

Calibration and Performance of the CMS Electromagnetic Calorimeter in LHC Run2

Francesca Cavallari^{1,*} and Chiara Rovelli^{1,**}

(on behalf of the CMS Collaboration)

¹Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Italy

Abstract. Many physics analyses using the Compact Muon Solenoid (CMS) detector at the LHC require accurate, high resolution electron and photon energy measurements. Excellent energy resolution is crucial for studies of Higgs boson decays with electromagnetic particles in the final state, as well as searches for very high mass resonances decaying to energetic photons or electrons. The CMS electromagnetic calorimeter (ECAL) is a fundamental instrument for these analyses and its energy resolution is crucial for the Higgs boson mass measurement. Recently the energy response of the calorimeter has been precisely calibrated exploiting the full Run2 data, aiming at a legacy reprocessing of the data. A dedicated calibration of each detector channel has been performed with physics events exploiting electrons from W and Z boson decays, photons from π^0 and η decays, and from the azimuthally symmetric energy distribution of minimum bias events. This talk presents the calibration strategies that have been implemented and the excellent performance achieved by the CMS ECAL with the ultimate calibration of Run2 data, in terms of energy scale stability and energy resolution.

1 Introduction

The Compact Muon Solenoid (CMS) experiment [1] at the CERN Large Hadron Collider is equipped with a scintillating crystal electromagnetic calorimeter (ECAL) [2], made of 75848 lead tungstate scintillating crystals. The calorimeter is divided into a barrel part and two endcaps. A lead and silicon strip preshower is located in front of the endcaps. The ECAL is the largest crystal calorimeter ever built for a high energy physics experiment. Its energy resolution is fundamental for many CMS physics analyses, in particular for the mass resolution of the Higgs boson decay into two photons [3]. This channel requires excellent photon energy resolution to observe a tiny mass peak over a large background.

Crystal calorimeters provide excellent energy resolution, but require constant monitoring to correct for environmental effects and radiation induced light output change. Periodic channel-to-channel energy calibrations are also required. In LHC Run2 (2015-2018) high instantaneous (up to $2 \cdot 10^{34} \text{cm}^{-2} \text{s}^{-1}$) and large integrated luminosities (almost 145fb^{-1}) have provided a challenging environment for the detector. The ECAL team has exploited many

*e-mail: Francesca.Cavallari@roma1.infn.it

**e-mail: Chiara.Rovelli@roma1.infn.it

different data sources to cope with these effects, both during data-taking and in offline analysis.

The ECAL is equipped with a laser light injection system, that monitors crystal transparency every 40 minutes. Channel-to-channel calibrations are computed exploiting physics events, but require more time to accumulate the necessary amount of data. CMS collision data are promptly processed after 48 hours from data-taking using the laser monitoring corrections and the existing channel-to-channel calibrations. A refined set of calibrations is computed at the end of each year, using the full statistics of the data, and is made available for the end of year data reprocessing, which typically takes place soon after the end of data-taking. The reprocessing of the CMS proton-proton collision data for one year takes a few months, using the available computing resources. Therefore the experiment can afford only one or two such reprocessings per year.

An improved ECAL calibration for the full Run 2 dataset (2016-2018) was performed during 2019 to achieve optimal performance. The CMS experiment has used these calibration sets for the full Run2 data reprocessing, which has been carried out during Long shutdown 2. This reprocessing constitutes the “ultra-legacy” data set, which will be used for analyses requiring optimal energy resolution and will be preserved for future analyses on Run2 data.

This paper presents the ECAL calibration methods used for the Run2 “ultra-legacy” reprocessing and results, with a particular emphasis on the computing aspects. Many low-level parameters are used to compute the crystal energy deposits and the time of arrival (pedestals, pulse shapes, alignment and synchronization). Although these parameters are not described in this paper, all of them have been recomputed or checked for the purpose of the “ultra-legacy” reprocessing. In addition Monte Carlo simulation conditions have been derived from data, for each year of Run2, in order to achieve an improved description of the detector response and particularly noise.

2 Laser Monitoring Data

The ECAL crystal transparency is monitored with laser light [4] during data taking. Laser light is injected in each crystal in monitoring regions of about 800 crystals. A set of PN diodes are mounted on the detector and are used to normalize the light seen by each channel. Two wavelengths are used: blue, close to the lead tungstate scintillation emission peak, and green. Blue and orange LED light is also injected in the endcap crystals. Laser calibration triggers are issued at a rate of 100 Hz in the empty part of the LHC orbit (LHC abort gap), and 600 events are recorded for each crystal, cycling through the different regions. A full sequence through all the ECAL regions with all wavelengths takes about 40 minutes.

