WED-SSS-CCT-06

DATA PROCESSING AND ANALYSIS FOR AGING DETECTION OF CMS RPC CHAMBERS USING CERN ROOT

Tsvetelina Stefanova – Student
Department of Computer Systems and Technologies,
University of Ruse "Angel Kanchev"
E-mail: tsveti.stefanov@gmail.com

Assoc. Prof. Roumyana Hadjiiska, PhD
Institute for Nuclear Research and Nuclear Energy,
Bulgarian Academy of Sciences
E-mail: roumyana.mileva.hadjiiska@cern.ch

Abstract: *The Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) at CERN relies heavily on Resistive Plate Chambers (RPCs) for muon detection and triggering. As the LHC transitions to higher luminosity phases, keeping the optimal performance of the detector system poses a significant challenge. This study focuses on the data processing and analysis techniques used to detect eventual aging in the CMS RPC chambers using the CERN ROOT software. Detector performance data from 2018 and 2023 were compared, revealing efficiency drops in certain regions of the chambers. These findings underscore the necessity for ongoing technical support and monitoring to maintain optimal operation and avoid damage to the chambers. The results also highlight the potential for using this data, coupled with machine learning models, to predict future aging patterns.*

Keywords: *Resistive Plate Chambers (RPCs), efficiency monitoring, data cleaning, data analysis, non-parametric hypothesis testing*

INTRODUCTION

The Compact Muon Solenoid (CMS) experiment at CERN's Large Hadron Collider (LHC) is a crucial tool for advancing our understanding of particle physics. CMS investigates phenomena ranging from the Higgs boson to potential new particles that could explain dark matter. Among its subsystems, Resistive Plate Chambers (RPCs) play a critical role in muon detection and timing due to their fast response and reliability (Abbrescia, M. et al., 2005; CMS Collaboration, Amsler, C., V. Chiochia, S. De Visscher., 2010; CMS collaboration, Chatrchyan, S. et al., 2008).

RPCs might be subject to aging effects due to prolonged exposure to high particle fluxes and radiation (Aly, R. et al., 2020). With the High Luminosity LHC (HL-LHC) phase set to begin in 2029, the need for robust monitoring and predictive maintenance systems becomes even more pronounced. This paper explores the use of CERN ROOT software to process and analyze RPC performance data, focusing on detecting signs of aging.

METHODOLOGY

For the purposes of our study, we used the ROOT framework, developed by CERN, along with Python packages for data processing and analysis like NumPy and SciPy.

Overview of the ROOT Framework

CERN ROOT is an open-source data analysis framework extensively used in particle physics. It provides tools for storing, processing, and visualizing large datasets generated by experiments such as CMS. ROOT files (.root) are hierarchically organized, allowing efficient access to multidimensional data structures. The framework also includes statistical libraries and visualization tools for creating histograms, scatter plots, and other graphical representations (ROOT, 2024-a; ROOT, 2024-b).

The flexibility of ROOT, combined with its compatibility with programming languages such as Python through PyROOT and JupyROOT, makes it an ideal choice for analyzing RPC chamber performance. The following subsections detail the specific steps taken to leverage ROOT in this study.

Datasets

Efficiency data obtained using the segment extrapolation method (Petkov, P., 2007; Shah, M. A., 2016) from 1020 RPC chambers of the CMS detector for the years 2018 and 2023 were used. These data were stored in ROOT files, each containing two-dimensional histograms of efficiency values at a spatial resolution of 1 cm². The chambers analyzed were from the barrel (central) section (CMS collaboration, 2018), covering both high and low radiation zones. The dataset encompassed over 1 million individual measurements, reflecting the comprehensive scope of the analysis.

Preprocessing

To ensure accurate comparisons, preprocessing steps included:

- **Ratio Calculation:** For each cm^2 of an RPC, efficiency ratios were calculated as the efficiency value for 2018 divided by the efficiency value for 2023.

