Comput Softw Big Sci (2024) 8:11
https://doi.org/10.1007/s41781-024-00118-z

Research

Autoencoder-Based Anomaly Detection System for Online Data Quality Monitoring of the CMS Electromagnetic Calorimeter

The CMS ECAL Collaboration*

Received: 20 September 2023 / Accepted: 19 March 2024

Abstract The CMS detector is a general-purpose apparatus that detects high-energy collisions produced at the LHC. Online data quality monitoring of the CMS electromagnetic calorimeter is a vital operational tool that allows detector experts to quickly identify, localize, and diagnose a broad range of detector issues that could affect the quality of physics data. A real-time autoencoder-based anomaly detection system using semi-supervised machine learning is presented enabling the detection of anomalies in the CMS electromagnetic calorimeter data. A novel method is introduced which maximizes the anomaly detection performance by exploiting the time-dependent evolution of anomalies as well as spatial variations in the detector response. The autoencoder-based system is able to efficiently detect anomalies, while maintaining a very low false discovery rate. The performance of the system is validated with anomalies found in 2018 and 2022 LHC collision data. In addition, the first results from deploying the autoencoder-based system in the CMS online data quality monitoring workflow during the beginning of Run 3 of the LHC are presented, showing its ability to detect issues missed by the existing system.

Keywords Anomaly detection · Machine learning · Autoencoders · Data quality monitoring · Calorimeter

1 Introduction

The CMS experiment has been taking high-quality proton–proton (pp) collision data produced by the LHC at CERN for over a decade. Figure 1 shows a schematic view of the CMS detector. The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [1].

The energy of photons and electrons reconstructed in CMS events is obtained from the ECAL, the detector element most relevant for this paper. In the barrel section of the CMS ECAL, an energy resolution of about 1% is achieved for unconverted or late-converting photons in the tens of GeV energy range [2]. An excellent ECAL detector resolution is the basis for precision physics measurements and the detection of new particles, as was the case for the discovery of the Higgs boson H [3,4]. For example, the diphoton mass resolution measured in the Higgs boson decays, H$\rightarrow \gamma\gamma$, is typically in the 1–2% range, depending on the measurement of the photon energies in the ECAL and the topology of the photons in the event [5]. Precision measurements of particle decays such as H$\rightarrow \gamma\gamma$ do not only rely on excellent detector calibrations but start with the recording of high-quality raw data.

The CMS data quality monitoring (DQM) system [6] is a crucial operational tool to record high-quality physics data. Presently, the DQM consists of a software system that produces a set of histograms that are based on a preliminary analysis of a subset of data collected by the CMS detector. Conventional cut-based thresholds are used to define quality flags on these histograms which are monitored continuously by a DQM shifter in the CMS control room who reports on any apparent irregularities observed. While this system has proven to be dependable, the changing running conditions and increasing collision rates, together with aging electronics, bring forth failure modes that are newer and harder to predict. Machine learning (ML) techniques are nowadays widely used in high-energy physics [7] and provide an excellent tool for anomaly detection in particle physics searches (see, e.g., Ref. [8]). Thus, ML is a natural choice for an automated system monitoring the data quality of an experiment. Previ-

* e-mail: aharilal@andrew.cmu.edu

Published online: 24 June 2024

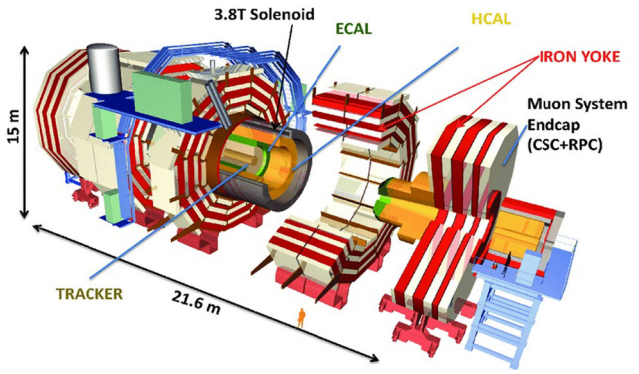


Fig. 1 Schematic view of the CMS detector and its various subdetectors

ous efforts in ML for DQM within the CMS collaboration have explored such techniques [9,10]. In this paper, a semi-supervised method of anomaly detection for the online DQM of the CMS ECAL is presented, exploiting an autoencoder (AE) [11] on ECAL data processed as two-dimensional (2D) images. In a novel approach, correction strategies are implemented to account for spatial variations in the ECAL response as well as the time-dependent nature of anomalies in the detector.

While preliminary studies performed for the ECAL barrel [10] showed the potential of detecting anomalies using AE for DQM, significant improvements in AE reconstruction and resolution are presented in this work together with novel post-processing strategies based on physics insights that enable the detection and localization of various anomalies with a very low false detection rate. The ML-based system is deployed in the online DQM for the ECAL during LHC Run 3 collisions, complementing the existing DQM plots. The first results indicate that the AE-based anomaly detection system is a highly valuable diagnostic tool for ECAL experts involved in real-time data taking operations.

This paper is organized as follows: "Experimental Environment" section introduces the CMS ECAL and its current DQM system, and "Machine Learning-Based Anomaly Detection Strategy" section presents the AE network and the AE-based anomaly detection strategy, including the data sets used and their pre-processing. "Training, Evaluation, and Post-processing" section discusses the strategy of training and validation of the network, as well as the corrections that are applied to account for the deviations of the ECAL response as a function of position and time. A metric to assess the performance of the AE-based anomaly detection method is also described, and a comparison to a baseline scenario similar to the existing ECAL DQM system is explained. "Results" section presents the AE performance on validation and test data sets with anomalies. "Deployment during LHC Run 3" section discusses the deployment of the real-time AE-based anomaly detection system in the Run 3 online ECAL DQM operation, followed by a summary in "Summary" section.

2 Experimental Environment

2.1 Proton–Proton Collisions at the LHC

The LHC has provided pp collisions at center-of-mass energies of 7 and 8 TeV (2009–2012), rising to 13 TeV (2015–2018), and further to 13.6 TeV for the ongoing Run 3 that started in 2022. Each LHC beam consists of about 2500 tightly packed bunches of $\sim 10^{11}$ protons with 25 ns bunch spacing. The beams travel in opposite directions in the beam pipe and collide at four intersection points, one of them being the center of the CMS detector. LHC operations involve "fills", where a fill is defined as a period during which the same proton beams are circulating in the LHC, and it typically consists of ten or more hours of collisions. The collision rate expressed through the instantaneous luminosity varies during each fill, and it decreases with time as the number of particles in the proton bunches and the beam intensity decay. Additional pp interactions within the same bunch crossing, referred to as "pileup" (PU), can contribute to additional tracks and calorimetric energy depositions, increasing the event activity in the detector. The PU is correlated with the instantaneous luminosity and is thus higher at the beginning of a fill than at the end.

An LHC fill is often divided into CMS "runs" that correspond to a start and stop of the CMS data acquisition system. Each run is further divided into time intervals called "lumisections" (LS) of an approximate time duration of 23 s corresponding to 2^{18} LHC orbits, over which the instantaneous luminosity is considered to remain approximately constant.

2.2 The CMS Electromagnetic Calorimeter

The CMS electromagnetic calorimeter provides homogeneous coverage in pseudorapidity $|\eta| < 1.48$ in a barrel region (EB) and $1.48 < |\eta| < 3.0$ in two endcap regions (EE+ and EE−), as shown in Fig. 2. Preshower detectors consisting of two planes of silicon sensors interleaved with three radiation lengths of lead are located in front of each endcap detector. The ECAL consists of 75 848 lead tungstate ($PbWO_4$) crystals. The barrel granularity is 360-fold in ϕ and (2×85)-fold in η provided by a total of 61 200 crystals, with each crystal having a dimension of 0.0174×0.0174 in $\Delta\eta\times\Delta\phi$ space, while each endcap is divided into two halves, with each comprising 3662 crystals.

Light signal from the ECAL crystals is detected, amplified, and digitized every 25 ns. The collected data is stored in on-detector buffers awaiting an Accept/Reject signal from the first stage of the global CMS trigger system [12,13]. Upon

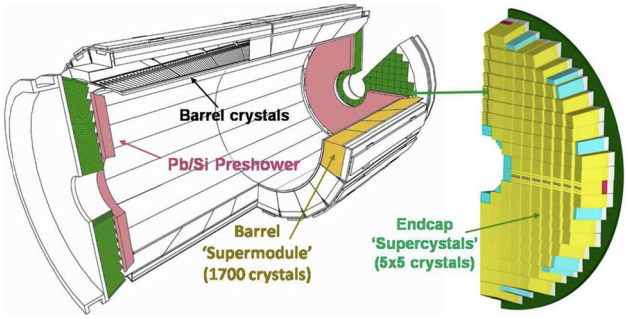


Fig. 2 Schematic view of the ECAL showing the cylindrical barrel closed by the two endcap regions with one half endcap displayed

receiving the Accept signal, a series of ten consecutive digitized samples are read out for each channel (or crystal) to measure the signal pulse amplitude and timing, after applying certain predetermined thresholds, to get an optimal pulse reconstruction as described in [14]. These ten samples for each channel are referred to as digitized hits or "digis". When a digi is registered for a given crystal, an occupancy value of 1 is counted for the crystal in any given DQM histogram. A "readout tower" (henceforth referred to as a tower) is defined as a set of 5×5 crystals ("supercrystals"), and 68 of these towers form a "supermodule" in the barrel (see Fig. 2). An unrolled projection of the ECAL barrel as well as one endcap is displayed in Fig. 3. Each single square in Fig. 3a, b represents a tower. The numbered rectangular regions in Fig. 3a represent the supermodules in the barrel, while the numbered regions in Fig. 3b indicate sectors in the endcaps.

The transverse energy of a trigger tower is computed by the front-end electronics and trigger concentrator cards (TCC) within the off-detector electronics in the ECAL. A classification flag is assigned by the TCC based on the threshold level that has been passed by the energy in each trigger tower. A dedicated hardware device, called the Selective Read-out Processor [15], receives from the TCCs the trigger tower energy maps and the corresponding flags, and it executes a selection algorithm [16] that classifies the trigger tower as one of the following: "suppressed" (energy is below a low threshold), "single" (energy is in-between the low and high thresholds), and "central" (energy is above the high threshold). For a "central" trigger tower, 3×3 or 5×5 regions around it are classified as "neighbors". The Selective Read-out Processor transmits the selective read-out flag for each tower to the Data Concentrator Cards, which read out the crystals as follows: for crystals that form "central", "neighbor", or "single" trigger towers, all energy samples are kept, while crystals belonging to "suppressed" trigger towers are ignored or read out with a high zero suppression threshold.

2.3 Data Quality Monitoring in ECAL

During data taking, monitoring the data quality is a crucial, time-sensitive task to ensure optimal detector performance and the recording of high quality data suitable for physics analyses. The CMS DQM [6] has two main modes of operation: offline and online. The offline DQM gives a retrospective view of data processed with the full statistics, passing various offline processing chains. It is mainly used in CMS for data quality certification [17]. The data are manually certified as good or not-good by comparing against several reference distributions, expert knowledge of running conditions, and known issues.

The online DQM offers a real-time snapshot of a subset of the raw data by populating a set of histograms after a preliminary analysis of the data, followed by a data quality interpretation step. These histograms are updated every LS and are accumulated over a run. They are monitored continuously by a DQM shifter who reports on any apparent irregularities observed and informs detector experts to identify and act in real-time on any related issues with the detector. The current online DQM has many built-in alarms based on the set of histograms and indicates the presence of errors in a way that makes it easy to spot any ECAL issues at a glance.

