

# Validation of Physics Models of Geant4 Versions 10.4.p03, 10.6.p02 and 10.7.p01 using Data from CMS Experiment

*Sunanda Banerjee*<sup>1</sup> and *Vladimir Ivanchenko*<sup>2,3</sup> on behalf of the CMS Collaboration

<sup>1</sup>Fermi National Accelerator Laboratory, Batavia, IL, USA

<sup>2</sup>Tomsk State University, 634050 Tomsk, Russia

<sup>3</sup>CERN, CH1211, Geneva 23, Switzerland

**Abstract.** CMS tuned its simulation program and chose a specific physics model of Geant4 by comparing the simulation results with dedicated test beam experiments. Test beam data provide measurements of energy response of the calorimeter as well as resolution for well identified charged hadrons over a large energy region. CMS continues to validate the physics models using the test beam data as well as collision data from the Large Hadron Collider. Isolated charged particles are measured simultaneously in the tracker as well as in the calorimeters. These events are selected using dedicated triggers and are used to measure the response in the calorimeter. Different versions of Geant4 (10.2.p02, 10.4.p03, 10.6.p02) have been used by CMS for its Monte Carlo production and a new version (10.7.p01) is now chosen for future productions. A suitable physics list (collection of physics models) is chosen by optimizing performance against accuracy. A detailed comparison between data and Geant4 predictions is presented in this paper.

## 1 Introduction

The CMS experiment [1] at the Large Hadron Collider uses the Geant4 toolkit [2-4]. The simulation software as well as the Geant4 code is evolving since the CMS simulation code has been put into operation. In our previous reports [5-7], the full simulation scheme for Run 2 was described. Several variants of Geant4 version 10 have been used for Run 2. Initially, Geant4 10.0.p02 was adopted and a new multi-threaded mode was made operational [7]. CPU performance of the CMS simulation was substantially improved compared to the Run 1 simulation. CMS switched to the Geant4 version 10.2.p02 for 2017 Monte Carlo (MC) production. CMS used Geant4 version 10.4.p03 for its 2018 MC production and 10.6.p02 for the 2020 production. CMS is getting ready for the MC production for the Run 3 which is scheduled to start in 2022 and chose to use the Geant4 version 10.7.p01. There has been an active effort within CMS to evaluate the quality of simulation as one changes either the application software or the version of Geant4.

Two sources of data are used for validating the software. Data exist from a test beam campaign during 2006 where a prototype of the CMS hadron calorimeter (HCAL) and one super-module of the electromagnetic calorimeter (ECAL) was exposed to hadron beams of

different types and of different energies. Results from this test beam are published [8] and are used in earlier tuning of the CMS simulation program [9-10]. As a second source, collision data from the CMS experiment are used. Here low luminosity runs with zero bias or minimum bias triggers are analysed from the 2016B run period. Similar analysis was done earlier in CMS [11].

These validation results consolidate the recent adaptation of the new Geant4 versions and the physics list used for Monte Carlo production.

## 2 CMS Simulation Software

CMS simulation software (a component of the overall CMS software CMSSW) started with the physics list [12] QGSP\_FTFP\_BERT\_EML during the early Run 2 operation of the LHC. Since its 2017 MC production with Geant4 version 10.2.p02, CMS moved to the physics list FTFP\_BERT\_EMM.

The hadronic physics model is specified by the component FTFP\_BERT. This is the recommended physics list from the Geant4 collaboration. In this physics list, the transition between FTFP and Bertini cascade model is modified in different Geant4 versions. Till the version 10.4.p03, Bertini cascade model is valid for all particles for energies below 12 GeV and the FTFP model is valid above 3 GeV. In the versions 10.6.p02 and 10.7.p01, the same transition energies are used for pions but for other particles, Bertini Cascade model is assumed to be valid only for energies below 6 GeV.

The electromagnetic physics model is specified by the term EML or EMM. EML uses a simplified multiple scattering model in all parts of the detector. The EMM physics list makes use of the default Geant4 multiple scattering model in regions for the sampling calorimeter. This gives rise to an optimization between performance and accuracy of the simulation.

The effect of saturation on scintillation light emission is handled by incorporating Birk's law [13] to lead tungstate crystals as well as to plastic scintillators. After the Geant4 release 10.6, Geant4 experts have shown, that the traditional value of the Birks saturation coefficient was inconsistent with Geant4 simulation [14] and requires retuning for concrete experiment. In CMS different parameters as used for the ECAL crystals and the HCAL scintillators. Analysis of the test beam data and isolated hadrons from Run 2 allowed to conclude, that the reduction coefficient of plastic scintillator is increased from 0.0052 to 0.006. The saturation coefficients for the ECAL crystals are kept unchanged.