The laser raw data occupies about 40 kB/event for each monitoring region, and amounts to about 40 TB for one year. A dedicated computer farm in the CMS online network processes the laser data and computes transparency corrections for each crystal. A reduced data format is produced with all necessary information for each crystal for each laser monitoring sequence (typically signal mean amplitude divided by the PN normalization, and laser peak analysis information). The reduced data format occupies about 200 GB per year. The processing software is written in c++ and root libraries [5]. In 2018 the laser farm was upgraded and optimized: the laser farm PCs were replaced by one node with 40 CPUs, model Intel Xeon CPU E5-2640 v4 @ 2.40GHz, mounting a set of raid disks of 7 TB in order to keep all active data locally during processing. A second identical spare node is ready in case of faults. A synchronization script mirrors the disks of the production node to the spare one, in order to support quick switch to the spare. This configuration is able to cope with the data processing without creating backlog: the processing takes approximately three-four hours for a typical twelve hours fill.

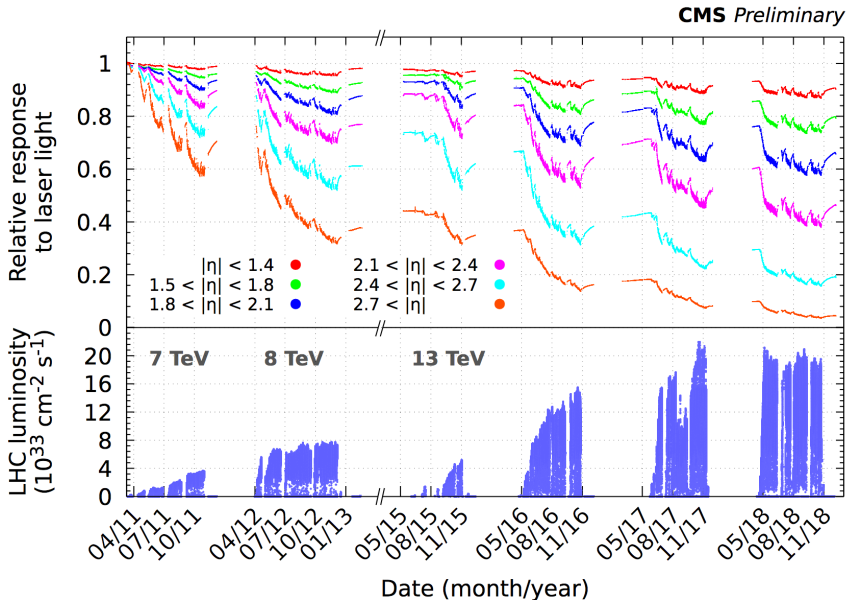


Figure 1. Evolution of the ECAL channel response to laser light versus time since 2011. The channels are subdivided in ranges of η . The bottom panel shows the instantaneous luminosity delivered by the LHC in the same time period.

The laser transparency corrections are inserted in the CMS conditions database [6], and then used for the prompt processing of CMS data, which starts 48 hours after data taking. Several key monitoring plots are produced and checked periodically by the ECAL operations team. In case of issues in the laser hardware or in the laser correction computation, the CMS prompt data processing may be halted, while the problem is corrected in the database. During 2019 the laser data have been reanalysed to correct for anomalous readings in a few channels and achieve the best possible monitoring corrections.

Figure 1 shows the evolution of the ECAL response to laser light since 2011. The crystals in the barrel and endcap regions until $\eta = 2.4$ have suffered a moderate transparency loss (e.g. up to 60%), which can be precisely corrected using the laser data to achieve an excellent energy resolution. The crystals in the forward region, closer to the beam pipe, have suffered larger transparency loss. These are outside the CMS tracker coverage and the requirement on energy resolution is less demanding. Nevertheless, it is important to ensure good energy scale stability in this region, for the triggering and energy reconstruction of forward photons and jets.

3 Channel to channel inter-calibration

Channel-to-channel inter-calibration exploits several physics signals:

- photons from π^0 and η decays,
- electrons from W boson decay,
- electron pairs from Z boson decays,
- azimuthal symmetry of minimum bias events.

