Development of Equipment Failure Prognostics Model Based on Logical Analysis of Data (LAD)

Alireza Ghasemi, Sasan Esmaeili, Soumaya Yacout

Abstract—**This paper develops an equipment failure prognostics model, in the context of Condition Based Maintenance (CBM), in order to predict the equipment's chance of survival, using Logical Analysis of Data (LAD). LAD has the advantage of not relying on any statistical theory, which enables it to overcome the conventional problems concerning the statistical properties of the datasets. LAD's main advantage is its straightforward procedure and self-explanatory results. In this paper, our main objective is to develop methods to calculate equipment's survival probability at a certain future moment, using LAD. We employ LAD's pattern generation procedure. Then, we introduce a guideline to employ the generated patterns to estimate the equipment's survival probability. The proposed methods are applied on Prognostics and Health Management Challenge dataset, a condition monitoring dataset collected from some mechanical equipment, provided by NASA Ames Prognostics Data Repository. Analysis of performance of the proposed methods reveals that the methods provide comprehensible results that are greatly beneficial to maintenance practitioners. Prognostics results obtained by the proposed methods are compared with that of Proportional Hazards Model (PHM). The comparison reveals that the proposed methods are promising tools that compare favorably to the PH Model. Since the proposed prognostics model is at its beginning phase, some future directions are presented to improve the performance of the model.**

Index Terms— **Condition Based Maintenance (CBM), Logical Analysis of Data (LAD), Prognostics Condition Monitoring.**

I. Introduction

Widely applied in maintenance, Condition Based Maintenance (CBM) [Jardine et al. (2006)] is a maintenance program that engages the equipment's health condition in optimizing or improving the maintenance activities. The equipment's age and health condition indicators are the factors based on which CBM diagnoses a failure in equipment or prognosticates an imminent failure.

Logical Analysis of Data (LAD), first introduced in [Crama et al. (1988)], is a combinatorics, optimization and Boolean logic based methodology for the analysis of datasets. The typical aim of LAD is to extract knowledge hidden in observations of a dataset in order to detect the sets of causes that would lead to certain effects. In the context of maintenance, a cause can be the monitored equipment's age or any health condition indicator value, while an effect can be the equipment's survival or failure. Each cause is called an Attribute. A literal is either an attribute or its Negation. Negation of an attribute contradicts the attribute. Based on certain effects, observations are categorized into two classes: observations that fail during the coming period, referred to as the Positive Class, and observations that survive at least until the end of the coming period, referred to as the Negative Class. A Positive (Negative) Pattern is a set of literals that is reflected in one or more of the observations of the positive (negative) class while not reflected in any (many) of the observations of the negative (positive) class. The number of literals forming the pattern is called the degree of pattern. A pattern cannot be formed of an attribute and its negation.

Since its introduction, LAD has been widely applied for the analysis of datasets from different fields such as medicine, biotechnology, economics, finance, politics, properties, oil exploration, manufacturing and maintenance. [Abramson et al. (2005), G. Alexe et al. (2006), S. Alexe et al. (2003), and Lauer et al. (2002)] applied LAD in medical fields such as cell growth, breast cancer, coronary risk, and electrocardiography in order to predict behavior of medical models. [G. Alexe et al. (2005), and G. Alexe et al. (2004)] used LAD in medical fields such as B-cell lymphoma, and ovarian cancer in order to diagnose medical diseases. [G. Alexe et al. (2008), Boros et al. (2000), A. B. Hammer et al. (1999), P. L. Hammer et al. (2006), P. L. Hammer et al. (2004), and Kim et al. (2008)] applied LAD in various fields such as voting, credit card scoring, housing, labor productivity, country risk, composition of soil in the oil, genotyping, and psychometric in order to discover knowledge from the data and estimate the behavior of the models. [Yacout (2010), Bennane et al. (2012), Mortada et al. (2012), Mortada et al. (2011), Ghasemi et al. (2013)] applied LAD on industrial equipment such as power transformer, oil transformer, aircraft, and rotor bearing in order to diagnose equipment failure. LAD has proved to be a promising technique that provides interpretable results that are comparable to most pioneer techniques in the field of diagnostics in CBM.

Manuscript revised Aug. 19th, 2013.

Dr. Alireza Ghasemi is an assistant professor at Department of Industrial Engineering, Dalhousie University, 5269 Morris Street, Room 300, Halifax NS. e-mail: alireza.ghasemi@dal.ca).

Sasan Esmaeili is a graduate MASc student from Department of Industrial Engineering, Dalhousie University, 5269 Morris Street, Halifax NS.

Dr. Soumaya Yacout is with the Department of Industrial Engineering and Mathematics, École Polytechnique of Montréal.

(Advance online publication: 29 November 2013)

Unlike the earlier applications of LAD in the filed of CBM, our focus is not on the detection of the failure, which is the diagnostics objective of CBM, but on calculation of the probability of failure at certain moment in future, which is the prognostics objective of CBM, and has been comparatively untested. In the following section, we will improve the LAD methodology to predict equipment's chance of survival at each observation moment when new data on attributes of the equipment is available.

II. METHODOLOGY

In this section, we will illustrate LAD's basic steps in the context of CBM. Patterns, LAD's outcomes that characterize the failure and survival characteristics of equipment, will be generated. Then, we will present a guideline to use the generated patterns for equipment's failure prognostics. The failure prognostics model, which represents the relation between the equipment's age and health condition indicators with its failure, will be constructed based on a given historical dataset, called the Train Set. Quality of the failure prognostics model will be examined by applying it on another part of the historical dataset, called the Test Set. The former process is Train Phase, while the latter one is Test Phase.

