Utilizing Telematics Data to Support Effective Equipment Fleet-Management Decisions: Utilization Rate and Hazard Functions

Article *in* Journal of Computing in Civil Engineering · January 2015

DOI: 10.1061/(ASCE)CP.1943-5487.0000444

Utilizing Telematics Data to Support Effective Equipment Fleet-Management Decisions: Utilization Rate and Hazard Functions

Hisham Said, Ph.D., A.M.ASCE[1]; Tony Nicoletti[2]; and Peter Perez-Hernandez, S.M.ASCE[3]

Abstract: Contractors and equipment rental companies have started to acknowledge and use the telematics technology as a reliable solution for timely collection of their equipment fleet data. Telematics is the integration of wireless communications, vehicle monitoring systems, and location devices to provide real-time spatial and performance data of the fleet machines. Despite the large amount of real-time equipment data made available by telematics, fleet managers still try to identify ways to use such data to make informed fleet-management decisions. This paper presents the development of novel telematics-based computational methodologies to support two major equipment fleet management tasks: fleet use assessment and equipment health monitoring. First, a description of the telematics system and data used are presented. Second, a computational algorithm is proposed to quantify the fleet-wide equipment used, based on basic telematics data. Third, a health-monitoring framework is developed to estimate equipment failure events using telematics-based hazard functions, which were developed using survival analysis techniques. Finally, the telematics data sets of large equipment fleets (dozers, excavators, backhoes, and dump trucks) from two companies are used to verify and validate the proposed research developments by providing insightful fleet-management information. **DOI: 10.1061/(ASCE)CP.1943-5487.0000444.** *© 2014 American Society of Civil Engineers.*

Introduction

Equipment fleet management is a critical function in construction companies as it greatly contributes to the business profitability by managing the lifecycle operations of owned expensive assets. Fleet management is not simply about operating and maintaining companies' equipment and machines. Rather, it includes a wider scope of interdependent processes, like equipment investment justification, models specification, acquisition, assignment, and disposal. One of the critical processes of fleet management is the collection and control of equipment data related to use, costs, and condition diagnosis. Fleet data collection imposes great challenges to fleet managers to effectively and efficiently collect, store, and process equipment performance data in a timely manner.

In recent years, contractors and equipment rental companies started to acknowledge and use the telematics technology as a reliable solution for timely collection of their equipment fleet data. Telematics is the integration of wireless communications, vehicle monitoring systems, and location devices to provide real-time spatial and performance data of the fleet machines. Because of its great benefit to contractors and heavy equipment rental houses, the telematics market has witnessed significant growth, which resulted in installing the technology in 5.8 million equipment units and achieving more than $2 billion revenue in 2009 in the United States. (Fletcher and Lauron 2009). Telematics provides various advantages over other equipment-tracking technologies, such as global position systems (GPS) and radio frequency identification (RFID) (Lu et al. 2006). Telematics provides more than spatiotemporal data (location and time), as it communicates equipment performance and condition data such as fuel consumption, engine hours, and oil pressure. In addition, telematics does not require dedicated infrastructure of receivers like RFID, as it transmits data through the regular wireless communication networks used for mobile phones, or through satellite communication. Accordingly, telematics provides a more holistic and cost-effective technology for monitoring large equipment fleets. One limitation of telematics is its default reduced frequency of transmitting data compared with RFID, which limits its application for detailed operation tracking. Such frequency can be increased, but would require more data storage infrastructure to maintain the larger expected data volumes.

Telematics can be provided by two possible suppliers: the original equipment manufacturer (OEM) or a third-party telematics service provider (TSP). Newer equipment is usually manufactured with a telematics system already installed by the original equipment manufacturer. Telematics data are continuously sent to the OEM's Web system, which stores, organizes, and presents these data to the fleet managers using visual and user-friendly interfaces. In contrast, older equipment tiers that do not have an OEM telematics system can be equipped with TSP-rigged units that are connected to the equipment's mechanical and electrical subsystems to obtain telematics data. TSP companies maintain their own Web systems to archive and report telematics data to the fleet manager, similar to OEM systems.

Despite the growth and great potential of telematics, fleet managers are challenged by the unclear possible uses of the significant amount of data available by the technology. For middle and large fleet sizes, telematics is considered a source of massive data overload to the fleet managers, who face difficulties in linking telematics to business functions and performance metrics (Monnot and Williams 2011; Trimble and Bowman 2012; Jackson 2012;

[1]Assistant Professor, Dept. of Civil Engineering, Santa Clara Univ., Santa Clara, CA 95053; and Adjunct Lecturer, Faculty of Engineering, Cairo Univ., Giza, Egypt (corresponding author). E-mail: hsaid@scu.edu

[2]Director, Sales and Business Development, DPL America, Menlo Park, CA 94025. E-mail: tony@dplamerica.com

[3]M.Sc. Student, Dept. of Civil Engineering, Santa Clara Univ., Santa Clara, CA 95053. E-mail: pperezhernandez@scu.edu

Note. This manuscript was submitted on December 2, 2013; approved on September 5, 2014; published online on November 18, 2014. Discussion period open until April 18, 2015; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Computing in Civil Engineering***, © ASCE, ISSN 0887-3801/04014122(11)/$25.00.

Sutton 2013). For example, there were no previous reported experiences or methodologies of possible uses for the continuous reporting of equipment geographic locations in managing the whole fleet. Use challenges of telematics can be referred to the factual difference between data and information. Telematics provides a huge amount of data that report the individual machine's location and performance, but does not directly provide useful information about the machine's operational efficiency and business return. Identifying such useful information requires critical investigation of the typical operations of heavy construction and fleet-based companies, and accordingly establishing a link to related telematics data. As a result, current fleet-management practices and critical decisions (e.g., fleet use, maintenance, condition assessment) depend on either extensive paper-based tracking systems or ad-hoc subjective rules. This observation is valid for heavy construction companies, equipment rental houses, and logistics/distribution firms (Beach 2013), which individually have different business objectives but are all very similar in managing their equipment fleets.

Accordingly, this paper presents the development of novel telematics-based computational methodologies to support two major equipment fleet-management tasks: fleet use assessment and equipment health monitoring. The objective of this paper is to propose one hypothesis: Telematics data already provides a detailed and large record of equipment performance and condition, which requires basic processing computational operations to obtain higher-level fleet-management information. Thus, this paper is organized into five major sections: (1) a brief review of previous studies on telematics implementation in construction operations and fleet management; (2) a description of the telematics system and data used in this research; (3) a description of the developed computational algorithms to generate telematics-based fleet use metrics; (4) an explanation of the proposed health-monitoring framework that generates equipment failure hazard functions from obtained telematics data; and (5) future research directions and recommendations. This is a pilot study that considers only two of the possible applications of telematics (fleet use and equipment health condition), which is part of a larger research activity to develop a holistic telematics-based, fleet-management system. These two applications were selected for two main reasons: (1) their critical role in managing equipment fleets in terms of effectively using fleet assets and maintaining them; and (2) their need for two different types of telematics data (basic and CAN-bus), as explained later. It is envisioned that the system would include other modules that are related to critical equipment fleet and project-management tasks, such as preventive maintenance planning, cost control, productivity assessment, and fleet replacement.