There are two kinds of histograms present in the ECAL DQM: "Occupancy-style" histograms shown in Fig. 3a, b filled with critical quantities from the real-time detector data and "Quality-style" histograms displayed in Fig. 3c, d. Quality-style histograms are obtained by applying predefined thresholds and requirements to the Occupancy-style histograms, where the thresholds are derived from typical detector response. In the examples shown in Fig. 3, the Occupancy-style histograms are plotted at a tower-level granularity while the Quality-style histograms are plotted at a channel-level granularity. The quality histograms are drawn in easily identifiable colored maps, and the color code scheme used is as follows: green for "good", red for "bad", brown for "known problems", and yellow for "no data" (that may or may not be problematic depending on the context). Information about known bad channels and towers is displayed in a channel status map with an example given in Fig. 4. Here, the different colors correspond to the status values as: 0—channel OK, 1—channel with pedestal not in range, 2—channel with no laser, 3–7—various types of noisy channels, 8–9—channels in fixed gain, 10–14—various types of dead channels, and status > 14—channels with issues in low voltage or high voltage. Depending on the severity of the problem, these channels are either masked in the Data Acquisition system of the detector or in the offline reconstruction, and the data from them are ignored and may or may not show up in the DQM. This information is stored and regularly updated in a database and is used to mark the towers in the DQM quality plots in dark colors, e.g., dark brown.

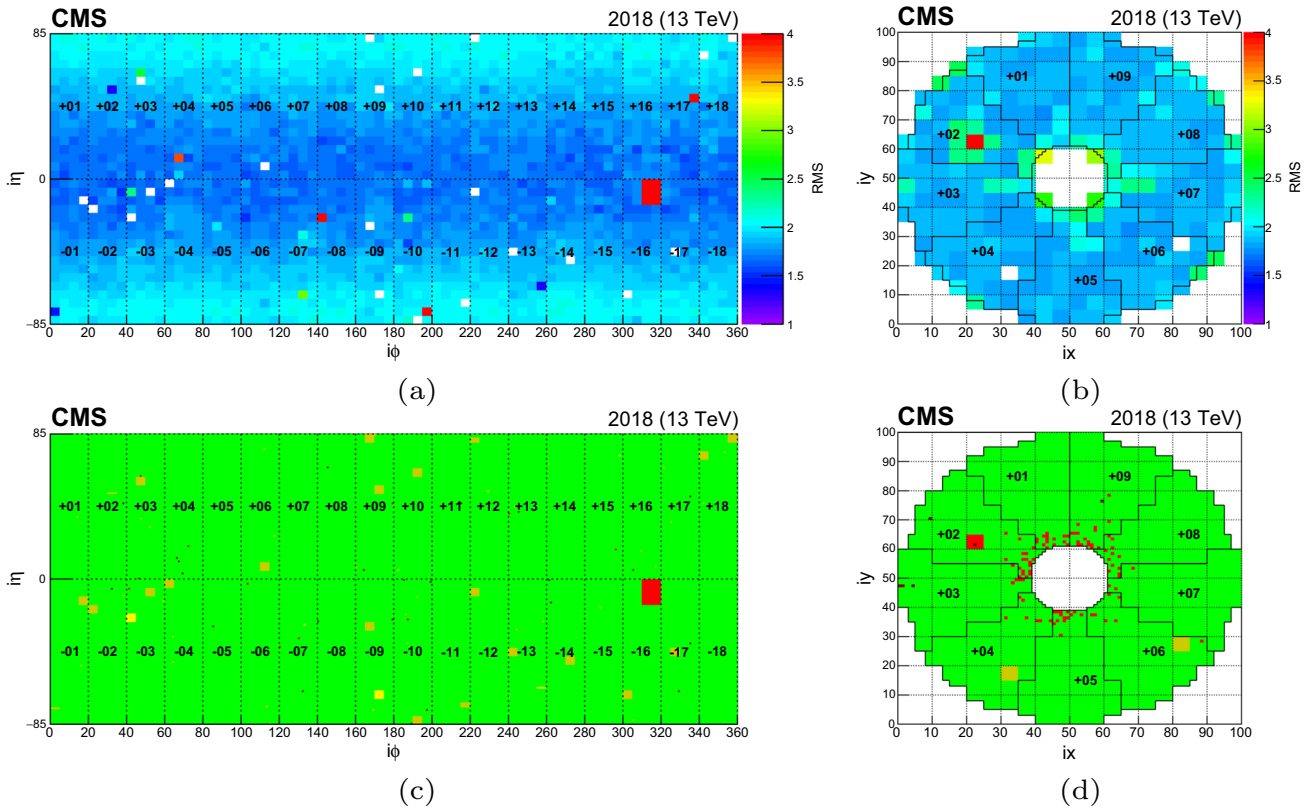


Fig. 3 Example histograms from the ECAL DQM with **a** and **b** showing the distribution of RMS of the pedestal values in the barrel and EE+, respectively, drawn at a tower-level granularity. Diagrams **c** and **d** show the corresponding quality map for the two regions, drawn at a channel-level granularity, after a set of cuts is applied on the noise values shown in **a** and **b**

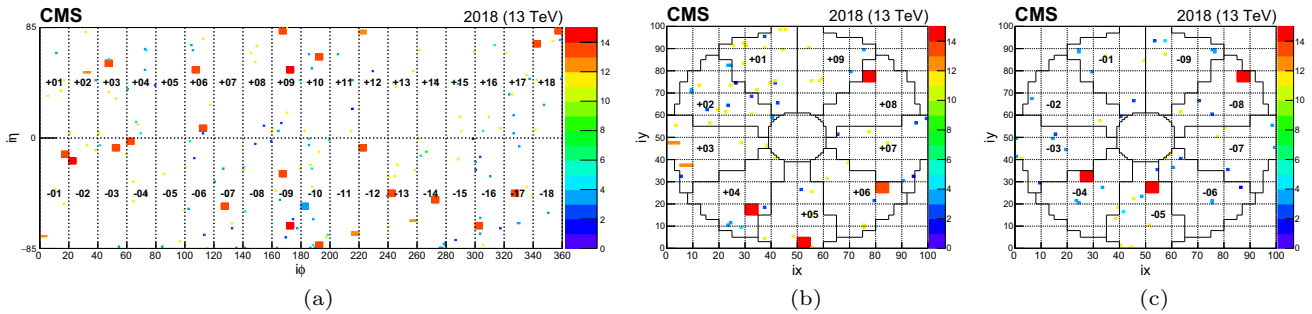


Fig. 4 Channel status maps used in the ECAL DQM indicating the known problematic channels color coded for various types of errors for **a** EB, **b** EE+, and **c** EE−

As an example of existing DQM plots, Fig. 3a, b shows the distribution of the pedestal Root Mean Square (RMS) in EB and EE+, respectively. Regions with high noise could indicate, for instance, a potential problem with the high voltage in the detector. Figure 3c, d indicates the corresponding quality maps. Here, the crystals are shown in red if their RMS values are greater than the set thresholds.

Anomalies can come in different shapes and sizes, as illustrated in Fig. 5, attributed to various sources such as underlying hardware components. Furthermore, the ever-changing LHC and CMS running conditions can often result in failure modes that are hard to predict. Although continuous improvements of the DQM have allowed the existing system to be updated and respond to new problems, e.g., issues with electronic components, it can become challenging to define hard-coded rules and thresholds manually for every failure mode. To overcome such challenges, an automated anomaly detection system using machine learning is developed to complement the existing DQM.

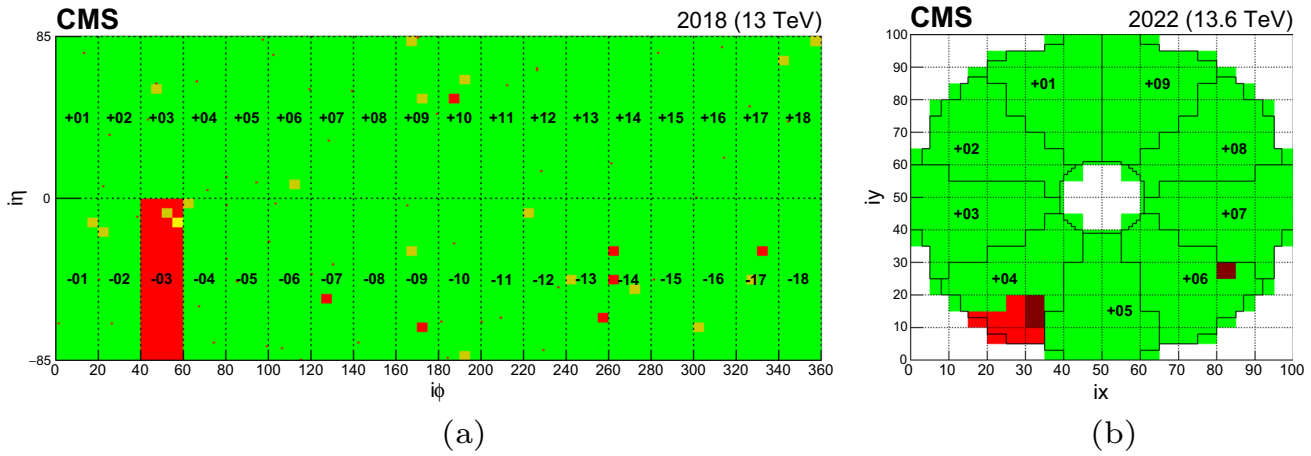


Fig. 5 DQM quality plots with different anomalies shown in red, while the towers with known issues show up as dark brown or dark yellow. **a** EB−03 turned off due to a voltage failure and seen in red. **b** Anomaly in EE+04 (marked in red) originating from an electronics failure affecting 200 channels

3 Machine Learning-Based Anomaly Detection Strategy

Unsupervised or semi-supervised ML methods used for anomaly detection are an excellent choice to supplement the ECAL DQM system. With the existing manual data certification procedure in CMS and ECAL, the recorded offline data is certified to be good or bad for physics analyses and/or for detector calibration based on various defined markers. Using a semi-supervised approach, the network is trained exclusively on the certified good physics data set, so that it learns the patterns of good data and is able to detect anything that differs from the nominal patterns it has learned. The network is able to detect anomalies without the need to explicitly see the anomalous data during training. A semi-supervised anomaly detection and localization method for the online ECAL DQM is developed using an AE network based on a computer vision technique. The AE is built with a convolutional neural network (CNN) architecture [18] exploiting ECAL data processed as 2D images. Corrections that take into account the spatial variations in the ECAL response and the time-dependent nature of anomalies in the detector are implemented in order to effectively maximize the anomaly detection efficiency while minimizing the false positive detection probability.

3.1 Autoencoder Network

An AE is used with a Residual Neural Network (ResNet) [19] CNN architecture implemented with a PyTorch [20] backend. The encoder part of the AE takes the input data and compresses it into a lower dimensional representation, called the latent space, which contains a meaningful internal representation of the input data. The decoder part then decompresses the encoded data back to the original image of the same dimensions, or reconstructs the image. To measure how well the output matches the input, or the goodness of the AE reconstruction, a reconstruction loss ($\mathcal{L}$) [1] is computed using Mean Squared Error between the input (x) and the AE-reconstructed output (x') as defined in Eq. (1):

$$\mathcal{L}(x, x') = ||(x - x')||^2 \tag{1}$$

A network trained on good images will learn to reconstruct them well by minimizing this loss function. When fed with anomalous data, the AE returns higher loss in the anomalous region, forming the basis of the anomaly detection strategy discussed in "Anomaly Detection Strategy" section.