Table 1 shows a sample train set. The set is composed of the monitored data, including different observation moments, associated with different pieces of equipment, and their corresponding attributes. Each row corresponds to an observation moment, for which the equipment identification and the observation time are respectively shown in the first and the second columns. The third column shows the class of each observation. The last observation moment of each piece of equipment, referred to as the observation that will fail during the current period, is shown with the dark background. The forth and the fifth columns respectively show the measurements of age and condition of equipment. Unlike the earlier applications of LAD in the field of CBM [Yacout (2010), Mortada et al. (2012), and Mortada et al. (2011)], we consider both age and condition of equipment as the equipment's attributes, and use both of them as LAD attributes.

This section is structured as follows: First, we will describe two data binarization methods. Then, we will describe two pattern generation methods. Finally, we will introduce two methods to employ the generated patterns to calculate the survival probability of the equipment from which a new observation is collected.

TABLE I. SAMPLE TRAIN SET

Observations			Attributes	
Equipment ID.	Observation Time	Class	Age	Condition Indicator
1	0	-	0	14
1	1	-	1	16
1	2	-	2	20
1	3	-	3	18
1	4	+	4	20
2	0	-	0	12
2	1	-	1	18
2	2	+	2	22
3	0	-	0	16
3	1	-	1	18
3	2	-	2	20

A. Data Binarization

LAD deals with Boolean attribute values, while in many real life problems, the attribute values may appear in numerical form (e.g. temperature), nominal form (e.g. color), or ordered form (e.g. color describing a traffic light). The binarization procedure transforms each non-binary attribute value into several binary ones, by comparing attribute values to certain thresholds called Cut-Points. According to [Boros et al. (2000)], for each numerical attribute, a binary attribute is associated with every cut-point as following:

$$b_{a,c} = \begin{cases} 1 & ;if \quad a \geq c \\ 0 & ;if \quad a < c \end{cases} \tag{1}$$

Where a is the numerical value of attribute, c is the cut-point value, and $b_{a,c}$ is the binary value of attribute, associated with a and c. As a result, each numerical attribute is converted to n binary attributes, where n is equal to the number of cut-points. In the literature, there are several approaches to define cut-points, from which we will employ Sensitive Discriminating method and Equipartitioning method.

The sensitive discriminating method begins by sorting the attribute values in increasing order. A cut-point is defined as average of two consecutive attribute values, each belonging to different classes. The outcome cut-point represents a threshold, which is able to differentiate between positive and negative classes.

The equipartitioning method also begins by sorting the attribute values in increasing order. The cut-points are defined in such a way that all the attribute values are approximately equally divided into a pre-defined number of intervals. Appropriate number of intervals is selected by comparing the quality of results associated with different values.

B. Pattern Generation

A pattern discriminates one or more of the observations of its class from all or most of the observations of the opposite class. The basic pattern generation algorithms are mainly based on generating all combinations of literals, and examining whether each of the combinations can be considered as a pattern. This results in a huge computational effort.

Recently, some heuristic methods have been introduced that require less computational effort while providing equivalent performance, from which we will employ Mixed Integer Linear Programming (MILP) method and Hybrid Greedy method.

MILP-based pattern generation method, first introduced by [Ryoo (2009)], develops a Mixed Integer Linear Programming and formulates a linear set-covering problem to generate Strong Pure patterns. A pattern is strong if the set of observations covered by the pattern is not a subset of that covered by other patterns. A pattern is pure if it does not cover any observation from the opposite class. The objective of the model is to generate a pattern that leads to the minimum number of observations in certain class, which are not covered by the generated pattern. Then, by reconstructing the previous model, different patterns are generated one by one, up to a point that all the observations are covered by at least one pattern. Although each generated pattern differs from previously generated ones, yet it might cover some or all of the previously covered observations while some remaining observations are still uncovered. This will result in generating redundant pattern while no more uncovered observation gets covered. In order to avoid generating redundant patterns, all the observations that were previously covered by generated patterns will be removed before reconstructing the model. In order to prevent the model from generating the same pattern twice, a new constraint is required to be added to the model, after a pattern is generated.

[Boros et al. (2000)] introduced a heuristic algorithm, called hybrid greedy method, to obtain optimal Prime pure patterns. A pattern is prime if removal of any of its literals results in coverage of observations from the opposite class. The restriction on the generation of pure patterns can be relaxed by allowing the algorithm to cover observations from the opposite class. In this case, a pattern will be defined as a combination of literals covering at least a minimum number of observations of the pattern's class, and at most a maximum number of observations of the opposite class. The numbers are called Coverage and Fuzziness parameters, respectively. The hybrid greedy method is composed of two phases: the first and also the favored phase is the Bottom-Up phase. If any observation is left uncovered by the end of the first phase, the second phase, which is the Top-Down phase, is performed. The bottom-up algorithm starts with only one literal. Then it tries to add as many literals as required up to a point that the combination of literals forms a pattern. The top-down algorithm starts with a combination of literals that certainly is a pattern. Then it tries to remove as many literals as possible from the pattern up to a point where the removal of the remaining literals will result in coverage of observations from the opposite class more than specified fuzziness parameter.

C. Prognostics Model Formulation

Prognostics aim at the detection of the failure at certain moments in the future, which to the author's knowledge has been relatively untested, when using LAD. We will introduce two methods to calculate the conditional survival probability of the equipment, based on the estimated survival functions using Kaplan-Meier (KM) estimation [Kaplan et al. (1958)]. Table 2 shows the generated positive and negative patterns along with their corresponding covered observations based on the sample train set provided in the Table 1.

