

# Data mining and statistical analysis of construction equipment failure

Hongqin Fan \*

*Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong, China*

*\* Corresponding author (bshfan@inet.polyu.edu.hk)*

**Purpose** Construction equipment is a key resource, and contractors that own a large equipment fleet take all necessary measures to maximize equipment utilization and minimize equipment failures. Although most contractors implement scheduled maintenance programs and carry out periodic inspections and repairs on their construction equipment, it is still difficult to predict the occurrence of a specific failure of a piece of equipment in the short or long term. According to a survey in the United States, approximately 46% of the major equipment repairs was undertaken as a result of an unexpected failure. Although it is not possible to predict all failure events, a slight improvement in their prediction represents a significant saving in time and cost for a large contractor. Statistical power law models and data-mining models were compared to investigate their pros and cons in predicting critical failure events of heavy construction equipment. **Method** With large amounts of equipment failure data accumulated in a surface mining project, two different types of failure models were created for comparative analysis from a practical point of view. For selected equipment units, failure data were collected along with the relevant factors which may cause variations of equipment failure rate (or mean time to failure). In a classical approach, Power law models of equipment failure rates are fitted using RGA 7.0; while in the data-mining approach, the mean time to failure is modeled using a data-mining algorithm-decision tree induction, establishing logical, mathematical, and statistical relations between MTTF (Mean Time Between Failures) and its various factor of impact (equipment conditions, failure history, environmental conditions, etc.). Both models are used for validation tests on randomly selected time periods and compared in terms of their performance. **Results & Discussion** The two types of models were compared.

**Keywords:** *equipment management, failure analysis, data mining, decision support*

## INTRODUCTION

Construction equipment is a key resource, and contractors owning large equipment fleet take all necessary measures to maximize the equipment utilization and minimize the equipment failures. Although most contractors implement scheduled preventive maintenance programs and carry out periodic inspections/repairs on their construction equipment, it is still a difficult task to predict the occurrence of a specific failure event for a piece of equipment in the short or long term. According to a survey in the United States, approximately 46% of the major equipment repairs followed upon an unexpected failure. As a result of such breakdowns, the equipment unit has to be pulled out of production and repaired on site, or brought to a shop for repair. In addition to the impact on the project, other problems arise from these unexpected failures, include high costs for emergency repairs on a remote jobsite, and high storage-costs for a large number of spare parts. Although it is not possible to predict all failure events, a slight improvement in their prediction represents a significant saving in time and costs for a large contractor.

This paper addresses the predictive analysis on the major failures (critical and catastrophic ones) of construction equipment for a contractor. With large amounts of equipment failure data accumulated in a surface mining project, two different types of failure models are created for comparative analysis from a practical point of view, i.e. classical time-dependent power law models, versus generic time series models. For selected equipment units, their failure data are analyzed along with the relevant influencing factors which may cause variations of equipment failure intensity (or mean time between failures). Through a large number of experimental tests on equipment reliability analysis, it is concluded that classical power law models are easy to apply and are capable of predicting reliability metrics at both the system and subsystem levels of an equipment system with fair results, yet time series models based on predictive data mining algorithms are more flexible, comprehensive, and accurate by taking various influencing factors into account in reliability analysis.

The contributions of the paper are two folded: first, relevant issues are discussed on applying generic

predictive data mining models to time series analysis of equipment reliability, its advantages and disadvantages; second, a systematic comparison is made between classical power law models and generic time series models in terms of their performance and usability in forecasting equipment reliability metrics..

#### **RELATED WORKS**

According to Vorster (2004; 2005), construction equipment involved in any civil engineering and mining works must be managed to minimize unscheduled downtime. Equipment age, reliability, and the repair/maintenance costs are closely related and should be balanced constantly; repair before failure is more cost effective than crisis-based run-to-failure; scattered breakdowns at random inconvenient times have larger impact on planning and activity. If reliability metrics can be predicted with a fair level of accuracy, the decisions on equipment maintenance and repairs can be optimized to reduce on-shift emergency repairs. Smith and Oren (1980) also points out that system reliability estimate strongly influences predicted profitability and customer acceptance.

Reliability is the probability that a component or system will perform a required function for a given period of time when used under stated operating condition (Ebeling 1997), although it is difficult to predict the time at which a piece of equipment fails due to the inherent uncertain nature of failure events and multiple factors of impact, the time-dependent failure events demonstrate some statistical rules and the patterns of trend. Duane proposed the power law model on the failures of a complex repairable system; the accumulated MTBF is linearly related to the operating time on log-log scale [Duane 1964]. Barabady and Kumar (2008) used various statistical distributions including Weibull, exponential, normal, and log normal distribution to analyze the reliability of a crushing plant, in order to identify the bottlenecks in the system and to find the components or subsystems with low reliability for a given designed performance.

Time series is a series of sequenced observations of event data, usually taken in equally spaced time intervals. The theories used for time series analysis have been used for reliability analysis and forecasting of a complex system. For examples, Ho and Xie (1998) used the classical time series analysis method of ARIMA for predicting the number of failures of a mechanical system; Hong and Pai (2006) used Support Vector Machine (SVM), a machine learning algorithm for predicting engine reliability, comparisons were made with power law models, ARIMA, General Regression Neural Network (GRNN) models in terms of their prediction performance. The researchers concluded that, compared with the power

law model, time series models can depict the nonlinear complex relationship among the reliability metrics and these other observations in reliability performance.

#### **PROBLEM STATEMENT**

A contractor's equipment fleet is working on an oil sand mining project on 3-shift schedule around the clock. Among the equipment fleet are dozers, graders, trucks, backhoes etc. The contractor has a team of operators, superintendents, project managers working on the jobsite and keeping full working records of downtime, uptime, failure events, and repair details on each unit. Apart from the preventive maintenance and scheduled overhauls, there are unscheduled random failures on each equipment unit. The contractor is keen to predict the reliability of each unit so that better decisions on allocations of equipment and maintenance resources can be made for scheduling purpose. Although traditional reliability theory can be applied to the heavy equipment in service, there are practical obstacles which make it difficult to apply these reliability modeling techniques originally developed from manufacture industry; the construction environment is highly uncontrollable with constantly changing weather conditions, job natures, and operating conditions, all of which have an impact on the equipment reliability. Each unscheduled critical failure leads to an emergency repair case and causes interruptions to construction works with various financial impact; under some critical failure circumstances, the equipment cannot be repaired on the jobsite and must be brought to a distant shop for extensive repairs.

The contractor has accumulated many years of equipment reliability data along with their history of maintenance and repairs, failure data contains such information as(1) Equipment description: equipment identification, type, model, sub-systems, year of manufacture, odometer and hour meter readings; (2) Equipment downtime and uptime: equipment shut-downs for emergency repairs, scheduled preventive maintenance and overhaul events; (3) Equipment repair details: class of failures, reason down, work done, maintenance personnel (mechanics, electricians, welders, etc. ), working hours, locations. Sample reliability data of a piece of equipment is shown in Table 1.

Construction equipment is a complex system comprising of various subsystem: engine, braking system, hydraulic system, undercarriage, etc., these subsystems and components have different economic lives and different reliability metrics; they are not completely independent and must be kept in working conditions and work in coordination for the equipment to function properly. For each equipment unit,

the contractor is interested in predicting the equipment reliability metrics for the planning period, such as rate of failures, reliability level for the scheduled mission, availability, time between failures, length of uninterrupted working hours without failure given a

minimum reliability level. Predictions at both system level and subsystem levels are desired for management decisions for the upcoming planning periods.

Table. 1. Sample reliability data of an equipment unit in the field

Time Down	Name	Trade	Class	Reason Down	Status	Time Up	Work Done	Down Time	Location	Skill Set	Workforce	Contractor
01/07/01 05:01	HISCO	MECH	Steam	STEAM FOR P2 SERVICE	UP	01/07/01 08:22	3401 COMPLETED STEAMING.	3.35	Steam Bay			Steam Bay Contractor
01/07/01 08:23	ACTIN	MECH	Service	P2 SERVICE	UP	01/07/01 13:09	3302 SERVICE COMPLETED.HRS.4782./WO.3 97905 REPLACE	4.77	Shop	HD Mechanic	2	
01/07/03 10:25	HERRI	MECH	Air Conditioning	AIR CONDITIONING - POLAR AIR	UP	01/07/03 11:48	PRESSURE TESTED & RECHARGED	1.38	Field			Polar Air
01/07/04 01:00	MCCAN	MECH	Drive System	ENGINE OIL DIPSTICK	UP	01/07/04 01:15	RPD ENGINE OIL DIPSTICK, TRANS OIL FILTER DIPSTICK	0.25	Field	HD Mechanic	1	
01/07/17 13:00	KOSTI	MECH	Field Service	FIELD SERVICE	UP	01/07/17 13:20	COMPLETED HOURS 5165	0.33	Field	HD Mechanic	1	
01/07/17 21:50	ANTHO	MECH	Repair Light	HEADLIGHTS NOT WORKING	UP	01/07/17 22:00	REPAIRED WIRING FOR HEADLIGHTS	0.17	Field	HD Mechanic	1	
01/07/19 09:00	RYAN	MECH	Drive System	CHANGE OILS FINAL DRIVES	UP	01/07/19 12:14	COMPLETED	3.23	Shop	HD Mechanic	2	
01/07/19 14:00	FOY	MECH	Air System	NO POWER	UP	01/07/19 14:20	REPAIR ENGINE AIR FILTER	0.33	Field	HD Mechanic	1	

In addition, the project manager needs to identify these frequent failures (failures occur frequently and periodically), cascading failures (one failure causes another), underlying failure causes, and opportunities for improvement.

The predicted reliability metrics can help to optimize the scheduled maintenance of equipment. For examples, preventive maintenance can be rescheduled to reduce failures in service; overhaul decisions can be made to avoid frequent or major failures; maintenance crew and other resources can be properly allocated; equipment can be assigned to the projects according to their predicted performance and project characteristics.

**POWER LAW MODELS AND EQUIPMENT RELIABILITY ANALYSIS**

A piece of construction equipment is considered as a fielded system comprised of many subsystems, components or assemblies with different reliability performance and life cycles. Critical failures of any of them can lead to equipment shutdowns on shift, and

failed components must be fixed to bring the equipment back to work as soon as possible. The overall

equipment system follows a “failure-fix-failure” cycle during operations.

The failure rate of a piece of construction equipment follows a typical “bathtub” curve: the new equipment experiences a burn-in stage with decreased rates of failure in the first half or one year, and then becomes stable in reliability; when the equipment ages, the wear-out stage is entered with increasing rates of failures. Although it is possible to judge the stages of the equipment based on the past experience and recommendations from the manufacturer, the transition point of stages or even the whole life cycle of the equipment varies to a large degree depending on such conditions as the design and manufacturing reliability, equipment use, degree of care and maintenance, repair history etc. As a result, the most reliable approach to identify the three stages of an equipment unit and make predictions on its performance is to perform reliability analysis on its life-to-date failure data.

A power law model indicates that the failures of a complex system are time dependent and follow a Non Homogeneous Poisson Process (NHPP). The power law model was first proposed by Duane in

1962 to describe the failures of a complex system at the stage of development (repeated design tests and improvement in reliability). Duane found that the accumulated MTBF of a system (if the repair time of the system is small compared with MTBF) has a linear relationship with the time if plotted on a log-log paper, with the slope indicating the trend of changes in failure intensity. The power law model could also be used to describe the change of reliabilities of a fielded system in service and make predictions on the failure rates in the upcoming decision periods.

For a unit of construction machine under the policy of minimum repair (just conduct minimum repair to bring the machine back to work), the system failure intensity function can be expressed by a power law model as below:

$$u(t) = \lambda \beta t^{\beta-1} \dots \dots \dots [1]$$

If  $\beta=1$ , the instantaneous failure intensity is a constant, the equipment has stable reliability; if  $\beta < 1$  the equipment is in the burn-in stage, and if  $\beta > 1$  the equipment is in the wear-out stage. Therefore the power law model is a generalization of the homogeneous Poisson process (HPP, Weibull distribution) and allows for change in the intensity function as a repairable system ages [Reliasoft 2012].

For each equipment unit, MTBF is calculated as the accumulated equipment operating time  $t$  divided by the accumulated number of failures up to time  $t$ :  $MTBF = t/N(t)$ . MTBF is then plotted against the operating time  $t$  on a log-log scale paper, which should be approximately a straight line according to the principle of a power law model. An expert tool, RGA 7 by Reliasoft (2012) is used for calculation of best fit line and plotting. Figure 1 shows the MTBF versus time plot for a D11 dozer. It is noticed that although a straight line can be used to fit failure data at the system level, some noisy data exists, due to influences on the arrival pattern of equipment failures from some external factors. This power law plot also shows clearly the equipment MTBF is decreasing with time as this piece of new equipment grows in reliability.

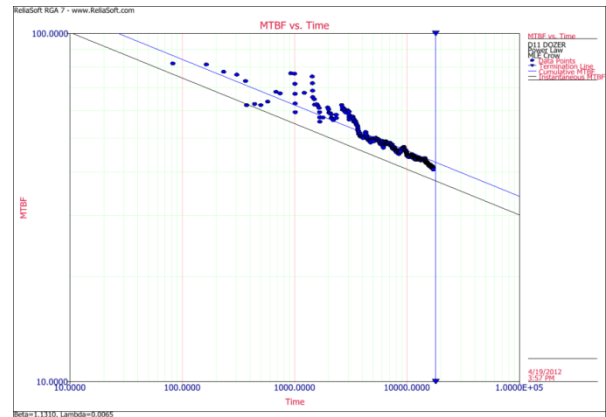


Fig. 1. MTBF versus operating time plot of a D11 Dozer

The same plots can also be made on different sub-systems, such as engine, hydraulic system, air systems, undercarriage etc. as shown in Figure 2. As a result, the MTBF of the equipment system and sub-systems can be predicted using the fitted power law plots for the planning periods. Other reliability metrics including expected number of failures, reliability, expected operating time given reliability level, can also be estimated using the RGA 7 tool.

Prediction results of MTBF and number of failures of D11 Dozer using power law model are shown in Tab. 2 in part. The upper and lower values of MTBF and number of failures with upper and lower confidence of 90% are also shown in the table.

Although it is desirable to apply the power law model into lower level components of the equipment system, say the starter of the engine, to make better decisions on the replace/repair of individual components, data on these components may not be sufficient to have statistical significance. In the meantime, it may not be reasonable to assume complete independence among components for a complex system like construction equipment.

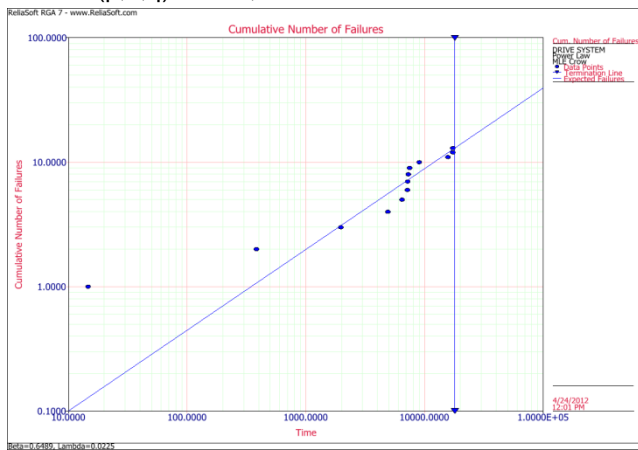
**TIME SERIES MODELS AND EQUIPMENT RELIABILITY ANALYSIS**

Time series data is a sequence of observations taken at equal intervals of time. The purpose of collecting and modeling time series data is to identify the change patterns in data and to forecast the future values assuming the current trend continues. Many reliability metrics of construction equipment can be modeled as time series data, including reliability (%), number of failures, mean time between failures; all these metrics change with consecutive time periods of equipment operations. If reliability variation of a piece of equipment can be captured in a time series model, its future values can be forecasted based on the history of reliability data and related factors of impact.

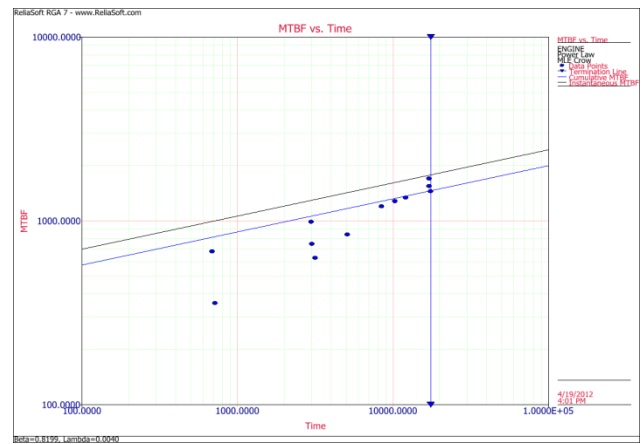
### ARIMA models

Traditionally time series data is analyzed by breaking down into four components (Box and Jenkins 1994): (1) Trend movement, which is the general direction in which a time-series is moving over a long interval of time; (2) Cyclic variations, being the long-term oscillations of the time series about trend; (3) Seasonal variations, which are the seasonal movements of the time series; (4) Irregular variations, which are variations of the time series due to random shocks. AutoRegressive Integrated Moving Average (ARIMA) model is the traditional forecasting method for time series analysis first proposed by Box and Jenkins. In ARIMA (p,d,q) model, a time series data is decom-

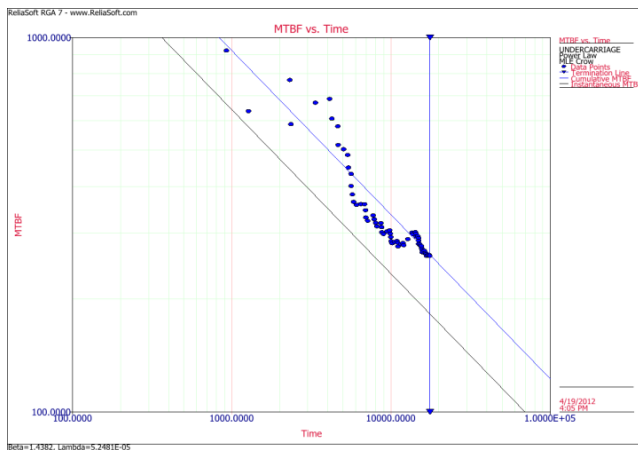
posed into autoregressive (the current value is correlated to the previous p lagged values of the time series), and moving average (the current observation shows random shock from the previous q lagged values of the time series) after integration (differentiation of a time series to make it stationary) if necessary. Seasonal variations and influences from other relevant time series or some intervention variables can also be modeled yet their use is difficult due to requirements on visual judgment and complex statistical tests.



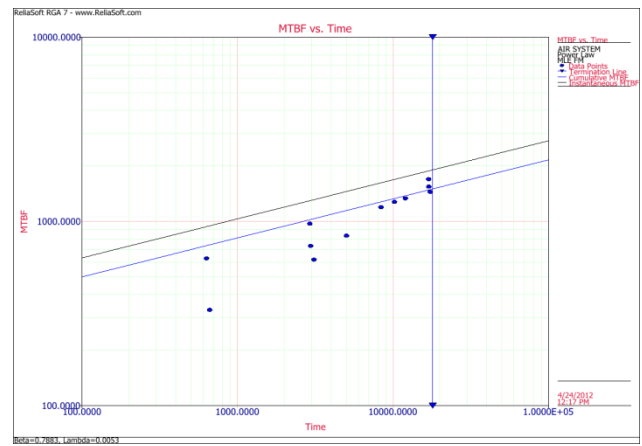
(a) Drive system



(b) Engine



(c) Undercarriage



(d) Air system

Fig.2. MTBF versus time plot of selected subsystems of a D11 Dozer

Table 2. Prediction results of MTBF and number of failures of D11 Dozer using power law model

Period	Actual MTBF	Predicted MTBF			Error	Actual number of failures	Predicted number of failures			Error
		Upper (90%)	Estimated	Lower (90%)			Upper (90%)	Estimated	Lower (90%)	
95	107.13	49.77	44.15	38.87	-62.99	1.00	2.40	2.14	1.91	1.14
96	110.37	49.69	44.08	38.81	-66.28	5.00	2.86	2.54	2.28	-2.46
97	73.42	49.62	44.02	38.75	-29.40	11.00	3.06	2.72	2.44	-8.28
98	194.38	49.55	43.96	38.70	150.43	1.00	3.20	2.84	2.55	1.84
99	188.68	49.48	43.89	38.64	144.79	3.00	3.30	2.94	2.63	-0.06
100	9.52	49.41	43.83	38.59	34.31	7.00	3.39	3.02	2.70	-3.98

**Time series analysis using predictive data mining models**

As a substitute to the classical ARIMA family of models, predictive data mining algorithms can be used to explore the relationship among the data in a time series, relevant time series, and intervention variables in a generic expression:

$$Y_t = f \left( \begin{matrix} Y_{t-1}, Y_{t-2}, \dots, Y_{t-n} \\ X_{1(t-1)}, X_{1(t-2)}, \dots, X_{1(t-n_1)} \\ X_{2(t-1)}, X_{2(t-2)}, \dots, X_{2(t-n_2)} \\ \dots \end{matrix} \right) + \epsilon_t \dots\dots\dots[2]$$

Where

- Y<sub>t</sub> — Current observation
- Y<sub>t-i</sub> — Previous n observations, i=1,2,...,n
- X<sub>i</sub> — Related time series or invention variable i
- X<sub>i(t-j)</sub> — historical observations of related time series or invention variable at (t-j)
- n<sub>i</sub> — correlated lagged values of related time series or invention variable i
- ε<sub>t</sub> — residual of the fitted model

Eqn [2] is a generic expression of using predictive data mining models for time series analysis; there are many data mining algorithms that can be used to explore this underlying relationship in time series data including decision tree, multilayer back propagation neural network, general regression neural network, gene expression programming, etc., and ARIMA (p,d,q) model can be considered as a special case of assuming linear relationship among variables. In general situations, the model in eqn [2] can be represented in combined logical, mathematical, and statistical forms in order to best describe the knowledge hidden in reliability data. See Han and Kamber (2006), Larose (2005) etc., for details on various predictive data mining algorithms.

Compared with the general prediction problems, time series analysis using predictive data mining algorithm has an important feature of autoregression and dependence with the historical observations of relat-

ed time series and intervention variables, as shown in Eqn [2]. The lagged values of both the time series and relevant time series are used as surrogate variables in the model, and the numbers of these lagged variables are selected using information criteria such as how much the inclusion of a lagged variable in the model helps to improve the overall model fit to the time series data. The algorithm of regression tree induction is presented in this paper to compare with the power law models.

**Regression trees models**

Regression trees are learned from data to reflect the postulated relationship between the predicted values and their predictors. This is a supervised learning process: in time series data, if the current observation is to be predicted, all the other determinants, including previous “n” observations, correlated time series, correlated factors or perturbation events, are used as a collection of data space for computer to learn the tree structure.

The regression tree algorithm works as below: the algorithm searches over the data space and recursively partitions it into subspace, where more pure information or promising relations can be found. For example, the regression tree can use a measurement such as information gain or chi-square test to search for most information-rich splitting of data space by an input variable as well as a split-on value so that the partitioned data space contains purer information on the prediction results. At each partitioned data space a regression model is built to predict the outcome.

**Model training, validation and forecasting**

Mean Time between Failures (MTBF) of equipment is computed on a weekly basis, and the regression tree model is learned automatically from the data collected over a two year period. Apart from the MTBF data series, Mean Time to Repair (MTTR), preventative maintenance (PM) data series are used

as predictor series, and overhaul event is modeled as an intervention variable (“1” for overhaul, and “0” for none) in the predictive data mining model. Regression tree model of a D11 dozer is presented here for illustration purpose. Fig. 3 shows the first three levels of the derived regression tree model, each node at a lower level containing more consistent information on MTBF. Tree splitting criteria is learned from data and attached to the bifurcation of tree branches, for example, “Whether the previous MTBF is more than 132.987 hrs” is the first splitting criterion, and “whether time is before Week #74” is used as second criterion for further splitting of nodes into child nodes, and so on. Each leaf node contains a regression formula for MTBF forecasting.

The MTBF time series model is validated by reserving 10% of the collected data. Forecasted values and actual values are compared with part of the results shown in Fig. 4 and Tab. 3. As seen in Fig. 4, the trend of MTBF variation with time is also detected by the algorithm.

Time series forecasting results on MTBF and number of failures are shown in Tab. 3 for comparison with these results from the power law model in Tab. 2. The accuracy of forecasting is improved substantially by using predictive data mining models. The power law model tends to take the mean values with a little consideration to the overall trend, however the time series model can follow both the long term trend and short term variations.

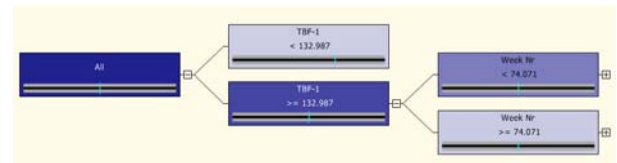


Fig.3. Regression Tree model for forecasting Mean Time between Failures (MTBF) of a D11 Dozer

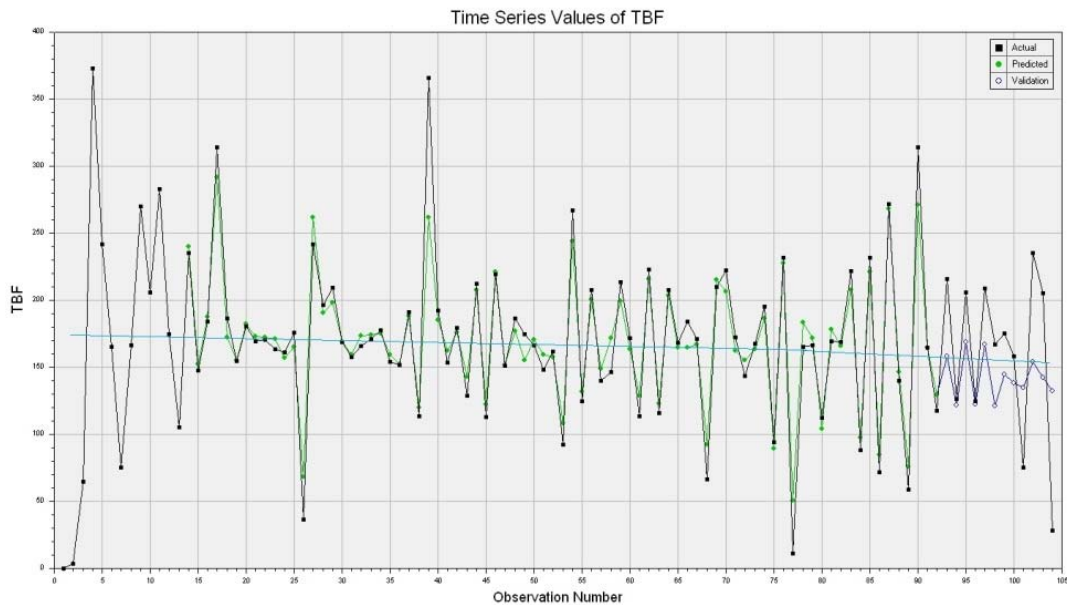


Fig. 4. Forecasting MTBF time series of a D11 dozer

Table 3. Validation on forecasting of MTBF (part of results)

Period	Mean Time between Failures (MTBF)			Number of Failures		
	Actual (hrs)	Predicted (hrs)	Error (hrs)	Actual	Predicted	Error
95	215.78	158.34	57.44	4	4.37	-0.37
96	126.6	121.74	4.86	5	4.12	0.88
97	205.63	169.05	36.59	7	4.79	2.21
98	124.65	122.22	2.43	1	4.63	-3.63
99	209.03	166.98	42.06	5	4.58	0.42
100	166.78	120.98	45.8	11	5.6	5.4

## COMPARISON OF POWER LAW MODELS AND PREDICTIVE DATA MINING MODELS

Both power law models and time series models can be used for forecasting of reliability metrics of construction equipment, their pros and cons are summarized in Tab. 4 from different perspectives.

Table 4. Comparison of power law models and time series models in equipment reliability forecasting

	Power law models	Time series models
Data requirements	Medium to large	Very large
Ability to account for factors of impact	No	Yes
Assumption	Non Homogeneous Poisson Process, random failure process with different intensities at different stages of equipment life	Data series with underlying patterns caused by both randomness and a large number of influencing factors, both internal and external
System and sub-system level modeling	Easy	Difficult to model at subsystem level due to sparse data
Modeling	Fit data into a NHPP process model (extended Weibull distribution)	Use complex computer algorithm to find potential trends, rules and patterns from reliability data
Detecting changes of failure patterns	Yes, to a limited degree	Yes
Complexity of model	Low	Medium to high
Accuracy	Moderate to relatively high	High

## CONCLUSIONS

A reliable fleet of construction equipment is critical to the success of heavy construction projects. Contractors can make informed decisions on equipment maintenance, repairs, replacement if the reliability of equipment can be well understood and future metrics can be forecasted with a satisfactory level of accuracy. Analysis on the field reliability data provides firsthand fact-based information on equipment failure trends, regular and irregular patterns, as well as underlying causes. Classical power law models

and predictive data mining based time series models are compared in forecasting equipment reliability metrics in this paper.

The classical power law models can be applied conveniently to reliability analysis with solid statistical foundations, the simple time-dependent NHPP model is able to identify the changing trends in equipment reliability and predict the reliability metrics of equipment in the planning horizon. However the time series models based on predictive data mining algorithms are more flexible and powerful in creating forecasting models with due consideration to the influencing factors of reliability. Although the data mining algorithms are complex in implementation, there are commercial data mining tools available for explorative analysis through their user-friendly interfaces, model visualization, and interactive features.

Although predictive data mining models can automatically sift through reliability data for pattern recognition, it requires large amount of reliability data in order to produce valid results, which might be difficult for a single piece of equipment within its history of operations in some situations. On the contrary, power law models can make good prediction results under such circumstances, what is more, reliability of the subsystems of the equipment, more desirable by the maintenance crew, can also be predicted using power law models if there is a minimum number of data ensuring statistical significance.

## ACKNOWLEDGEMENT

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 515710)

## References

1. Barabady, J., Kumar, U., "Reliability analysis of mining equipment: A case study of a crushing plant at Jajarm Bauxite Mine in Iran", *Reliability Engineering and System Safety*, Vol. 93, pp. 647–653, 2008
2. Box, G., Jenkins, G., Reinsel, G., "Time series analysis: Forecasting and control", 3<sup>rd</sup> edition, Prentice Hall, 1994.
3. Duane, J.T., "Learning Curve Approach To Reliability Monitoring," *IEEE Transactions on Aerospace*, Vol. 2, pp. 563-566, 1964.
4. Ebeling C. E., "An Introduction To Reliability and Maintainability Engineering", McGraw-Hill (Asia) Co., pp. 5, 1997.
5. Han, J., Kamber, M., "Data mining: concepts and techniques", 2nd edition, Elsevier Inc., 2006



6. Ho, S.L., Xie, M., "The use of ARIMA models for reliability forecasting and analysis", Vol. 35 (1-2), pp. 213-216, 1998
7. Hong, W.C., Pai, P.F., "Predicting engine reliability by support vector machines", *International Journal of Advanced Manufacturing Technology*, Vol. 28 (1-2), pp. 154-161, 2006.
8. Larose, D.T., "Discovering knowledge in data: an introduction to data mining", John Wiley&Sons, 2005
9. ReliaSoft Corporation, "Repairable Systems Analysis",  
[http://www.weibull.com/RelGrowthWeb/Repairable\\_Systems.htm](http://www.weibull.com/RelGrowthWeb/Repairable_Systems.htm), accessed on 14 April, 2012.
10. ReliaSoft Corporation, "RGA: Reliability Growth Analysis and Repairable System Analysis Software Tool",  
[http://www.weibull.com/RelGrowthWeb/Repairable\\_Systems.htm](http://www.weibull.com/RelGrowthWeb/Repairable_Systems.htm), accessed on 14 April, 2012
11. Smith, S.A., Oren, S. S., "Reliability growth of repairable systems", *Naval Research Logistics Quarterly*, vol. 27(4), pp. 539-547, 1980.
12. Vorster, M., "The Critical Balance", *Construction Equipment*, vol. 107(8), pp.69-70, 2004.
13. Vorster, M., "Inspection and Prevention Improve Reliability", *Construction Equipment*, 108(1), pp.91-92, 2005.
14. Bell, J., "A Mixed-use Renaissance", *Mortgage Banking*, Vol. 5(8), pp. 66, 2004.