Figure 6 shows the architecture used for the AE. Each "ResBlock" consists of two convolutional layers, with a Rectified Linear activation function (ReLU) [21] in between and a residual mapping. The input image (shape of 36×72 for EB and 22×22 for EE) is passed through the encoder network that consists of a CNN, followed by a maxpool layer that aggregates the maximum values of the feature maps. It is then sent through sequential layers of ResBlocks, where the feature maps are up-sampled progressively. This is followed by a global maxpool that creates a compressed dense layer of the encoded space. The encoded layer is then passed to the decoder network as the input, which reverses these operations and gives out a reconstructed image. Three separate models are trained with this architecture: one for the barrel and one for each of the two endcaps. The choice of training the separate models for the barrel and the endcaps is attributed to the differences in their shape, granularity, and response.

[1] Note that the reconstruction loss in this paper always refers to the AE reconstruction loss and has no relation to CMS particle reconstruction.

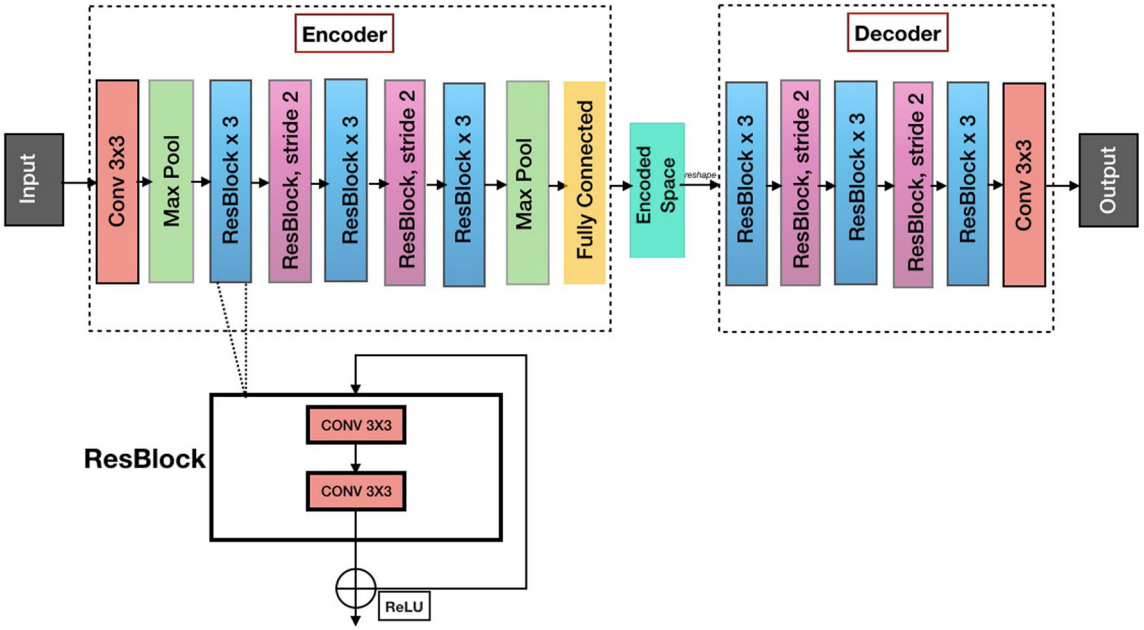


Fig. 6 AE architecture showing the encoder and decoder networks, with the ResBlock structure displayed

3.2 Data Set and Event Selection

The data set used for training and validation of the AE network is taken from CMS runs collected in 2018 during LHC Run 2. It contains runs that are manually certified as good by the data certification procedure in CMS and ECAL. Each input image for the AE is the digi occupancy map from a single LS that can be seen from Fig. 7. It is to be noted that the current DQM checks a much larger phase space of detector quantities. Digi occupancy maps are chosen for the ML-based DQM as they are a good indicator of most detector problems. The occupancy maps are processed offline to include 500 events per LS, which is approximately the number of data received per LS in the online DQM. Although the actual number of events varies per LS in real-time data taking, this approximation is used to ensure that the network is sensitive to variations in the occupancy due to anomalies and not due to differences in the collected statistics. It is demonstrated in "Deployment during LHC Run 3" section that the AE-based system performs very well in the online DQM deployment on real-time data with varying numbers of events per LS.

3.3 Detector Images

Occupancy histograms using digis from the ECAL barrel and endcap sections are fed to the network as 2D images for each LS. The occupancy images are drawn at a tower-level granularity with the image shapes of 34×72 for the barrel and 20×20 for the endcaps. Typical occupancy images for a single LS are shown in Fig. 7. Different from the usual crystal-level indexing used in most ECAL DQM plots, the occupancy maps displayed here have a modified tower-level indexing for convenience, $i\eta_{tow}$-$i\phi_{tow}$ for the barrel and ix_{tow}-iy_{tow} for the endcaps. For the empty regions in the endcap images, e.g., (ix_{tow},iy_{tow})=(0,0), their occupancy values are set to zero during training and the loss values of these regions are not taken into account during inference.

The input images to the network are padded at the edges to mitigate edge effects, since learning at the boundaries becomes sub-optimal due to under-representation of the edge values during convolution. After padding by duplicating the first and last rows of the image for the barrel, the input image shape is 36×72. For the endcaps, padding both the first and last rows and columns gives a 22×22 input image shape. During the inference, however, the original shape of the images without the padding is used.

3.4 Pre-processing

To ensure consistent data quality interpretation across different LHC running conditions, it is important to find a coherent way of normalizing the occupancy to make it independent of factors such as the LHC instantaneous luminosity, which is correlated with PU. As shown in Fig. 8a, the data set used for training indicates a linear relation between occupancy and PU at first order. A linear fit is performed to the distribution and a correction factor is derived from the fit parameters. After the correction is applied to each occupancy map per LS as pre-processing, an almost flat relation between occupancy and PU is obtained, as shown in Fig. 8b. After removing the PU dependence across the data set with the correction,

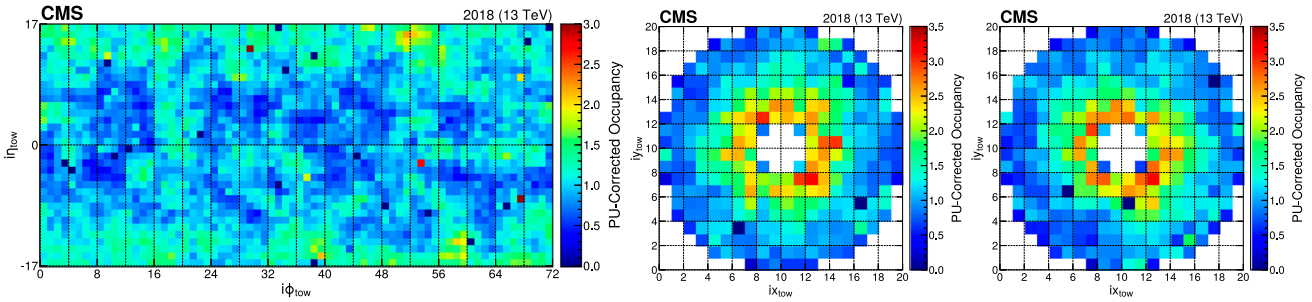


Fig. 7 Typical occupancy maps for a single LS from the training data set for **a** EB, **b** EE+, and **c** EE−. During training, the edges of these maps are padded

each occupancy map is then re-scaled such that the average occupancy across the towers is around one, giving typical occupancy maps as in Fig. 7.

3.5 Anomaly Detection Strategy

Figure 9 illustrates the anomaly detection strategy for the ECAL DQM using endcap images as an example. The input occupancy image (top left) is fed to the AE, which outputs a reconstructed image (top right). Then the Mean Squared Error on each tower is calculated and plotted as a 2D loss map in the same coordinates ($i\eta_{tow} - i\phi_{tow}$ for the barrel and $ix_{tow} - iy_{tow}$ for the endcaps). As shown in the bottom-right panel, the anomalous region is highlighted with the loss higher than the rest of the image. After applying some post-processing steps explained in "The ECAL Spatial Response Correction" and "Time Correction" sections, a threshold to flag the anomaly is calculated based on the anomalous loss values. The threshold is applied to the post-processed loss map to create a quality plot (bottom left), where towers with the loss above the threshold are tagged as anomalous and are shown in red. Towers with the loss below the threshold are identified as good and are shown in green. The final quality plot can be easily and quickly interpreted by a DQM shifter.

4 Training, Evaluation, and Post-processing

4.1 Training and Validation

The available data set consists of 100 000 good images processed offline with each image corresponding to a single LS. This data set is split into a training and a validation set with a ratio of 9:1. In addition to the validation set with good images, another validation set is obtained that comprise the same good images but with "fake" anomalies introduced. Three kinds of anomalies are explored:

- Missing supermodule/sector: Entire barrel supermodules and endcap sectors are randomly set to have zero occupancy values in each LS.
- Single zero occupancy tower: A single tower is set to have zero occupancy at random in each LS. Such low-occupancy single towers are usually harder to detect.
- Single hot tower: A single tower is set to be "hot", or having higher-than-nominal occupancy. For a tower with 25 crystals and 500 events per LS, the average occupancy is of the order:

$$\text{occupancy} = 25 \times 500 \times f, \tag{2}$$

where f is the frequency of the readout. For the barrel, nominal f ranges from 0.01 to 0.03, while for the endcaps it ranges from 0.02 to 0.05. For $f = 1$, the readout is said to be in "full-readout". Hot towers with $f = 1$ are easier to detect as their values stand out clearly from the nominal value. Thus for the validation, a more challenging borderline scenario is targeted with $f = 0.1$ for the barrel and $f = 0.2$ for the endcaps. The readout frequency target is chosen to be higher for the endcaps, as the nominal occupancy values for the endcaps are larger in the higher $|\eta|$ region compared to the barrel.

These fake anomalies are used to derive thresholds on the loss maps for efficient anomaly tagging with the AE. While fake anomalies are representative examples of real anomalies that occur in the detector, the AE model is further tested on real anomalous data from the 2018 and 2022 runs as discussed in "Testing on Real Anomalies" section.

4.2 The ECAL Spatial Response Correction

Since the multiplicity of particle production in a fixed rapidity interval is constant at a hadron collider, the number of particles per geometric interval increases for higher $|\eta|$, which is related to rapidity. Due to this effect, it is observed that ECAL crystals in regions of high $|\eta|$ exhibit higher occupancy than

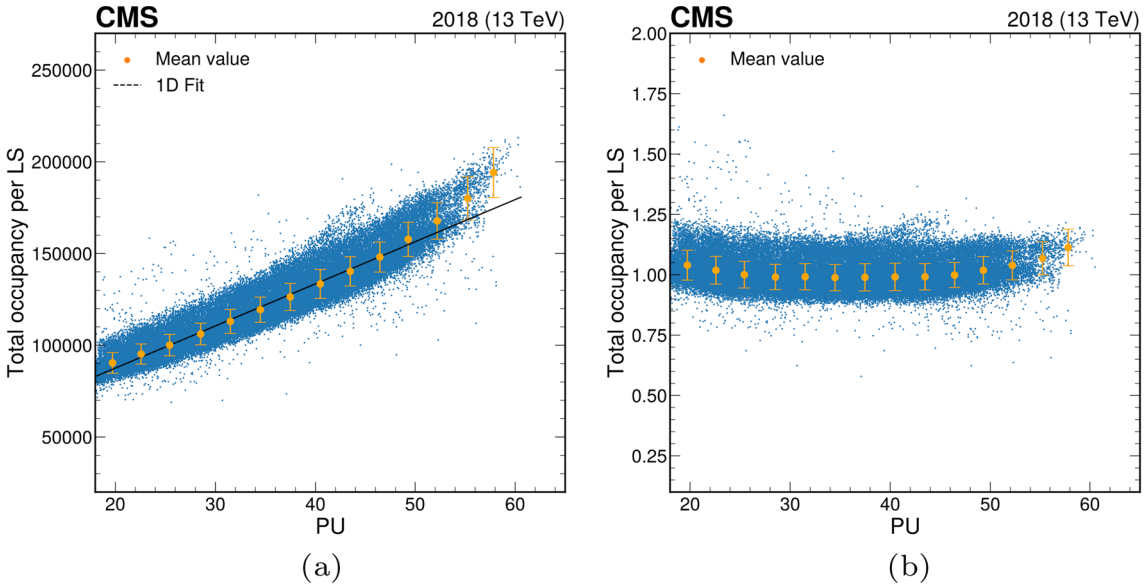


Fig. 8 Sum of digi occupancy for each LS occupancy map versus PU (**a**) before and (**b**) after the PU correction is applied. Overlaid on the scatter plot are the mean values with standard deviation for each PU bin