TABLE II. LIST OF GENERATED PATTERNS BASED ON THE SAMPLE TRAIN SET

+ Pattern	Covered Observations	- Pattern	Covered Observations
PP1	1-3 , 1-4 , 3-3	NP1	1-0 , 1-1 , 1-2 , 2-0 , 2-1 2-0 , 2-2 , 3-0 , 3-1 , 3-2
PP2	2-2 , 3-3	NP2	1-0 , 1-1 , 1-3 , 2-0 , 2-0 2-1 , 3-0 , 3-1

We associate to each pattern p, Pattern Conditional Survival Probabilities $SP_{p(i)}$ for $\forall i \in \{1,2,\ldots,T\}$, which represent the pattern's survival estimation of a piece of equipment for at least i periods, when the equipment's observation is covered by the pattern. T is the maximum available survival period within train set. Since LAD bases its pattern generation on the train set, its ability to prognosticate is limited to the attributes that it has observed in the train phase. In other words, T represents the LAD's maximum perception. KM estimation of pattern conditional survival probability is defined as the proportion of the number of observations covered by pattern P whose corresponding pieces of equipment survived at least i periods after being covered by the pattern, to the total number of observations covered by pattern P.

$$SP_P(i) = \frac{\#(P \cap S; \tau > \tau_0 + i\Delta)}{\#(P \cap S; \tau > \tau_0)} \tag{2}$$

Where S is the set of observations in the train set, and P∩S represents the subset of observations in the train set S that are covered by the pattern P. Function #(N) counts the number of members of the set N. τ is the actual failure time of the corresponding equipment, and $\tau 0$ is the current age of the corresponding equipment at the observation moment when it is covered by pattern P. Δ is the observation period length. Due to the fact that both age and condition of equipment are considered as the equipment's attributes in our study, the above-mentioned survival probability contains the prognostics information based on both age and condition of the equipment. Table 3 shows KM estimation of conditional survival probability of the patterns in the Table 2, based on their corresponding covered observations. For example, SPPP1(1) is equal to 0.333 because PP1 covers observations 1-3, 1-4, and 3-3, but only observation 1-3 has corresponding equipment (i.e. equipment 1) that survives more than one period after being covered by PP1. Both corresponding equipment of observation 1-4 and 3-3 have failed during next period as soon as they are covered by PP1.

TABLE III. KM ESTIMATION OF CONDITIONAL SURVIVAL PROBABILITY OF GENERATED PATTERNS

$i\Delta$	1	2	3	4
PP1	0.333	0	0	0
PP2	0	0	0	0
NP1	0.889	0.667	0.333	0.111
NP2	1	0.714	0.428	0.143

We also defined the Baseline Conditional Survival Probability to indicate time-based survival function, regardless of the equipment's condition. This is taken into consideration for the probable case where no pattern covers an observation. This way, we will be able to calculate the conditional survival probability of the equipment at observation moments, which are not covered by any of the patterns. KM estimation of baseline conditional survival probability is calculated as the proportion of the number of pieces of equipment that survived at least i periods, to the number of all the pieces of equipment in train set.

$$SP_b(i) = \frac{\#(E; \tau > i\Delta)}{\#(E)} \tag{3}$$

Where E is the set of all pieces of equipment in the train set. Table 4 shows KM estimation of baseline conditional survival probability based on all the observations in the train set. SPb(3) equal to 0.667 means that two out of three pieces of equipment in the train set have survived more than 3 periods.

TABLE IV. KM ESTIMATION OF BASELINE CONDITIONAL SURVIVAL PROBABILITY

$i\Delta$	1	2	3	4
$SP_b(i)$	1	1	0.667	0.333

Considering the mentioned conditional survival probabilities, we introduce two methods to calculate the conditional survival probability of the equipment from which a new observation is collected. Table 5 shows a sample test set along with the list of patterns that cover each observation.

TABLE V. SAMPLE TEST SET

Observations		Attributes		Covering Patterns
Equip. ID.	Observation Time	Age	Condition Indicator	
1	0	0	14	NP1 , NP2
1	1	1	16	NP1 , NP2
1	2	2	20	NP1
1	3	3	22	PP1 , PP2

The first method favors the Pattern Conditional Survival Probability (SP_p), while it takes into account the ones that were calculated for the equipment based on observations at previous observation moments (SP_{former}), less weightily. It also contains the Baseline Conditional Survival Probability (SP_b). Defining n as the number of patterns that cover an observation, the conditional survival probability of the equipment for i periods is calculated as follows:

$$SP_{obs}(i) = \begin{cases} \dfrac{\sum_{p=1}^{n} SP_p(i)+SP_b(i)}{n+1} & ;if \quad t=0 \\ \dfrac{\sum_{p=1}^{n} SP_p(i)+SP_{former}(i+1)}{n+1} & ;if \quad t>0 \end{cases} \qquad (4)$$

Using the 1st method, as introduced in eq. (4), the conditional survival probabilities of the equipment at different observation moments are shown in Table 6. $SP_{obs(2)}$ for 1-0 is equal to 0.794 because the observation 1-0 is covered by patterns NP1 and NP2 for which $SP_{NP1(2)}$ and $SP_{NP2(2)}$ are equal to 0.667 and 0.714 respectively (see Table 3), and SPb(2) is equal to 1 (see Table 4). As a result $SP_{obs(2)}$ for 1-0 is equal to (0.667 + 0.714 + 1) / 3 = 0.794. $SP_{former(1)}$ for 1-1 is also equal to 0.794 because its corresponding equipment was formerly predicted to survive for at least 2 periods with the probability of 0.794 ($SP_{obs(2)}$ for 1-0 is 0.794). Since all the train data failed before the fifth period, the fourth period is considered as the LAD's maximum perception, and the probability of survival more than four periods is equal to zero.

