Reduction of Power Consumption in Sensor Network Applications using Machine Learning Techniques

G M Shafiullah[1], Adam Thompson[2], Peter J Wolfs[2], Shawkat Ali[3]
[1] Centre for Railway Engineering, Faculty of Sciences, Engineering & Health,
[2] College of Engineering & Built Environment, Faculty of Sciences, Engineering & Health
[3]School of Computing Sciences, Faculty of Business & Informatics
Central Queensland University, Rockhampton, QLD-4702, Australia
Phone: +61749309313, Email: g.shafiullah@cqu.edu.au

Abstract— Wireless sensor networking (WSN) and modern machine learning techniques have encouraged interest in the development of vehicle monitoring systems that ensure safe and secure operations of the rail vehicle. To make an energy-efficient WSN application, power consumption due to raw data collection and pre-processing needs to be kept to a minimum level. In this paper, an energy-efficient data acquisition method has investigated for WSN applications using modern machine learning techniques. In an existing system, four sensor nodes were placed in each railway wagon to collect data to develop a monitoring system for railways. In this system, three sensor nodes were placed in each wagon to collect the same data using popular regression algorithms, which reduces power consumption of the system. This study was conducted using six different regression algorithms with five different datasets. Finally the best suitable algorithm have suggested based on the performance metrics of the algorithms that include: correlation coefficient, root mean square error (RMSE), mean absolute error (MAE), root relative squared error (RRSE), relative absolute error (RAE) and computation complexity.

Key Words – Wireless sensor networking; machine learning techniques; railway wagons; regression analysis

I. INTRODUCTION

Recent emergence of micro-electro-mechanical systems (MEMS) technology, wireless communications and integrated circuit design have enabled the development of low-cost, low-power, multipurpose sensor networks. These low-power sensor networks provide a new monitoring and control capability in the architectural infrastructure, vehicle infrastructure, environmental management, and safety and security systems. Sensor network applications require long lifetimes, data accuracy, and energy efficiency. Energy efficiency is the major concern issue to design an efficient WSN application [1-3].

With the increased demand for railway services, railway monitoring systems continue to advance at a remarkable pace to maintain reliable, safe and secure operations. If a security-related incident has occurred, a monitoring system may support the operator in taking the appropriate action, communicating to the right authorities, checking the availability of rescue teams and providing all necessary information. Typical dynamic behaviors of railway wagons are responsible for safe, cost-effective and reliable operations of freight railways. The performance of rail vehicles running on tracks is limited by the lateral instability inherent to the design of the wagon's steering and the response of the railway wagon to individual or combined irregularities. Railway track irregularities need to be kept within safe operating margins by undertaking appropriate maintenance programs [4-6].

Predicting vehicle characteristics online from track measurement data has been addressed in various studies [6-12]. Wireless sensor networks are widely used to monitor railway tracks and irregularities, detect abandoned objects in railway stations and develop intrusion detection systems, secure railway operations, monitor tunnels [13-15]. Machine learning techniques have been introduced in different research projects to predict the typical dynamic behavior of railway wagons running on the track [16-20]. Raw data collection, data pre-processing, and formatting are essential parts of developing any monitoring systems including the above mentioned research works.

Matthias Seifert envisages [13] that a network of smart sensors will be used as a means to monitor public spaces for potential intrusions and accordingly alert the operators at a control centre about the incident. The added advantage of WSN is to monitor large areas with greater efficiency in video-based intrusion detection systems. Aboelela et al. [14], introduces a new approach to reduce the occurrence rate of accidents and improve the efficiency of railroad maintenance activities by developing a system based on WSN.

Central Queensland University (CQU), in association with the Centre for Railway Engineering (CRE) [12], has been investigating a Health Card device for railways. This Health Card system is an autonomous device used for analysis of car body motion signals that can detect track conditions and monitor derailment conditions. The Health Card is capable of resolving car body motions into six degrees of freedom. To do this the Health Card uses accelerometers and angular rate sensors with a coordinate transform. Two prototypes have been developed based on wired and wireless solutions. The Health Card system uses fast Fourier transforms (FFT) to efficiently convert the signal into a time-frequency spectrograph so that events can be detected according to their

short-term spectral content. From spectral analysis, it has been found that small residual responses exist in the pitch and yaw degrees of freedom and the wagon was not laterally constrained [6, 12].