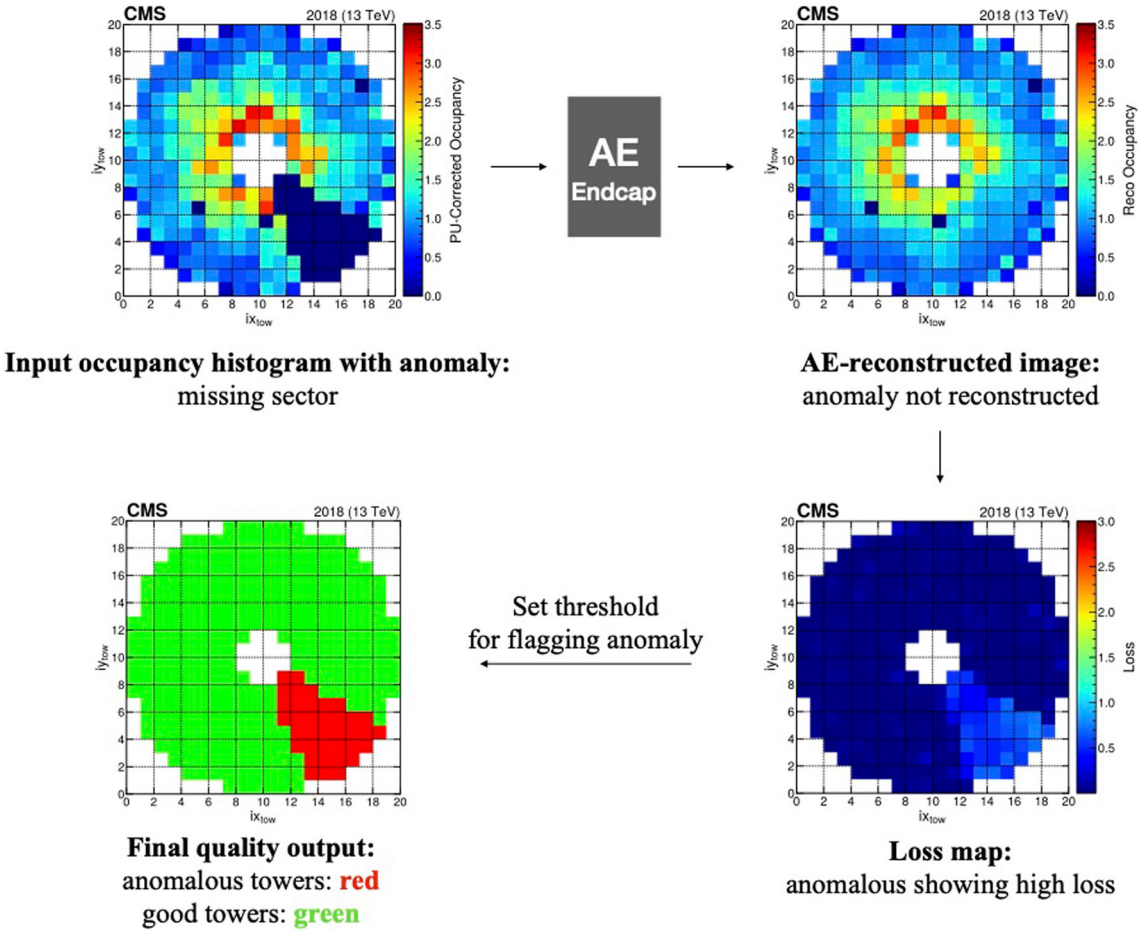


Fig. 9 Illustration of the AE-based anomaly detection strategy

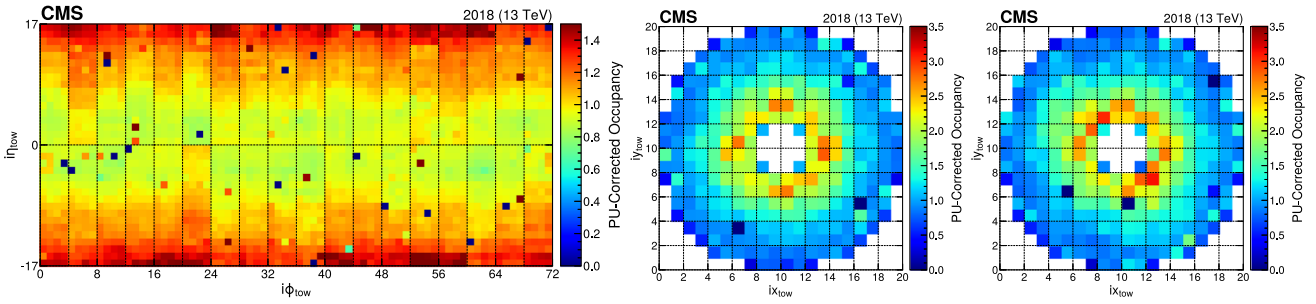


Fig. 10 Average occupancy map for **a** the barrel, **b** EE+, and **c** EE− from the 2018 data set containing runs manually certified as good by the CMS data certification

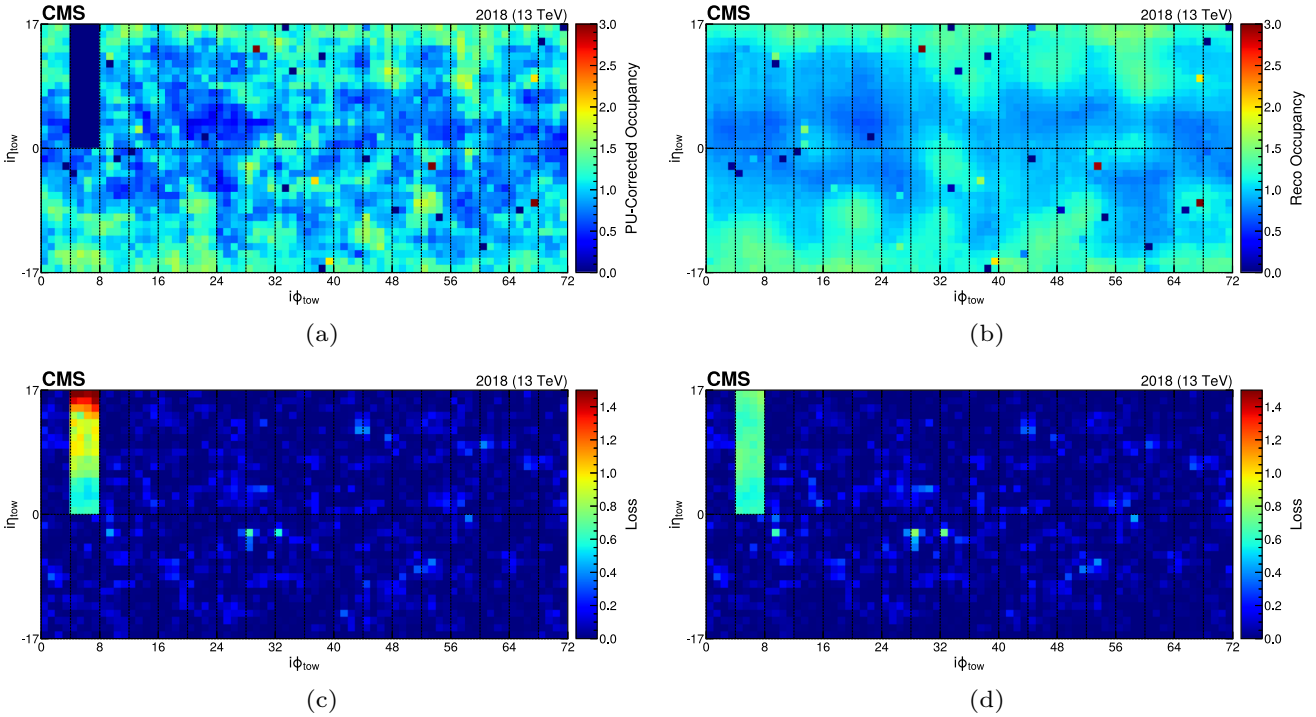


Fig. 11 **a** Occupancy map with a missing supermodule in the barrel. **b** AE-reconstructed occupancy map. **c** Loss map showing the missing supermodule, indicating higher loss at high $|\eta|$ owing to differences in the detector response. **d** Loss map after the spatial correction is applied

those of low $|\eta|$ in both the barrel and the endcaps, as can be seen in Fig. 10.

This difference in detector response is also visible in the AE loss map for specific anomalies as illustrated in Fig. 11 with a missing supermodule. Figure 11a shows the PU-corrected occupancy map with one supermodule having zero occupancy, and Fig. 11b reflects the corresponding AE-reconstructed output where the AE fails to reconstruct the anomaly. Figure 11c shows the tower-level loss map calculated between the input and output, exhibiting high loss in the anomalous supermodule region; the towers at the highest $|\eta|$ tend to have a higher loss than those at lower $|\eta|$ due to the higher average occupancy in these regions. To mitigate this effect and obtain uniform loss in the anomalous region, the loss is normalized by the average occupancy shown in Fig. 10a. After this "spatial response correction", flat loss is observed in the anomalous region, as shown in Fig. 11d, where all towers in the supermodule are interpreted as equally anomalous.

4.3 Time Correction

Real anomalies persist with time in consecutive LSs, while random fluctuations average out. A correction is implemented to exploit the time-dependent nature of the anomalies in the detector, named "time correction", which brings a significant improvement in the AE performance. Figure 12 shows the time correction strategy that is applied. Spatially

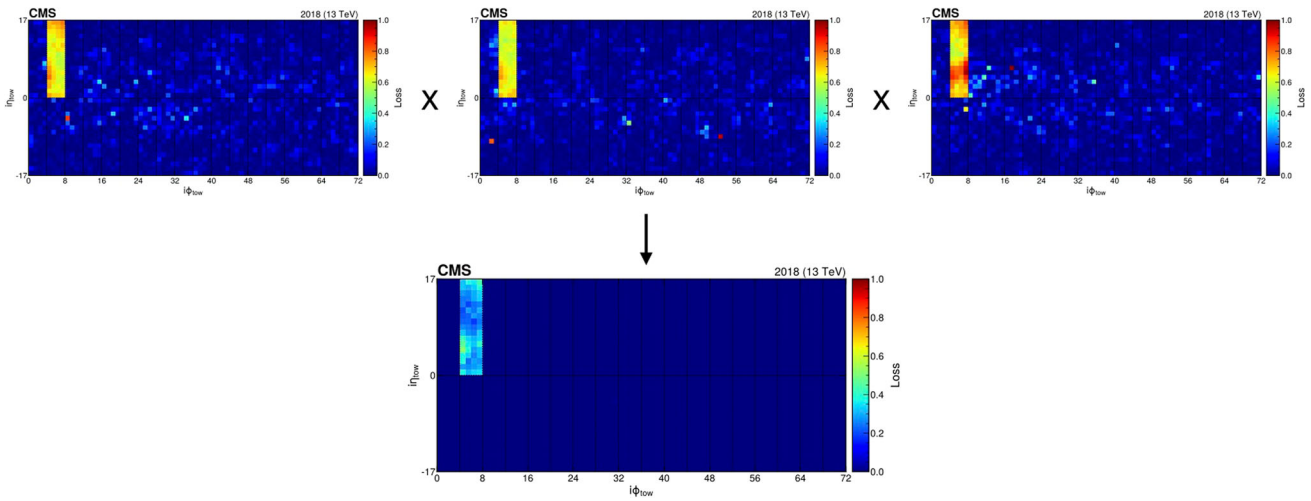


Fig. 12 Time correction strategy: multiplication of loss maps from three consecutive LSs in the barrel (top) to obtain a final loss map (bottom), where only the real anomaly of the missing supermodule is highlighted with high loss, while random towers with false positives are suppressed

corrected loss maps from three consecutive LSs (top panel) are multiplied together at a tower level. The resulting time-multiplied loss map at the bottom shows that the persistent anomaly of the missing supermodule is enhanced and random fluctuations visible in the loss maps from each LS are suppressed, reducing false positives. It is observed that multiplication rather than averaging is a better strategy for enhancing and suppressing the resulting loss values. Multiplication results in the low loss of good towers being smaller and the high loss of bad towers being larger, widening the gap between both and thus enhancing the distinction between both scenarios.

