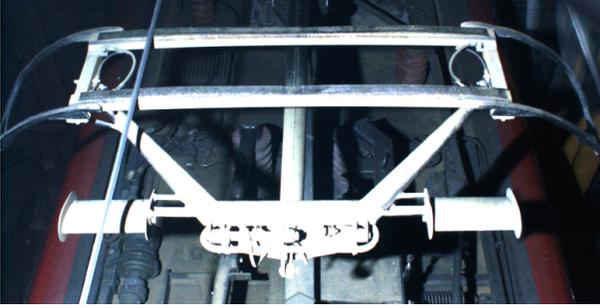


Figure 2. British Rail Class 390 pantograph showing the two carbon strips. One of them leads the contact with the catenary.

systems. In order to draw the current while the train is in motion, the pantograph equips two carbon strips that are in constant sliding contact with the overhead line, also known as the catenary, see Figure 2. Given the permanent friction regime of this means of power transfer, each carbon strip is subject to wear. And in addition to this main degradation mode, there are many other factors that may impact the condition of this asset, such as the amount of current flow (Bucca, G., and Collina, A., 2015; Ding, T., Xuan, W., He, Q., Wu, H., and Xiong, W., 2014), the irregular contact height relative to the rails (Shing, A. W. C., and Wong, P. P. L., 2008), the specific carbon material (Auditeau, G., Bucca, G., Collina, A., and Tanzi, E., 2011; Auditeau, G., 2016), and the ambient temperature (Ocoleanu, C. F., Popa, I., Manolea, G., Dolan, A. I., and Vlase, S., 2009). The combined effect of all these phenomena may produce chips and cracks on the surface of the carbon strip, although the most critical degradation factor that can be directly observed is the season.

This work conducts a thorough analysis of the pantograph carbon strip degradation at the fleet level in order to enhance the performance of its thickness prediction at 30,000 km into the actual operating life of each asset, which is expected to show a great deal of variation according to the seasonal weather. Given the intense mission profile of the fleet, this horizon for the prediction is assumed to provide enough notice time for the maintenance team to schedule the depot resources effectively. The proposed model of the degrading carbon thickness sequence exploits its diversity in time (or distance) through a set of spreading convolutions. Finally, the prognosis evaluation is performed with a rolling window prediction technique, focusing on the uncertainty of the predicted error, which is given by the maximum variability of the error distribution for a given confidence interval.

The article is organised as follows: Section 2 describes the analysis procedure that has been explored, including the description of the data, the evaluation technique, and the prognosis enhancements, along with their preliminary results. Section 3 discusses the overall outcomes and the limitations

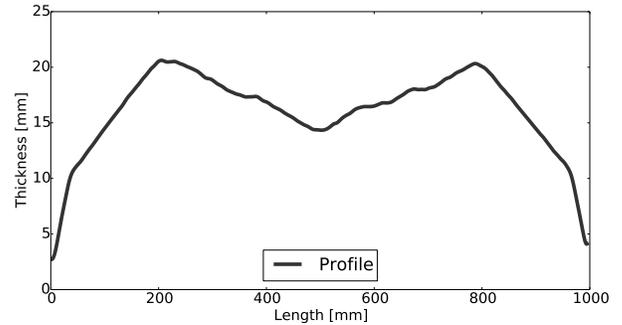


Figure 3. Acquired pantograph carbon strip profile. Note that the figure is not to scale: the carbon strip is a very wide asset.

of the approach, and Section 4 concludes the manuscript and reflects on its impact on the current maintenance plan.

2. METHODS AND RESULTS

This section describes process that has been followed in order to obtain a robust pantograph carbon strip prognosis method. Thus, the development is incremental and preliminary results are provided.

2.1. Carbon Strip Data Preprocessing

The carbon strip is a rectangular piece of carbon material that is mounted at the top of the pantograph. It is 20 mm thick, 30 mm wide and 1,000 mm long. Each pantograph equips two of these strips, and the leader always precedes the contact with the overhead line. Additionally, there are two cars on each train that equip a pantograph, although only one of them is active at a time (i.e., in contact with the catenary). Its rated operating voltage is 25kV AC.

The TrainScanner acquires a cloud of points for each pantograph carbon strip. Based on this data, the carbons are reconstructed with a triangulation technique, and a thickness profile is extracted for each asset, see Figure 3. It can be observed that the degraded area spans from 200 mm to 800 mm, and the most critical part is at the centre, from 400 mm to 600 mm. The system automatically identifies this region and extracts the minimum thickness value for further wear analysis.

This article evaluates the effectiveness of carbon strip prognostics with a dataset of thickness measurements at the fleet level, acquired between June 1 2016 and June 1 2017 at irregular intervals (the monitoring operations are not scheduled). It comprises an amount of 224 strip elements, and each sequence of carbon thickness needs to be preprocessed to add robustness to the prediction. To this end, the following issues are taken into account:

1. **Asset replacement:** steep positive thickness increments (greater than 5 mm) with a final value close to a new asset

measure, i.e., 20 mm, need to be segmented and treated as different assets.

2. **Acquisition failures:** extreme values out of strip range (over 20 mm) or zeroes are regarded as invalid data and thus they need to be discarded from the analysis by removing them from the carbon thickness sequence.
3. **Stability/Monotonicity:** each thickness segment needs to be asserted an overall monotonic negative trend according to the nature of the carbon material erosion, and a monotonic positive progression regarding the accumulated mileage. To this end, a monotonicity index is useful to quantify the amount of regularity in the evolution, which is based on the difference between the number of positive and negative increments (Davydov, Y., and Zitikis, R., 2017).
4. **Sensor precision:** TrainScanner's rated measurement precision is 0.5 mm. The prediction method needs to be robust to this inherent data acquisition system variability.

The resulting set of data should be smooth enough to be subject to further analysis following the ISO 13374 standard (ISO, 2003), which is the main PHM development guideline considered in this work, although similar structured approaches have also been developed for overhead monitoring systems (Brahimi, M., Medjaher, K., Leouatni, M., and Zerhouni, N., 2016). Obviously, the primary interest here is focused on the Prognosis module and the dynamic properties of the carbon strip degradation.

2.2. Rolling Window Prediction Evaluation

A rolling window is a prediction performance estimation procedure that is essentially based on the idea that “the past is used to predict the future”. It is an iterative process that frames a history window at some point in the evolution, learns the trend from it in order to make a prediction over a given horizon frame, and finally scores the error difference with the remaining coming data (Hota, H. S., Handa, R., and Shrivastava, A. K., 2007), see Figure 4.

Ultimately, the distribution of the resulting error score is used to estimate the performance of the prediction method, which is mainly driven by the amount of variability (Trilla, A., Dersin, P., and Cabré, X., 2018). To this end, the maximum deviation of the error distribution around its mean value is determined for a confidence interval of 95%. This quantity is here referred to as the “uncertainty”. Obviously, the error increases as the prediction horizon is extended into the future.

2.3. Robust Online Linear Regression

The Class 390 tilting Pendolino trains run a steady mission profile on the West Coast Main Line in the UK, featuring a very high availability (running 1,000 miles a day on average), which leads to expect a uniform degradation behaviour. In

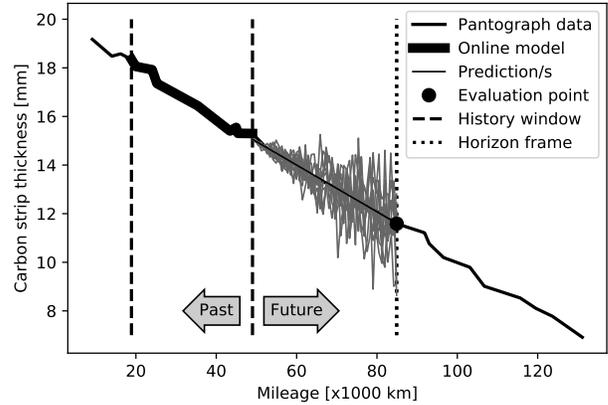


Figure 4. Diagram of the rolling window prediction evaluation for carbon strip data.

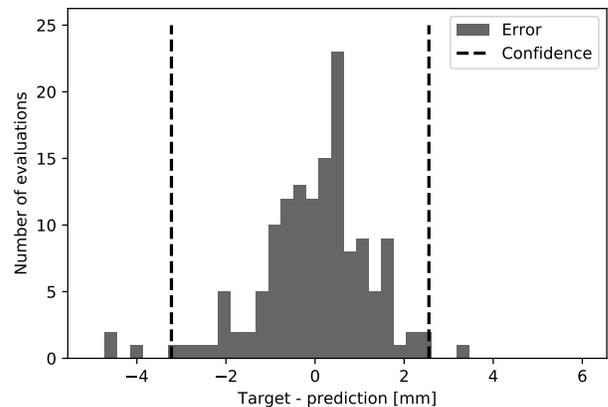


Figure 5. Histogram of the prediction error with robust online linear regression (ROLR). The 95% confidence interval indicates the uncertainty.

order to get a baseline for this study, the model linearity is assumed for the carbon strips in this high-speed rail scenario, following other carbon-based degradations like the brake pads (Trilla, A., Dersin, P., and Cabré, X., 2018). Therefore, a robust ordinary linear regression approach (ROLR) based on weighted least-squares fitting is evaluated. The regression is applied to each window of carbon thickness history after the aforementioned robust data-weighting process, and the prediction is obtained by extrapolating the evolution over the horizon frame. It is to note that the squared-error cost function of use here is very convenient to deal with the data-acquisition precision instability, which may be positive or negative. Finally, given the limited amount of data that is available at the sequence level, the history window is set to be equal to the prediction horizon, i.e., 30,000 km. Figure 5 shows the resulting distribution of this prediction error.

It can be seen that the linear method for the baseline shows an uncertainty of 2.89 mm. However, the resulting distribution shape is asymmetric because it displays a skewed centrality,

instead of the normal Gaussian distribution that would be expected with the least-squares optimisation procedure of use. This might be indicative that the linear assumption is not adequate and perhaps it needs to be questioned. The following sections, though, first delve into the particular bits of information that may be obtained from external context variables, and how they may be used to enhance the prediction.

2.4. Potential Improvement with Seasonal Context

One of the main extrinsic factors that may affect the degradation of the pantograph is the season. Variations of temperature (Ocoleanu, C. F., Popa, I., Manolea, G., Dolan, A. I., and Vlase, S., 2009), humidity, rain, wind... may cause an unsteady wear on the surface of the carbon material of the strip. It is well known that in the winter the contact wire freezes with the icing temperatures, possibly causing abnormal degradation. The spring, instead, is the driest period (although the rain is fairly well distributed throughout the year in the UK).

Further insight into these issues may be displayed through the seasonal wear rates, which grossly indicate the dynamic behaviour of the carbon degradation (i.e., the pace of the deterioration) due to these factors. In order to capture this indicator, the slope parameter of the linear regression on the strip thickness sequence is taken. Figure 6 shows the distribution of wear rates throughout the year using a Gaussian kernel density estimation procedure. It is to note that the winter and spring seasons are located on the extremes of the overall multimodal density. Winter shows the highest rates (over $12 \cdot 10^{-5}$ mm/km), whereas spring shows the lowest rates (under $5 \cdot 10^{-5}$ mm/km). Given that the prediction method of use here is linear (this may be interpreted as the derivative of the wear function), the extremely different error values related to these two sequential seasons prove that a non-linearity is inherently present as seasons gradually change. Therefore, this justifies the specific consideration of the seasonal factor as discrete context variables corresponding to the three modes of wear rate: winter, spring, and summer/autumn (note that their centrality conflates into the same value). The representation of the season as a nominal one-hot encoded vector (instead of a scalar ordinal encoding) is a convenient and effective solution with neural networks (Hancock, J. T., and Khoshgoftaar, T. M., 2020), the use of which is explored further in the following sections.

2.5. Data Blending through Neural Networks

In order to take advantage of the seasonal non-linear context variables discussed in Section 2.4, this section explores blending these different sources of information with a neural network ensemble. Regardless of the difficulty of the prediction task, the neural technique unifies the way of approaching this problem.

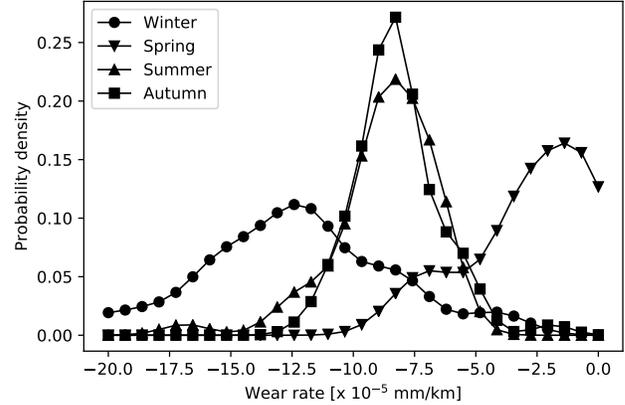


Figure 6. Density distribution of the wear rates according to season throughout the year.

2.5.1. Feature Ensemble with a Multilayer Perceptron

The Multilayer Perceptron is a general-purpose neural network architecture that can seamlessly integrate extrinsic data from different sources in order to refine a prediction (Trilla, A., Dersin, P., and Cabré, X., 2018). It is based on a feed-forward structure with a hidden layer in the middle, which provides the capacity to learn non-linear relationships between the inputs (i.e., the present features) and the output (i.e., the future thickness value). Moreover, its industrialisation is straightforward through a series of matrix multiplications that any platform can efficiently implement with a standard linear algebra library.

For the pantograph carbon strip scenario presented in this work, the baseline prediction result with linear regression is provided as a real-valued feature along with the rest of the aforementioned seasonal context variables (as binary flags with one-hot encoding). Moreover, the strip thickness value within the 30,000 km horizon is provided as the supervised output target prediction, see Figure 7. The hidden neurons are designed with a Rectified Linear Unit activation function to learn the non-linearities (Nair, V., and Hinton, G. E., 2010). The neural network is ultimately trained with a stochastic gradient descent protocol using backpropagation, an adaptive learning rate with momentum (Kingma, D. P., and Ba, J. L., 2015), and considering a squared-error cost function.

In order to get the network to learn effectively, its expressiveness (i.e., the capacity to represent the learnt knowledge) needs to match the complexity of the data within the objective prediction problem. To do so, the number of hidden units H needs to be adjusted because they modulate this learning ability. Note that the input dimensionality of this network is 4 (i.e., 3 context variables plus the result of the linear prediction), therefore, every hidden unit adds 6 new parameters to the model (4 inputs, 1 output, and 1 bias). In order to determine the optimum size of the hidden layer so that underfitting and overfitting learning problems may be

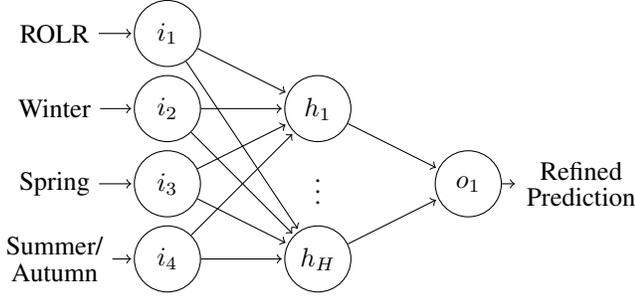


Figure 7. Multilayer Perceptron architecture blending the Robust Online Linear Regression (ROLR) with the set of three seasonal context variables to obtain a better refined prediction.

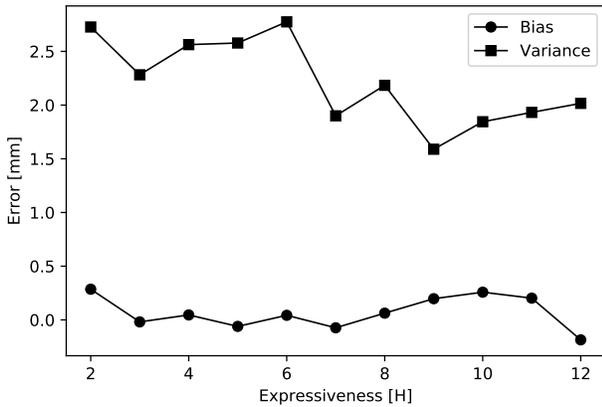


Figure 8. Expressiveness analysis of the neural feature ensemble. The bias represents the mode of the error distribution, and the variance represents its uncertainty.

avoided, a range of values are evaluated with Monte Carlo cross-validation (Dubitzky, W., Granzow, M., and Berrar, D., 2007), applying 10 rounds of repeated random sub-sampling with a train/test split of 95%/5%. This procedure yields over 70 evaluation points, which is a sufficient sample size to reliably estimate the prediction uncertainty. Figure 8 shows the results of this study through a bias/variance tradeoff analysis using the mode and the uncertainty values of the expected skewed error distributions, following customary descriptive statistics tools.

It can be seen that the most interesting performance score (i.e., the variance, or uncertainty) shows a randomly decreasing evolution as the expressiveness of the network grows (i.e., H increases), until the amount of hidden neurons reaches 9. From that point forward, the uncertainty rises, so the network stops generalising and begins to memorise the data, which is a sign of overfitting. Therefore, the optimum size for the hidden layer is of 9 units (it is to note that any residual bias can be corrected a posteriori with this estimation). It can be seen that the resulting system outperforms the previous linear approach as it now shows an uncertainty of 1.59 mm. This improve-

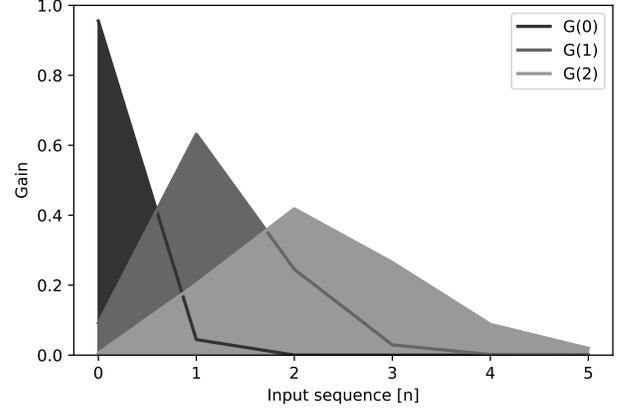


Figure 9. Impulse response of the spreading filters $G(s)$ (with $\alpha = 10$) for the time-delay convolution.

ment is mainly due to modelling the inherent non-linearities in the extrinsic seasonal context variables. Nevertheless, this result is still driven by the assumed linear evolution of the carbon thickness, which is a clear point of improvement that is explored in the next section.

2.5.2. Time-Delay Neural Network Embedding

This section builds upon the former feature ensemble approach, drops the questionable linearity assumption that drives the baseline prediction from Section 2.3, and proposes integrating the carbon thickness data directly through a neural structure known as a Time-Delay Neural Network (TDNN) (Peddinti, V., Povey, D., and Khudanpur, S., 2015). This approach maps the decreasing dynamic evolution of the data into a fixed spatial pattern using a weighted average operation in time with a set of spreading filters $G(s)$ defined by Eq. (1), where L is the size of the delay line (input data buffer), α is the spreading factor, and s is the spatial shift. Note that S is a normalisation factor that ensures that all shifts may deliver the same amount of energy, see Figure 9.

$$G_s = G(s) = \frac{1}{S} \sum_{n=0}^L x[n] \left(\frac{n+1}{s+1} \right)^\alpha \exp \left(-\alpha \frac{n}{s+1} \right) \quad (1)$$

In addition to empowering the system to deal with the thickness data evolution directly (i.e., an autoassociation that does not assume any specific behaviour, like the linearity), the convolution with the spreading filters exploits the local features of the data and reduces the searchable weight space for the learning stage. Furthermore, it increases the robustness to uneven sampling, which is to be taken into account as the inspections through the TrainScanner are not scheduled. This, in turn, enables the neural network that follows to handle sequences with different lengths, which is a clear limitation of

the ordinary multilayer perceptron (where the input dimensionality is fixed). Also, the use of variable history lengths may be of help to reduce the high uncertainty of strips showing a faster wear rate (Trilla, A., Dersin, P., and Cabré, X., 2018; Greitzer, F. L., and Ferryman, T. A., 2001).

The enhanced solution that this work suggests first builds the time-series embedding by applying the filters over the thickness sequence to obtain three spatial shifts (i.e., the high, middle, and low parts of the evolution). It uses the spreading factor α as a modulator to adjust the bandwidth of the filters to the length of any given sequence (applying the first filter $G(0)$ to the newest thickness sample to deal with a most unweighted value close to the prediction result), see Figure 9. And then, it assembles the resulting physical features with the former set of seasonal context variables that have proven to be useful in this modelling approach. Figure 10 shows this architecture.

At this point, the expressiveness of the new multilayer perceptron needs to be adjusted to the new embedded features following the cross-validation procedure described in Section 2.5.1. Now, each hidden unit adds 8 new parameters to the model. Figure 11 shows the result of this expressiveness analysis, which indicates that with 6 hidden neurons, the uncertainty of the prediction drops to 1.39 mm. Note that for this richer input representation (6 variables instead of 4), the model has become somewhat simpler (6 hidden units instead of 9), which makes perfect sense regarding the complexity tradeoff between the features and the predictive learning capacity.

3. DISCUSSION

This work exposes the gradual performance enhancement of pantograph carbon strip prognosis, initially relying on linear regression (resulting in 2.89 mm of uncertainty), then refining this prediction by accounting for non-linearities through the seasonal context information (1.59 mm), and finally dealing with the thickness evolution data directly with a set of spreading filters (1.39 mm). What is more, if these error results are assumed to belong to a normally distributed random variable, their incremental differences are statistically significant with a confidence interval of 95% using an Independent Samples t-test. In this case, the powerful Student hypothesis test with the Gaussian normality assumption is preferred over weaker non-parametric approaches like the Mann-Whitney U test, in spite of its apparent appropriateness to compare skewed distributions.

Despite the nice interpretability of the initial linear behaviour that emulates the prominent uniform physical degradation of this asset, every step taken toward dropping this linear assumption has led to increasingly better results in terms of prediction uncertainty. However, the resulting neural model has also increased its complexity, thus becoming more difficult to

interpret. Neural networks are typically regarded as “black boxes” because of their intricate nested inner functions.

In order to shed some light into the internal behaviour of the best-performing TDNN ensemble model, Figure 12 shows an input-standardised sensitivity analysis based on the profile method (Shojaeefard, M. H., Akbari, M. Tahani, M., and Farhani, F., 2013). It can be seen that there are three correlated patterns of behaviour:

- The three physical features (low, middle, and high-parts of the thickness sequence) show a rather linear increasing pattern along 8 mm of the whole output dynamic range. Their likelihood can be explained by their common source of information (i.e., the carbon evolution), which is already expected to be linearly uniform.
- The winter and spring seasons show a convex function (first negative, and then positive), with the inflection point around 0.5σ , and also covering 8 mm of the output dynamic range. These two seasons display the most extreme wear rates, see Figure 6, and the neural network seems to use them in a similar way for the refined predictions. In the end, it's in the transition from winter to spring that the main nonlinearity occurs.
- The summer/autumn seasonal variable exhibits a kind of offset rectifier function with the inflection point located at -0.5σ . This variable stretches up to 12 mm of the output dynamic range. It is to note that the associated wear rate applies to six months and it is represented by one single variable, thus maybe this explains its extended range.

While it may be difficult to assess the contribution of each variable in terms of importance, the amount of dynamic range in the output may be indicative of their rank, leaving the summer/autumn flag as the most critical variable. Further testing with an ablation study would be needed to derive stronger statements.

The current approach conducts a rough discretisation of the seasonal factor with three mutually-exclusive binary variables, but seasons change gradually, and the mid-season nuances are possibly missed with this solution. Nevertheless, conducting a seasonal information blending, e.g., at the month level, increases the number of extrinsic variables from three to twelve, and this in turn may enlarge the amount of weights in the neural network to an excess of expressiveness, increasing the potential risk of overfitting the data.

In addition to the principal seasonal information, other external sources of potential prognosable input have also been informally studied. On the one hand, there is the particular location of the pantograph. Each Class 390 train equips two pantographs, and the decision of using one or the other depends on the driver. This arbitrary factor may affect the degra-

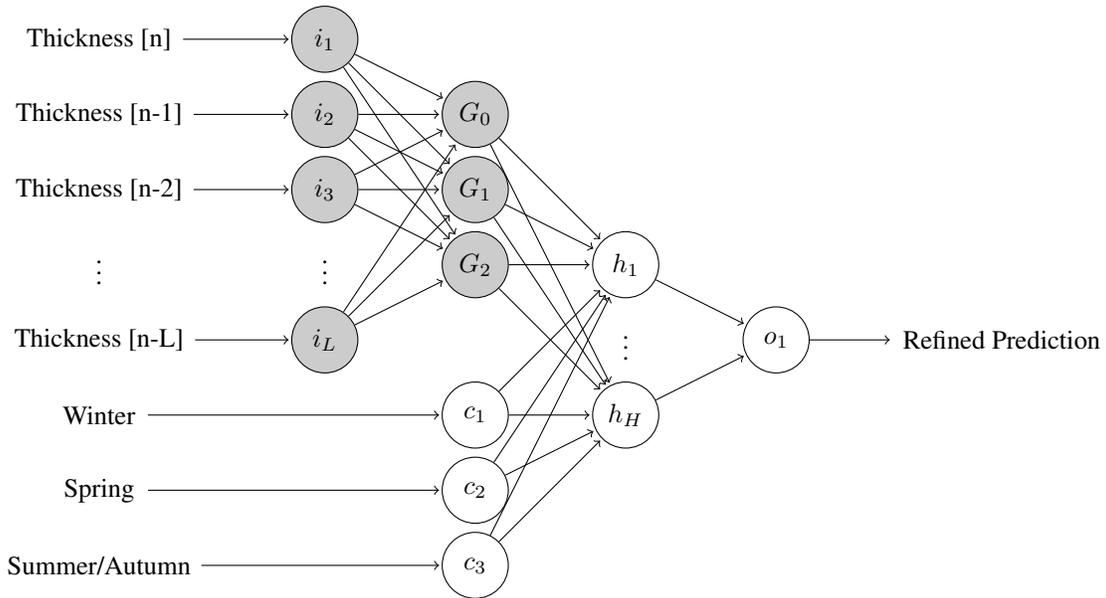


Figure 10. Architecture of the Time-Delay Neural Network ensemble. Strip thickness data is convoluted with the spreading filters (shown as shaded units), and blended into the set of context variables with a multilayer perceptron.

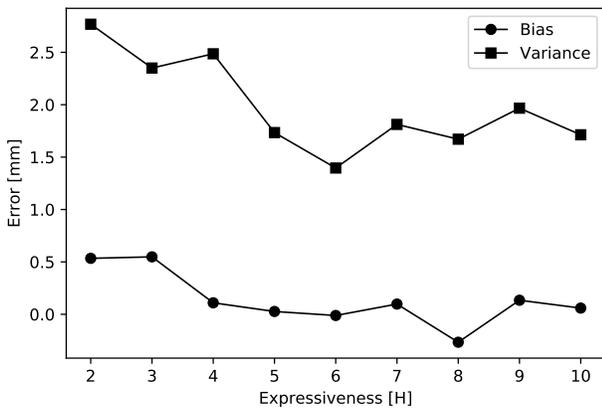


Figure 11. Expressiveness analysis of the TDNN ensemble. The bias represents the mode of the error distribution, and the variance represents its uncertainty.

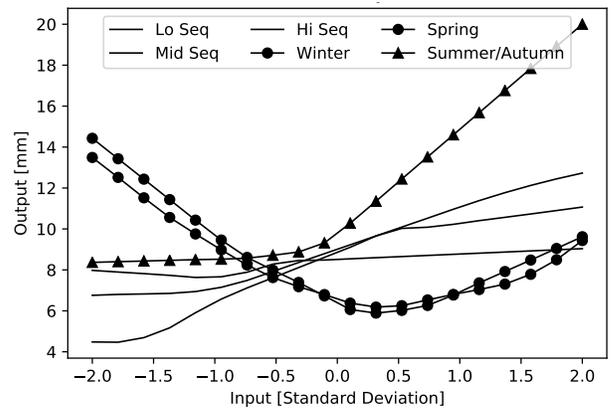


Figure 12. Sensitivity analysis of the best-performing TDNN ensemble ($H=6$). The dynamic range of the inputs is normalised to their standard deviation. The variables that display the same pattern share the same line style.

dation of the carbon strips, although particular behaviours seem unlikely to be displayed because driver rota is the common way of operating the rolling stock.

On the other hand, there is the position of the carbon strip in the pantograph. Depending on the sense of the trip (upwards to Scotland, or downwards to England), different strips lead the contact with the catenary. But again, it's the driver's decision to use one pantograph or the other, so for the same rotation reason, a singular behaviour is unlikely to show. In

the end, neither the pantograph location nor the strip position have proven to be of much use in the prognostication of future carbon strip thickness.

4. CONCLUSION

At present, the carbon strip replacement criterion for the Class 390 pantographs is based on a single thickness threshold value. This inefficient approach does not take into account

the rate of wear that the different strips display, which varies significantly throughout the year with the seasons. Thus, the same thickness value can lead to different operating mileages before the asset reaches its actual end of life (i.e., when there is no carbon material left on the strip).

This article presents the most sophisticated technique for TrainScanner pantograph carbon strip prognostics, which is based on a Time-Delay Neural Network that blends a spread sequence of carbon thickness values with the seasonal context information. This approach yields a prediction error uncertainty around 1.39 mm at the asset level and for a projected horizon of 30,000 km, which is related to the planning time that is necessary for scheduling the maintenance resources at the depot. Therefore, if the expected mileage to the next visit is under this distance frame, the strip threshold scrap limit could be safely extended up to this performance value.

The future work that is currently envisaged may further deal with other extrinsic context variables in order to add more robustness to the prognosis method. The neural network has proven to be a very versatile approach for assembling different data sources. In this regard, we may exploit the temporal persistence of large amounts of other nominal (i.e., non-parametric) data provided by related onboard subsystems (Hu, X., Eklund, N., and Goebel, K., 2007), e.g., from traction. Alternatively, we also expect to explore other sequence learning approaches through the Long Short-Term Memory units (Hochreiter, S., and Schmidhuber, J., 1997), and seek the complementary characteristics that may help the current approach attain a better effectiveness.

ACKNOWLEDGMENT

We would like to show our gratitude to Professor Xavier Vilasís-Cardona for his valuable support and comments that greatly improved the manuscript.

REFERENCES

- Atamuradov, V., Medjaher, K., Dersin, P., Lamoureux, B., and Zerhouni, N. (2017). Prognostics and Health Management for Maintenance Practitioners - Review, Implementation and Tools Evaluation. *International Journal of Prognostics and Health Management*, 8(60), 1–31.
- Auditeau, G. (2016). Effect on wear of contact strips and wire of materials used for pantograph sliding strips (COSTRIM project). *Proc. of the World Congress on Railway Research*.
- Auditeau, G., Bucca, G., Collina, A., and Tanzi, E. (2011). Experimental analysis of effect of plain carbon and impregnated carbon contact strips on contact wire wear. *Proc. of the World Congress on Railway Research*.
- Brahimi, M., Medjaher, K., Leouatni, M., and Zerhouni, N. (2016). Critical Components Selection for a Prognostics and Health Management System Design: an Application to an Overhead Contact System. *Annual Conference of the Prognostics and Health Management Society (ISBN: 978-1-936263-22-6)*, 7, 1–8.
- Bucca, G., and Collina, A. (2015). Electromechanical interaction between carbon-based pantograph strip and copper contact wire: A heuristic wear model. *Tribology International (ISSN: 0301-679X)*, 92, 47–56.
- Davydov, Y., and Zitikis, R. (2017). Quantifying non-monotonicity of functions and the lack of positivity in signed measures. *Modern Stochastics: Theory and Applications*, 4(3), 219–231.
- Ding, T., Xuan, W., He, Q., Wu, H., and Xiong, W. (2014). Study on Friction and Wear Properties of Pantograph Strip/Copper Contact Wire for High-Speed Train. *The Open Mechanical Engineering Journal (ISSN: 1874-155X)*, 8, 125–128.
- Dubitzky, W., Granzow, M., and Berrar, D. (2007). Fundamentals of data mining in genomics and proteomics. *Springer Science & Business Media*, 178.
- Greitzer, F. L., and Ferryman, T. A. (2001). Predicting Remaining Life of Mechanical Systems. In *Intelligent Ship Symposium IV*.
- Hancock, J. T., and Khoshgoftaar, T. M. (2020). Survey on categorical data for neural networks. *Journal of Big Data*, 7(28), 1–41.
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(1), 1735–1780.
- Hota, H. S., Handa, R., and Shrivastava, A. K. (2007). Time Series Data Prediction Using Sliding Window Based RBF Neural Network. *International Journal of Computational Intelligence Research*, 13(5), 1145–1156.
- Hu, X., Eklund, N., and Goebel, K. (2007). A Data Fusion Approach for Aircraft Engine Fault Diagnostics. *Proc. of ASME Turbo Expo*, 1(GT2007-27941), 767–775.
- ISO. (2003). *Condition monitoring and diagnostics of machine systems: Data processing, communication and presentation* (Tech. Rep. No. 13374-1:2003). International Organization for Standardization.
- Kingma, D. P., and Ba, J. L. (2015). Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs.LG]*, 1–15.
- Nair, V., and Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *Proc. of the 27th International Conference on Machine Learning*, 1–8.
- Ocoleanu, C. F., Popa, I., Manolea, G., Dolan, A. I., and Vlase, S. (2009). Temperature investigation in contact pantograph AC contact line. *International Journal of Circuits, Systems and Signal Processing (ISSN: 1998-4464)*, 3(3), 154–163.
- Peddinti, V., Povey, D., and Khudanpur, S. (2015). A Time Delay Neural Network Architecture for Efficient Mod-

eling of Long Temporal Contexts. In *Interspeech* (pp. 3214–3218).

- Shing, A. W. C., and Wong, P. P. L. (2008). Wear of pantograph collector strips. *Proc. of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* (ISSN: 2041-3017), 222(2), 169–176.
- Shojaeefard, M. H., Akbari, M. Tahani, M., and Farhani, F. (2013). Sensitivity Analysis of the Artificial Neural Network Outputs in Friction Stir Lap Joining of Aluminum to Brass. *Advances in Materials Science and Engineering*, 1–7.
- Trilla, A., and Cabré, X. (2018). Determining the Equivalent Conicity for Railway Wheelset Maintenance with Deep Ensembles. *Proc. of the Annual Conference of the Prognostics and Health Management Society*, 10(1), 1–6.
- Trilla, A., Dersin, P., and Cabré, X. (2018). Estimating the Uncertainty of Brake Pad Prognostics for High-speed Rail with a Neural Network Feature Ensemble. *Proc. of the Annual Conference of the Prognostics and Health Management Society*, 10(1), 1–7.
- Tsui, K. L., Chen, N., Zhou, Q., Hai, Y., and Wang, W. (2015). Prognostics and Health Management: A Review on Data Driven Approaches. *Mathematical Problems in Engineering*, 2015(793161), 1–18.

BIOGRAPHIES

Alexandre Trilla graduated from La Salle University of Barcelona with a M.Sc. in Electronics and Telecommunications Engineering in 2008, and a M.Sc. in IT Management in 2010. He has an academic research background in spoken language processing, and an industrial research background in PHM. He has authored several publications in scientific conferences and journals (IEEE Transactions on Audio, Speech, and Language Processing, Chemical Engineering Transactions, and the Journal of Rail and Rapid Transit). At present, he is a Senior Data Scientist and R&D Program Manager at Alstom, working on the deployment of PHM to the railway environment. He leads the development of predictive maintenance based on Machine Learning, and he is especially interested in the solutions with artificial neural networks.

Verónica Fernández graduated from Universitat Autònoma of Barcelona with a B.Sc. in Chemical Engineering in 2016, and a M.Sc. in Automatic Systems and Industrial Electronics Engineering from Universitat Politècnica of Catalonia in 2019. At present, she is a Data Scientist working on PHM in the railway environment.

Xavier Cabré was born in Barcelona. He graduated from La Salle University of Barcelona with a M.Sc. in Electronics and Telecommunications Engineering in 1998. He has an industrial research background in high-technology environments. In 2015, he co-authored a patent in image processing and computer vision for the detection of railway components. At present he is the Project/Program manager of TrainScanner.