Cen et al. [17] investigated a machine learning approach to automate the identification process of railroad wheels using collected data from wheel inspections. Decision trees and SVM based classification scheme has used to analyze the railroad wheel inspection data. With tenfold cross validation, C4.5 algorithm achieved an average classification accuracy of 76.2 percent with extracted decision rules and 75.5 percent with the pruned decision tree, while with SVM algorithm 76.589 percent accuracy was obtained. Cen et al. [17] introduced Bagging classification ensemble approach specially for imbalanced data which boosted the prediction accuracy to 81 percent. The experimental results prove that the proposed approach is very efficient, producing a classifier ensemble that has high sensitivity, specificity and gMeans values during classification [17-18].

Marco et al. [20] have introduced a data set extracted from a real-life vehicle tracking sensor network using popular classification algorithms. This data set has extracted based on the sensor data collected during a real world wireless distributed sensor network (WSDN) experiment carried out at Twenty-nine Palms, CA. The WSDN vehicle classification problem comprises with local classification and global decision fusion. Maximum Likelihood, *k*-Nearest Neighbor, and Support Vector Machine algorithms were used in this experiment. It has been seen that although the classification rates for the available modalities are only acceptable, methods used in multisensor networks such as data fusion will enhance the performance of these tasks.

In this study, we have developed prediction models using popular regression algorithms to reduce the power consumption of an existing railway monitoring technique, developed by Central Queensland University. We have developed models with six popular regression algorithms and applied them to a unified platform. We have assessed the performance of different models and proposed the most suitable algorithm. This paper is organized as follows: Section II discusses the background to the study. Section III presents an overview of the regression algorithms. The development of the model with different algorithms is discussed in Section IV. Results and analyses are described in Section V. Section VI concludes the article with future directions.

II. BACKGROUND OF THE STUDY

The "Health Card" system developed by a team of Engineers at Central Queensland University [6, 12] aims to monitor every wagon in the fleet using low cost intelligent devices. Solid-state transducers including accelerometers and angular rate sensors with a coordinate transform to resolve car body motions into six degrees of freedom was used in Health Card. Popular spectrogram techniques were used to obtain a time-frequency representation of the car body motion. An algorithm was developed to analyse signals from accelerometers mounted on the wagon body, to identify the dynamic interaction of the track and the rail vehicle. The algorithm has validated using collected field data including accelerations measured at strategic points on the wagon body and the bogies.

A set of four prototypes "Health Cards" [12] has been developed by a team at Central Queensland University. Each prototype "Health Cards" incorporates a 27 MHz microcontroller with 256kB of onboard RAM, four dual-axis accelerometers, a GPS receiver, two low power radios, lithium ion batteries and a solar panel. Data was collected from a ballast wagon which was a conventional three piece bogie, spaced l_b = 10.97m apart. Dual axis accelerometers were fitted to each corner of the body and each side frame. The accelerometers were spaced l = 14.4m apart. The test run was a normal ballast lying operation, starting with a full load of ballast, traveling to the maintenance site, dropping the ballast on the track, and returning empty via the same route. A PC based data acquisition system was used to store data. The main purpose of the data acquisition was to provide real data that represented to the Health Card device. Data was to be used to validate and demonstrate the effectiveness of signal analysis techniques and finally develop a model to monitor typical dynamic behavior and track irregularities [6, 12].

Steven et al. [6, 12] placed dual-axis accelerometers at each corner of the body and each side frame. All of the axes measured in the vertical and lateral conditions. The aim of the sensing arrangement was to capture roll, pitch, yaw, vertical and lateral accelerations of the wagon body. The ADXL202/10 dual-axis acceleration sensor measured 16 channel acceleration data in *g* units, with 8 channels for the wagon body and 8 for the wagon side frame. Four sensor nodes were placed in each wagon body and locations of sensors in the wagon body are front left body, front right body, rear left body and rear right body.