TABLE VI. CONDITIONAL SURVIVAL PROBABILITY OF SAMPLE TEST EQUIPMENT AT DIFFERENT OBSERVATION MOMENTS – FIRST CALCULATION METHOD

Obs	Covering Patterns	$\Sigma\, SP_p(t)$				$SP_b(t)$			
		1	2	3	4	1	2	3	4
1-0	NP1,NP2	1.89	1.38	0.76	0.25	1	1	0.67	0.33
1-1	NP1,NP2	1.89	1.38	0.76	0.25	-	-	-	-
1-2	NP1	0.89	0.67	0.33	0.11	-	-	-	-
1-3	PP1,PP2	0.33	0	0	0	-	-	-	-

Obs	Covering Patterns	$SP_{former}(t)$				$SP_{obs}(t)$			
		1	2	3	4	1	2	3	4
1-0	NP1,NP2	-	-	-	-	0.96	0.79	0.48	0.19
1-1	NP1,NP2	0.79	0.48	0.19	0	0.89	0.62	0.32	0.08
1-2	NP1	0.62	0.32	0.08	0	0.76	0.5	0.21	0.06
1-3	PP1,PP2	0.5	0.21	0.06	0	0.28	0.07	0.02	0

The second method also prefers the latest observation to older observation. But, it considers equal weight for Pattern and Baseline Conditional Survival Probabilities. The conditional survival probability of the equipment at current observation moment is calculated as follows:

$$SP_{obs}(i) = \begin{cases} \dfrac{\frac{\sum_{p=1}^{n} SP_p(i)}{n}+SP_b(i)}{2} & ;if \quad t=0 \\ \dfrac{\frac{\sum_{p=1}^{n} SP_p(i)+SP_{former}(i+1)}{n+1}+SP_b(i)}{2} & ;if \quad t>0 \end{cases} \qquad (5)$$

Using the 2nd method, as introduced in eq. (5), the conditional survival probabilities of the equipment at different observation moments are shown in Table 7. $SP_{b(1)}$ for 1-3 is equal to 0.5 because one out of two pieces of equipment that have survived more than 3 periods, has survived more than 4 periods.

TABLE VII. CONDITIONAL SURVIVAL PROBABILITY OF SAMPLE TEST EQUIPMENT AT DIFFERENT OBSERVATION MOMENTS – SECOND CALCULATION METHOD

Obs	Covering Patterns	$\Sigma\, SP_p(t)$				$SP_b(t)$			
		1	2	3	4	1	2	3	4
1-0	NP1,NP2	1.89	1.38	0.76	0.25	1	1	0.67	0.33
1-1	NP1,NP2	1.89	1.38	0.76	0.25	1	0.67	0.33	0
1-2	NP1	0.89	0.67	0.33	0.11	0.67	0.33	0	0
1-3	PP1,PP2	0.33	0	0	0	0.5	0	0	0

Obs	Covering Patterns	$SP_{former}(t)$				$SP_{obs}(t)$			
		1	2	3	4	1	2	3	4
1-0	NP1,NP2	-	-	-	-	0.97	0.85	0.53	0.23
1-1	NP1,NP2	0.85	0.53	0.23	0	0.96	0.65	0.33	0.04
1-2	NP1	0.65	0.33	0.04	0	0.72	0.42	0.09	0.03
1-3	PP1,PP2	0.42	0.09	0.03	0	0.38	0.02	0	0

III. EXPERIMENTS

We applied the LAD methodology on *Prognostics and Health Management Challenge* dataset, a condition monitoring dataset provided by *NASA Ames Prognostics Data Repository*. The dataset consists of approximately 46,000 observations associated with 218 pieces of mechanical equipment. For each observation, 3 operational settings and 21 measurements associated with the equipment's attributes are provided.

In order to model the system in a reasonable time, we had to decrease the dataset size. To do so, we extracted every 10th observation and reduced the number of observations to about 4,600. Correlation analysis reveals that most of the attributes are highly correlated. Involving correlated attributes in the model is not appropriate due to two main reasons: First and foremost, correlated attributes do not provide any additional information. Second, the more attributes involved, the more modeling time is required. In order to remove the effect of involving trivial attributes, we applied the Principal Component Analysis (PCA). PCA [Pearson (1901)] is a mathematical method that converts a multi-attribute dataset into a dataset of linearly uncorrelated attributes by extracting only the most informative attributes. PCA shows that the first two principal components convey more than 97% of characteristics of all the attributes before conversion. So, we constructed the model based on these two attributes as a substitute for system's 21 attributes.

From the dataset of 218 pieces of equipment, a dataset of 15 pieces of equipment was extracted to test the performance of the model, and from the remaining pieces of equipment, 10 datasets of 70 randomly extracted pieces of equipment were generated to train the model. We constructed 10 models based on different train datasets, and for each model, we calculated the conditional survival probabilities based on the same test dataset. This enables us to compare the conditional survival probability of each matched pair observation obtained by different models while eliminating the difference (error) due to randomness of different random test data. Finally, the prognostics results provided by all the 10 models were averaged over the models. The final analysis of the performances was performed using the averages.

The proposed LAD prognostics model was entirely coded in the Python programming language, and all of the steps were carried out automatically. The inputs and outputs interfaced through the Excel spreadsheets. The MILP model was solved, using the CPLEX optimization software package module for Python.