Dedicated reduced data streams are saved for each type of events. The CMS trigger uses a two-tiered system [7]. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a simplified version of the full event reconstruction software, and reduces the event rate to around 1 kHz before data storage. Depending on the HLT result, events may be saved in different streams with different event content, to optimize the bandwidth.

The following sections describe in detail each calibration method. Section 3.4 describes the combination of the inter-calibrations from the various methods.

3.1 Azimuthal symmetry of minimum bias events

This method exploits the intrinsic symmetry of the energy deposits in rings of fixed azimuthal angle in minimum bias collision events. In the HLT farm, the energy deposits in each crystal are saved in a dedicated stream with minimum event content, at a fixed rate. ECAL hits between a minimum and a maximum threshold are summed for each crystal for a given period. The energy sums are equalized between crystals in the same ring. This method can provide a measurement point per crystal every few tens of minutes. However it is very sensitive to the energy thresholds. Since the noise evolves in the ECAL, due to transparency changes in the crystals and photodetector radiation damage, the accuracy of this method was worse than the other methods in Run2. Therefore it was used mostly to monitor the single crystal response and to spot anomalous channels, rather than to derive crystal inter-calibrations.

3.2 π^0 and η calibrations

π^0 and η mesons are produced in abundance in LHC proton-proton collisions and their decay to two photons can be used for calibration purposes. However the photon energies in these decays are quite low, typically a few GeV, while the ECAL readout stores scattered low energy hits with a strong zero suppression, and only higher energy electromagnetic showers are acquired in “full readout” mode. A special readout mode has been used for the π^0 s, to record the ECAL hits in a 5x5 crystal readout matrix, whose energy sum is above a threshold of 2-3 GeV.

The π^0 data are recorded in parasitic mode with respect to the normal CMS data. In the HLT farm, a dedicated algorithm processes hadronic events searching for photon pairs whose mass is in the π^0 and η meson range. A reduced data stream is filled with minimal event information: only the ECAL hits around the two candidate photons are saved. One event occupies 2 kB (compared to the average CMS event size of about 1 MB). As a consequence of the small data size, π^0 data can be recorded at ~ 7 kHz, and the data for one year occupies about 20-30 TB.

These data are used to monitor the ECAL response stability and to verify the quality of the laser monitoring corrections before prompt processing begins. For a region as large as half barrel, one measurement point every 5 minutes is available with a precision of few permill, allowing detailed investigations of ECAL response stability to be carried out.

At the end of each year the full statistics of the π^0 data is used to compute channel-to-channel intercalibrations. Figure 2 shows the mass distribution for one crystal in the barrel for one year of Run2. The calibration algorithm proceeds iteratively correcting each crystal calibration by the observed mass peak shift, equalizing the response in rings of fixed azimuth. At each iteration the histograms of the di-photon mass of all the events that hit each crystal are made, using the calibrations of the previous step. The calibration requires about 10 iterations

to converge, and these take about 2 weeks for one year data on the CAF [13]. The processing jobs are subdivided by about ten-twenty data files, there are approximately 20000 files per year, and approximately 1000 jobs can run in parallel at the CAF on the calibrations and alignment queues. It takes about 10 hours for all the jobs to complete. The software is based on the CMSSW framework [8, 9] for raw data reconstruction and on the root software [5] for data merging and fitting. Approximately 50% of the CPU time is spent in the low level ECAL reconstruction, described in [10], which computes the pulse amplitudes in the crystals. This could eventually be optimized by saving the uncalibrated pulse amplitudes of each crystal at the first iteration, but it would require a large disk buffer. Once the jobs are completed, the histograms are merged and the fits to the distributions are done, as shown in Figure 2.

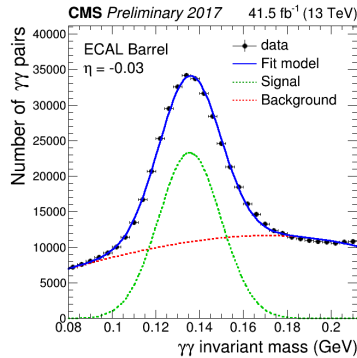


Figure 2. Reconstructed invariant mass distribution of photon pairs around the π^0 mass peak for one barrel crystal in 2017.