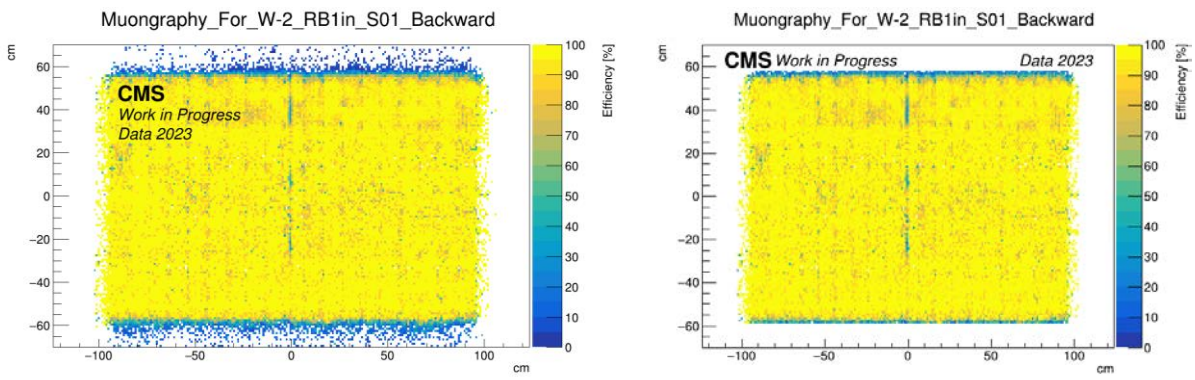
$$Ratio = \frac{Efficiency_{2018}}{Efficiency_{2023}} \tag{1}$$

- **Difference and Relative Difference:** Absolute (non-normalised) and relative differences were computed to highlight deviations in performance.

$$Difference = Efficiency_{2018} - Efficiency_{2023} \tag{2}$$

$$Relative\ difference = \frac{Efficiency_{2018} - Efficiency_{2023}}{Efficiency_{2018}} \tag{3}$$

- **Artifact Removal:** Border artifacts from extrapolated efficiency values were identified and excluded.



a) b)

Fig. 1. Border artifacts removal: a) before removing artifacts; b) after removing artifacts

These preprocessing steps were implemented using ROOT's TTree and TH2D classes, enabling efficient handling of large datasets, as well as the ROOT methods Add() and Divide().

 Copyrights© 2024 ISBN 978-954-712-951-1 (Print); e-ISBN 978-954-712-952-8

Visualization of the data obtained after completing the preprocessing steps further guided the identification of regions exhibiting significant changes.

Statistical Analysis

While for the data preprocessing we used entirely the ROOT framework, for further analysis we relied on the seamless integration of ROOT with external Python packages. In order to evaluate the changes in the efficiency of RPC chambers in 2023 compared to 2018, we formulated three pairs of hypotheses which are as follows:

Table 1. Pairs of hypotheses

№	Null hypothesis	Alternative hypothesis
1	ΔEff(2018-2023) / Eff(2018) = 0	ΔEff(2018-2023) / Eff(2018) > 0
2	Eff(2018) / Eff(2023) = 1	Eff(2018) / Eff(2023) > 1
3	ΔEff(2018-2023) = 0	ΔEff(2018-2023) > 0

All of them state the following:

- Null Hypothesis (H_0): The efficiency values for 2018 and 2023 are equal.
- Alternative Hypothesis (H_1): The efficiency for 2023 is lower than for 2018 (right-tailed test).

We observed lack of normal distribution in the ratios and in the differences of the efficiencies for most of the RPCs (Fig. 2), so non-parametric right-sided statistical tests for dependent samples or for one sample are suitable for our task.

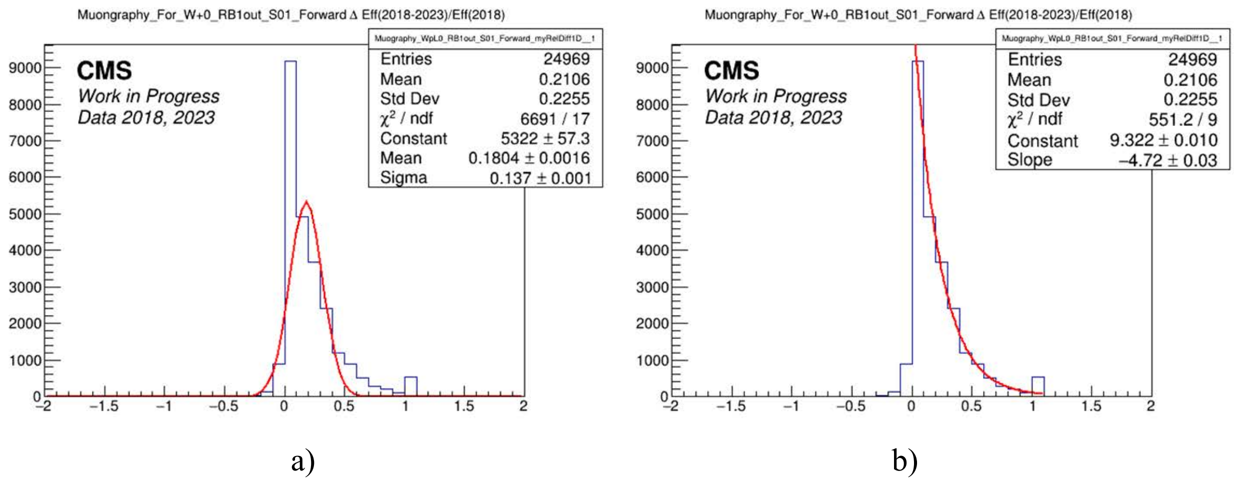


Fig. 2. Distribution of the relative differences of the efficiencies per cm^2 of a certain RPC: a) compared to the normal distribution; b) compared to the exponential distribution

As a non-parametric statistical test checking the changes in the overall level of efficiency of each RPC chamber we used the Wilcoxon signed-rank test. However, it is also important to know if there are concrete areas of the RPCs which are exposed to faster aging. For this reason, we also adapted the Z-test to evaluate the changes for each square centimeter.

Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is a non-parametric test that does not require normal distribution of the data (Conover, W. J., 1999). In this study, it was applied as a one-sample test to test changes in overall efficiency levels using the calculated non-normalized and relative differences with hypotheses 1) and 3) (Table 1). Data preprocessing ensured that all artifacts were removed, and the test focused only on significant negative changes.

Using the ROOT framework and Python's SciPy library, the test ranked differences between paired values, excluding bins with zero differences. The p-values were calculated by comparing the test statistic to its expected distribution under the null hypothesis. For the Wilcoxon signed-rank test, the sum of positive ranks was compared to critical values derived from the Wilcoxon distribution, as implemented in the SciPy library. A significance level of 0.01 was used to determine whether the negative changes were statistically significant.

Z-Test

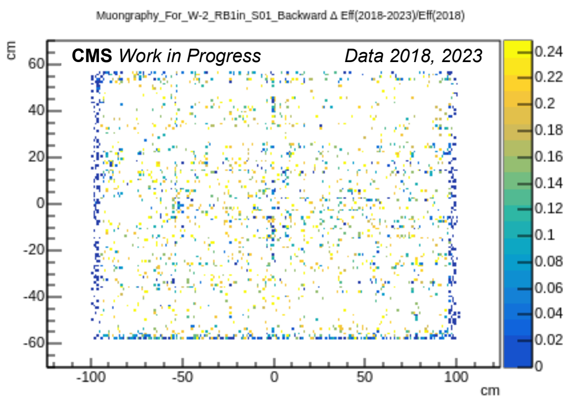
The Z-test was used to evaluate differences in efficiency ratios and relative differences for individual square centimeters of each chamber by testing hypotheses 1) and 2) (Table 1). This approach allowed for localized detection of aging effects. The Z-statistic was calculated as (Freedman, D., Pisani, R., & Purves, R., 2007):

$$z = \frac{x - \mu}{\sigma} \tag{4}$$

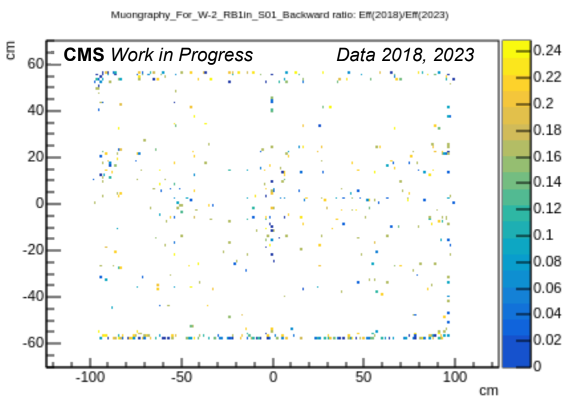
where x is the observed value, μ is the expected mean (set to 1 for ratios and 0 for relative differences), and σ is the standard deviation of the respective ratios and relative differences. ROOT libraries were used to process the histograms and compute these values. The p-values for the Z-test were determined by comparing the observed Z-statistic to the cumulative distribution function (CDF) of the standard normal distribution, representing the probability of observing values as extreme or more extreme under the null hypothesis.