Table 8 shows the prognostics results for a test piece of equipment based on the second conditional survival probability calculation method at 21 consecutive observation moments. Before getting to the 22nd observation moment, the equipment has failed. Each row corresponds to an observation moment, for which the conditional probabilities of survival up to 1 to 5 periods later are respectively shown in the columns 5 to 9. For instance, the 16th row shows the conditional probabilities of survival up to the period 17th to 21st, based on the equipment's attributes at the 16th observation moment. The conditional probabilities of survival up to the period 17th to 21st are respectively equal to 0.923, 0.878, 0.799, 0.732, and 0.665.

However, the set of conditional probabilities of survival for the future predictable periods is not meaningfully comparable to its matched pair set provided by other experiments. Therefore, we transformed the information of the set into a single comparable value, Mean Residual Life (MRL), so that we can compare performance of different experiments. MRL represents the expected value of equipment residual life, and is formulated as following [Banjevic et al. (2006)]:

$$MRL = \sum_{i=1}^{\infty} i\Delta \times Probability(\tau > \tau_0 + i\Delta | \tau > \tau_0) \quad (6)$$

Where Probability ($\tau > \tau0 + i\Delta | \tau > \tau0$) shows the probability of survival for at least i periods, knowing that the equipment has not failed until $\tau0$. This conditional probability is identical with conditional survival probability SPobs(i), introduced in this work. So, the MRL is formulated in terms of SPobs(i) as following:

$$MRL = \sum_{i=1}^{\infty} i\Delta \times SP_{obs}(i) \quad (7)$$

TABLE VIII. PROGNOSTICS RESULTS FOR A TEST PIECE OF EQUIPMENT

Observations				Conditional Survival Probability				
ID	Age	Cond. 1	Cond. 2	X > Δ	X > 2Δ	X > 3Δ	X > 4Δ	X > 5Δ
1	0	0.059	0.011	0.99	0.990	0.978	0.961	0.939
1	1	0.044	0.078	0.99	0.986	0.973	0.953	0.933
1	2	0.004	0.000	0.99	0.984	0.965	0.948	0.917
1	3	0.009	0.023	0.99	0.983	0.970	0.948	0.924
1	4	0.005	0.012	0.99	0.986	0.972	0.955	0.931
1	5	0.059	0.011	0.99	0.986	0.971	0.951	0.924
1	6	0.008	0.024	0.99	0.978	0.961	0.929	0.902
1	7	0.004	0.000	0.99	0.982	0.960	0.942	0.909
1	8	0.016	0.000	0.99	0.983	0.969	0.950	0.924
1	9	0.016	0.000	0.99	0.984	0.970	0.953	0.926
1	10	0.009	0.024	0.99	0.982	0.969	0.946	0.923
1	11	0.043	0.078	0.99	0.983	0.968	0.947	0.924
1	12	0.008	0.024	0.99	0.976	0.958	0.925	0.897
1	13	0.004	0.000	0.99	0.975	0.946	0.923	0.884
1	14	0.060	0.011	0.99	0.967	0.946	0.913	0.870
1	15	0.043	0.076	0.98	0.950	0.913	0.862	0.815
1	16	0.004	0.013	0.92	0.878	0.799	0.732	0.665
1	17	0.004	0.013	0.96	0.890	0.793	0.702	0.609
1	18	0.007	0.000	0.85	0.740	0.647	0.569	0.472
1	19	0.007	0.027	0.82	0.656	0.540	0.429	0.367
1	20	0.063	0.008	0.87	0.667	0.524	0.403	0.301

Observations				Residual Life		
Eqp. ID.	Age	Cond.1	Cond.2	MRL	RL	Diff.
1	0	0.059	0.011	14.80	20	- 5.20
1	1	0.044	0.078	13.87	19	- 5.13
1	2	0.004	0.000	13.46	18	- 4.54
1	3	0.009	0.023	13.43	17	- 3.57
1	4	0.005	0.012	13.56	16	- 2.44
1	5	0.059	0.011	13.49	15	- 1.51
1	6	0.008	0.024	12.96	14	- 1.04
1	7	0.004	0.000	13.27	13	0.27
1	8	0.016	0.000	13.41	12	1.41
1	9	0.016	0.000	13.43	11	2.43
1	10	0.009	0.024	13.33	10	3.33
1	11	0.043	0.078	13.45	9	4.45
1	12	0.008	0.024	12.81	8	4.81
1	13	0.004	0.000	12.65	7	5.65
1	14	0.060	0.011	12.14	6	6.14
1	15	0.043	0.076	10.91	5	5.91
1	16	0.004	0.013	8.22	4	4.22
1	17	0.004	0.013	7.37	3	4.37
1	18	0.007	0.000	6.24	2	4.24
1	19	0.007	0.027	4.82	1	3.82
1	20	0.063	0.008	4.17	0	4.17

In the Table 8, associated with each observation moment, the MRL and the actual *Residual Life* (*RL*) are calculated and shown in columns 10 and 11, respectively. The actual RL is not determined until the equipment failure moment. At this moment, the equipment lifetime is determined as the time difference between the failure moment and installation moment. In the above sample, the equipment lifetime is equal to 20. Then, for each observation moment, the actual RL is calculated by subtracting the observation moment from the equipment lifetime. For instance, the actual RL associated with the 16th observation moment is equal to 4. At each observation moment, the difference between the MRL and the actual RL is calculated and shown in column 12. The lower the difference, the better the performance of the model. Figure 1 shows a comparison between the MRL and the actual RL of the equipment. In early observation moments, the model underestimates the MRL. As time passes by, the MRL gets closer to the actual RL, and the model correctly estimates the MRL almost at the mid-age observation moments. Later, when getting closer to the actual failure moment, the model overestimates the MRL.