3.3 W/Z boson electron calibrations

W decays to an electron and neutrino are exploited to select electrons used for the ECAL calibration. Candidate electrons are formed from tracks reconstructed with a GSF algorithm, matched to calorimeter clusters with recovery of radiated photons [11]. Events with at least one candidate electron are saved in a dedicated stream. W/Z boson decays to electrons are saved in a reduced format, called ALCARAW or ALCARECO, which occupies about 50 TB per year. A software framework called ECALELF [12], based on the CMSSW framework [8, 9], is used to reconstruct the events, with the flexibility to reprocess only the electron ECAL hits from the raw samples applying different conditions (low level parameters, monitoring corrections and calibration constants). All other steps in the event processing are left untouched, including the clustering. The ECALELF framework provides the ability to reprocess the W and Z electrons recorded in one year in two hours on the CERN batch system, thus it is a powerful tool to check the effects of new laser corrections or calibration constants. It takes 0.18 s to process one event on a standard CPU of the CMS CAF [13]. The number of jobs has been optimized taking into account that approximately 1000 jobs can run in parallel at the CAF on the calibration and alignment queues, and every job takes about 20-30 minutes. Common root files [5] with summary information are used by both W and Z analyses, which occupy about 800 GB per year. It takes about 0.015 s/event to process these files to produce the energy resolution and stability plots.

The ratio of ECAL energy, E, to tracker momentum, p, for electrons is used to calibrate the crystals in rings of fixed azimuthal angle. A fit to the E/p distribution for the different

regions is performed using a template E/p distribution which depends on the detector azimuthal angle, to take into account the different energy loss in the upstream material. This method is sufficiently accurate to correct for residual drifts of the calorimeter response in different laser monitoring regions (groups of approximately 200 crystals). It provides one measurement point about every 2 fb^{-1} of integrated luminosity, corresponding to roughly one week of data-taking. With the full statistics of one year it is also possible to compute the inter-calibrations for each crystal using this method.

Since 2016, because of the larger integrated luminosities, the mass of electron-positron pairs from Z boson decays is used to inter-calibrate the ECAL channels at the single crystal level. This requires the full statistics of one year. This method maximizes the likelihood comparing the reconstructed mass distribution with that predicted by Monte Carlo simulation. The method can also be used to calculate the inter-calibrations in crystal rings, to estimate the energy resolution of electrons in different pseudorapidity regions and to equalize the response of the different rings (absolute energy calibration). The software has been developed for this application and is called iJAZZ [14].

The acceptance for electrons in CMS extends up to $|\eta| < 2.5$, the region covered by the Tracker. For the inner endcap region, $2.5 < |\eta| < 3$, beyond the tracker coverage, a special algorithm is used. It selects one ECAL cluster in this region and one electron elsewhere in the ECAL. This algorithm is able to select enough events to calibrate individual crystals. This is the only method currently able to calibrate the forward crystals.

3.4 Combination of the inter-calibrations and absolute energy scale

To estimate the precision of the intercalibration constants of each method, the effect of each calibration set on the energy resolution is tested using the $Z \rightarrow ee$ mass distribution. Then all methods are combined with a weighted mean to account for each method precision as a function of the azimuthal angle. Figure 3 shows the precision of the various methods and of the combination for the year 2017. In the barrel region, the E/p method has the best precision (0.4-0.6%). The π^0 and $Z \rightarrow ee$ methods have comparable performance (0.5-1%). In the endcaps, the π^0 method works well for $|\eta| < 2$ with an accuracy of $\sim 3\%$, the E/p method has an accuracy of 1.5% and the $Z \rightarrow ee$ method has the best accuracy. The $Z \rightarrow ee$ method with one electron and one ECAL cluster is used to calibrate the forward endcap crystals. Finally $Z \rightarrow ee$ events are used to set the absolute energy scale equalizing the response of the rings. Overall, it takes about 2 months per year to calculate the calibrations and combine them.

Figure 4 shows the energy resolution for electrons for the prompt and “ultra-legacy” calibrations for the year 2017. A large improvement is observed between prompt processing performance and “ultra-legacy” refined calibrations. The effect is particularly evident for the forward region, where significantly improved ECAL energy scale stability and resolution have been achieved.

4 Conclusions

Crystal calorimeters provide excellent energy resolution, but require constant monitoring to correct for environmental effects and radiation induced light output changes, and periodic channel-to-channel energy calibrations. In 2016-2018 the CMS experiment has recorded approximately 145 fb^{-1} with a maximum instantaneous luminosity of $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, a challenging environment for the detector. During data-taking the CMS ECAL has exploited laser light to monitor crystal transparency and correct for radiation induced response changes, and has utilised several physics signals to monitor the response stability, in order to achieve a

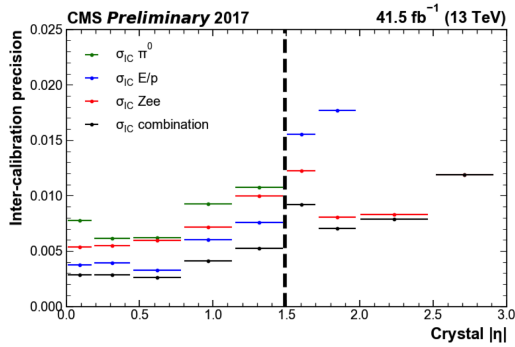


Figure 3. Residual mis-calibration of the ECAL channel inter-calibration, as a function of pseudo-rapidity using the full 2017 dataset, obtained with three different methods (green, blue, red dots), and combining the three methods (black).

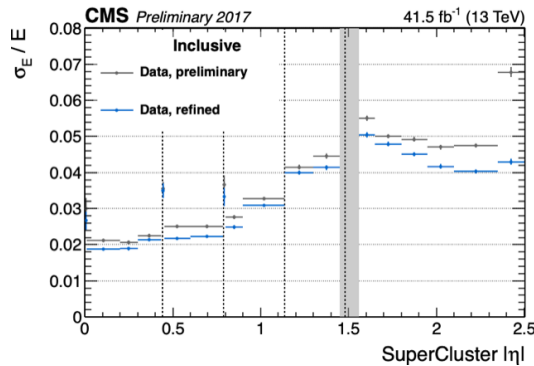


Figure 4. Relative electron energy resolution as a function of the pseudo-rapidity η for the ECAL Barrel and Endcaps, for the preliminary processing (grey) and for the “ultra-legacy” reprocessing (blue) of 2017 data.

good performance for the prompt CMS data processing. During Long Shutdown 2 (2019-2020) the ECAL crystals were recalibrated in order to achieve optimal energy resolution for the “ultra-legacy” CMS data reprocessing. A large improvement in response stability and energy resolution is observed between promptly derived derived calibrations and “ultra-legacy” refined calibrations, particularly in the forward regions of the endcaps.

References

- [1] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.
- [2] CMS Collaboration, *The electromagnetic calorimeter project*, Technical Design Report, CERN/LHCC 97-33, CMS TDR 4 (1997).
- [3] CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, CMS-PAG-HIG-2019-004.

- [4] M. Anfreville et al., *Laser monitoring system for the CMS PbWO4 e.m. calorimeter*, NIM **A594** (2008) 292-320, doi: 10.1016/j.nima.2008.01.104
- [5] I. Antcheva et al., *ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization*, Comp. Phys. Comm., **180-12** (2009), 2499-2512. <https://doi.org/10.1016/j.cpc.2009.08.005>
- [6] Cavallari F et al. , *Time-critical Database Condition Data Handling in the CMS Experiment During the First Data Taking Period*, J. Phys.: Conf. Ser. 331 (2011) 042007.
- [7] V. Khachatryan et al. [CMS Collaboration], *The CMS trigger system*, JINST **12**, no. 01, P01020 (2017) doi:10.1088/1748-0221/12/01/P01020
- [8] Lange D, *The CMS Reconstruction Software*, J. Phys.: Conf. Ser. 331 (2011) 032020.
- [9] *CMSSW software*, <https://github.com/cms-sw/cmssw>
- [10] CMS Collaboration, *Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions*, EGM-18-001, submitted to J. Inst.
- [11] CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JINST **10** (2015) P06005, doi: 10.1088/1748-0221/10/06/P06005
- [12] S. Nourbakhsh, *ECALELF*, <https://gitlab.cern.ch/shervin/ECALELF>, and the current release is <https://github.com/ECALELFS/ECALELF>
- [13] O. Buchmuller, *The CMS CERN Analysis Facility (CAF)*, Journ. Phys. Conference Series, Vol. 219, Number 5.
- [14] F. Couderc, *Quest for the Higgs boson(s) from D0 to CMS experiments*, Physics, Sorbonne Université, 2018, HAL Id: <https://tel.archives-ouvertes.fr/tel-02106984>