Given the duration of ~ 23 s for each LS with 500 events, time correction with three LSs yields a latency of approximately one minute. Including more LSs for the time correction is shown to bring no further significant reduction in the FDR in the AE performance during offline validation. Thus, one minute of latency is chosen to be an optimal trade-off for the time correction. For the online deployment where the LSs contain different number of events, however, this is changed to six LSs as discussed in "Deployment during LHC Run 3" section.

4.4 Anomaly Tagging Threshold and Performance Metric

The goal of the ML-based DQM system is to maximize the anomaly detection efficiency while minimizing the number of false positives. An anomaly is tagged using a threshold obtained from a validation set with fake anomalies. The threshold on the final post-processed loss map is chosen such that the loss values of 99% of anomalous towers are above the threshold, as illustrated in Fig. 13, which shows the loss distribution from a zero occupancy tower scenario. The bump

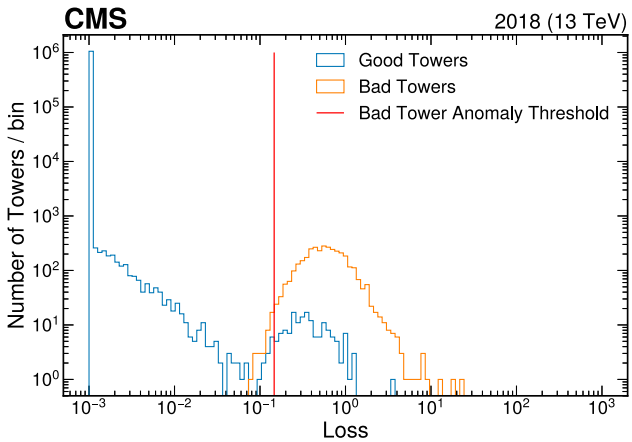


Fig. 13 Loss distribution for the zero occupancy tower scenario after spatial and time correction for EE−. The anomaly threshold is set as the lower 1% of the zero occupancy tower loss values

in the tail of the good tower loss distribution is discussed in "Testing on Fake Anomalies" section.

To assess the performance of the AE network, the false discovery rate (FDR) is used as a metric and is defined as

$$\mathrm{FDR} = \frac{\text{Number of good towers above } \epsilon}{\text{Number of good and bad towers above } \epsilon}$$

where, ϵ is the threshold for anomaly detection.

The FDR value for 99% anomaly detection represents the fraction of false detection in all anomalies detected, when using the threshold chosen to catch 99% of the anomalies present in the data set. In other words, the FDR is the ratio of good towers tagged as anomalous to all towers labeled as anomalous by the AE. A lower FDR indicates better performance and fewer false alarms during data taking. The FDR is

Table 1 Summary of FDR using 99% anomaly detection threshold for the ECAL barrel fake anomaly scenarios

	FDR for 99% anomaly detection		
	Missing Supermodule (%)	Zero Occupancy Tower (%)	Hot Tower (%)
Baseline (no correction)	14	90	5.2
Baseline (after time correction)	5.9	80	< 0.01
AE (no correction)	3.6	51	2.8
AE (after spatial correction)	3.1	49	2.9
AE (after spatial and time corrections)	0.13	4.1	< 0.01

Table 2 Summary of FDR using 99% anomaly detection threshold for fake anomaly scenarios in the endcaps

	FDR for 99% anomaly detection					
	Missing Sector		Zero Occupancy Tower		Hot Tower	
	EE+ (%)	EE− (%)	EE+ (%)	EE− (%)	EE+ (%)	EE− (%)
AE (no correction)	29	28	86	86	< 0.01	< 0.01
AE (after spatial correction)	1.8	2.2	11	14	0.02	0.04
AE (after spatial and time corrections)	0.06	0.18	1.4	4.4	< 0.01	< 0.01

calculated for each anomaly scenario using the corresponding anomaly tagging threshold during validation.

4.5 Comparison with Baseline Anomaly Detection Algorithm

A baseline study is carried out comparing the performance of the AE with a traditional cut-based approach. In this approach, baseline loss for a tower at a given (ϕ, η) position is calculated for the barrel using the occupancy of the tower ($t_{\phi\eta}$) and the average occupancy of all towers within the same η-ring ($\langle t_\eta \rangle$) as

$$\text{Loss of tower}_{\phi\eta} = |t_{\phi\eta} - \langle t_\eta \rangle|. \tag{3}$$

This approach is equivalent to the way most anomaly thresholds are defined in the standard ECAL DQM. A threshold is derived on this baseline loss using the same criteria as that for the AE, of achieving 99% anomaly detection from the fake anomaly validation. A similar baseline loss is not attempted for the endcaps, since the gradient of occupancy across the towers is much larger for the endcaps even within the same η-ring, and thus such baseline loss would not be a good method for detecting bad towers.

5 Results

5.1 Testing on Fake Anomalies

The performance of the AE-based DQM method is studied first on three distinct anomaly scenarios—missing supermodule/sector, single zero occupancy tower, and single hot tower—where artificial (fake) anomalies are added onto good images as outlined in "Training and Validation" section. Tables 1 and 2 summarize the FDR values calculated with anomaly tagging thresholds determined for each scenario for 99% anomaly detection. For the barrel where the baseline scenario is studied for comparison, it can be seen that the AE outperforms the baseline for all anomaly scenarios considered. For both the barrel and the endcaps, the FDRs for the single zero occupancy tower scenario are observed to be always higher than that for the single hot tower case. This is because hot towers are in general easier to spot, as they stand out with much higher occupancy compared to neighboring towers of average occupancy.

The effect of each consecutive correction on the FDRs can be seen from the tables. The AE spatial correction reduces the FDRs in the missing supermodule/sector and the single zero occupancy tower scenarios, where the occupancy values are set to zero for the barrel/endcaps. Without the correction, the loss values for the towers with zero occupancy anomalies are proportional to the towers' nominal occupancy, which indicates that the loss is biased to be larger in the higher $|\eta|$ region (see, e.g., Fig. 11c). The corresponding loss map exhibits an effective gradient across the map, following that of the average occupancy map shown in Fig. 10. It can be seen that the spatial correction has a greater effect for the endcaps than for the barrel, as the gradient in occupancy values across the towers is more pronounced for the endcaps.

In the case of the hot tower anomaly, the FDRs increase after the spatial correction. This is because the hot tower loss is biased to be higher in the opposite direction, towards the lower $|\eta|$ region. This leads to different effects of the

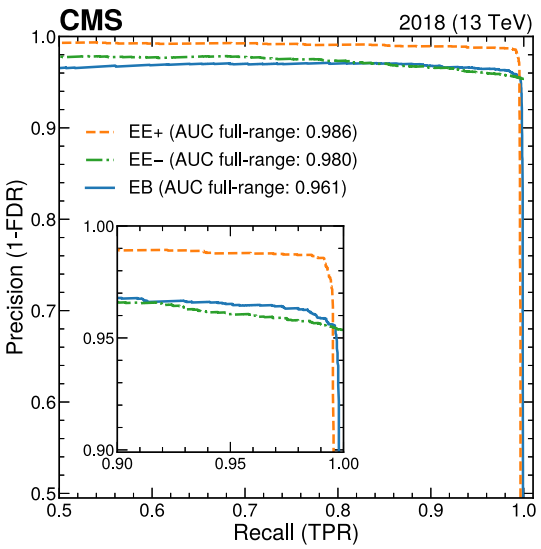


Fig. 14 Precision (1-FDR) versus Recall (TPR) curve for the zero-occupancy scenario after spatial and time corrections are applied. The inlay plot shows a zoomed-in view. The full-range area under the curves (AUC) are also indicated

spatial correction for different anomaly scenarios. For zero occupancy towers, spatial correction flattens out the gradient in the loss distribution and improves their detection. For the hot towers, the gradient is enhanced and the AE performance slightly drops. However, this effect is mitigated by the time correction that greatly improves the FDRs for all anomaly scenarios, with excellent final performance scores for both the barrel and the endcaps.

The remaining false positives, which show up as apparently "good towers" above the anomaly threshold after the time correction, are very likely to be actual anomalous towers that have been undetected so far in the data set of good LSs. These towers show up with the higher loss in the tail of the good tower loss distribution (see Fig. 13) and contribute to overestimating the FDR.

While 99% anomaly tagging efficiency is chosen as the working point, FDR values at different working points can be seen from Fig. 14 for the most challenging scenario of zero-occupancy tower after the spatial and time corrections are applied. The observed difference in the performance of the different parts of the detector is attributed to the different amounts of contamination of actual anomalous towers in the validation data set. In both barrel and endcaps, high-precision high-recall of the AE-based anomaly detection is clearly displayed.

5.2 Testing on Real Anomalies

Following the tests on fake anomalies, the AE performance is studied on known anomalous data from LHC runs in 2018 and 2022. The input occupancy images with anomalies and the final quality plots from the AE loss maps are illustrated in Figs. 15 and 16. Figure 15a shows the barrel occupancy map with the supermodule EB−03 in error due to a data unpacking issue from a 2018 run. The AE quality output in Fig. 15b correctly highlights the anomalous supermodule region in red. Figure 15c shows an occupancy map with a region of hot towers and a zero occupancy tower in the center from a 2018 run, and Fig. 15d correctly identifies all the anomalous towers shown in red. It is interesting to note that this error was not detected in the online DQM global quality plots at the time of data taking, while the AE is able to detect it.

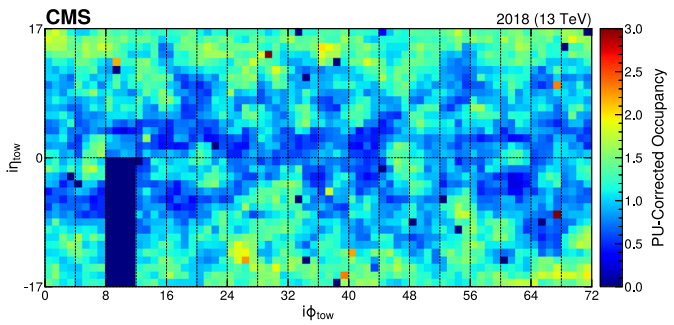
A similar test on real anomalies for the endcaps is illustrated in Fig. 16. An anomaly of more than half of the EE+ turned off in a 2018 run (see Fig. 16a) is spotted by the AE quality plot shown in Fig. 16b. Some of the green towers in the red region come from the masked, known problematic towers that the AE has learned during training. Figure 16c shows a region of towers turned off in the upper left quadrant of EE+ from a run in 2022, and the AE quality plot in Fig. 16d correctly identifies these towers in red.