For this study, a significance level of 0.25 was chosen to balance sensitivity and specificity, highlighting areas of potential efficiency loss.

After calculating the test statistics and the p-values for each cm^2 of a certain RPC chamber, the p-values which were less than 0.25 were stored in TH2D histograms using ROOT, so that the results could be visually interpreted. The following is an example of a histogram containing the p-values from the Z-test for a certain RPC chamber using the relative difference and the ratio between the efficiencies from 2018 and 2023:



a)



b)

Fig. 3. 2D histograms with p-values from the Z-test for an RPC: a) based on the relative difference of the efficiencies for 2018 and 2023; b) based on the ratio of the efficiencies for those two years

Those histograms (Fig. 3) show which areas of the chamber might be prone to faster aging. The parts which might be prone to faster aging have lower p-values and they are marked with darker colours.

Copyrights© 2024 ISBN 978-954-712-951-1 (Print); e-ISBN 978-954-712-952-8

RESULTS AND DISCUSSION

The analysis revealed efficiency declines within certain RPC chambers between 2018 and 2023, particularly in regions with varying radiation backgrounds. Higher radiation exposure was associated with more significant efficiency drops, as observed in efficiency ratios and relative differences.

This study demonstrates the efficacy of using CERN ROOT software for detecting aging in the RPC chambers. Statistical analysis and visual inspection corroborated these efficiency trends, emphasizing the importance of continuous monitoring and predictive maintenance to sustain chamber performance, particularly as the LHC transitions to its HL-LHC phase.

The findings highlight the critical need for enhanced monitoring and maintenance strategies as CMS transitions to the HL-LHC phase. Two key implications arise:

1. **Localized Monitoring**: Efficiency variations within individual chambers necessitate detailed spatial analysis to identify at-risk regions (Aly, R. et al., 2020).
2. **Predictive Modeling**: The incorporation of machine learning models, trained on historical data, could provide valuable insights into aging trends and inform maintenance schedules (Shumka, E. et al., 2023).

Future work should expand the dataset to include additional years and explore potential correlations between chamber positions and aging rates. Additionally, the integration of real-time monitoring tools with ROOT could facilitate dynamic analyses, further enhancing detector performance.

REFERENCES

Abbrescia, M. et al. (2005). Cosmic ray tests of double-gap resistive plate chambers for the CMS.

Aly, R. et al. (2020). Aging study on Resistive Plate chambers of the CMS muon detector for HL-LHC. Journal of Instrumentation, 15(11), C11002.

CMS Collaboration, Amsler, C., V. Chiochia, S. De Visscher. (2010). Performance study of the CMS barrel resistive plate chambers with cosmic rays. Journal of Instrumentation, 5(3):T03017.

CMS collaboration. (2018). Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$= 13 TeV. Journal of Instrumentation, 13(6), P06015.

CMS collaboration, Chatrchyan, S. et al. (2008). The CMS experiment at the CERN LHC. Journal of instrumentation.-Bristol, 2006, currens, 3, S08004.

Conover, W. J. (1999). Practical nonparametric statistics (Vol. 3). Wiley.

Freedman, D., Pisani, R., & Purves, R. (2007). Statistics. International student edition. Pisani, R. Purves, 4th edn. WW Norton & Company, New York.

Petkov, P. (2007). Tests of the RPC for the CMS Barrel, Second National Conference on Particle Physics, Primorsko, 5 – 10 June 2007.

ROOT. (2024-a). https://en.wikipedia.org/wiki/ROOT (Accessed: December 2024)

ROOT. (2024-b). Analysing petabytes of data, scientifically, https://root.cern.ch/ (Accessed: December 2024)

Shah, M. A. (2016). Comparison of CMS Resistive Plate Chambers performance during LHC RUN-1 and RUN-2. arXiv preprint arXiv:1605.09366.

Shumka, E. et al. (2023). Machine Learning based tool for CMS RPC currents quality monitoring. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1054, 168449.