Previous Research

Because of the novelty of the telematics technology, few research studies have been performed to investigate its use in construction-equipment performance tracking and fleet management. Monnot and Williams (2011) briefly presented a general overview the types of data collected by telematics and their possible benefits in equipment fleet management, such as reporting of machines hours, locations, fuel consumption, and health. However, this study did not suggest detailed methodologies for transforming telematics data into useful information that would benefit the fleet managers. Trimble and Bowman (2012) performed a detailed market survey of available telematics services in terms of integration and usability features. Integration features are related to the operation of the equipment and vehicles in terms of vehicle location, safety, diagnostics, communication, and interactivity. Usability features were also studied to evaluate the usefulness of the solution to different customer classes (small fleets and specialized fleets), data-reporting capabilities (Web interfacing and data accessing), and integration ability into company operations (staff management and risk management). The outcome of the performed survey was a consumer's guide to company planning, in adopting telematics technology to help in evaluating and selecting the solution provider that best serves the company's operations. The performed survey revealed that most of the capabilities of existing telematics systems are for data communication, not fleet performance analysis (Trimble and Bowman 2012). Aslan and Koo (2012) proposed an implementation plan for the use of telematics technology in improving the productivity of roadway maintenance operations. A data collection system layout was proposed to gather telematics data and report operation productivity on a real-time basis with geospatial locations. Future work was suggested to perform field tests of the system and to develop productivity improvement metrics to enhance roadway maintenance operations.

Despite the contribution of previous research studies, no effort was made in developing methodologies to transform telematics data into useful information that can be used and integrated in the management operations of heavy equipment fleets. The development of these methodologies requires the identification of relevant telematics data and the development of efficient algorithms to generate required equipment performance information.

Telematics System and Data

This section presents a brief technical description of the components of a typical telematics system and the types of collected equipment data. As shown in Fig. 1, a typical OEM or TSP telematics system involves three basic components: (1) transponder units on fleet assets to collect data; (2) communication medium to transmit data; and (3) a user interface to allow fleet managers to view real-time data and even communicate with the asset.

- First, a transponder unit is installed on each fleet asset that is physically connected to the engine and equipment control system to collect operation and location data on a real-time basis.

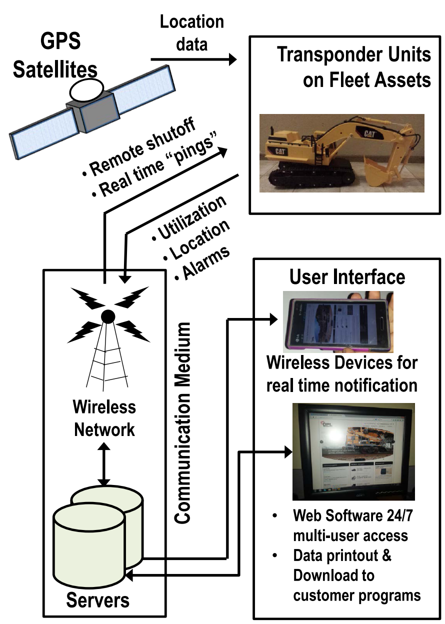

Fig. 1. Telematics system components and data flow

The next subsection describes in more detail the transponder unit, its hardware connection to the equipment, and collected data.

- Second, the transponder unit transmits the collected data through a standard wireless communication medium, which includes code division multiple access (CDMA) and global system for mobile communications (GSM) networks. The transponder unit literally calls the system server to transmit current values of telematics data once a change occurs to the equipment status, like turning on the ignition or incurring an engine health problem. The servers store and archive the telematics data that are collected for all fleet assets for an extended period of time, and can be accessed by fleet managers for real-time or periodical reporting and analysis.
- Third, the telematics system uses multiple user-interface alternatives to enable friendly and reliable access to the stored fleet operational data. Fleet managers can access the data through an online system that provides various data view options and basic reporting functions. As shown in Fig. 1, one-way notifications can be sent from the system server to wireless mobile devices of the fleet management department to report serious occurrences related to unauthorized equipment use or breakdown. A more interactive and advanced user interface can be provided through an online software package, in which the fleet manager can view the data and generate basic generic reports related to equipment runtime hours and use. As shown in Fig. 2, fleet managers can use graphical tools within the system to draw *geo fences* and *geo zones* to represent distinct geographic zones that can be used to refer to company construction sites or service yards. After these geo fences and zones are defined, the telematics system relates the telematics data to the geo zones from where the data were sent, using the unit's GPS latitude and longitude. As described in later sections, mapping the geographic presence of the company in terms of its construction sites and yards would greatly help in the proposed methodology of fleet use analysis. Telematics online tools can also provide some remote-control capabilities by calling the fleet unit through the communication medium to either request an update of specific data or even remotely disabling the equipment.

The rugged, metal telematics transponder unit is approximately the size of a large ashtray and houses the GPS receiver, wireless radio, and internal circuitry. As shown in Fig. 3, the transponder unit has four types of connecting cables to receive different telematics data and transmit them to the wireless network and then to the system server. First, the location data are obtained from a GPS antenna that is placed on top of the equipment using magnets or adhesive. Second, the main interface cables are used to connect the transponder to multiple points in the equipment to (1) power the transponder unit itself; (2) identify ignition events of the engine; and (3) obtain odometer readings. Third, other health diagnostic data can be obtained through analog or digital connections that connect to the equipment's controller area network communication (CAN-Bus). CAN-Bus is an internal communication system in heavy construction equipment that interconnects its electronic control units (ECUs), which control the equipment's subsystems and sensors, such as fuel, oil, fans, and engine. (Romans et al. 2000, Wan et al. 2009). The most recent equipment models are designed based on the J1939 standard, which is specified by the Society of Automotive Engineers (SAE) to define protocols of transferring equipment data between ECUs across the CAN-bus network. Older tiers of equipment and/or their engines are not manufactured with an internal monitoring system, which results in disabling the CAN-bus connection. Accordingly, only basic data can be obtained for older equipment tiers through the main interface cables, which

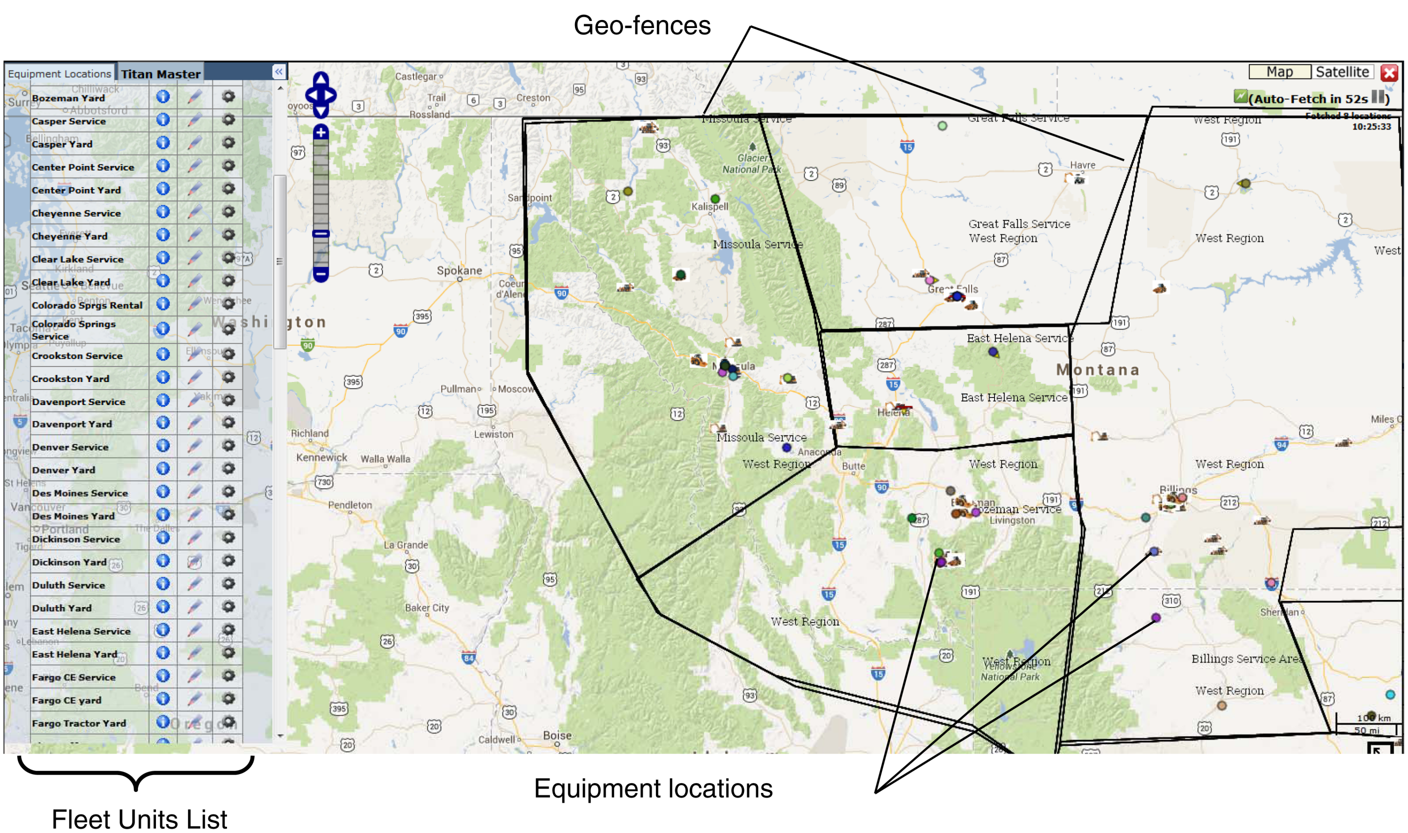

Fig. 2. Telematics user interface, locations map, and use of geo fences (image courtesy of DPL America)

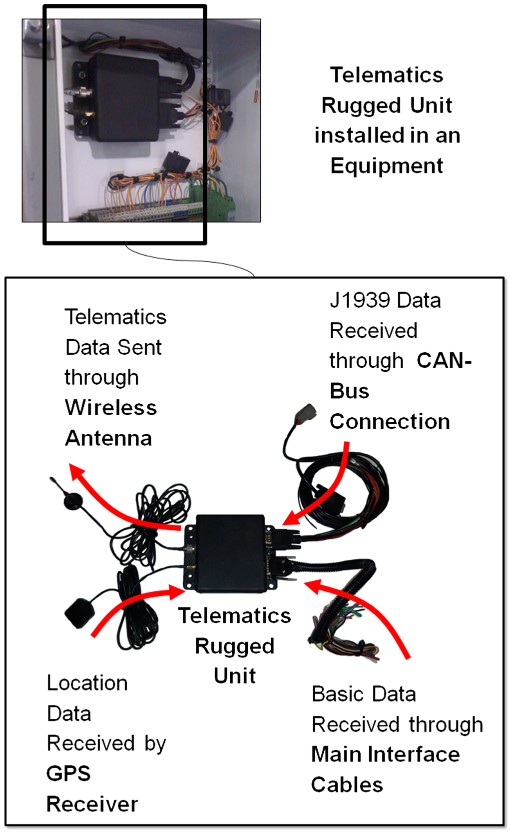

Fig. 3. Telematics transponder unit installation and connections (images courtesy of DPL America)

include ignition events and odometer readings. Other advanced data can be obtained in newer equipment tiers through the CAN-bus connection, such as fuel level, oil pressure, coolant temperature, engine average speed in rpm (round/minute), and error codes for the engine, transmission, or brakes.

Integration of Telematics Data into Fleet Management

Although telematics technology provides a real-time reporting of the equipment fleet data, there was no clear methodology for integrating such data in the major functions and decisions of fleet management (Monnot and Williams 2011). As shown in Fig. 4, each piece of equipment in the fleet is tracked by receiving a massive amount of time-stamped data entries that report the values of the equipment's monitored variables. Each data entry consists of the following: (1) time stamp of the event, when the data was received from the equipment's telematics transponder; (2) basic operation data that include the equipment's location within the defined geo zones (jobsites or yards), engine run time, GPS-based speed value, temperature, and battery voltage; (3) CAN-bus data that are available only for newer equipment tiers and include total fuel used, engine speed (rpm), machine lamps status (engine and transmission), engine oil pressure, temperatures (oil and coolant), fuel rate (gallons per hour, GPH), and calculated engine run and idle times based on the reported rpm.

Fleet managers can export these data from the telematics system in the form of different reports and spreadsheets, but are faced with different challenges when using them in the routine and strategic management decisions. First, fleet managers lack efficient computational algorithms to transform large amounts of raw telematics data into more useful higher-level information on their fleet condition. Second, telematics technology providers do not provide a clear methodology of how to integrate the collected data and information into typical fleet managerial tasks, such as health condition assessment, maintenance, replacement, operations control, and cost control. Third, it is not possible to justify the return on investment of procuring telematics technology to higher company management because of the lack of established-use methodologies of telematics data in a company's operations.

The following sections provide a detailed description of two possible uses of telematics data in fleet management: identification of overall fleet use and assessment of equipment health condition. The proposed methodologies of equipment use and health assessment are applied to the telematics data of a large equipment rental company and a trucking company. Similar application of the proposed methodology can also be performed to construction companies, as implied assumptions are valid in both cases. These developed methodologies are envisioned to be part of a larger framework that provides comprehensive integration of telematics data into a company's fleet and asset management program.

Fleet Use Assessment

A new algorithm was developed to calculate and assess the company's fleet use based on equipment locations that are reported by the telematics system. Fleet use is defined in this study as the percentage of time the equipment spent out of the company yards relative to the total reporting time. This definition of fleet use is based on the assumption that having the equipment out of its base yard implies that it is either being used by a construction company or is rented/leased by the rental company. Accurate assessment of fleet use is a very critical prerequisite for other managerial tasks and decisions related to strategic business planning and fleet

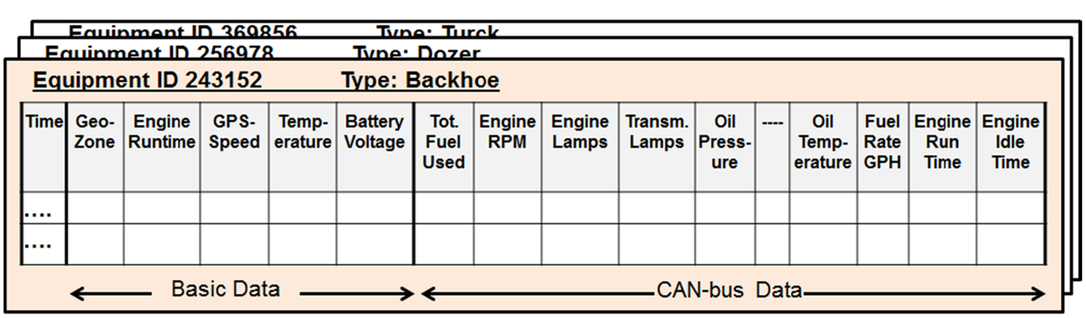

Fig. 4. Telematics time-stamped data entries

replacement. Telematics technology is proposed as an accurate, reliable, and efficient methodology of collecting necessary data for the assessment of fleet use. To quantify fleet use, two values are retrieved from each telematics data entry of every piece of equipment: (1) the location of the equipment in terms of the geo zone where it exists; and (2) the time of the reported location. These two values are analyzed using the newly developed telematics-based fleet-use assessment algorithm, which involves the following steps (Fig. 5):