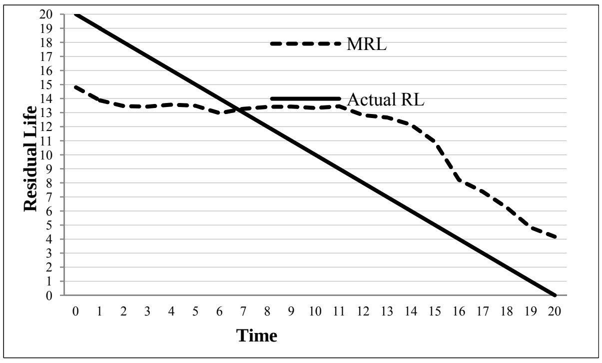

Fig. 1. A Comparison between MRL and Actual RL

Our prognostics model formation generally has two phases: At the first phase, data binarization, using either the sensitive discriminating or the equipartitioning method, is performed. The performance of the equipartitioning method depends on a pre-defined number of cut-points. At the second phase, survival analysis is performed based on the patterns generated by either the MILP or the hybrid greedy method. The performance of the hybrid greedy method depends on the pre-defined degree, coverage and fuzziness parameters. Table 9 shows our *Design of Experiments* (*DOE*). We have examined 5 parameter settings for the data binarization phase, and 13 parameter settings for the survival analysis phase. We have also included the Weibull PHM [Banjevic et al. (2006)] to calculate conditional survival probabilities to compare different LAD settings performances with that of PHM. As a result, (5×14=) 70 experiments were designed in order to compare the performance of the model based on different parameters and methods. Each experiment was performed 10 times based on different train sets and the results were averaged over all the 10 runs.

The absolute value of differences between the MRL and the actual RL indicates the accuracy of the experiment. The lower the difference, the more accurate the experiment. So, the measurement under study in the DOE is the absolute value of differences between the MRL and the actual RL.

Let $X = \{1,2,\ldots,U\}$ be defined as the test set, where U is the number of observations in the test set. In order to compare the performance of different experiments, first we associate the set $Z^e = (z_1^e, z_2^e, \ldots, z_U^e)$ with experiment e. The n^{th} member of the set, z_n^e is formulated as $|MRL_n^e - RL_n|$, where MRL_n^e is the estimated MRL by experiment e for observation n, and RL_n is the actual RL for observation n. Then, experiments $1,2,\ldots,m$ are compared based on the sets $Z^1, Z^2, \ldots, Z^m$. To do so, members of the sets are compared pair by pair, using the *Friedman Matched-pair Test* (*Dunn's Multiple Comparison Test*) [Friedman (1940)].

The comparison is structured as follows: First, we compare two conditional survival probability calculation methods introduced in *eq*. (4) and *eq*. (5). Second, different hybrid greedy methods are compared. Third, the best hybrid greedy method is compared with the MILP and the PHM methods. Comparison of the two methods of conditional survival probability calculation reveals that the second method that equally prefers the baseline and pattern survival probability (*eq*. (5)), statistically outperforms the first method that prefers the pattern survival probability (*eq*. (4)), in all 70 experiments.

Comparison of the hybrid greedy methods reveals that the one with the parameters *coverage* ≥ 10% and *fuzziness* = 0 outperforms the other methods. Table 10 shows the comparison between the mean values of different hybrid greedy methods. The method with the parameters *coverage* ≥ 10% and *fuzziness* = 0 provides the lowest mean value although the differences are not statistically significant.

TABLE IX. DESIGN OF EXPERIMENTS (DOE)

Data Binarization	Parameters
Sensitive Discriminating	-
Equipartitioning	# cut-points = 20
	# cut-points = 30
	# cut-points = 40
	# cut-points = 50
Survival Analysis	**Parameters**
PHM	-
MILP	-
Hybrid Greedy	d = 3 coverage > 10% fuzziness <=0
	d = 3 coverage > 10% fuzziness <=1
	d = 3 coverage > 10% fuzziness <=2
	d = 3 coverage > 20% fuzziness <=0
	d = 3 coverage > 20% fuzziness <=1
	d = 3 coverage > 20% fuzziness <=2
	d = 3 coverage > 30% fuzziness <=0
	d = 3 coverage > 30% fuzziness <=1
	d = 3 coverage > 30% fuzziness <=2
	d = 3 coverage > 40% fuzziness <=0
	d = 3 coverage > 40% fuzziness <=1
	d = 3 coverage > 40% fuzziness <=2

TABLE X. HYBRID GREEDY #1 VS. … VS. HYBRID GREEDY #12

	C>0.1 F <=0	C>0.1 F <=1	C>0.1 F <=2	C>0.2 F <=0	C>0.2 F <=1	C>0.2 F <=2
Sensitive Discriminating	3.751^{+}	3.774	3.78	3.783	3.789	3.758
Equipartitioning (n=20)	3.748^{+}	3.762	3.826	3.778	3.816	3.835
Equipartitioning (n=30)	3.731^{+}	3.753	3.754	3.77	3.792	3.835
Equipartitioning (n=40)	3.723^{+}	3.727	3.751	3.738	3.749	3.753
Equipartitioning (n=50)	3.728^{+}	3.739	3.788	3.75	3.734	3.818

$^{+}$ *Minimum mean value*

	C>0.3 F <=0	C>0.3 F <=1	C>0.3 F <=2	C>0.4 F <=0	C>0.4 F <=1	C>0.4 F <=2
Sensitive Discriminating	3.788	3.775	3.783	3.796	3.775	3.787
Equipartitioning (n=20)	3.812	3.839	3.822	3.833	3.842	3.857
Equipartitioning (n=30)	3.778	3.778	3.829	3.786	3.795	3.785
Equipartitioning (n=40)	3.758	3.731	3.732	3.779	3.738	3.775
Equipartitioning (n=50)	3.751	3.766	3.754	3.763	3.772	3.753

Comparison of the best hybrid greedy, the MILP, and the PHM reveals that the PHM statistically outperforms the LAD methods. While the hybrid greedy and the MILP methods are not statistically different. Table 11 shows the comparison between the mean values of the three methods.

TABLE XI. HYBRID GREEDY VS. MILP VS. PHM

	Hybrid Greedy	MILP	PHM
Sensitive Discriminating	3.751	3.811	3.507^{+}
Equipartitioning (n=20)	3.748	3.867	3.51^{+}
Equipartitioning (n=30)	3.731	3.826	3.509^{+}
Equipartitioning (n=40)	3.723	3.801	3.51^{+}
Equipartitioning (n=50)	3.728	3.801	3.51^{+}

$^{+}$ *Minimum mean value*

Figure 2 shows the difference between the MRL and the actual RL of a test piece of equipment, using the best models provided by the hybrid greedy, the MILP, and the PHM methods. It shows that the LAD methods underestimate the MRL at the early observation moments (pessimistic outlook about the equipment future). As time passes by, the estimations get closer to the actual RL, and they correctly estimate the MRL between 4th and 7th observations. Later, when getting closer to the actual failure moment, they overestimate the MRL (optimistic outlook about the equipment future). It can be concluded that the LAD methods have neither a constant optimistic outlook nor a constant pessimistic outlook about the equipment future, whereas they adjust their outlook over the equipment lifetime. Contrary to the LAD methods, the PHM method always overestimates the MRL by at least one period (optimistic outlook about the equipment future). It also reveals that the PHM method is stable at the early observation moments (see zone A), while the LAD methods are stable when getting closer to the actual failure moment (see zone B).

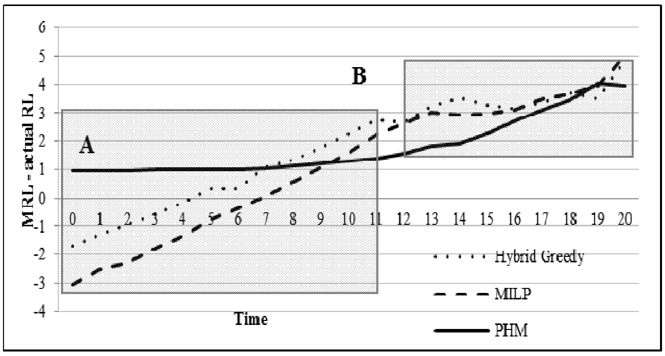

Fig. 2. Difference between MRL and Actual RL using Hybrid Greedy, MILP, and PHM Methods

Table 12 shows the comparison between the average run-time of the three methods. It is shown that the run-time, as was expected, increases as the number of cut-points increases. In the train phase, the PHM method runs faster in comparison with the LAD methods. In the test phase, the LAD methods run faster than the PHM method. In total, the PHM method comparatively runs faster in 4 out of 5 cases, while in the remaining case, the MILP method runs faster. Among the LAD methods, the MILP method runs, as was expected, much significantly faster than the hybrid greedy method.

IV. CONCLUSION

In this paper, we developed an equipment failure prognostics model by employing the Logical Analysis of Data (LAD). We improved the LAD methodology to predict equipment's chance of survival at each observation moment when new data on the equipment health condition indicators is collected. The LAD model was applied on the Prognostics and Health Management Challenge dataset, a condition monitoring dataset provided by NASA Ames Prognostics Data Repository. Analysis of performance of the LAD model revealed that it provides comprehensible results that are greatly beneficial to maintenance practitioners. Prognostics results obtained using the LAD model, were compared with that using PH Model. Following results are only based on one example and need to be investigated further.

Comparison with respect to the accuracy of estimated MRL showed that: The conditional survival probability calculation method that equally favors the baseline and pattern survival probabilities statistically outperformed the one that prefers the pattern survival probability. The hybrid greedy method with the parameters coverage >10% and fuzziness= 0 statistically outperformed other hybrid greedy methods. The PHM method statistically outperformed the both LAD methods. Also, it is noticed that the performances of the LAD model is highly sensitive to its defined survival function. However, the LAD's results are highly interpretable and easy to understand, which is of great value for maintenance practitioners.

Comparison with respect to the run-time showed that: Fewer cut-points is preferred due to the fact that the accuracy of prognostics did not significantly depend much on the number of cut-points at the tested levels. In the train phase, the PHM method ran faster than the LAD methods, while in the test phase, the LAD methods ran faster than the PHM method. In 4 out of 5 cases, the PHM method ran faster in total. Among the LAD methods, the MILP method ran much significantly faster than the hybrid greedy method. Since the LAD methods were not statistically different, the MILP is preferred to the hybrid greedy, due to faster result achievement.

Our results also showed that the PHM method has an optimistic outlook about the equipment's survival. The LAD methods have neither constant optimistic nor constant pessimistic outlooks about the equipment's survival, whereas their outlooks change gradually from pessimistic to optimistic, as the equipment health deteriorates over its lifetime. The PHM method is more stable at the early observation moments, while the LAD method stabilizes when the equipment gets older.

The LAD model has the advantage of not relying on any statistical theory, which enables it to overcome the conventional problems concerning the statistical properties of the datasets. Its main advantage is its straightforward process and self-explanatory results, which are greatly beneficial to maintenance practitioners.