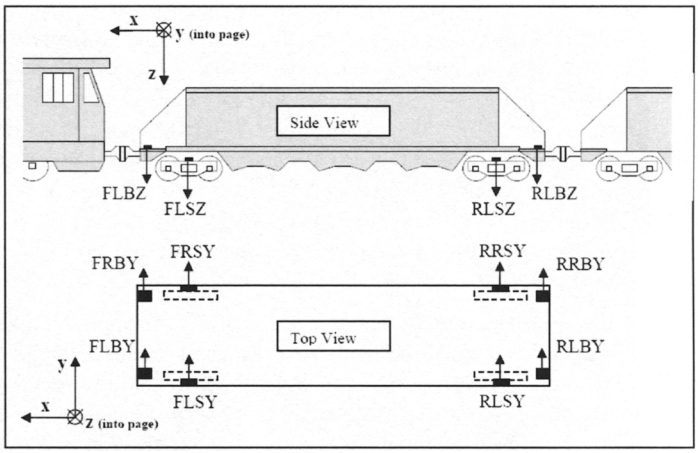

Fig. 1: Accelerometer locations and Axis naming convention [7]

Data collected from these four sensors are front left body vertical (FLBZ), front left body lateral (FLBY), front right body vertical (FRBZ), front right body lateral (FRBY), rear left body vertical (RLBZ), rear left body lateral (RLBY), rear right body vertical (RRBZ), rear right body lateral (RRBY). Sensor locations and naming convention are illustrated in Figure 1.

In this paper, we have introduced a field data acquisition method for the wagon body using popular machine learning techniques which is more energy-efficient than the existing data acquisition method. Dual axis accelerometers placed on each corner of the wagon body measured vertical and lateral condition data of that individual location. Prediction models have been developed using the collected data. The model predicted the vertical and lateral conditions of the fourth sensor nodes', i.e., sensor node located at the rear right corner of the wagon body. The prediction model replaces the use of fourth sensor nodes in the rear right corner of the wagon body. This prediction method reduces the use of one sensor node in each wagon which reduces power consumption of the application significantly.

III. REGRESSION ALGORITHMS

Regression analysis is the most significant and popular learning area for future decision making or forecasting of data. Researchers already have introduced different types of regression algorithms, including popular regression analysis for time series data forecasting, tree based algorithm, rule-based learning, lazy learning, multilayer perception, and statistical learning [21-28]. Currently various statistical forecasting and regression approaches are used to monitor railway wagons to ensure safety and security. This section describes the popular regression algorithms that used to develop an energy-efficient model for sensor network applications. We have considered Tree-based learning reduced error pruning tree (REPTree) and M5Prime, Lazy-based learning IBK, Regression-based learning linear regression, Statistical learning based algorithm support vector machine (SVM) regression, and Neural Network based multilayer perception (MLP).

REPTree: REPTree is a fast regression tree that uses information gain/variance reduction and prunes it using reduced-error pruning. REPTree deals with missing values by splitting instances into pieces. Optimized for speed it only sorts values for numeric attributes once [21].

M5 Prime: M5 Prime is useful for numeric prediction. It is a rational reconstruction of Quinlan's M5 model tree inducer. Decision trees were designed for assigning nominal categories. M5 Prime extended decision trees by adding numeric prediction by modifying the leaf nodes of the tree [22, 23].

IBK: Instance-based learning algorithms are derived from the nearest neighbor machine learning philosophy. IBK is an implementation of the k-nearest neighbor's algorithm. The number of nearest neighbors (k) can be set manually, or determined automatically. Each unseen instance is always compared with existing ones using a distance metric. WEKA's default setting is k = 1 [21, 24].

Linear Regression: Regression analysis [25-26] is a statistical forecasting model that addresses and evaluates the relationship between a given variable (dependent) and one or more independent variables. The major goal in regression analysis is to create a mathematical model that can be used to predict the values of a dependent variable based upon the values of an independent variable. Regression algorithm does this by finding the line that minimizes the sum of the squares of the vertical distances of the points from the line. The goodness of fit and the statistical significance of the estimated parameters are a matrix of regression analysis.

SVM Regression: SVM is a statistical based learning, which has been used for binary classification for the first time. SVM model can usually be expressed in terms of a support vector and applied to nonlinear problems using different kernel functions. Based on the support vector's information, SVM regression produces the final output function. WEKA by default considers sequential minimal optimization (SMO) for SVM and polynomial kernel with degree 1[21, 27].

Multilayer Perception: A multilayer neural network (NN) consists of three layers: input, hidden and output. After receiving an input pattern, the NN based architecture passes the signal through the network to predict the output in the output layer. Output compares with actual value and calculated error to modify the weights. WEKA uses the back propagation (BP) algorithm to train the model, though it is slower than a few other learning techniques [27-28].