The AE is thus able to spot various kinds of anomalies at a tower-level granularity using a single threshold. It does not require any set definitions or rules on the type of anomalies that can be detected, which underlines the importance of unsupervised or semi-supervised ML as a powerful and adaptable anomaly detection tool.

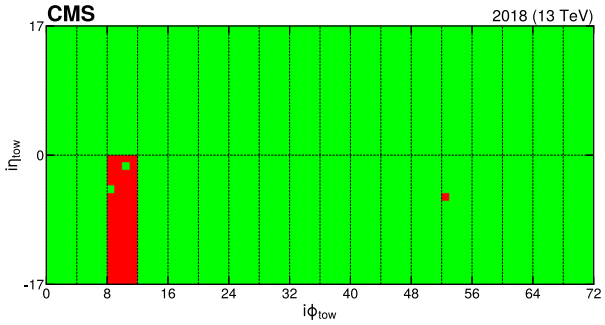
6 Deployment During LHC Run 3

The AE-based anomaly detection system labeled MLDQM has been deployed in the ECAL online DQM workflow in CMSSW for the barrel starting in LHC Run 3 in 2022 and for the endcaps in 2023. New ML quality plots from the AE (see Figs. 17a and 18a) have been added to the ECAL DQM. The model inference is accomplished using the trained Pytorch models exported to ONNX [22], which is implemented in the CMS software framework using ONNX Runtime [23].

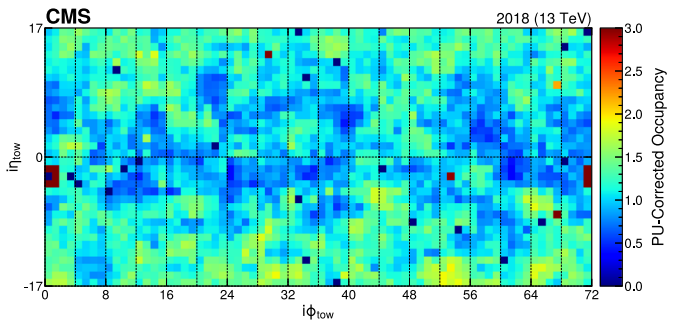
The MLDQM models deployed for the barrel and endcaps have shown so far very good performance with Run 3 data. Figure 17a illustrates the new quality plot obtained from the inference of the trained AE model for the barrel, using the digi occupancy histogram shown in Fig. 17b as input to the model. Similar plots are shown for the endcaps in Fig. 18a, b, which are obtained from the inference of the endcap models with the digi occupancy histograms in Fig. 18c, d as inputs. The number of events per LS in the digi occupancy map received by the DQM in Run 3 is approximately 100 – 150, smaller than the 500 events per LS used for training. Accordingly, the occupancy maps are summed over four consecutive LSs to collect sufficient statistics. The summed occupancy map is then fed as an input to the AE model, after necessary corrections are applied with respect to the number of



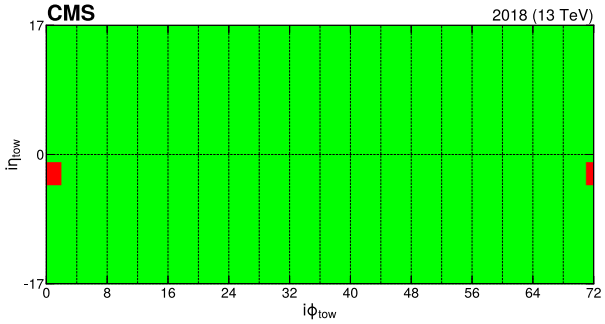
(a) From a 2018 run: Input occupancy map with an error in the supermodule EB−03



(b) From a 2018 run: Final AE quality plot



(c) From a 2018 run: Input occupancy plot with a group of hot towers



(d) From a 2018 run: Final AE quality plot

Fig. 15 Input occupancy images with real anomalies and corresponding AE quality plots for the barrel

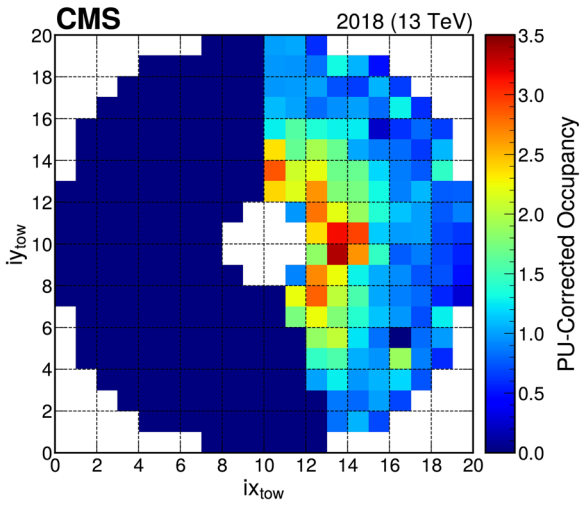
events and PU. The resulting loss map then undergoes both the spatial and time corrections. Six consecutive loss maps are used for the time correction during online deployment to minimize false alarms, with the final quality plot essentially accumulated over nine LSs.

From the AE quality plot in Fig. 17a, two circled red towers can be seen in the supermodules EB+06 and EB−06, both corresponding to zero occupancy towers in the input occupancy map, as shown in Fig. 17b. The AE quality plot also contains two brown towers that correspond to a zero occupancy tower in EB−09 and a hot tower in EB−13. Both are known problematic towers.

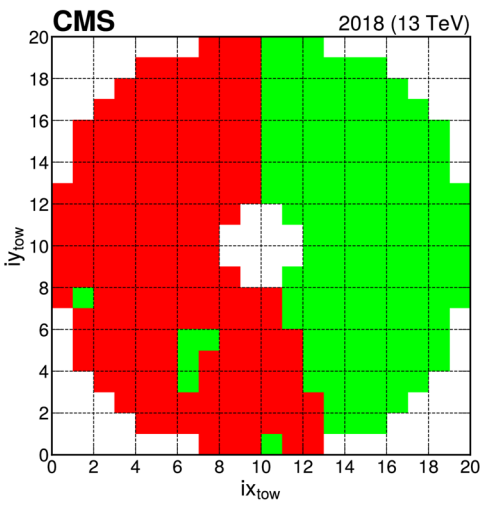
Similarly, in the AE quality plot for the endcap in Fig. 18a, a red tower can be seen in the sector EE−02, corresponding to a zero occupancy tower in the input digi occupancy map in Fig. 18c. From the quality plot shown in Fig. 18b, a brown tower in the sector EE+05 corresponds to a zero occupancy tower in the input digi occupancy map in Fig. 18d, which is a known problematic tower. Other zero occupancy towers in the input occupancy maps that do not show up in the AE quality plots correspond to other known problematic towers that have been present since Run 2. These known bad towers are learned by the AE during the training process.

6.1 Detecting Degrading Towers

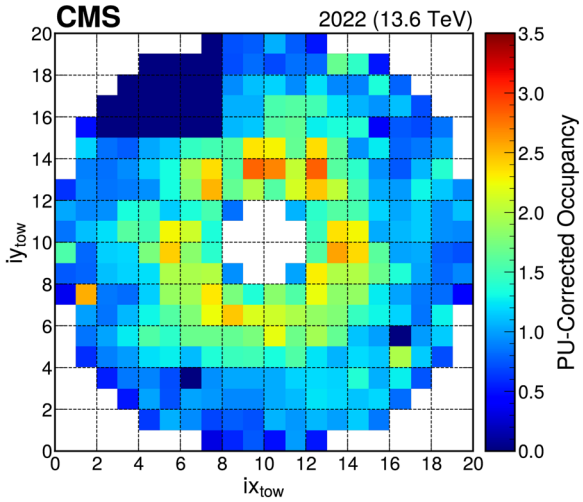
During the MLDQM deployment, it has been observed that the AE can catch new problematic towers with transient anomalous behaviors, which are hard to detect and can be missed by the existing DQM software and plots. Figure 19a shows an AE quality plot with two towers in EB+08 marked in red. Figure 19b shows the total digi occupancy map accumulated over all LSs from the entire run in the online DQM. Here, Tower 1 is visible with very low occupancy compared to other towers, indicating that it is a persistent zero occupancy tower. On the other hand, the faint visibility of Tower 2 reflects that it likely had zero occupancy in several LSs but not in all, possibly corresponding to a transient anomaly. This feature is also observed in the occupancy map averaged over several runs in Run 3, e.g., see Fig. 19c. The low occupancy of Tower 2 in this offline-produced plot implies that the tower indeed had zero occupancy for many LSs in these runs. It is not a random occurrence as in the case of single-event upsets [24] that frequently happen in the detector and are recovered quickly. The ability of the AE to identify degrading channels can be of immense use to detector experts when monitoring the health of the ECAL. MLDQM can be used to keep track of how often a particular tower is flagged as bad



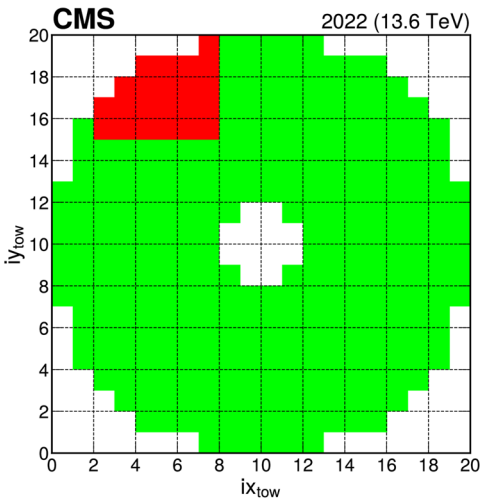
(a) From a 2018 run: Input occupancy plot with more than half of the EE+ turned off



(b) From a 2018 run: Final AE quality plot

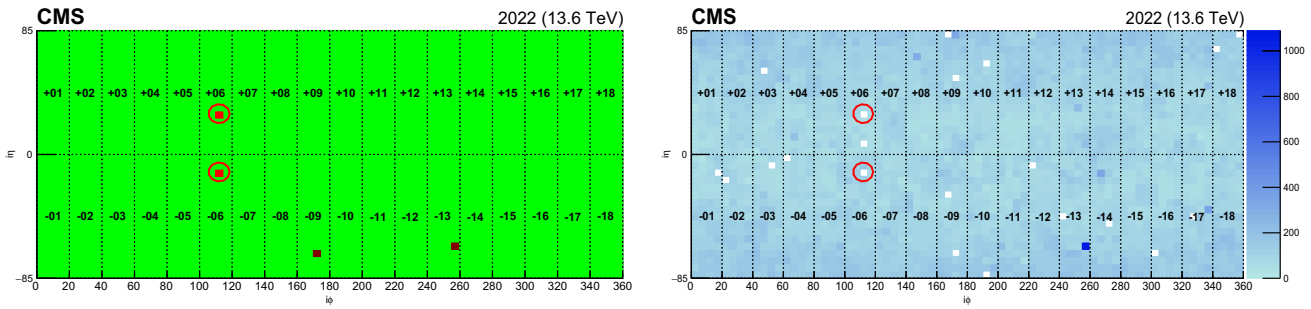


(c) From a 2022 run: Input occupancy plot with a region of towers turned off in EE+



(d) From a 2022 run: Final AE quality plot

Fig. 16 Input occupancy images with real anomalies and corresponding AE quality plots for the endcaps



(a) (b)