Since the proposed LAD model is at its beginning phase, further research is required to improve the performance of the model. Due to the fact that the performances of the proposed calculation methods are highly sensitive to the defined survival function, a future research direction is to improve the survival function to reflect equipment's probable failure better. Due to the fact that the PH Model and the LAD model are stable at the early and the late observation moments, respectively, another future research direction is to investigate a hybrid LAD-PHM Model to benefit from both models' advantages. Another future research direction is to develop a technique to calibrate the LAD model to adjust for both underestimation and overestimation.

REFERENCES

[1] S. Abramson, G. Alexe, P. L. Hammer, D. Knight and J. Kohn. A computational approach to predicting cell growth on polymeric biomaterials. *Journal of Biomedical Material Research 73A(1),* pp. 116-124. 2005.

[2] G. Alexe, S. Alexe, D. E. Axelrod, T. O. Bonates, I. Lozina, M. Reiss and P. L. Hammer. Breast cancer prognosis by combinatorial analysis of gene expression data. *Breast Cancer Research 8(4),* pp. R41. 2006.

[3] G. Alexe, S. Alexe, D. E. Axelrod, P. L. Hammer and D. Weissmann. Logical analysis of diffuse large B-cell lymphomas. *Artificial Intelligence in Medicine 34(3),* pp. 235-267. 2005.

[4] G. Alexe, S. Alexe, P. L. Hammer and A. Kogan. Comprehensive vs. comprehensible classifiers in logical analysis of data. *Discrete Applied Mathematics 156(6),* pp. 870-882. 2008.

[5] G. Alexe, S. Alexe, L. A. Liotta, E. Petricoin, M. Reiss and P. L. Hammer. Ovarian cancer detection by logical analysis of proteomic data. *Proteomics 4(3),* pp. 766-783. 2004.

[6] S. Alexe, E. Blackstone, P. L. Hammer, H. Ishwaran, M. S. Lauer and C. E. Pothier S. Coronary risk prediction by logical analysis of data. *Annals of Operations Research 119(1),* pp. 15-42. 2003.

[7] D. Banjevic and A. K. S. Jardine. Calculation of reliability function and remaining useful life for a markov failure time process. *IMA Journal of Management Mathematics 17(2),* pp. 115-130. 2006.

[8] A. Bennane and S. Yacout. LAD-CBM; new data processing tool for diagnosis and prognosis in condition-based maintenance. *Journal of Intelligent Manufacturing 23(2),* pp. 265-275. 2012.

[9] E. Boros, P. L. Hammer, T. Ibaraki, A. Kogan, E. Mayoraz and I. Muchnik. An implementation of logical analysis of data. *Knowledge and Data Engineering, IEEE Transactions on 12(2),* pp. 292-306. 2000.

[10] Y. Crama, P. L. Hammer and T. Ibaraki. Cause-effect relationships and partially defined boolean functions. *Annals of Operations Research 16(1),* pp. 299-326. 1988.

[11] M. Friedman. A comparison of alternative tests of significance for the problem of *m* rankings. *The Annals of Mathematical Statistics 11(1),* pp. 86-92. 1940.

[12] Alireza Ghasemi, Sasan Esmaeili, Optimal Condition-Based Maintenance Replacement based on Logical Analysis of Data (LAD), *Lecture Notes in Engineering and Computer Science: Proceedings of The International Multi Conference of Engineers and Computer Scientists 2013, 13-15 March, 2013, Hong Kon, pp830-833*.

[13] A. B. Hammer, P. L. Hammer and I. Muchnik. Logical analysis of chinese labor productivity patterns. *Annals of Operations Research 87(0),* pp. 165-176. 1999.

[14] P. L. Hammer, A. Kogan and M. A. Lejeune. Modeling country risk ratings using partial orders. *European Journal of Operational Research 175(2),* pp. 836-859. 2006.

[15] P. L. Hammer, Y. Liu, B. Simeone and S. Szedmak. Saturated systems of homogeneous boxes and the logical analysis of numerical data. *Discrete Applied Mathematics 144(1-2),* pp. 103-109. 2004.

[16] A. K. S. Jardine, D. Lin and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing 20(7),* pp. 1483-1510. 2006.

[17] E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association 53(282),* pp. 457-481. 1958.

[18] K. Kim and H. S. Ryoo. A LAD-based method for selecting short oligo probes for genotyping applications. *OR-Spectrum 30(2),* pp. 249-268. 2008.

[19] M. S. Lauer, S. Alexe, P. Snader Claire E., E. H. Blackstone, H. Ishwaran and P. L. Hammer. Use of the logical analysis of data method for assessing long-term mortality risk after exercise electrocardiography. *Circulation 106(6),* pp. 685-690. 2002.

[20] M. A. Mortada, T. Carroll, S. Yacout and A. Lakis. Rogue components: Their effect and control using logical analysis of data. *Journal of Intelligent Manufacturing 23(2),* pp. 289-302. 2012.

[21] M. A. Mortada, S. Yacout and A. Lakis. Diagnosis of rotor bearings using logical analysis of data. *Journal of Quality in Maintenance Engineering 17(4),* pp. 371-397. 2011.

[22] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6 2(11),* pp. 559-572. 1901.

[23] H. S. Ryoo. MILP approach to pattern generation in logical analysis of data. *Discrete Applied Mathematics 157(4),* pp. 749-761. 2009.

[24] S. Yacout. Fault detection and diagnosis for condition based maintenance using the logical analysis of data. Presented at Computers and Industrial Engineering (CIE), 2010 40th International Conference on. 2010.