Submission code: 0PL5YRWD8R

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Node quality control automation for more efficient decision-making

Gilmario Maciel (Petrobras), GLAUCIA KIMIE SAKAI (Fundação Gorceix), Adriano Escobar (Fundação Gorceix), Pablo De Carvalho (Fundação Gorceix), Gabriel Barata (Fundação Gorceix), Stephanie Tavares (Fundação Gorceix), Leandro da Silva Sadala Valente (Cenpes-Petrobras)

Node quality control automation for more efficient decision-making

Abstract Summary

This study presents automation strategies for onboard quality control of seismic nodes. Our goal is to overcome the limitations of subjective analyses performed by technicians, thereby optimizing the speed, accuracy, and reliability of identifying nodes unsuitable for reuse. The proposed methodology analyzes battery voltage decay, tilt measurements, and the amplitude spectrum of geophone components. Battery decay patterns are assessed using logistic regression models trained with geometric features extracted from voltage time series. Amplitude spectral consistency, in turn, is verified through divergence metrics applied in the frequency domain. We validated the methodology with 2,479 nodes—approximately 15% of the complete dataset—manually labeled based on field records. The results showed that 99% of the nodes were correctly identified, with only three genuine anomalous nodes not being recognized. The high classification performance and operational feasibility of the methodology reinforce its potential in obtaining more effective quality control workflows.

1 Introduction

In deepwater seismic surveys, seismic nodes are deployed on the seafloor and remain operational for periods that can exceed 100 days. Due to the high number of required recording positions relative to the limited number of available nodes, a roll-along deployment scheme is employed to optimize seismic data acquisition. This cycle of recovery and repositioning intensifies the need for rigorous quality control to ensure the continuous functionality of the equipment. Inadequate quality control of operational node data can lead to the reuse of malfunctioning equipment, compromising seismic acquisition and resulting in incomplete records, which frequently need re-surveying in certain areas. This quality control process involves critical assessments of battery performance, node tilt, spectral recording behavior, and data integrity. This control is performed manually and due to the large volume of data to analyze, sampling is employed, not encompassing a complete analysis of all operational data. The objective of this work is to implement automation strategies to mitigate the limitations of subjective evaluations of technicians in quality control, optimizing speed, accuracy, and reliability in identifying unsuitable nodes for reuse.

2 Methodology

The methodology employed consists of three node feature analyses: battery voltage decay, tilt relative to the ocean bottom, and amplitude spectrum.

2.1 Battery Voltage Decay Analysis

In the battery quality control process, nodes are expected to exhibit voltage decay characteristics consistent with the specifications provided by the manufacturer. This evaluation primarily considers the total acquisition elapsed time and the variation in voltage over time.

The recording elapsed times of the nodes are evaluated collectively to detect outliers. These times are analyzed considering three aspects: (a) the acquisition end times should be ordered chronologically, (b) the differences in elapsed times between neighboring nodes cannot exceed

3 days, and (c) each elapsed time must be within the median ± standard deviation of all elapsed times belonging to the same receiver line.

In addition to the collective analysis of recorded elapsed times, individual evaluations of node voltage variations aim to identify atypical decay patterns. These patterns are handled by two binary logistic classifiers (HOSMER et al., 2013), an approach well-suited for distinguishing between two groups, normal and anomalous nodes. These classifiers are driven by geometric parameters obtained from the voltage curves, including the average voltage decay rate (i.e., the slope between the first and last measurements), and the deviation of the observed decay from an ideal linear regression. Combining these classifiers substantially increases the accuracy of detecting anomalous voltage decay behavior.

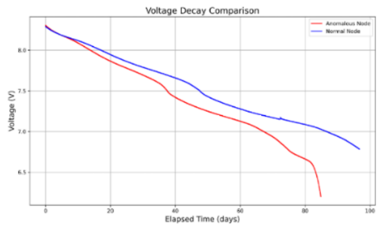

Figure 1 - The graph shows the voltage decay over time of a normal node (blue) and an anomalous one (red).

2.2 Equipment Tilt Analysis

This analysis aims to provide technicians with an additional parameter for more assertive judgments regarding data integrity. The reported information includes: (a) identification of equipment tilts above 10 degrees (b) detection of abrupt equipment tilt variations over time (c) flagging of potentially inconsistent tilt values. Equipment tilt variation is estimated based on the difference between two average of tilt values calculated from segments of the time series corresponding to the 5th to 10th and 85th to 95th percentiles, respectively, to capture potential trend changes from the beginning to the end of the curve. The indication of inconsistent tilt values is based on the application of the z-score method, which uses statistical standardization and a predefined threshold to detect significant deviations from expected behavior (MONTGOMERY & RUNGER, 2018). The results of this analysis are not considered in the automatic identification of anomalous equipment, since unusual tilt values are more often related to deployment characteristics rather than equipment malfunction.

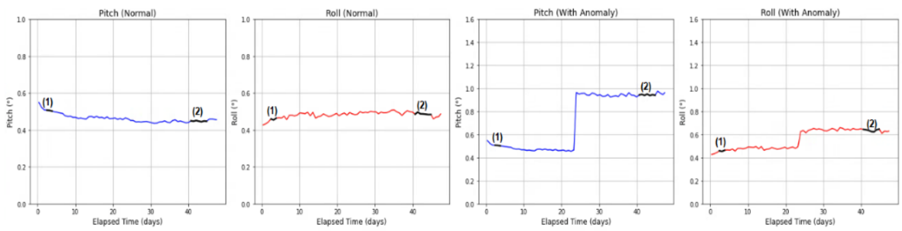

Figure 2 – Variations in equipment tilt and rotation over time for normal and anomalous nodes. The intervals from the 5th to the 10th and from the 85th to the 95th percentiles for elapsed time are identified by (1) and (2), respectively.

2.3 Amplitude Spectrum Analysis

This analysis compares the frequency domain amplitude spectra of the three node measurement components. These measurements are performed by geophones arranged in a Galperin configuration, whose main advantage is that they provide identical responses for all components. Significant deviations in these measurements may indicate coupling problems or equipment malfunction.

To evaluate the similarity between the amplitude spectra of the components, each was considered a Power Spectral Density (PSD) function and treated as a probability distribution. The Jensen-Shannon divergence was then applied to these functions. This metric, a symmetric and smoothed version of the Kullback-Leibler divergence (KULLBACK & LEIBLER, 1951), quantifies the difference between a pair of probability density curves. This technique was chosen based on three factors: stability and robustness. Additionally, the sum of pairwise divergences was used as an extra indicator for anomalous responses

To increase classification accuracy and reduce false negatives, a threshold value was introduced for the sum of the divergence coefficients between a pair of components. This threshold was determined empirically based on the exploratory data analysis and the experience of the analyzing technician. Divergence values above this threshold indicate loss of similarity between the amplitude spectra in at least one of the components.

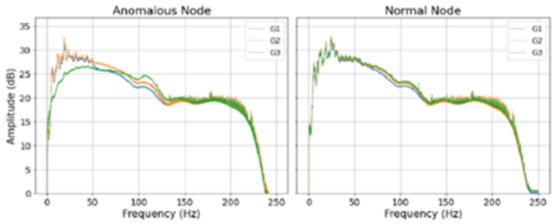

Figure 3 - Amplitude spectrum of anomalous (left) and normal (right) nodes.

3 Results

The experiment was conducted using three types of data from 2,479 nodes. These nodes were acquired during operational campaigns and represent approximately 15% of the total available data sets. The data were processed by automated routines developed to identify anomalous patterns. The validation dataset contained labels previously defined based on manual inspections and field records. Of these, five nodes exhibited battery-related failures, and one showed amplitude spectral inconsistency, attributed to a malfunction in at least one component, totaling six anomalous nodes.

As a result of applying the methodology, all six anomalous nodes were correctly identified, demonstrating the routine's ability to detect actual failures. Among the nodes considered normal, 2,470 were correctly classified, while three were mistakenly identified as anomalous. The classification decision considered a node anomalous if it exhibited issues in at least one of the two analyses: battery voltage decay or amplitude spectrum. Table 1 presents the relationship between genuinely anomalous and normal nodes, and those predicted as such or not.

To verify the effectiveness of the results, Figure 4 presents three representative cases from the validation dataset: one false negative, a genuinely anomalous node that was not recognized as such, based on battery voltage decay; and two true negatives, genuinely anomalous nodes that were correctly identified, one based on the amplitude spectrum and the other on battery voltage decay.

	Normal (P)	*Anomalous (P)*
Normal (G)	***2470***	***3***
Anomalous (G)	***0***	***6***

Tabela 1 – Model classification performance, where G and P indicate genuinely and predicted classes, respectively.

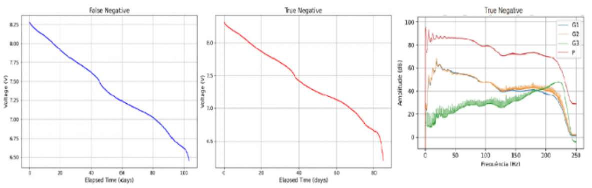

Figure 4 – A false negative on the left and a true negative in the middle, both based on the battery voltage decay, and a true negative based on the amplitude spectrum on the right.

Additionally, Figure 5 shows graphical representations of decision-making analysis. These figures provide a comprehensive visual understanding of node behavior for each analyzed characteristic and how the integration of both analyses enhances the robustness of the final classification.

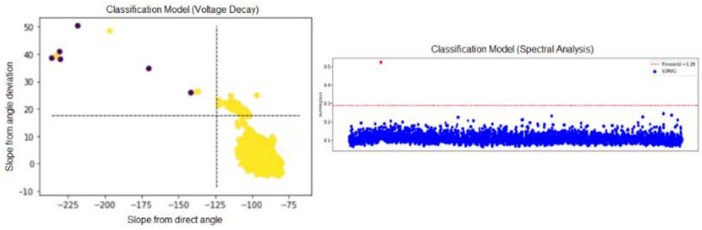

Figure 5 - The classification based on battery voltage decay (left) uses purple dots for anomalous nodes and yellow dots for normal nodes. In contrast, the classification for the amplitude spectrum (right) represents anomalous nodes with red dots and normal nodes with blue dots. The black traced line on the first graph and the red dotted line on the right graph indicate the classification thresholds.

4 Conclusions

This study introduced a methodology to automate the quality control process for seismic nodes. The approach focused on analyzing battery voltage decay, equipment tilt, and the spectral consistency of geophone signals. For validation, we conducted numerical experiments using real-world data from 2,479 nodes acquired during operational campaigns, which represented approximately 15% of the total available data.

The results demonstrate that the proposed methodology successfully achieved its objectives. It mitigated the limitations of subjective evaluations by objectively defining classification thresholds directly through the model. Furthermore, the approach maintained high levels of accuracy and reliability, correctly identifying 99% of the nodes during validation, with a minimal occurrence of erroneous classifications. The methodology also proved to be operationally efficient: the validation dataset was analyzed in approximately three hours, a significant contrast to the current manual process, which assesses about 10 nodes per day in the operational workflow.

Currently, we are expanding the methodology to include additional quality control processes not addressed in this study, such as the analysis of clock drift in the equipment's internal timing systems. Automating this step is expected to complement the analyses already implemented, contributing to a more comprehensive and robust system for the operational validation of seismic nodes.

References

HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X**.** *Applied Logistic Regression*. 3. ed. Hoboken: Wiley, 2013.

KULLBACK, Solomon; LEIBLER, Richard A. *On information and sufficiency*. The Annals of Mathematical Statistics, v. 22, n. 1, p. 79–86, 1951. DOI: 10.1214/aoms/1177729694.

MONTGOMERY, D. C.; RUNGER, G. C. *Applied Statistics and Probability for Engineers*. 7. ed. Hoboken: Wiley, 2018.

PAPOULIS, A.; PIL, Lai, S. U. Probability, Random Variables, and Stochastic Processes. 4. ed. New York: McGraw-Hill, 2002.