Fig. 17 From a 2022 Run: **a** ML quality plot in the ECAL DQM from the AE model, with the new bad towers circled. **b** Digi occupancy plot of 1 LS, with the same circled towers with zero occupancy. Four such occupancy plots from four consecutive LSs are summed to make the input to the AE model which results in the quality plot in a

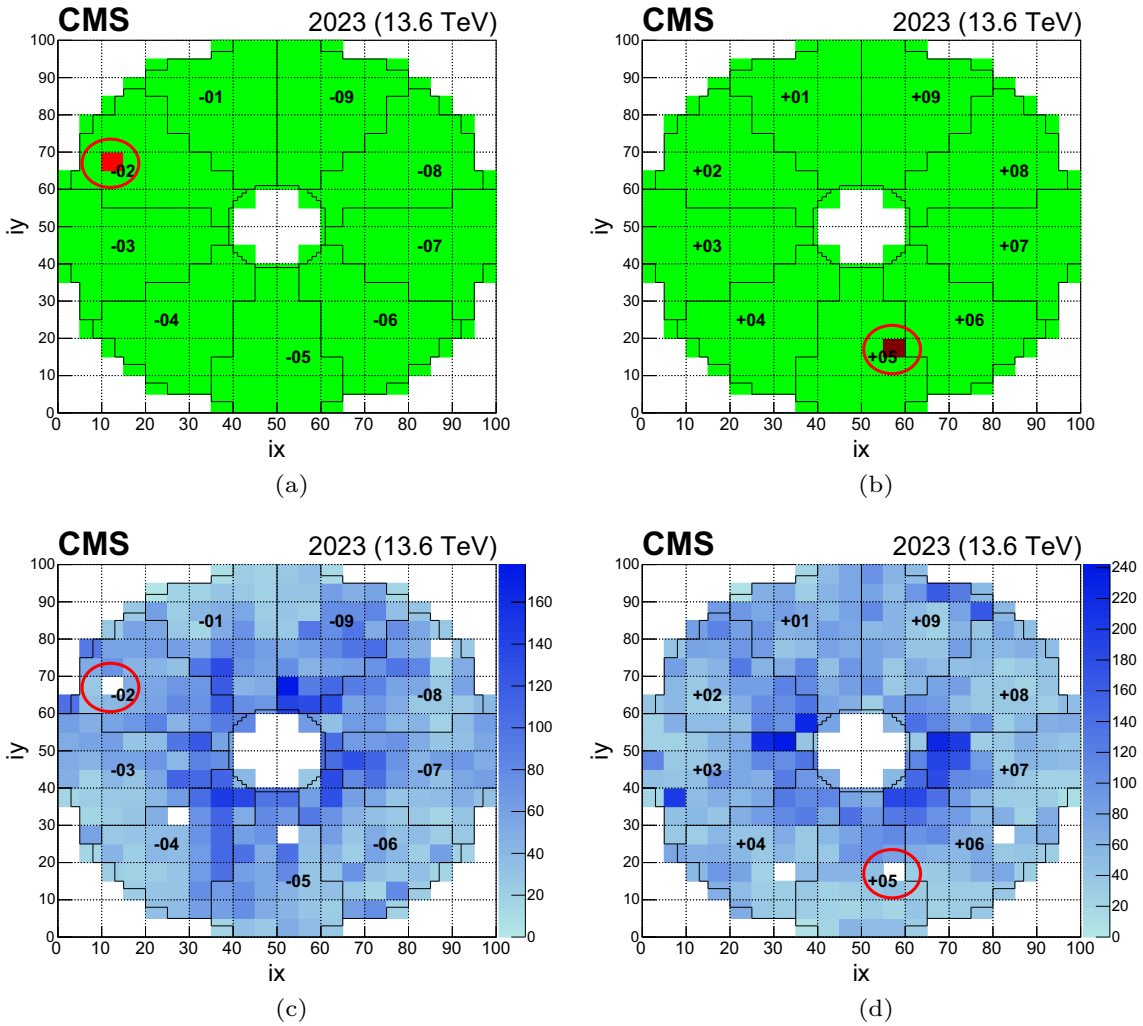


Fig. 18 From a 2023 Run: ML quality plot in the ECAL DQM from the AE models a for the Endcap EE−, with **a** new bad tower circled and **b** for the Endcap EE+ with a known bad tower circled. Digi occupancy plots of 1 LS **c** of EE− and **d** of EE+ with the corresponding circled towers with zero occupancy. Four such occupancy plots from four consecutive LSs are summed to make the input to the AE model which results in the quality plots in a and b

by the AE, and a threshold can be defined with this frequency to mask the transient tower proactively.

7 Summary

An Autoencoder (AE)-based anomaly detection and localization system has been successfully developed, tested, and deployed for the CMS electromagnetic calorimeter (ECAL) barrel and endcaps using semi-supervised machine learning.

Occupancy histograms from the ECAL processed as images are used as the input to the network after normalizing the histograms with respect to pileup. Correction strategies are implemented that utilize the variations in the detector response and the time-dependent nature of anomalies. A novel application of the spatial and time corrections yields an order of magnitude improvement in the AE performance. Anomaly tagging thresholds chosen at an estimated anomaly tagging rate of 99% are obtained from validation data sets with fake anomalies introduced. In further validations on real anomalies from 2018 and 2022 CMS runs, the AE-based anomaly detection system is able to spot anomalies of various shapes, sizes, and locations in the detector at a tower-level granularity using a single threshold for each part, e.g., the barrel, of the ECAL. The deployment of the AE-based system in the Data Quality Monitoring (DQM) workflow for LHC Run 3 shows that the system performs well in detecting anomalies, as well as identifying degrading channels that are missed by the existing DQM plots. The AE-based DQM system complements and strengthens the existing DQM by

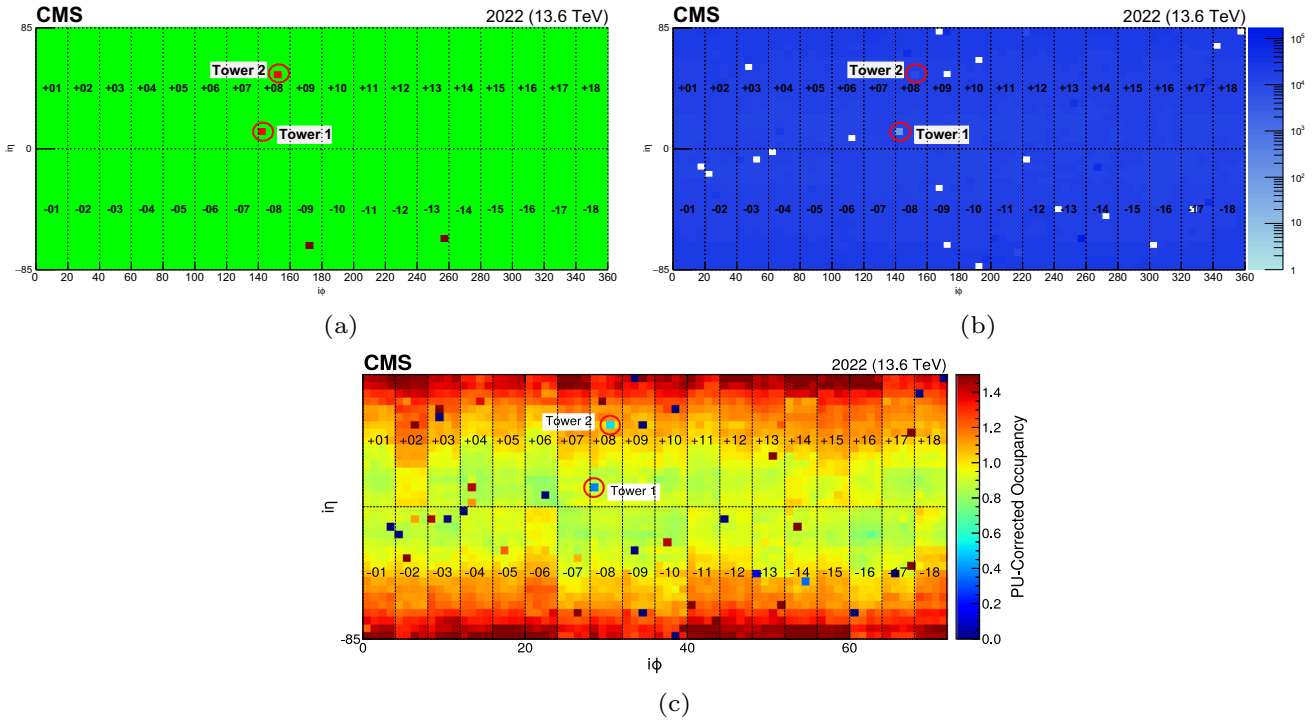


Fig. 19 From a 2022 Run **a** ML quality plot, accumulated over 9 LSs, in the ECAL DQM from the AE model, **b** The digi occupancy plot accumulated over all the LSs in the full run from the online DQM, and **c** PU-corrected average occupancy map for several runs in 2022

helping detector experts make more accurate decisions and reduce false alarms. The anomaly detection system using machine learning described in this paper can be generalized and adapted not only to other subsystems of the CMS detector but also to other particle physics experiments for anomaly detection and data quality monitoring.

Acknowledgements We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid and other centres for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: SC (Armenia), BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENESCYT (Ecuador); MoER, ERC PUT and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP–FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MES and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); MCIN/AEI and PCTI (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); MHESI and NSTDA (Thailand); TUBITAK and TENMAK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA). The principal authors of this study are grateful for the work of Amy Germer on an initial version of the autoencoder-based system for the ECAL endcap sections of the CMS detector. T. Dimova affiliated with an institute or an international laboratory covered by a cooperation agreement with CERN

Author contributions A. Harilal, M. B. Andrews, K. Park and M. Paulini contributed to the design, study and implementation of the research. A. Harilal, K. Park and M. Paulini prepared the manuscript. All members of the ECAL editorial board and in particular the reviewers E. Marco, D. Petyt, N. Marinelli, R. Paramatti, C. Rovelli, P. Gras, F. Cavallari, as well as members of CMS DQM-DC group contributed to the review of the manuscript. All of the authors are responsible for the operation and maintenance of the ECAL detector and the CMS data quality monitoring system.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: Release and preservation of data used by the CMS Collaboration as the basis for publications is guided by the CMS policy as stated in "CMS data preservation, re-use and open access policy".]

Declarations

Conflict of Interest The authors declare no Conflict of interest.