1. Export fleet data from the telematics system in the form of a standard location history spreadsheet report for each piece of equipment. This report helps to focus only on the needed data and ignore other telematics data that are not relevant to the current task of fleet use assessment.
2. Identify the number of equipment pieces in the fleet (NE) and the number of telematics data entries in the location history report of each equipment (ND_i). Set the counter values of equipment (i) and data entries (j) to 1, where the entries are ordered from recent to oldest. Therefore the time stamps ($T_{i,j}$) of the first entries will be larger than those later in the list (i.e., $T_{i,j} > T_{i,j+1}$).
3. Check whether the equipment i is reported by data entry j as in-yard ($\mathrm{Zone}_{i,j}$ = "Yard"). This is done by examining the value of the geo-zone parameter and looking for the word "yard" as an indication that its GPS location is within the defined geo fences of one of the company yards. If the equipment is in-yard, set the tracking variable $\mathrm{IN}_{i,j}$, equal to "True"; otherwise set it equal to "False." For the first equipment in the fleet, go to step 8. For the next equipment, go to step 4 if the equipment is in yard;otherwise go to step 5.
4. In case equipment i was in-yard for data entry j, check whether it was also in-yard for previous data entry $j-1$. If this condition is true ($\mathrm{IN}_{i,j} = \mathrm{IN}_{i,j-1}$ = "True"), update the time in-yard variable of the equipment (TI_i) by calculating the total time difference between data entries j and $j+1$ ($T_{i,j}$ and $T_{i,j+1}$, respectively), as shown in Eq. (1). If this condition is false ($\mathrm{IN}_{i,j} \neq \mathrm{IN}_{i,j-1}$), go to step 6

$$TI_i = TI_i + (T_{i,j} - T_{i,j+1}) \tag{1}$$

5. In case equipment i was out of yard for data entry j, check whether it was also out of yard for previous data entry $j-1$. If this condition is true ($\mathrm{IN}_{i,j} = \mathrm{IN}_{i,j-1}$ = "False"), update the time out-of-yard variable of the equipment (TO_i) by calculating the total time difference between data entries j and $j+1$ ($T_{i,j}$ and $T_{i,j+1}$, respectively), as shown in Eq. (2). If this condition is false ($\mathrm{IN}_{i,j} \neq \mathrm{IN}_{i,j-1}$), go to step 6

$$TO_i = TO_i + (T_{i,j} - T_{i,j+1}) \tag{2}$$

6. In case equipment i had different location values (inside and outside yard) in data entries j and $j+1$, update both time variables TI_i and TO_i using Eqs. (3) and (4), which assume that the time difference is equally divided between the in-yard and out-of-yard statuses

$$TI_i = TI_i + (T_{i,j} - T_{i,j+1})/2 \tag{3}$$

$$TO_i = TO_i + (T_{i,j} - T_{i,j+1})/2 \tag{4}$$

7. Repeat steps 3 through 6 for the next data entry ($j = j+1$) until the last data entry ND is analyzed.
8. Repeat steps 3 through 7 for the next equipment in the fleet ($i = i+1$) until the complete location history report of the last equipment NE is analyzed. Calculate the use rate U_i for each equipment using Eq. (5) as the ratio between the out-of-yard time (TO_i) and the total analyzed time ($TO_i + TI_i$)

$$U_i = TO_i/(TO_i + TI_i) \tag{5}$$

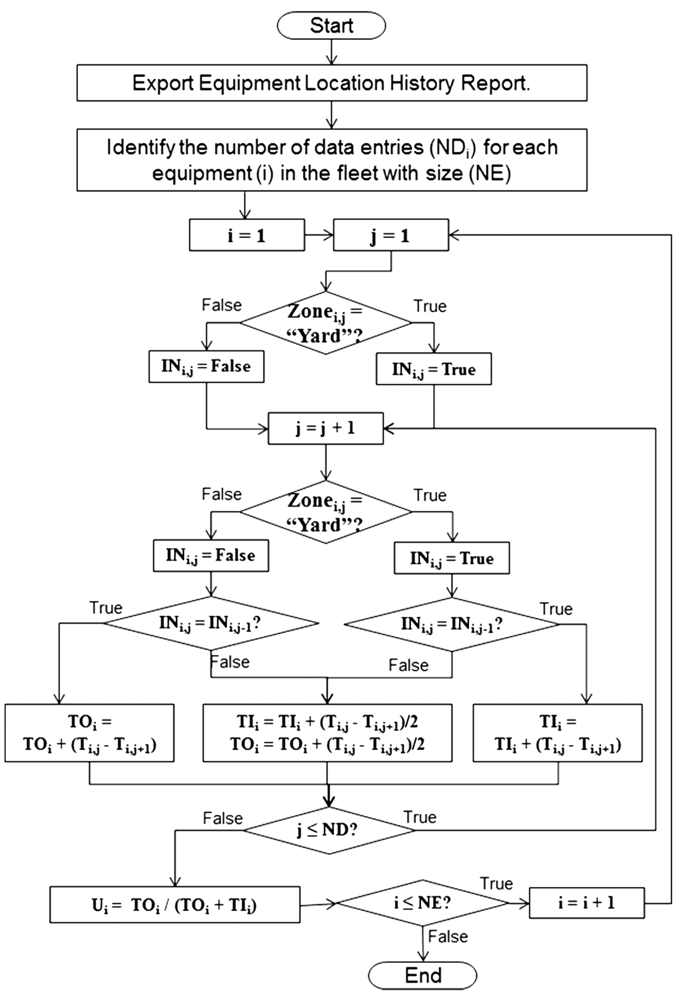

Fig. 5. Telematics-based fleet-use assessment algorithm

Visual plotting of the fleet equipment use rates can be useful in providing valuable insight to the managers for effective fleet replacement, expansion, and disposal decisions. As shown in Fig. 6, the use rate of all fleet pieces of the rental company's equipment is plotted with an average use of approximately 38.7% during the analyzed time of 6 months. A targeted overall use rate was set by the fleet manager to be 35%, which indicates that the fleet is over-used. Such a low targeted use rate is imposed by the manager to account for possible peak periods of equipment rental and use demand. More insight can be obtained by plotting the use rates of each equipment type, such as backhoes and excavators, as shown in Fig. 7. The backhoes were found to be the most used equipment type, with an average use rate of 49%, whereas the excavators' average use was approximately the same as the whole fleet (37.5%). Per the use information obtained from the telematics data, fleet managers would consider buying more pieces of over-used equipment types, and disposing some pieces of under-used equipment types. Despite the simplicity of the proposed use rate algorithm, its novelty comes from its ability to transform large dispersed data volume into a fleet-wide use rate. To the best of the authors'

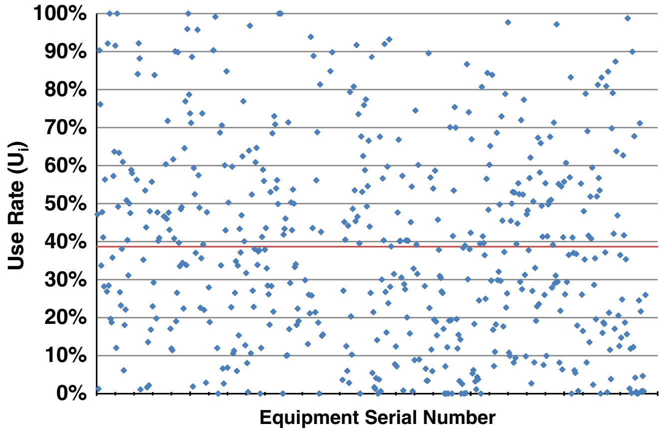

Fig. 6. Plot of equipment individual-use rates and average fleet-use rate

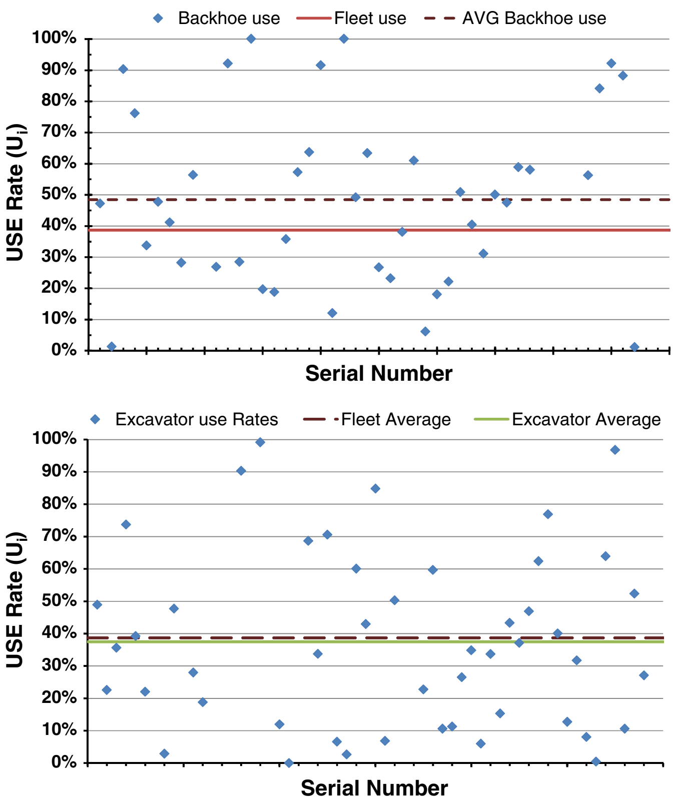

Fig. 7. Use of backhoes and excavators

knowledge, no such metric has been proposed before for equipment fleet based on telematics data.

Equipment Health Monitoring

Equipment health monitoring (EHM) is the process of collecting vital equipment performance parameters to continuously assess the condition of the equipment and detect signs of possible failure. EHM provides a proactive approach to equipment asset maintenance by fixing the equipment just before a severe failure event happens, instead of the current reactive approach after a failure occurs. EHM is an essential facilitator of predictive maintenance program (Gransberg et al. 2006, p. 237), in which maintenance tasks are scheduled just before failures are expected to happen based on the monitored performance of the machine. Previous research on equipment and machine health monitory focused on monitoring the condition (primarily the vibrations) of stationary mechanical machines or electrical micro-machines (Dutta and Giurgiutiu 2000; Yan and Gao 2007; Da et al. 2011). Equipment manufacturers have encouraged research and development efforts to develop health-monitoring systems that integrate remote sensing and equipment oil sampling to diagnose the equipment's condition and estimate its life expectancy (Murakami et al. 2002). However, no recent research studies are known to have investigated the use of telematics data in equipment health monitoring and diagnostics.

A telematics-based equipment health-monitoring (T-EHM) framework was developed and is proposed in this paper to support fleet service managers in using telematics data in their predictive maintenance programs. The developed T-EHM framework consists of two modules: (1) the health parameters processing and visualization (HPPV) module; and (2) the equipment failure hazard estimation (EFHE) module. The following subsections describe each of the framework modules.

Health Parameters Processing and Visualization Module

The objective of this module is to identify, retrieve, process, and visualize telematics data that represent vital equipment health parameters. The HPPV module currently uses 10 CAN-bus values for each telematics data entry, which are identified by the research team, the company service manager, and the limited research available on heavy equipment health monitoring (Murakami et al. 2002; Dekate 2013). Other possibly related parameters, such as intake manifold temperature, were not selected because of observed inconsistencies or incompleteness in their data. Such inconsistencies in some telematics data occur as a result of some equipment manufacturers not precisely following the J1939 standard, which is adopted by most telematics technology providers. However, the proposed module and underlying methodologies can be expanded flexibly to include additional parameters in case their complete telematics data are not available. Accordingly, the selected parameters include:

1. Maximum coolant temperature (MCT_i) in degrees Fahrenheit, which is observed on the day when the telematics data entry is received;
2. Maximum engine oil pressure (MOP_i) in pounds per square inch (PSI);
3. Maximum engine oil temperature (MOT_i) in degrees Fahrenheit;
4. Maximum engine speed (MES_i), in rounds per minute (rpm);
5. Maximum engine percent torque (MPT_i), which indicates the load on the engine as a percentage value;
6. Engine lights, which include red stop light (RSL), amber warning light (AWL), engine protection light (EPL), and malfunction indicator light (MIL). RSL is used to relay trouble code information that is of a severe enough condition that it warrants stopping the vehicle. AWL is used to relay trouble code information that is reporting a problem with the vehicle system but the vehicle does not need to be stopped immediately. In contrast, EPL and MIL are used to report less severe hazard conditions related to the vehicle system and emissions-related issues. These engine light variables are modeled as binary values (1 or 0);
7. Maximum fuel rate (MFR_i) in gallons/hour;

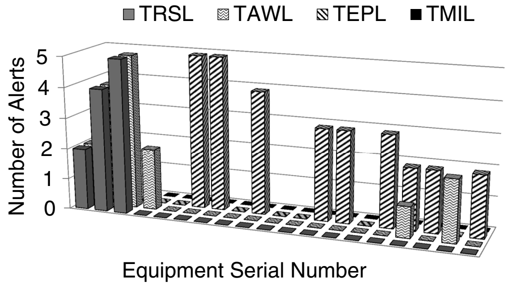

Fig. 8. Visualization of equipment health parameter (engine lamps) for dozers

8. Engine working hours (HW_i), which reports the cumulative number of hours the engine ran with a speed (rpm) above a specified threshold, set by the fleet service manager;
9. Engine Idling hours (HI_i), which reports the cumulative number of hours the engine ran with a speed (rpm) less than the specified threshold; and
10. Engine total run hours (HT_i).

The health parameter values of a piece of equipment can be visualized and compared with other equipment of the same type to enable fleet service managers to qualitatively assess its performance and health. Fig. 8 shows an example visualization of the total counts of RSL, AWL, EPL, and MIL engine lamps (i.e., TRSL, TAWL, TEPL, and TMIL) for all dozers in the fleet during a 6-month period, which primarily highlights three unhealthy machines because of the high number of reported engine lamps. Accordingly, fleet and service managers can either increase the maintenance commitment to these machines or select them for potential replacement or sale. However, visualizing already obtained telematics data represents a reactive approach to equipment health monitoring (Gransberg et al. 2006, p. 237). The second proposed module, in the next subsection, attempts to provide a proactive approach to predicting equipment health issues based on ongoing collected telematics data.

Equipment Failure Hazard Estimation Module

The EFHE module is designed to integrate telematics data into predictive maintenance programs by deriving a failure hazard function for every equipment type. Hazard functions are a class of survival models that statistically analyze a company's data timeline to relate failure events' dependency on one or more covariates over time (Ma and Krings 2008; Gu et al. 2011). Cox's proportional hazards model (Cox 1972), one of the fundamental survival models, is used in this research to relate equipment failure hazard to the previously identified telematics health parameters. Cox's model is one of the most used survival function techniques as it captures the survival/failure determinates more than nonparametric models and is less restrictive than the parameter models (Baily et al. 2006).

It is hypothesized that the failure hazard probability of a piece of equipment follows a time-varying dynamic hazard Cox function $h(t)$ (Martinussen and Scheike 2006), as shown in Eq. (6). The hazard function is controlled by its baseline failure rate $h_0(t)$, covariates vector $X(t)$ (i.e., values of telematics health parameters), and their coefficients vector $\beta(t)$. This function assumes that covariates are multiplicatively related to the hazard, which indicates that a change in any element of the covariates vector results in proportional scaling of the hazard. For example, an increase in the engine oil temperature is expected to result in an increase in the equipment's failure hazard. The function is made dynamic by allowing its baseline failure rate and the coefficients of its covariates to change over time

$$h(t) = h_0(t) \times \exp[\beta(t) \cdot X(t)] \tag{6}$$

The EFHE module follows a newly developed methodology to construct and assess the fitting of the hazard function to a selected set of health parameters to estimate the failure hazard of an equipment type. The methodology is generic and can be applied to any equipment type and health parameters that are initially suggested by the fleet manager. The outcome of the methodology is the time-varying survival function to the provided telematics data, and the statistical fitting metrics of the parameters and their coefficients. The methodology is iterative as the fleet manager can consider different groups of health parameters to obtain the most fitted survival function. The methodology involves the following nine steps:

1. The telematics data time series of each equipment is broken down into successive survival lives that are divided by the equipment's failure events, as shown in Fig. 9. The engine's RSL and AWL lights were used to refer to sever failure events of the equipment, which eliminates the fleet manager's need to track actual maintenance records and integrate with the already

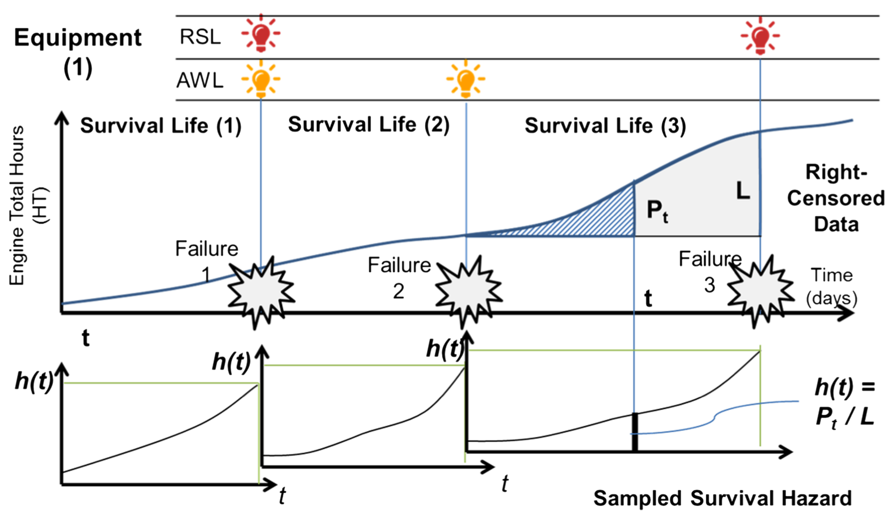

Fig. 9. Telematics-based sampling of equipment failure and survival lives

available telematics data. The last open-ended time interval of equipment, which represents a recovery from a failure with no observed failure event during the analysis period, is a powerful technique in survival analysis, called data right-censoring (Ma and Krings 2008); it considers only relevant and complete failure and recovery cycles;

2. The time-varying periods of the survival function are defined by the fleet manager or service technician to represent the dynamically changing equipment's failure hazard over time. The number and durations of these time-varying periods need to be carefully set as they affect the performance of the proposed methodology. Shorter periods would provide more dynamic and accurate survival function, but would result in fewer telematics data observations to statistically generate function coefficients with sufficient fitting;
3. The failure hazard $h(t)$ for each telematics entry at time t experienced by the equipment over a survival life period is quantified and sampled as the ratio between (1) the net engine total hours consumed since the survival life start time to the corresponding time t; and (2) the survival life length L as the difference between net engine total hours experienced during the survival period, as shown in Fig. 9. Accordingly, the total hours parameter cannot be considered as a survival function covariate as it is used in calculating the function dependent variable (i.e., failure hazard);
4. The outliers in the telematics data entries are identified and eliminated if the absolute studentized residual of any of its fields is bigger than 3.0 (Montgomery et al. 2001). Outliers represent low or high inconsistent extremes that are observed in the collected data that would prohibit the development of effective regression models of the survival function;
5. The telematics entries are distributed between the time-varying periods of the survival function, based on their time (t) values;
6. The telematics entries of each time-varying period are randomly divided into two equal groups (Lucko and Rojas 2010): (1) an estimation group that is used to estimate the hazard function regression coefficients (see step 7); and (2) a prediction group that is used to validate the hazard function and its estimated coefficient (see step 8);
7. For each time period, the survival function baseline failure rate $h_0(t)$ and coefficients vector $\beta(t)$ are estimated by applying the data linearization regression technique to the estimation group of telematics entries. The regression population includes all of the telematics data estimation group within the corresponding time-varying survival function period. The fitness of the generated survival function and its coefficients can be evaluated using (1) the p-value for the constant and each covariate coefficient as generated from by regression analysis, which is used to test the hypothesis of survival function dependency on each of the covariates; (2) coefficients of determination (R-square, Multiple R-square, and adjusted R-square) to test the fit of the resulting survival function to the observed data; and (3) analysis of variance (ANOVA) significance level F, which quantifies the probability that the proposed function does not explain the variation in the equipment hazard (Field 2005);
8. For each time period, the survival function coefficients are validated by using the survival function to calculate the failure hazard estimate values for every telematics entry in the prediction group and comparing the estimates with the observed values. The validity is assessed using the Pearson coefficient of correlation (R_{corr}) and the student t-test to examine the hypothesis that no relation exists between the observed and estimated hazard values (Lucko et al. 2006). In addition, the variance between estimated and observed failure hazard values is quantified using the root mean square error (RMSE), where its smaller values reflect higher prediction accuracy (Montgomery et al. 2001); and
9. Repeat steps 7 and 8 to experiment with different combinations of the proposed covariates to find the survival function coefficients that provide (1) the maximal fit to the estimation telematics data group (p-value, R-square, F); and (2) the minimal variance with the prediction data group (i.e., RMSE) with validated correlation (R_{corr} and t-test).

The health-monitoring methodology of the proposed EFHE module was tested by applying it to the telematics data time series of two fleets (dozers and backhoes) in the rental company, in addition to the data of dump trucks in another hauling company.

Table 1. Final Regression Results for the Dozers' Hazard Function

		Survival intervals (days)				
		$0 < t < 50$	$50 < t < 100$	$100 < t < 150$	$150 < t < 300$	$300 < t$
Class	Variables/parameters	Value (P-value)	Value (P-value)	Value (P-value)	Value (P-value)	Value (P-value)
Parameters	Constant (C)	1 (N/A)	1 (N/A)	1 (N/A)	1 (N/A)	1 (N/A)
	X1 (MCT)	0 (N/A)	−0.00848 (0.00002)	0 (N/A)	0 (N/A)	0 (N/A)
	X2 (MOP)	−0.04917 (0.0)	−0.0206 (0.0)	−0.01491 (0.0)	−0.00728 (0.0)	0 (N/A)
	X3 (MOT)	0.01101 (0.0003)	0.0096 (0.00016)	0 (N/A)	0 (N/A)	0 (N/A)
	X4 (MES)	0 (N/A)	0 (N/A)	0 (N/A)	0 (N/A)	−0.0001 (0.0)
	X5 (MPT)	0.02792 (0.0007)	0.0135 (0.0061)	0.01336 (0.00035)	0.00659 (0.00093)	0 (N/A)
	X6 (MFR)	−0.35908 (0.0)	−0.1408 (0.00001)	−0.13776 (0.0)	−0.09602	0 (N/A)
	X7 (HW)	0.00364 (0.0)	0.00195 (0.0)	0.00193 (0.0)	0.00129	0.00016 (0.09811)
	X8 (HI)	0 (N/A)	0 (N/A)	0 (N/A)	0 (N/A)	0.00101 (0.07346)
Estimation	Observations	338	185	133	198	24
	Multiple R	0.800193	0.854431	0.857417	0.845453	0.906493448
	R square	0.640309	0.730053	0.735165	0.714791	0.821730372
	Adjusted R square	0.633072	0.716926	0.721035	0.705226	0.757133264
	Significance F	1.054E-72	3.71E-48	1.46E-35	1.3E-51	7.37E-08
Prediction	Observations	338	185	133	198	27
	RMSE	0.2967	0.2696	0.223	0.1691	0.0499
	R_{corr}	0.3896	0.3745	0.4375	0.4965	0.7845
	Observed t-test	7.707	5.5236	5.6319	8.0085	6.3251
	Critical t-test	1.64912	1.65304	1.6563	1.65221	1.70562

The analyzed fleets included 21 dozers, 29 backhoes, and 17 trucks. The collected data from each fleet represented the observation values of the previously mentioned 10 parameters, in which (1) engine lights were used to model failure events and survival period boundaries; (2) the engine's total run hour was used to calculate the failure hazard; and (3) the remaining eight parameters were proposed as the covariates of the survival function, which include coolant temperature (MCT), oil pressure (MOP), oil temperature (MOT), engine speed (MES), percent torque (MPT), fuel rate (MFR), working hours (HW), and idling hours (HI). The initial telematics data collected included 1,836 sample data observations for the dozers, 3,315 observations for the backhoes, and 3,880 observations for the trucks. After removing the data outliers, the data populations were reduced to 1,767, 3,016, and 3,485 for the dozers, backhoes, and trucks, respectively. To obtain the time-varying survival function, the data of each fleet were divided into five possible survival intervals: less than 50 days, between 50 and 100 days, between 100 and 150 days, between 150 and 300 days, and more than 300 days. The final step in data preparation is to split the data in every period between the two groups of survival function coefficients estimation and survival function prediction validation.

Regression analyses were performed for all five survival intervals of the three fleets to fit their observed pairs of telematics health parameters and the survival metric with the proposed Cox's survival function. Tables 1–3 give the final regression results for the survival functions of the dozers, backhoes, and trucks, respectively. For each fleet, the regression results include the used metrics of coefficients estimation quality (multiple R, R-square, adjusted

Table 2. Final Regression Results for the Backhoes' Hazard Function

		Survival intervals (days)				
		$0 < t < 50$	$50 < t < 100$	$100 < t < 150$	$150 < t < 300$	$300 < t$
Class	Variables/parameters	Value (P-value)	Value (P-value)	Value (P-value)	Value (P-value)	Value (P-value)
Parameters	Constant (C)	1 (N/A)	1 (N/A)	1 (N/A)	1 (N/A)	1 (N/A)
	X1 (MCT)	0 (N/A)	0 (N/A)	0 (N/A)	0 (N/A)	0 (N/A)
	X2 (MOP)	−0.03958 (0.0)	−0.0165 (0.0)	−0.01449 (0.00001)	−0.01442 (0.0)	−0.01318 (0.)
	X3 (MOT)	0.007714 (0.0)	0 (N/A)	0 (N/A)	0 (N/A)	−0.00093 (0.023)
	X4 (MES)	0 (N/A)	0 (N/A)	0.000467 (0.001295)	0 (N/A)	0 (N/A)
	X5 (MPT)	0 (N/A)	0.01835 (0.00196)	−0.00913 (0.03857)	0.005372 (0.0595)	0 (N/A)
	X6 (MFR)	0 (N/A)	−0.18099 (0.00614)	0 (N/A)	0 (N/A)	0 (N/A)
	X7 (HW)	0.001505 (0.0)	0 (N/A)	0.00156 (0.0)	0 (N/A)	0.007903 (0.0)
	X8 (HI)	0 (N/A)	0.000929 (0.0)	−0.0013 (0.01273)	0.001227 (0.0)	−0.00515 (0.0)
Estimation	Observations	664	305	236	241	62
	Multiple R	0.768780949	0.811130642	0.767930736	0.789077673	0.951865503
	R square	0.591024147	0.657932919	0.589717615	0.622643574	0.906047937
	Adjusted R square	0.588273842	0.651201353	0.578284153	0.615270831	0.883946968
	Significance F	8.52E-128	9.4E-69	1.03E-42	5E-50	9.16E-29
Prediction	Observations	665	305	236	241	61
	RMSE	0.374626	0.23463	0.22687	0.26875	0.1
	R_{corr}	0.119655	0.3971456	0.32333	0.49067	0.878159
	Observed t-test	3.103263	7.5326	5.22682	8.7055	10.7
	Critical t-test	1.647	1.65	1.651	1.651	1.671

Table 3. Final Regression Results for the Dump Trucks' Hazard Function

		Survival intervals (days)				
		$0 < t < 50$	$50 < t < 100$	$100 < t < 150$	$150 < t < 300$	
Class	Variables/parameters	Value P-value	Value P-value	Value P-value	Value P-value	$300 < t$
Parameters	Constant (C)	1 (N/A)	1 (N/A)	1 (N/A)	1 (N/A)	No failure recorded for this period
	X1 (MCT)	0 (N/A)	−0.00326 (0.0108)	−0.00292 (0.0)	0 (N/A)	
	X2 (MOP)	−0.05807 (0.0)	0 (N/A)	0 (N/A)	0 (N/A)	
	X3 (MOT)	0.00394 (0.0139)	0.00287 (0.00353)	0 (N/A)	0 (N/A)	
	X4 (MES)	0 (N/A)	−0.00024 (0.00669)	0 (N/A)	−0.00004 (0.0)	
	X5 (MPT)	0 (N/A)	0 (N/A)	0 (N/A)	0 (N/A)	
	X6 (MFR)	0.0724 (0.00006)	0 (N/A)	0.023205 (0.000078)	0 (N/A)	
	X7 (HW)	0.000463 (0.0)	0.000394 (0.0)	0.000066 (0.065468)	0 (N/A)	
	X8 (HI)	0 (N/A)	−0.0004 (0.0012)	0 (N/A)	0 (N/A)	
Estimation	Observations	1,064	411	232	37	
	Multiple R	0.8083898	0.8909896	0.8238763	0.824544	
	R square	0.653494	0.79386	0.6787724	0.679874	
	Adjusted R square	0.65157	0.78937	0.6716	0.652096	
	Significance F	4.74E-242	1.366E-136	4.25498E-56	2.51E-10	
Prediction	Observations	1,063	410	232	36	
	RMSE	0.3468	0.174413	0.1307	0.066	
	R_{corr}	0.22514	0.376173	0.2435	0.286	
	Observed t-test	7.5267	8.2	3.807519	1.74	
	Critical t-test	1.64629	1.6486	1.6515	1.69	

R-square, and significance F) and survival function prediction accuracy (RMSE, R_{corr}, observed t-test, and critical t-test). The strong correlations between the considered telematics parameters and the equipment failure hazard were clearly highlighted by (1) the low values of significance F, which refer to the probability that the survival function does not explain the equipment failure hazard; and (2) the values of the observed t-Test values that are greater than the critical t-Test values, which lead to the rejection of the hypothesis of no correlation existing between observed and predicted failure hazard values. Accordingly, the equipment failure hazard failure function of each fleet type can be constructed using the generated Cox's function covariate coefficients, such as the survival function of the dozers shown by Eq. (7). The survival function of the trucks fleet includes only four intervals because of the absence of recorded failures beyond 300 days of the observed trucks, as shown in Table 3.

$$h(t) = \begin{cases} \mathrm{EXP}[-0.04917 \cdot \mathrm{MOP} + 0.01101 \cdot \mathrm{MOT} + 0.02792 \cdot \mathrm{MPT} - 0.35908 \cdot \mathrm{MFR} + 0.00364 \cdot \mathrm{HW}] & 0 \leq t < 50 \\ \mathrm{EXP}[-0.00848 \cdot \mathrm{MCT} - 0.0206 \cdot \mathrm{MOP} + 0.0096 \cdot \mathrm{MOT} + 0.0135 \cdot \mathrm{MPT} - 0.1408 \cdot \mathrm{MFR} + 0.00195 \cdot \mathrm{HW}] & 50 \leq t < 100 \\ \mathrm{EXP}[-0.01491 \cdot \mathrm{MOP} + 0.01336 \cdot \mathrm{MPT} - 0.13776 \cdot \mathrm{MFR} + 0.00193 \cdot \mathrm{HW}] & 100 \leq t < 150 \\ \mathrm{EXP}[-0.00728 \cdot \mathrm{MOP} + 0.00659 \cdot \mathrm{MPT} - 0.09602 \cdot \mathrm{MFR} + 0.00129 \cdot \mathrm{HW}] & 150 \leq t < 300 \\ \mathrm{EXP}[-0.0001 \cdot \mathrm{MES} + 0.00016 \cdot \mathrm{WH} + 0.00101 \cdot \mathrm{HI}] & 300 \leq t \end{cases} \tag{7}$$

The generated hazard functions of the three observed fleets provide useful insight into telematics-based prediction accuracy of equipment failure hazard and its dynamic dependency on telematics health parameters. As presented in Tables 1–3, better survival function fitting and prediction accuracy to the data were achieved for the last survival life interval compared with earlier intervals. For example, a higher R-square value and a lower RMSE value were achieved for the last interval ($t > 300$) of the dozers compared with their previous intervals, even with fewer data observations. This can be attributed to the consistency of these later observations for reporting a higher number of failure occurrences at such long survival periods, as a result of the equipment's accumulated tear and decay. In general, the data show that failure hazard can be statistically predicted with higher accuracy and fewer information requirements (number of covariates) with an increase in the equipment's survival time. This can be shown by a failure prediction error of 29.67% in the first survival interval of the dozers, which decreases to 4.99% error in the last modeled survival interval.

Summary and Conclusions

This paper presents novel methodologies to support heavy equipment fleet managers in using and integrating basic and CAN-bus telematics data into two major managerial tasks: fleet use assessment and equipment health monitoring. First, the telematics system is described in terms of the hardware, computational infrastructure, and collected data. Second, a telematics-based fleet-use assessment algorithm was developed to use received equipment spatiotemporal data to calculate its use rate based on its times outside and inside the idling yards. Third, a telematics-based equipment health-monitoring (T-EHM) framework was developed as a proactive maintenance tool to estimate the equipment failure probability. The framework applies survival analysis techniques to the collected telematics health parameters to generate a dynamic hazard function for each equipment type. The developed methodologies were validated through their application to large equipment fleets from two different companies (rental house and hauling company). Although the presented results are specific to the analyzed test beds and their telematics data, the proposed research methodologies (fleet use and health monitoring) are generic and can be generalized to other fleet-based business types. The merit of the developed methodologies was acknowledged by the telematics manager of the analyzed equipment rental company in terms of providing critical insights about the fleet use and equipment health and failure assessment. The telematics manager suggested integrating the developed methodologies into existing telematics systems to provide value-adding service and information to equipment fleet owners.

Future research efforts are planned to expand the developed methodologies, improve their accuracy, and provide intelligent decision-support capabilities. First, optimization models can be developed to maximize the fleet's overall use by moving equipment assets from low-use locations to higher-use locations. This would require the collection and integration of fleet operational data related to the expected generated revenue and the transportation cost of relocating fleet equipment assets. Second, the equipment's health-monitoring framework can be improved by using the health standards of equipment original manufacturers to identify the ideal operational ranges of equipment health parameters. Third, further research is needed to improve the accuracy of the predicted failure hazard in earlier survival time intervals by integrating additional data from other fleet data sources, such as the engine oil tests from commercial maintenance management software (Murakami et al. 2002). Fourth, automated modules can be developed based on the proposed use and health-assessment methodologies and integrated into available telematics data to provide value to heavy-equipment fleet owners, as suggested by the telematics manager of the case study equipment rental company.

References

Aslan, B., and Koo, D. (2012). "Productivity enhancement for maintenance equipment operations using telematics technology." *Construction Research Congress*, ASCE, West Lafayette, IN, 971–980.

Bailey, W. J., Weir-Jones, I., Couet, B., and Hogan, J. R. (2006). "Survival analysis: The statistically rigorous method for analyzing electrical submersible pump system performance." *SPE Annual Technical Conf. and Exhibition*, Society of Petroleum Engineers, Dallas, TX.

Beach, J. (2013). “Telematics unchained.” ⟨http://www.truckinginfo.com/channel/fleet-management/article/story/2013/05/telematics-unchained.aspx⟩ (Jan. 3, 2014).
Cox, D. R. (1972). “Regression models and life-tables.” *J. R. Stat. Soc. Ser. B*, 34(2), 187–220.
Da, Y., Shi, X., and Krishnamurthy, M. (2011). “Health monitoring, fault diagnosis and failure prognosis techniques for brushless permanent magnet machines.” *Proc., Vehicle Power and Propulsion Conf. (VPPC)*, IEEE, Chicago, 1–7.
Dekate, D. A. (2013). “Prognostics and engine health management of vehicle using automotive sensor systems.” *Int. J. Sci. Res.*, 2(2), 244–251.
Dutta, S., and Giurgiutiu, V. (2000). “Health monitoring and quality assurance for rotary micro-machines and active sensors.” *8th Int. Symp. on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-8)*, Pacific Center of Thermal-Fluids Engineering, Honolulu.
Field, A. (2005). *Discovering statistics using SPSS*, 2nd Ed., Sage, London.
Fletcher, L., and Lauron, G. (2009). “How can telematics help your fleet?” ⟨http://www.government-fleet.com/channel/mobility/article/story/2009/02/how-can-telematics-help-your-fleet.aspx?prestitial=1⟩ (Nov. 27, 2013).
Gransberg, D., Popescu, C., and Ryan, R. (2006). *Construction equipment management for engineers, estimators, and owners*, CRC Press, Taylor and Francis Group, Boca Raton, FL.
Gu, Y., Sinha, D., and Banerjee, S. (2011). “Analysis of cure rate survival data under proportional odds model.” *J. Lifetime Data Anal.*, 17(1), 123–134.
Jackson, T. (2012). “Off-road telematics: Why the disconnect?” ⟨http://www.equipmentworld.com/maintenance-23/⟩ (Nov. 27, 2013).
Lu, M., Chen, W., Shen, X., Lam, H., and Liu, J. (2006). “Positioning and tracking construction vehicles in highly dense urban areas and building construction sites.” *J. Autom. Constr.*, 16(5), 647–656.
Lucko, G., Anderson-Cook, C. M., and Vorster, M. C. (2006). “Statistical considerations for predicting residual value of heavy equipment.” *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2006)132:7(723), 723–732.
Lucko, G., and Rojas, E. M. (2010). “Research validation: Challenges and opportunities in the construction domain.” *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000025, 127–135.
Ma, Z., and Krings, A. W. (2008). “Survival analysis approach to reliability, survivability and prognostics and health management (PHM).” *Aerospace Conf.*, AIAA, Reston, VA, 1–20.
Martinussen, T., and Scheike, T. H. (2006). *Dynamic regression models for survival data: Statistics for biology and health*, Springer, New York.
Monnot, J., and Williams, R. (2011). “Construction equipment telematics.” *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000281, 793–796.
Montgomery, D. C., Peck, E. A., and Vining, C. G. (2001). *Introduction to linear regression analysis*, 3rd Ed., Wiley, New York.
Murakami, T., Saigo, T., Ohkura, Y., Okawa, Y., and Taninaga, T. (2002). “Development of vehicle health monitoring system (VHMS/WebCARE) for large-sized construction machine.” *Komatsu’s Technical Rep.*, Komatsu, 15–21.
Romans, W., Poore, B., and Mutziger, J. (2000). “Advanced instrumentation for agricultural equipment.” *Instrum. Meas. Mag.*, 3(1), 26–29.
Sutton, R. (2013). “Turbo telematics.” ⟨http://www.constructionequipment.com/blog/turbo-telematics⟩ (Nov. 27, 2013).
Trimble, T., and Bowman, D. (2012). “Market guide to fleet Telematics services.” *Rep. 12-UT-028*, National Surface Transportation Safety Center for Excellence, Virginia Tech Transportation Institute, Blacksburg, VA.
Wan, X., Xing, Y., and Cai, L. (2009). “Application and implementation of CAN bus technology in industry real-time data communication.” *2009 Int. Conf. on Industrial Mechatronics and Automation (ICIMA)*, IEEE, Chengdu, China, 278–281.
Yan, R., and Gao, R. (2007). “Approximate entropy as a diagnostic tool for machine health monitoring.” *Mech. Syst. Sig. Process.*, 21(2), 824–839.