To evaluate the prediction accuracy of the above mentioned algorithms for the data, percentage split test options were used. Prediction metrics considered in this study are given below with their mathematical expression [21]:

Correlation Coefficient (CC)	$CC = \frac{\frac{1}{n-1}\sum_i (Y_i - \bar{Y}_i)(Y^*_i - \bar{Y}^*_i)}{\sigma_{Y_i}\sigma_{Y^*_i}}$ where Y_i is the observation value and Y_i^* is the predicted value. σ_{Y_i} and $\sigma_{Y^*_i}$ are the standard deviation for Y_i and Y_i^*
Mean Absolute Error (MAE)	$MAE = \frac{1}{n}\sum_{i=1}^{n} \lvert Y_i - Y_i^* \rvert$
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Y_i - Y_i^*)^2}$
Relative Absolute Error (RAE) in %	$RAE = \left\lvert \frac{Y_i - Y_i^*}{Y_i^* - \bar{Y}^*} \right\rvert x100$
Root Relative Squared Error (RRSE) in %	$RMSE = \sqrt{\sum_{i=1}^{n} \left(\frac{Y_i - Y_i^*}{Y_i^* - \bar{Y}^*} \right)^2} x100$ where $\bar{Y}^* = \frac{1}{n}\sum_i Y^*$

IV. EXPERIMENTAL SETUP

In our experiments we have used five data sets from the collected data in [6]. We have predicted the data of the sensor node located in the rear right corner of the wagon body. As each sensor node collected both lateral and vertical condition data so we have predicted both the rear right vertical (RRBZ) and rear right lateral (RRBY) conditions using the collected data in [6]. For this experiment we have used REPTree, IBK, and M5Prime, linear regression, SVM and MLP regression algorithms. Correlation coefficient, RMSE, MAE, RRSE, RAE and computation complexity has been measured to evaluate the prediction accuracy. Percentage split test options were considered to evaluate the datasets for each of the algorithms. We have used 90 percent data for training and the remaining 10 percent for testing. The computational complexity includes both the model train period and the test set evaluation time. Few of the algorithms need more time to classify the test set than training the model. First we have developed the model with the stated five learning algorithms to predict rear right body vertical (RRBZ) condition. Later we have developed models to forecast rear right body lateral (RRBY) condition. Finally we have calculated the relative weighted performance for a given algorithm based on correlation coefficient, RMSE, MAE, RRSE, RAE and computational complexity and proposed the best suitable algorithm for data acquisition method. With the help of this system we need to place three sensors in each wagon instead of four sensor nodes in each wagon in the existing system. This prediction model reduces the use of one sensor node in each wagon hence, reduces power consumption of the system.

After necessary pre-processing and formatting we have passed the data into the learning algorithms to predict rear right body vertical and lateral conditions of railway wagons. For initial data pre-processing, and formatting we have used MATLAB [29] and WEKA [30] learning tools. WEKA includes a comprehensive set of data pre-processing tools, learning algorithms and evaluation methods, graphical user interfaces and environment for comparing learning algorithms [31]. With the help of WEKA [30] learning tools we have developed six models using the above stated learning techniques to predict RRBZ and RRBY of the rear right corner of the wagon. For our experiments we have used a unified platform. The configuration of the PC used in the experiments was Pentium IV, 3.0 GHz Processor, 1GB RAM. We have used WEKA release 3.5.7 for all of the experiments. A stop watch has been used to count computational time. Experiments have demonstrated that different algorithms predicted the value with minor to negligible errors. Computation complexity also differs with the learning techniques.

We have calculated the ranking performance for a given algorithm based on correlation coefficient, RMSE, MAE, RRSE, and RAE. The best performing algorithm on each of these measures is assigned the rank of 1 and the worst is 0. Thus, the rank of the *j*th algorithm on the *i*th dataset is calculated as stated in [27]:

$$R_{ij} = 1 - \frac{e_{ij} - \max(e_i)}{\min(e_i) - \max(e_i)} \qquad (1)$$

where e_{ij} is the percentage of correct classification for the *j*th algorithm on dataset *i*, and e_i is a vector accuracy for dataset *i*. A detailed comparison of algorithm performance can be evaluated from this equation.

We have evaluated the performance of all the regression algorithms using the total number of best and worst performances. The total number of the best and worst ranking for correlation coefficient, RMSE, MAE, RRSE, RAE and computational complexity for all the classifiers are evaluated by using the following equation:

$$C_i = \frac{1}{\rho}\left(\frac{s_i - f_i}{n}\right) + \frac{1}{\rho} \qquad (2)$$

where $\rho = 2$ is the weight shifting parameter, s_i is the total number of success or best cases for the *i*th classifier, f_i is the total number of failure or worst cases for the same classifier, and n is the total number of datasets.

Finally, we have measured the relative weighted performance for all the classifiers with two different weights for ranking average accuracy and computational complexity using the following equation:

$$Z = \alpha a_i + \beta t_i \qquad (3)$$

α and β are the weight parameters for ranking average accuracy against computational complexity. The average accuracy and computational complexity are denoted by a_i and t_i. By changing the values of β we have observed the effect of the relative importance of accuracy and computational complexity.

From detailed analysis of the results we have proposed the best suitable learning technique that reduces power consumption of the application significantly.

V. RESULTS AND ANALYSIS

Proposed algorithms with percentage split test options were used to predict the rear right body condition of a railway ballast wagon. We have used five sets of data in different instances, i.e. different times and locations. To cover a large experimental area, data sets were selected both from loaded and unloaded conditions and the number of data records also varies. Initially we have developed models to predict rear right body wagon conditions (both lateral and vertical) for five data sets with the six selected regression algorithms. We have measured *correlation coefficient, RMSE, MAE, RRSE, RAE and computation complexity* for each algorithm. We have run our models using the WEKA learning tools [30-31].

From initial experiments it has been observed that accuracy of the above mentioned metrics varies based on algorithms, data quality and number of records. From the experimental results it is very difficult to come to a conclusion and decide a best suitable algorithm to predict rear right body wagon conditions. Therefore, we have calculated the ranking

performance, classifier performance and computational complexity as stated in [27].

Table 1: Ranked algorithm performance based on correlation coefficient for the six algorithms

Algorithm	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5
IBK	1.0	1.0	1.0	1.0	0.0
REPTree	0.0	0.42212	0.22889	0.75505	1.0
Multilayer Perception	0.62318	0.47068	0.07833	0.09888	0.69632
SVM Regression	0.65098	0.01653	0.00610	0.00376	0.65406
M5Prime	0.65336	0.32327	0.06815	0.64754	0.67515
Linear Regression	0.65761	0.0	0.0	0	0.64293

Table 2: Ranking average across test set classification problems based on different performance metrics

Algorithm	IBK	REP Tree	MLP	SVM Reg.	M5P	Linear Reg.
Correlation Coefficient	0.8	0.5	0.5	0.5	0.5	0.2
MAE	0.1	0.4	0.8	0.5	0.5	0.7
RMSE	0.6	0.5	0.6	0.5	0.3	0.5
RAE	0.2	0.4	0.8	0.5	0.5	0.6
RRSE	0.4	0.5	0.5	0.5	0.4	0.7

Initially we have calculated ranking performance of all the above stated metrics using equation-1. Table 1 represents the ranked performance of correlation coefficient. The best performing algorithm on each of these measures is assigned the rank of 1 and the worst is 0. Then classifier performance have calculated from the total number of best (1.0) and worst (0.0) rankings for correlation *coefficient, RMSE, MAE, RRSE, RAE* and *computation complexity* using equation (2). Classifier performances for all of the algorithms are given in Table 2. We observed that for correlation coefficient measure IBK was the best performing algorithm, while it was the worst to measure MAE. For MAE and RAE measurement MLP was the best performing algorithm. Linear regression is the second choice to measure RAE and best performing to measure RRSE. Both IBK and MLP are the first choice to measure RMSE. Based on various accuracy measures it is observed that MLP is the best choice. Fig. 2 represents the performance of different algorithms to predict rear right body wagon condition.

Finally we have calculated relative weighted performance using equation (3), assuming $\alpha=1$ and β is from 0.4 to 2. Average classification accuracy of the classifiers was very close to each other, however, MLP was the best and IBK was the worst. With respect to computational time SVM was the worst algorithm. Considering computational complexity and average accuracy, linear regression was the best choice and SVM performed worst to predict rear right body wagon condition. Fig. 3 represents the overall weighted performances of the selected algorithms.

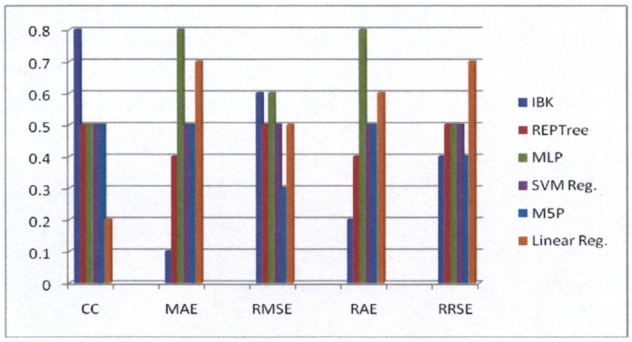

Fig.2: Algorithm performance for the six algorthms

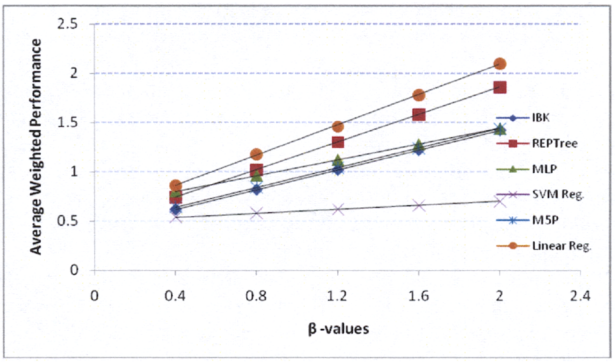

Fig 3.: Overall performance of the algorithms with respect to β, assuming $\alpha=1$

VI. CONCLUSION

Intelligent machine learning techniques play a key role in developing monitoring systems for both freight and passenger railway systems. In this paper, an energy-efficient data acquisition method for WSN applications has been investigating using popular regression algorithms. In an existing method four accelerometers were placed in each wagon body to collect necessary data to monitor typical dynamic behavior of railway wagons. Same data have collected in this study by placing three sensor nodes in each wagon body with the help of popular regression algorithms. A prediction model has developed to predict rear right wagon body lateral and vertical conditions. Ranking performance, average accuracy and average weighted performance has also evaluated to select a suitable algorithm for this application. From different analyses the experimental results showed that no individual algorithm performs best for all performance metrics. Considering average accuracy and computational complexity linear regression algorithm was the best suited

algorithm to predict rear right wagon body conditions. However, MLP was the most suitable if only average accuracy of performance metrics were considered. This data acquisition method reduces power consumption of the existing application significantly as it reduces use of one sensor node in each wagon. This also reduces computational complexity, development and maintenance cost both in hardware and human inspection.

This is the first time that modern machine learning techniques have been used in this context, specially in railway communication which still requires verification in different areas. Therefore, it deserves further investigation that focuses on some specific areas which are:

- introduce bagging techniques to improve the performance of the model
- investigate lateral acceleration of rail wagons
- predict front end rail wagon behavior from rear wagon collected data.

REFERENCES

[1] G. M. Shafiullah, A. Gyasi-Agyei, and P. Wolfs, "Survey of wireless communications applications in the railway industry," in *Conf. on Wireless Broadband and Ultra Wideband Comm.*, Sydney, Aug'07.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey of Sensor Networks." *IEEE Communications Magazine*, vol. 40, August, 2002, pp.102--114

[3] D. S. Sokwoo Rhee and S. Liu, "Techniques for minimizing power consumption in low data-rate wireless sensor networks." *Wireless Communications and Networking Conference*, vol. 3, March 2004, pp. 1727-1731.

[4] J. Smith, S. Russel, and M Looi, "Security as a safety issue in rail communications," in *8th Aus. Workshop on Safety Critical System and Software (SCS'03)*, Canberra, Australia, 2003, pp. 79-88.

[5] V. K. Garg and R. V. Dukkipati, *Dynamics of railway vehicle systems*. Academic Press, 1984.

[6] Steven S. Bleakley, "Time Frequency analysis of railway wagon body accelerations for a low-power autonomous device," Master's thesis, Faculty of Engineering and Physical Systems, Central Queensland University, Australia, October 2006.

[7] C. S. Bonaventura, J. W. Palese, and A. M. Zarembski, "Intelligent system for real-time prediction of railway vehicle response to the interaction with track geometry," in *Railroad Conf. 2000. Proc. of the 2000 ASME/IEEE Joint*, 2000, pp. 31-45.

[8] M. M., M. Dhanasekar, D. Skerman and J. Davey, " Monitoring the dynamics of freight wagons," in *Conf. on Railway Engineering, CORE 2002*, Wollongong, Australia, 2002, pp. 213-221.

[9] C. Cole and D. Roach, "Dynamic response of coal wagons during normal operation," in *1996 World Conf. on Rail Research*, Colorado, 1996, pp. 219-231.

[10] A. P., A. S., and Nejikovsky, "Remote ride quality monitoring and acle train set performance," in Proceedings of the 2002 ASME/IEEE Join Rail Conf., Washington DC, April'02, pp. 171-178.

[11] Y. Sato, and M. Miwa, "Measurement and analysis of track irregularity on super-high speed train –trips," *Heavy Vehicle Systems, A Series of the Int'l J. of Vehicle Design*, vol. 7, 2000, pp. 22-33.

[12] P. J. Wolfs, S. Bleakley, S. T. Senini, and P. Thomas, "An autonomous, low cost, distributed method for observing vehicle track interactions," in *Rail Conf. 2006*, Atlanta, USA, Apr'06.

[13] M. Seifert, "Using modern technology to enhance rail security," in *Conference on Railway Engineering*, Melbourne, Australia, April – May 2006, pp. 430-436.

[14] C. P. Emad Aboelela, William Edberg and V. Vokkarane, "Wireless sensor network based model for secure railway operations," *Performance, computing and communications conference 2006, 25th IEEE*, p. 6pp, April 2006.

[15] S. M. S. L. Cheekiralla, "Development of a wireless sensor unit for tunnel monitoring," Master's thesis, Civil and Environmental Engineering, Massachusetts Institute of Technology, February 2004.

[16] S. Nefti and M. Oussalah, "A neural network approach for railway safety prediction," in *2004 IEEE Int'l Conf. on Systems, Man and cybernetics*, October 2004, pp. 3915-3920.

[17] C. Li., B. Stratman, and S. Mahadevan, "Improving railroad wheel inspection planning using classification methods," in *Proc. of the 25th IASTED Int'l Multi-Conf.*, Innsbruck, Austria, Feb'07, pp. 366-371.

[18] C. Li, "Classifying imbalanced data using a bagging ensemble variation (BEV)," in *Proc. of the 45th annual southeast regional conf.*, USA, Mar'07, pp. 203-208.

[19] S. Kaewunruen and A. M. Remennikov, "Response and prediction of dynamic characteristics of worn rail pads under static preloads," in *14th Int'l Congress on Sound Vibration*, Cairns, Australia, July'07.

[20] M. F. Duarte and Y. H. Hu, "Vehicle classification in distributed sensor networks," *Journal of Parallel and Distributed Computing*, vol.64, pp.826-838, 2004.

[21] I. H. Witten, and E. Frank, "Data Mining: Practical Machine Learning Tool and Technique with Java Implementation" Morgan Kaufmann, San Francisco, 2000.

[22] G.H. John, and R. Kohavi, "Wrappers for feature subset selection" *Artificial Intelligence*, Volume 97, Issue 1-2, December 1997.

[23] S. Jo Cunningham and G. Holmes, "Developing innovative applications in agriculture using data mining," Tech Report, Dept. of Computer Science, University of Waikato, New Zealand.

[24] D. Aha, "Tolerating noisy, irrelevant, and novel attributes in instance-based learning algorithms," *Int'l Journal on Man-Machine Studies*, Vol. 36, pp. 267-287, 1992.

[25] "Linear Regression," GraphPad Software, Inc. San Diego, USA, Tech. Rep. [Online]. Available: http://www.curvefit.com/linear_regression.htm as at 25th Janury'08.

[26] A. O. Sykes, "An introduction to regression analysis," Chicago working Paper in Law and Economics, Tech. Rep. [Online].Available: http://www.law.uchicago.edu/Lawecon/WkngPprs_01-25/20.Sykes.Regression.pdf.

[27] Shawkat Ali and Kate A. Smith, "On learning algorithm selection for classification", *Journal on Applied Soft Computing, ELSEVIER*, vol-6, 2006, pp. 119-138.

[28] G. D. Magoulas, V. P. Plagianakos, and M. N. Vrahatis, "Neural network-based colonoscopic diagnosis using on-line learning and differential evolution", *Appl. Soft Computer*, vol-4, pp. 369-379, 2004.

[29] "The Math Works", The Math Works, Inc. [Online]. Available: http://www.mathworks.com/ as at 9th January'08.

[30] "Weka 3," The University of Waikato, NewZealand, Tech.Rep. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka as at 9th March'08.

[31] A. B. M. Shawkat Ali and Saleh A. Wasimi, *Data Mining: Methods and Techniques*. Victoria, Australia: Thomson, 2007.