References

1. Chatrchyan S et al (2008) The CMS experiment at the CERN LHC. JINST 3:S08004
2. Sirunyan AM et al (2021) Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. JINST 16(05):P05014
3. Aad G et al (2012) Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Phys Lett B 716:1–29
4. Chatrchyan Serguei et al (2012) Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. Phys Lett B 716:30–61
5. Sirunyan Albert M et al (2020) A measurement of the Higgs boson mass in the diphoton decay channel. Phys Lett B 805:135425
6. Azzolini V, Bugelskis D, Hreus T, Maeshima K, Fernandez MJ, Norkus A, Fraser PJ, Rovere M, Schneider MA (2019) The Data Quality Monitoring software for the CMS experiment at the LHC: past, present and future. EPJ Web Conf. 214:02003
7. Albertsson Kim et al (2018) Machine learning in high energy physics community white paper. J Phys Conf Ser 1085(2):022008
8. Benjamin Nachman (2020) Anomaly detection for physics analysis and less than supervised learning. 10
9. Alan Pol Adrian, Gianluca Cerminara, Cecile Germain, Maurizio Pierini, Agrima Seth (2019) Detector monitoring with artificial neural networks at the CMS experiment at the CERN Large Hadron Collider. Comput Softw Big Sci 3(1):3
10. Azzolini Virginia, Andrews Michael, Cerminara Gianluca, Dev Nabarun, Jessop Colin, Marinelli Nancy, Mudholkar Tanmay, Pierini Maurizio, Pol Adrian, Vlimant Jean-Roch (2019) Improving data quality monitoring via a partnership of technologies and resources between the CMS experiment at CERN and industry. EPJ Web Conf 214:01007
11. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504
12. Sirunyan Albert M et al (2020) Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV. JINST 15:P10017
13. Khachatryan Vardan et al (2017) The CMS trigger system. JINST 12:P01020
14. Sirunyan Albert M et al (2020) Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions. JINST 15(10):P10002
15. Almeida N, Busson P, Faure J-L, Gachelin O, Gras P, Mandjavidze I, Mur M, and Varela J (2004) The selective read-out processor for the CMS electromagnetic calorimeter. In IEEE Symposium Conference Record Nuclear Science, Vol. 3, pp 1721–1725
16. Scott R (2003) Study of the effects of data reduction algorithms on physics reconstruction in the CMS ECAL. Technical report, CERN, Geneva
17. Tuura L, Meyer A, Segoni I, Della Ricca G (2010) CMS data quality monitoring: systems and experiences. J Phys Conf Ser 219:072020
18. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
19. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. 12
20. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A et al (2019) Pytorch: An imperative style, high-performance deep learning library. In: H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, editors, Advances in Neural Information Processing Systems Curran Associates, Inc, 32, 8024–8035.
21. Nair V, Hinton GE (2010) Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10, Omnipress, USA, pp 807–814,
22. Bai J, Lu F, Zhang K, et al (2019) Onnx: Open neural network exchange. https://github.com/onnx/onnx
23. ONNX Runtime developers (2021) Onnx runtime. https://onnxruntime.ai/
24. Siddireddy PK (2018) The CMS ECAL Trigger and DAQ system: electronics auto-recovery and monitoring. 6

The CMS ECAL Collaboration

D. Abadjiev[1], T. Adams[2], P. Adzic[3,a], M. Ahmad[4], C. Amendola[5], M. B. Andrews[6], R. Arcidiacono[7a,7c], S. Argiro[7a,7b], A. Askew[2], E. Auffray[8], V. Azzolini[9], D. Bailleux[10], R. Band[10], D. Barney[8], P. Barria[11a], N. Bartosik[7a], C. Basile[11a,11b], D. Bastos[12], K. W. Bell[13], M. Besancon[5], R. Bianco[11a], C. Biino[7a], V. Blinov[14,b], C. Borca[7a,7b], A. Bornheim[15], R. M. Brown[13], M. Campana[11a,11b], S. Castells[10], F. Cavallari[11a], F. Cetorelli[16a,16b], R. M. Chatterjee[17,d], S. Chatterjee[18], G. Chaudhary[19], J. A. Chen[20], N. Chernyavskaya[8], H. Chung[21], M. Cipriani[8], L. Cokic[3], C. Cooke[13], F. Cossio[7a], F. Couderc[5], D. Cristoforetti[7a,7b], G. Cucciati[10], L. Cunqueiro Mendez[11a,11b], D. Da Silva Di Calafiori[22], I. Dafinei[11a], D. J. A. Cockerill[13], M. Dejardin[5], D. Del Re[11a,11b], G. Della Ricca[23a,23b], P. Depasse[24], J. Dervan[1], E. Di Marco[11a], M. Diemoz[11a], T. Dimova[14,b], G. Dissertori[22], M. Dittmar[22], A. Dolgopolov[10], M. Donegà[22], M. Dordevic[3], H. El Mamouni[24], F. Errico[25a,25b], F. Espinosa[9], J. L. Faure[5], J. Fay[24], J. Fernandez Menendez[26], F. Ferri[5], L. Finco[27], F. Fiori[28a,28b,28c], E. Frahm[17], W. Funk[8], T. Gadek[22], J. Gajownik[13], M. Galli[22], S. Ganjour[5], S. Gascon[24], A. Ghezzi[16a,16b], P. Ghose[2], S. Gninenko[14], S. Goadhouse[21], N. Godinovic[29], N. Golubev[14], P. Govoni[16a,16b], P. Gras[5], J. Hakala[21], G. Hamel de Monchenault[5], A. Harilal[6,*], N. Härringer[22], R. Hashmi[2], H. F. Heath[30], R. Hirosky[21], K. W. Ho[10], X. Hou[31,c], Q. Ingram[32], Sh. Jain[13,d], T. Javaid[33,c], C. Jessop[10], R. Jimènez[22], B. M. Joshi[17], E. Jourd'hui[24], K. Kaadze[34], Y.-W. Kao[20], L. Kardapoltsev[14,b], R. Khurana[20], J. King[35], A. Kirilovas[36], D. Konstantinov[14], M. Kovac[37], A. Krishna[1], C. M. Kuo[38], L. Lambrecht[39], G. Lavizzari[16a,16b], P. Lecoq[8], A. Ledovskoy[21], F. Legger[7a,7b], D. Lelas[29], Y. y. Li[20], Z. Liang[40], W. Lin[38], E. Longo[11a,11b], N. Loukas[10], R. -S. Lu[20], W. Lustermann[22], L. Lutton[10], A. -M. Lyon[22], K. Maeshima[41], J. Malcles[5], P. Mandrik[14], R. A. Manzoni[22], L. Marchese[22], N. Marinelli[10], A. C. Marini[8], L. Martin[1], B. Marzocchi[1], A. Mascellani[22,e], A. Massironi[16a], V. Matveev[14,b], G. Mazza[7a], P. Meridiani[11a], M. Mijic[3], J. Mijuskovic[11a,11b], P. Milenovic[3], J. Milosevic[3], M. Monteno[7a], F. Monti[31], F. Moortgat[8], J. Mousa[43], T. Mudholkar[6], F. Nessi-Tedaldi[22], C. Nicolaou[43], A. Nigamova[44], M. M. Obertino[7a,7b], G. Organtini[11a,11b], T. Orimoto[1], F. Orlandi[7a,7b], I. Ovtin[14,b], E. Paganis[20], D. Papagiannis[8], F. Pandolfi[11a], R. Paramatti[11a,11b], K. Park[6], N. Pastrone[7a], M. Paulini[6], F. Pauss[22], A. Petkovic,[37], E. Petraityte[36], V. Pettinacci[11a], D. Petyt[13], S. Pigazzini[22], B. S. Pinolini[16a], P. R. Prova[2], C. Quaranta[11a,11b], S. Ragazzi[16a,16b], S. Rahatlou[11a,11b], J. C. Rasteiro Da Silva[12], P. A. Razis[43], P. Rebello Teles[45], T. Reis[13], F. Riti[22], C. Rogan[35], T. Romanteau[46], A. Rosowsky[5], C. Rovelli[11a], M. Rovere[8], R. Rusack[17], G. Salvi[13], O. Sancar[17], A. Sanchez[6], C. Sandever[13], F. Santanastasio[11a,11b], R. Saradhy[17], U. Sarkar[46], M. Schneider[8], N. Schroeder[17], A. Sculac[24], T. Sculac[37], M. A. Shahzad[31,c], C. H. Shepherd-Themistocleous[13], P. Simkina[5], A. Singla[19], A. Singovsky[10], Y. Skovpen[14,b], V. J. Smith[30], L. Soffi[11a,11b], K. Stachon[22], A. Steen[8], J. Steggemann[22,e], M. Succar[47], J. Tao[31], A. Tishelman-Charny[1], P. C. Tiwari[18,f], M. Tornago[7a,7b], R. Tramontano[11a,11b], L. -S. Tsai[20], E. Usai[48], D. Valsecchi[22,g], A. Vagnerini[27], J. Varela[12], R. Venditti[21], P. Verma[18], E. Vlassov[7a], V. Wachirapusitanand[49], T. Wamorkar[1], C. Wang[31,c], J. Wang[31], M. A. Wadud[17], S. S. Yu[38], A. Zabi[46], A. Zghiche[46], L. Zhang[15], R. Y. Zhu[15], L. Zygal[10]

[1] Northeastern University, Boston, MA, USA
[2] Florida State University, Tallahassee, FL, USA
[3] VINCA Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia
[4] College Station, Texas A&M University, Texas, USA
[5] IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
[6] Carnegie Mellon University, Pittsburgh, PA, USA
[7] [(a)]INFN Sezione di Torino, Torino, Italy; [(b)]Università di Torino, Torino, Italy; [(c)]Università del Piemonte Orientale, Novara, Italy
[8] CERN, European Organization for Nuclear Research, Geneva, Switzerland
[9] Massachusetts Institute of Technology, Cambridge, MA, USA
[10] University of Notre Dame, Notre Dame, IN, USA
[11] [(a)]INFN Sezione di Roma, Roma, Italy; [(b)]Sapienza Università di Roma, Roma, Italy
[12] Laboratório de Instrumentação e Física Experimental de Partículas, Lisboa, Portugal
[13] Rutherford Appleton Laboratory, Didcot, UK
[14] Authors affiliated with an institute or an international laboratory covered by a cooperation agreement with CERN, Geneva, Switzerland

[15] California Institute of Technology, Pasadena, CA, USA
[16] [(a)]INFN Sezione di Milano-Bicocca, Milano, Italy; [(b)]Università di Milano-Bicocca, Milano, Italy
[17] University of Minnesota, Minneapolis, MN, USA
[18] Indian Institute of Technology Madras, Madras, India
[19] Panjab University, Chandigarh, India
[20] National Taiwan University (NTU), Taipei, Taiwan
[21] University of Virginia, Charlottesville, VA, USA
[22] ETH Zurich-Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland
[23] [(a)]INFN Sezione di Trieste, Trieste, Italy; [(b)]Università di Trieste, Trieste, Italy
[24] Institut de Physique des 1 Infinis de Lyon (IP2I), Villeurbanne, France
[25] [(a)]INFN Sezione di Bari, Bari, Italy; [(b)]Università di Bari, Bari, Italy
[26] Universidad de Oviedo, Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias (ICTEA), Oviedo, Spain
[27] University of Nebraska-Lincoln, Lincoln, NE, USA
[28] [(a)]INFN Sezione di Pisa, Pisa, Italy; [(b)]Università di Pisa, Pisa, Italy; [(c)]Scuola Normale Superiore di Pisa, Pisa, Italy
[29] Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia
[30] University of Bristol, Bristol, UK
[31] Institute of High Energy Physics, Beijing, China
[32] Paul Scherrer Institut, Villigen, Switzerland
[33] Beihang University, Beijing, China
[34] Kansas State University, Manhattan, KS, USA
[35] The University of Kansas, Lawrence, KS, USA
[36] Vilnius University, Vilnius, Lithuania
[37] Faculty of Science, University of Split, Split, Croatia
[38] National Central University, Chung-Li, Taiwan
[39] Ghent University, Ghent, Belgium
[40] Department of Physics, Tsinghua University, Beijing, China
[41] Fermi National Accelerator Laboratory, Batavia, IL, USA
[42] University of Cyprus, Nicosia, Cyprus
[43] University of Hamburg, Hamburg, Germany
[44] Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
[45] Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
[46] Universidad de Los Andes, Bogota, Colombia
[47] The University of Alabama, Tuscaloosa, AL, USA
[48] Chulalongkorn University, Faculty of Science, Department of Physics, Bangkok, Thailand

[a] Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia
[b] Also at another institute or international laboratory covered by a cooperation agreement with CERN, Geneva, Switzerland
[c] Also at University of Chinese Academy of Sciences, Beijing, China
[d] Also at Tata Institute of Fundamental Research-B, Mumbai, India
[e] Also at Ecole Polytechnique Fédérale Lausanne, Lausanne, Switzerland
[f] Also at Indian Institute of Science (IISc), Bangalore, India
[g] Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland