Risk Level Estimation for Electronics Boards in Drilling and Measurement Tools Based on the Hidden Markov Model

Jinlong Kang*[1,2], Christophe Varnier[1], Ahmed Mosallam[2], Noureddine Zerhouni[1], Fares Ben Youssef[3], and Nannan Shen[2]

[1]AS2M, FEMTO-ST, ENSMM, UBFC, Besançon, 25000, France
[2]Schlumberger, Clamart, 92140, France
[3]Schlumberger, Youngsville, 70592, United States

June 17, 2022

Abstract

The electronic boards in drilling and measurement (D&M) tools provide multiple functions, such as data acquisition, signal processing, operation control, and data storage. However, due to the harsh downhole operating conditions; i.e., high temperature, dynamic vibration, and extensive shocks, the boards are likely to suffer from complex failure modes and result in failed jobs. Estimating the risk level of the boards can tolerate and provide support for maintenance decision making and job planning, this paper presents a statistical method for risk assessment of the electronic boards. The method first selects relevant channels from D&M tool measurement data and extracts histogram features based on those selected channels. The histogram features are then enhanced based on a linear interpolation method and aggregated using weighted sum. Finally, hidden Markov models (HMMs) with different parameter settings are trained using the processed features. The best HMM is chosen according to the Akaike information criterion and Bayesian information criterion. The proposed HMM-based method is tested on a real-world data set of failed control processing unit boards that were assembled for a specific D&M tool. The experimental results show that this method is effective in estimating the risks as a sequence of events, which in turn, helps to achieve consistent risk estimation. The work presented in this paper is also part of a long-term project with the aim to construct a risk-based decision advisor for D&M tools used in the oil and gas industry.

Keywords— risk-level estimation, hidden Markov model, environmental exposure, electronic board, drilling and measurement tool

*jinlong.kang@femto-st.fr

1 Introduction

The bottomhole assembly (BHA) is an important device in an oil and gas industry drilling system. The BHA must provide power for the bit to rotate and crush the rock, survive a harsh operating environment, and provide accurate directional control of the well [1]. Frequently, a BHA consists of several drilling and measurement (D&M) tools; e.g., measurement-while-drilling, logging-while-drilling, and rotary steering system tools. A large number of electronic boards are built into D&M tools to facilitate multiple functions of a BHA. The reliability of the electronic boards is essential to the success of a drilling job [2] and one way to improve the reliability is tool maintenance prior to a failure [3].

Currently, tool maintenance is mainly time-based, which could result in either passive maintenance activities due to tool failures or unnecessary maintenance activities such as replacing a component too early. To date, limited research on advanced maintenance strategies such as condition-based maintenance, predictive maintenance, and risk-based maintenance have taken place to D&M tools. To fill some of the gaps, the authors of this paper initiated a project to develop risk-based maintenance decision approaches for D&M tools. Risk-based maintenance consists of three phases; i.e., risk estimate, risk evaluation, and decision making [4]. This paper presents only the first phase; i.e., a risk-estimate method for D&M tool electronic boards. In the context of maintenance, risk is usually defined as the probability of failure times the consequence of the failure [4]. In this paper, the authors assume the consequence of the failure is the same; i.e., the risk presented in this paper is equivalent to probability of failure unless otherwise specified.

Although a considerable number of studies on performance of electronics, summarized in [5] [6], have been conducted, most are focused on the electronic components performance, such as a capacitor. There are a few research efforts about electronic boards composed of multiple electronic components. The evaluation of electronic board failure in D&M tools is much more challenging because D&M tools are complex electronic-rich systems, which operate in dynamic and extreme environments. Moreover, environmental parameters such as temperature and vibration are usually collected for the entire tool, not collected for each board or electronic component separately.

Because D&M tools are designed for drilling and measurements in the oil and gas wells, research on the performance of D&M tool electronics are mostly conducted by the leading oil and gas service companies. For example, Mosallam et al. proposed a fault detection method for neutron generator subsystem in multifunction logging-while-drilling service based on empirical mode decomposition algorithm [7]. Bhatnagar et al. introduced a data-driven fault detection method for electronic boards in intelligent remote dual-valve system [8]. Kale et al. presented a probabilistic approach for risk estimates of D&M tool electronics based on reliability functions; e.g., Weibull, lognormal, and exponential distributions) with stress levels [9]. Zhan et al. proposed a cumulative damage model with Weibull distribution [3] [10]. The model is similar to that of Kale's work, but the stress levels are slightly different. Both methods assume a definite form of a reliability function based on prior knowledge. Furthermore, estimating the parameters of the reliability function usually requires a considerable amount of failure data. Based on the similarities between the cumulative stress paths between tools, Garvey et al. introduced a pattern recognition-based remaining useful life estimation for D&M tools [11]. This method is based on a strong assumption that the stress path is linear, which is difficult to hold. Moreover, the method cannot provide the uncertainty of the prognostic result. This paper will present a new method for risk estimates of electronic boards in D&M tools based on the Hidden Markov Model

(HMM). This method can output the uncertainty (probability) of risk estimates and does not use a definite form of a reliability function. Meanwhile, this method models the risks associated with an electronic board at different times as a function of evolving events, which closely approximates the real world.

This paper is organized into three sections. The first section presents the proposed method in detail. The following section presents a case study using actual data to confirm the method. The final section is a summary and proposes some research directions for the future.

2 Methodology

This section presents a brief introduction of the HMM fundamentals. Then, an overview of the underlying assumptions for the proposed method, together with a detailed description of the risk level estimate framework, will be presented.

2.1 Brief Introduction of the HMM Theory

An HMM consists of two parts, including an unobservable/hidden state sequence and a state-dependent observation sequence, as shown in Fig. 1. The unobservable state sequence $S_t : t = 1, 2, \ldots$ satisfies the Markov property; i.e., that the conditional probability of the state at time t (i.e., S_t) on previous states is equivalent to the conditional probability of S_t only on the most recent state S_{t-1}. Mathematically, this condition is expressed as $Pr(S_t|S_{t-1}, S_{t-2}, \ldots, S_1) = Pr(S_t|S_{t-1}), t = 2, 3, \ldots$ The observation sequence $\{O_t : t = 1, 2, \ldots\}$ is state-dependent. In other words, the distribution of O_t depends only on the current state St and not on historical states or observations. This can be expressed as $Pr(O_t|S_1, S_2, \ldots, S_t, O_1, O_2, \ldots, O_{t-1}) = Pr(O_t|S_t)$ [12].

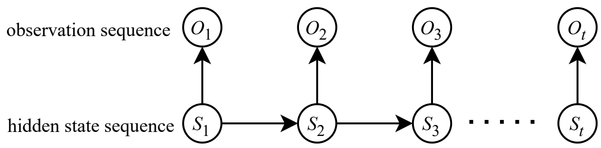

Figure 1: Schematic of HMM.

The HMM parameters can be characterized by a 3-tuple, $\mathbf{\Omega} = (\mathbf{M}, \mathbf{\Phi}, \boldsymbol{\pi})$, where $\mathbf{M}$ is the state transition probability matrix, $\mathbf{\Phi}$ is the observation probability distribution, and $\boldsymbol{\pi}$ is the state initial distribution.

Two problems can be formulated for HMM when applying the method to risk level estimating in this paper.

2.1.1 The Learning Problem

Given T consecutive observations $\mathbf{O} = \{O_1 = o_1, O_2 = o_2, \ldots, O_T = o_T\}$, the learning problem of HMM aims to find the model parameter $\mathbf{\Omega}$ that maximize the likelihood L

shown in Eq. (1).

$$
\begin{aligned}
L &= Pr\left(\mathbf{O}\right) = Pr\left(O_1 = o_1, O_2 = o_2, \ldots, O_T = o_T\right) \\
&= \sum_{s_1, s_2 \ldots, s_T = 1}^{m} Pr(O_1 = o_1, O_2 = o_2, \ldots, O_T = o_T, \\
& S_1 = s_1, S_2 = s_2, \ldots, s_T = S_T) \\
&= \boldsymbol{\pi}\mathbf{M}\boldsymbol{\Phi}(o_1)\mathbf{M}\boldsymbol{\Phi}(o_2)\cdots\mathbf{M}\boldsymbol{\Phi}(o_T)\mathbf{1}'
\end{aligned}
\tag{1}
$$

The Baum-Welch algorithm is commonly used to solve this problem [13].

2.1.2 The Decoding Problem

Given the HMM model parameter $\boldsymbol{\Omega}$ and T consecutive observations, $\mathbf{O} = \{O_1 = o_1, O_2 = o_2, \ldots, O_T = o_T\}$, as well the decoding problem of HMM is to determine the hidden state sequence $\mathbf{S} = \{S_1 = s_1, S_2 = s_2, \ldots, S_T = s_T\}$ that are most likely to generate the observation sequence, which is mathematically expressed as follows:

$$
\begin{aligned}
\max\ Pr(S_1 = s_1, S_2 = s_2, \ldots, S_T = s_T | \\
O_1 = o_1, O_2 = o_2, \ldots, O_T = o_T)
\end{aligned}
\tag{2}
$$

This problem can be solved by the Viterbi algorithm [14].

2.2 Risk Level Estimate Based on the HMM

As mentioned in Section I, the electronic boards are exposed to high temperatures, dynamic vibrations, and extensive shocks when the tool is drilling or measuring formation properties. These environmental conditions are the critical factors that cause the degradation or even the failure of the electronic boards. Thus, sensors are installed on the D&M tool to collect environmental data such as temperature, vibration, and shock during downhole operations. One could estimate the risk level of the electronic board through analyzing these environmental data.

2.2.1 Assumptions

To implement the HMM for risk level estimate of the electronic boards using the condition monitoring information, this paper makes the following assumptions:

- The observations; i.e., the final features used for training HMM, are Gaussian distributed. In addition, the features are mutually independent.
- The risk to the electronic boards is unobservable, but it is assumed that the risk has four levels, namely Level 1, Level 2, Level 3, and Level 4. Level 1 signifies the lowest risk while Level 4 is the highest risk. Moreover, the electronic boards are functional at the beginning. In other words, the initial state distribution of the HMM is $\boldsymbol{\pi} = [1, 0, 0, 0]$.
- The risk level of the electronic boards of the next time can either transit to the next risk level or stay in the same risk level as the current time. This means the HMM is left-to-right [15].

2.2.2 Proposed Risk Level Estimate Framework

Based on the previously stated assumptions, the framework of the proposed method is shown in Fig. 2. The framework includes four parts, namely, data collection, data preprocessing, learning, and decoding.

Specifically, in the data collection part, the tool measurement data are collected and stored in the memory chip during each downhole operation; i.e., each run. After each run, the tool is pulled up to the surface from the well, and the data stored in the memory chip are dumped and uploaded to cloud storage, from where one can retrieve the historical tool measurement data.

The data preprocessing portion consists of five steps; i.e., channel selection, histogram feature extraction, construction of cumulative exposure data sequence, data enrichment, and feature aggregation. These five steps are described as follows.

Channel Selection The channel selection is to choose the relevant signal channels for the risk-level estimate. Because the D&M tool acquires an enormous number of channels of information during running, many channels do not contain information concerning degradation of the electronic boards. Removing these channels does not only reduce the computation complexity in the following analysis, but also improves the estimation performance. Therefore, only the most important channels are selected for the risk-level estimate. In general, channels containing temperature, vibration, and shock data are selected because they are believed to have significant impacts on the electronic board lifetime [5].

Feature Extraction The histogram feature extraction computes exposure time under different environmental levels (occasionally referred to as stress levels) using the data from the selected channels of each run. Environmental levels are equivalent to histogram bins, and exposure time of these levels corresponds to the histogram frequencies multiplied by the data recording rate. The histogram feature of board b under environmental level i of channel j of run k is mathematically expressed as follows:

$$H_{ijk}^{(b)} = Frequency_{ijk}^{(b)} * RecordingRate \tag{3}$$

Construction of Cumulative Environmental Exposure Electronic board failures are usually caused by accumulative effects of environmental exposure during the numerous subsurface runs. Thus, environmental exposure data from a single run is not sufficient for determining electronic board risk. To use cumulative environmental exposure data, multiple subsurface runs are required. This step aims to construct cumulative environmental exposure data over runs based on the extracted histogram features. Specifically, the cumulative environment exposure data of board b under environmental level i of channel j after run K is denoted as follows:

$$C_{ijK}^{(b)} = \sum_{k=1}^{K} H_{ijk}^{(b)} \tag{4}$$

Data Enrichment Training the HMM commonly needs considerable data (or many observations). However, for each electronics board, only one observation can be obtained for each run after the collecting the cumulative exposure data. Moreover, it is difficult to obtain full life-cycle data from the electronic boards in a D&M

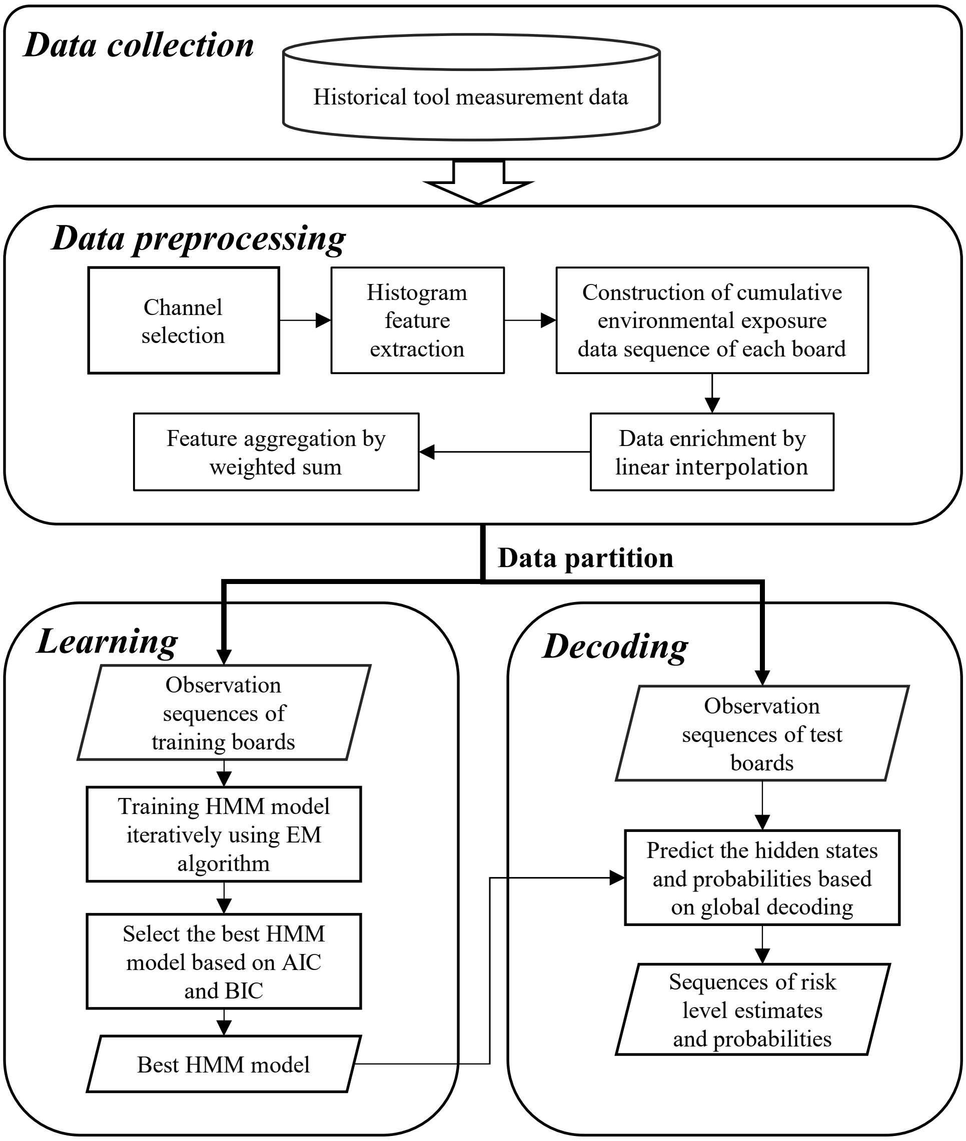

Figure 2: Framework of the proposed method.

tool. The reason is because the tool measurement data retrieved from the cloud storage have missing values for many runs due to numerous reasons, including tool lost in hole, failure of memory chips, and field technician failing to dump the data. Therefore, to augment the data volume for training the HMM, a linear interpolation method is adopted to generate synthetic data. The linear interpolation method is simple. Specifically, given two cumulative exposure data points from board b under environmental level i of channel j at time t_0 (end time of run $K-1$) and t_1 (end time of run K), the interpolated cumulative exposure data point at time t between t_0 and t_1 is given by:

$$C_{ijt}^{(b)} = C_{ij,K-1}^{(b)} + (t - t_0)\frac{C_{ijK}^{(b)} - C_{ij,K-1}^{(b)}}{t_1 - t_0} \tag{5}$$

In this paper, an equally spaced time sequence from time zero to the life end time of the board is first generated, and then new cumulative exposure data points are interpolated using Eq. (5).

Feature Aggregation The D&M tool is less likely to be exposed to high-temperature, high-vibration, and high-shock peak environments, which makes cumulative environmental exposure data at those levels not change significantly with time. Analyzing HMM parameters would fail if directly using these data. The reason is it would be impossible to estimate the state-dependent Gaussian distribution parameters when the observations are the same. In addition, it is difficult to determine which histogram features should be used. Nevertheless, the model complexity would substantially increase if all of the histogram features are used for HMM learning. Therefore, to simplify the model and achieve successful HMM parameter learning, the authors of this paper use weight sum to aggregate features from the same channel. The aggregated feature of board b of channel j at time t is formulized as follows:

$$AggF_{jt}^{(b)} = \sum_{i=1}^{n_j} w_{ij} C_{ijt}^{(b)} \tag{6}$$

where n_j represents the number of levels of channel j, and w_{ij} denotes the feature weight of level i of channel j.

After data preprocessing, the aggregated features can be used to build an observation of board b at time t as shown in Eq. (7). Then, it becomes possible to construct the observation sequence for each board. The boards are further divided into training boards and test boards. The observation sequences of the training boards are used to learn the HMM model parameter $\mathbf{\Omega}$. Because the Baum-Welch algorithm is a type of expectation maximization (EM) algorithm, the solution of EM is not unique. It is possible to train the HMM model many times and select the best model according to the Akaike information criterion and Bayesian information criterion as shown in Eq. (8). The less are the AIC and BIC, the better the model. With the trained HMM model and observation sequence of the test boards, the sequences of risk level and probabilities of those boards can be estimated. It should be noted that one can also estimate the same information for training boards.

$$O_t^{(b)} = [AggF_{1t}^{(b)}, AggF_{2t}^{(b)}, \ldots, AggF_{Nt}^{(b)}]^T \tag{7}$$

where N means the number of channels.

$$\begin{aligned} AIC &= -2\log(L) + 2p \\ BIC &= -2\log(L) + p\log(T) \end{aligned} \tag{8}$$

where L is the log likelihood of the HMM shown in Eq. (1), p is the number of model parameters, and T denotes the number of observations.

3 Case Study

In this section, historical tool measurement data from D&M tools will be used to validate the effectiveness of the proposed method. The data were collected during actual downhole oilfield operations. Fig. 3 shows the D&M tool studied in this paper. The tool uses a rotary steerable system that controls the direction in which a well is drilled while rotating the drillstring. Various types of electronic boards are installed in this tool to support its various functions. In this paper, the control process unit (CPU) board is selected. For this type of board only, the historical tool measurement data were collected from 21 boards whose failures were related to heat and vibration. As mentioned previously in the *Data Enrichment* section, it is difficult to collect full life-cycle environmental exposure data from failed boards. Thus, in this case study, some of the boards have missing data. The detailed data description of the 21 boards is shown in TABLE 1.

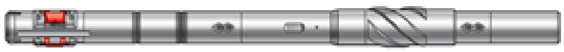

Figure 3: Rotary steerable system.

Additionally, four channels of tool measurement data, specifically, temperature, lateral vibration, axial vibration, and lateral shock peak were selected for this case study. The number of levels for the four channels and the corresponding weights for feature aggregation are summarized in TABLE 2. The time allowed for generating the time sequence for the Data Enrichment step is set to 10 hours.

The tool measurement data from the 21 boards were processed following the steps described in the *Data Preprocessing* section previously described. Then, the first 11 boards listed in TABLE 1 were selected as training boards, and the remaining 10 boards as test boards.

The software used for the HMM learning and decoding in this paper is the *depmixS4* [16] R package. The HMM model was trained 100 times using the training boards data. The best HMM model has an AIC of 52887.48 and a BIC of 53124.96. The learned state transition probability matrix of the HMM is as follows:

$$\mathbf{M} = \begin{bmatrix} 0.95 & 0.05 & 0 & 0 \\ 0 & 0.963 & 0.037 & 0 \\ 0 & 0 & 0.967 & 0.033 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

The learned observation probability distribution (Gaussian distribution) parameters are shown in TABLE III, where μ and σ denote the mean and standard deviation of the Gaussian distribution, respectively.

Using the trained HMM model, the risk level and its probability can be estimated on the condition of observing the sequence of each run. It should be noted that the risk estimation for the runs with missing data is not included. The sequences of risk level estimates and probabilities of training boards are shown in TABLE 4.

The sequences of risk level estimates and probabilities of test boards are shown in TABLE 5.

From these results, almost all of the boards (10/11 of training boards, 8/10 of test boards) are predicted to be at risk of Level 4 in their last run, which is consistent

Table 1: Data Description of The 21 Boards

Board ID	Runs	Missing Runs[a]	Data Coverage[b]	Total Pumping Time (h)[c]
x1005	18	4	92%	1104.5
x4028	20	2	87%	862.6
x1014	11	0	100%	714.4
x6054	19	1	93%	742.2
x9031	18	2	89%	1355.9
x9046	33	9	83%	1328.1
x0048	15	2	95%	1037.2
x5005	19	3	92%	1186.2
x5050	18	1	97%	1110.1
x5010	22	3	87%	1403.7
x0405	17	1	95%	1210.2
x1012	3	0	100%	125.7
x9045	43	8	80%	2067.2
x3007	11	1	94%	963.2
x9035	28	10	72%	1176.5
x7043	25	8	73%	1371.3
x7079	23	7	71%	1608.2
x9003	29	5	84%	1691.2
x9049	22	4	84%	1239.3
x6004	11	0	100%	450.3
x1260	20	2	90%	1324.2

[a] number of runs that have missing data.
[b] total pumping time of runs with data/total pumping time of all runs.
[c] total pumping times of all runs.

Table 2: Description of Channel Levels and Weights

Channel Name	Number of Levels	Level 1 Weight	Level 2 Weight	Level 3 Weight	Level 4 Weight
Temperature	4	0.5	1.5	2	10
Lateral vibration	3	0.5	1.5	10	–
Axial vibration	3	0.5	1.5	10	–
Lateral shock peak	3	0.5	1.5	10	–

Table 3: Learned Observation Probability Distribution Parameters

State (Risk Level)	**Aggregated Temperature**		**Aggregated Lateral Vibration**		**Aggregated Axial Vibration**		**Aggregated Lateral Shock Peak**	
–	μ	σ	μ	σ	μ	σ	μ	σ
Level 1	67.4	42.6	57.2	33.4	58.8	35.4	52.6	30.0
Level 2	288.5	127.9	182.7	46.7	184.0	46.4	68.0	41.1
Level 3	557.7	235.8	30.9	56.1	330.1	53.2	308.8	46.8
Level 4	803.2	299.6	546.6	116.0	533.1	102.5	491.1	81.1

Table 4: Sequences of Risk Level Estimates and Probabilities of Training Boards

Board ID	**Risk Level Sequence**	**Probability Sequence**
x1005	1-2-2-2-2-3-3-4-4-4-4-4-4-4	1-1-1-1-0.999-1-0.999-0.06-1-1-1-1-1-1
x4028	1-1-2-2-2-2-2-3-3-3-3-3-3-3-4-4	1-1-1-1-1-1-1-0.597-1-1-1-1-1-0.993-0.793-1
x1014	1-1-1-2-2-2-2-2-2-3-3	1-1-1-0.838-1-1-1-1-0.998-1-1
x6054	1-1-1-2-2-2-2-2-2-2-2-3-3-4-4-4-4	1-1-1-0.825-1-1-1-1-0.999-0.999-0.996-0.991-0.999-0.169-1-1-1
x9031	1-2-2-3-3-3-3-3-3-3-4-4-4-4-4-4	1-1-1-0.044-0.125-0.579-0.974-1-1-1-0.626-1-1-1-1-1
x9046	1-1-2-2-2-2-3-3-3-3-3-3-4-4-4-4-4-4-4-4-4-4-4	1-0.991-0.741-1-1-0.998-0.955-1-1-1-1-0.979-0.993-1-1-1-1-1-1-1-1-1-1
x0048	1-1-1-2-2-2-2-3-3-3-4-4	1-1-1-1-1-1-1-0.956-1-1-1-1
x5005	1-1-1-2-2-3-3-3-3-3-3-3-4-4-4-4	1-1-1-1-1-0.062-1-1-1-1-1-0.992-1-1-1-1
x5050	1-1-2-2-3-3-3-3-3-3-4-4-4-4-4-4-4	1-0.999-1-1-0.114-1-1-1-0.999-0.997-1-1-1-1-1-1-1
x5010	1-1-1-1-2-2-2-2-2-3-3-3-3-3-4-4-4-4-4	1-1-1-0.987-1-1-1-1-1-1-1-1-1-0.996-1-1-1-1-1
x0405	1-1-2-2-2-2-3-3-3-3-4-4-4-4-4-4	1-1-0.915-1-1-1-1-1-1-0.991-1-1-1-1-1-1

Table 5: Sequences of Risk Level Estimates and Probabilities of Test Boards

Board ID	Risk Level Sequence	Probability Sequence
x1012	1-1-1	1-1-1
x9045	1-1-1-1-2-2-2-3-3-3-3-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4	1-1-1-1-1-1-1-1-1-1-0.99-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1
x3007	1-2-2-2-2-2-3-3-3-4	1-1-1-1-1-1-1-1-1-0.21
x9035	1-1-1-1-2-2-2-3-3-3-3-4-4-4-4-4-4-4	1-1-1-1-1-0.98-1-1-1-1-0.05-1-1-1-1-1-1
x7043	1-2-2-2-2-2-3-3-3-4-4-4-4-4-4-4	0.97-1-1-1-1-1-1-0.99-1-1-1-1-1-1-1
x7079	1-1-2-2-2-3-3-3-4-4-4-4-4-4-4	1-0.99-1-1-1-1-1-1-0.08-1-1-1-1-1-1
x9003	1-1-1-1-2-2-2-3-4-4-4-4-4-4-4-4-4-4-4-4-4-4-4	1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1
x9049	1-1-2-2-3-3-3-3-3-4-4-4-4-4-4-4-4-4	1-1-0.98-1-0.04-1-1-1-1-1-1-1-1-1-1-1-1-1
x6004	1-1-1-1-1-1-2-2-2-2	1-1-1-1-1-0.95-1-1-1-1
x1260	1-1-2-2-2-2-3-3-3-3-3-4-4-4-4-4	1-1-0.06-1-1-0.99-1-1-1-1-1-0.2-1-1-1-1

with the actual situation because all of the boards failed. The three boards whose risk level in the last run is not estimated as Level 4 have much less pumping time (714.5 hr, 125.7 hr, and 450.3 hr) than the other boards according to the pumping time in TABLE 1. This result implies that the proposed method might not be suitable for boards with minor failures.

Considering that most boards do not have full-life cycle coverage, the proposed method is still capable of achieving acceptable results. Thus, the proposed method is effective for estimating the risk level of electronic boards.

4 Conclusion and Future works

A risk-level estimate method based on the HMM for electronic boards used in oil and gas industry D&M tools has been presented in this paper. The proposed method has a major advantage of unsupervised learning, that is, it can be trained using unlabeled data. This merit is of significance for D&M tools because it is costly or sometimes even impossible to label the true risk level of electronic boards. Overall, the results obtained from the failure data collect in the field show that the proposed method is effective for risk estimation.

Some shortcomings or challenges exist for our method that will be presented in the future. One major challenge is that the proposed method needs full-life cycle data because currently, many runs are missing data. The future work on this challenge might well be how to handle the missing values. Moreover, the weights for feature aggregation are defined based on experiences, which lacks theoretical support. The weights can be further optimized based on accelerated factors learned from highly accelerated life tests. Finally, the proposed method is not tolerant of early failure of boards because early-failure boards will add noise to the model. How to set a life-time threshold for determining early failure of the electronic boards is also worth studying. In this way, removing the early failure boards from our training boards will improve the model performance.

Acknowledgment

This work has been supported by the EIPHI graduate school (contract “ANR-17-EURE-0002”).

References

[1] Schlumberger, ”BHA,” 2021. [Online]. Available: https://glossary.oilfield.slb.com/en/terms/b/bha.

[2] K. Carter-Journet, A. Kale, D. Zhang and E. Pradeep, ”Estimating Probability of Failure for Drilling Tools with Life Prediction,” in SPE Asia Pacific Oil & Gas Conference and Exhibition, Adelaide, 2014.

[3] S. Zhan and J. Zhao, ”Advanced Prognostic Technique for Improving the Drilling Performance of Downhole Tools,” in The First International Conference on Reliability Systems Engineering, Beijing, 2015.

[4] F. I. Khan and M. M. Haddara, ”Risk-based maintenance (RBM): a quantitative approach for maintenance/inspection scheduling and planning,” Journal of Loss Prevention in the Process Industries, vol. 16, pp. 561–573, 2003.

[5] M. G. Pecht and M. Kang, Prognostics and Health Management of Electronics : Fundamentals, Machine Learning, and the Internet of Things, Glasgow: John Wiley & Sons, Incorporated, 2018.

[6] C. Bhargava, P. K. Sharma, M. Senthilkumar, S. Padmanaban, V. K. Ramachandaramurthy, Z. Leonowicz, F. Blaabjerg and M. Mitolo, ”Review of Health Prognostics and Condition Monitoring of Electronic Components,” IEEE Access, vol. 8, pp. 75163–75183, 2020.

[7] A. Mosallam, F. B. Youssef, J. Fulton and D. E. Viassolo, ”Data-Driven Fault Detection for Neutron Generator Subsystem in Multifunction Logging-While-Drilling Service,” in European Conference of The Prognostics and Health Management Society 2018, Utrecht, 2018.

[8] S. Bhatnagar, M. L. Cassou, Z. . A. Masry and A. Mosallam, ”Data-Driven Fault Detection Method for Electronic Boards in Intelligent Remote Dual-Valve System,” in European Conference of the Prognostics and Health Management Society 2021, Virtual, 2021.

[9] A. A. Kale, K. Carter-Journet, T. A. Falgout, L. Heuermann-Kuehn and D. Zurcher, ”A Probabilistic Approach for Reliability and Life Prediction of Electronics in Drilling and Evaluation Tools,” in Annual Conference of the PHM Society 2014, Texas, 2014.

[10] S. Zhan and A. Izhar, ”Real-time prognostic on downhole printed circuit board assembly of measurement-while-drilling/logging-while-drilling”. United States Patent US9051814B2, 9 June 2015.

[11] D. Garvey, J. Baumann, J. Lehr and J. W. Hines, "Pattern Recognition Based Remaining Useful Life Estimation of Bottom Hole Assembly Tools," in SPE/IADC Drilling Conference and Exhibition, Amsterdam, 2009.

[12] W. Zucchini, I. L. MacDonald and R. Langrock, Hidden Markov Models for Time Series - An Introduction Using R, 2nd ed., Boca Raton: CRC Press, 2016.

[13] C. Sammut and G. I. Webb, Eds., "Baum-Welch Algorithm," in Encyclopedia of Machine Learning, Boston, Springer, 2011.

[14] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," IEEE Transactions on Information Theory, vol. 13, no. 2, pp. 260–269, 1967.

[15] C. I. Oliver, Markov Processes for Stochastic Modeling, 2nd ed., London: Elsevier, 2013.

[16] I. Visser and M. Speekenbrink, "depmixS4: An R-package for hidden Markov models," Journal of Statistical Software, vol. 36, no. 7, pp. 1–21, 2010.