## 3 Validation versus test beam data

Dedicated measurements were carried out with prototypes of the CMS calorimeter in the test beam facility of CERN [8]. Two production wedges of the barrel hadron calorimeter (HB), one prototype module of the endcap (HE), and eighteen trays of the outer hadron calorimeter (HO) were exposed to hadron beams in the H2 beam line of the Super Proton Synchrotron (SPS). The HCAL was preceded by a super-module of the barrel electromagnetic calorimeter (EB). The platform holding the modules could be moved along the phi and eta directions allowing the beam to be directed onto any tower of the calorimeter. Monochromatic secondary and tertiary beams were used having a momentum between 2 and 350 GeV. Auxiliary beam counters were used to select pure beam interactions.

Both electromagnetic and hadron calorimeters were calibrated using 50 GeV electron beams. The energy response (ratio of measured calorimetric energy to the beam momentum), resolution, and shower profiles were measured for different beam momenta. This test beam data analysis allows selection of high purity samples for pions, kaons, protons and anti-protons. The same method of hit generation is used in test beam simulation as in the case of

full CMS. The digitisation chain of CMSSW is not used, instead hit energy was smeared using the Gaussian distribution with fixed widths for the ECAL (0.362 GeV) and the HCAL (0.64 GeV) obtained from the test beam analysis. Hits are collected in the  $7 \times 7$ -crystal matrix of the ECAL and in the  $3 \times 3$  HCAL tower matrix. Hits in the outer hadron calorimeter layer are excluded in both simulation and data analysis. In addition, a time cut corresponding to test beam data acquisition is applied. Visible energy response is computed using the following simple relation:

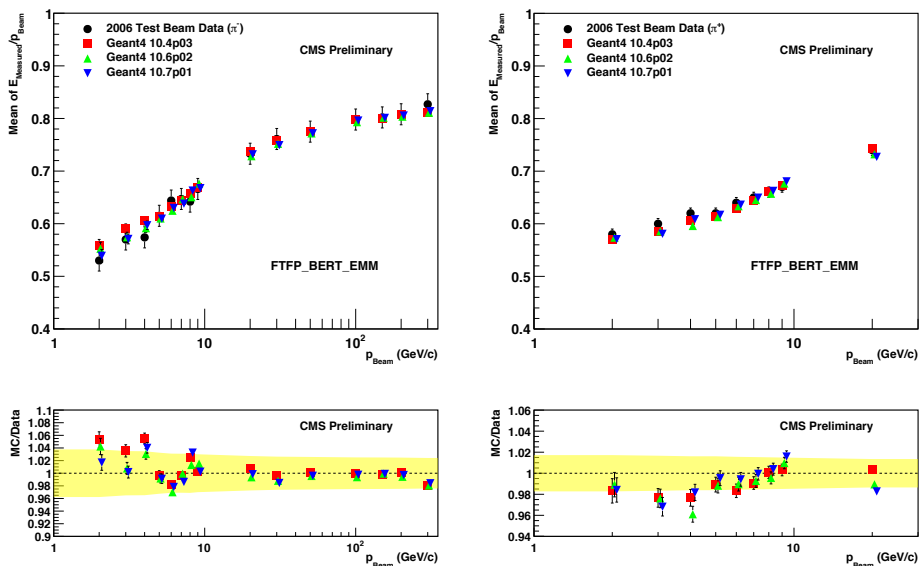
$$E_{\text{vis}} = E_{\text{ECAL}} \cdot f_{\text{ECAL}} + E_{\text{HCAL}} \cdot f_{\text{HCAL}},$$

where  $f_{\text{ECAL}} = 1.01$  and  $f_{\text{HCAL}} = 105.0$ . These factors are determined using the same procedure as has been done in the analysis of the test beam data.

The results of simulation with Geant4 10.4.p03, 10.6.p02 and 10.7.p01 are shown in the form of mean energy response as a function of beam momentum in Fig. 1 (for pions), Fig. 2 (for protons and anti-protons), and Fig. 3 (for kaons). The level of agreement between data and MC is indicated as chi-square per data point in Table 1. All three Geant4 versions provide compatible agreement for mean energy deposition for pions and protons within statistical accuracy of few %. The level of agreement is not good for kaons. Energy response for pions and kaons is similar in the data, but not in Monte Carlo. The predictions for Geant4 versions 10.6.p02 and 10.7.p01 show some improvements for kaons, some deterioration for positive pions, and acceptable agreement for negative pions, protons and anti-protons.

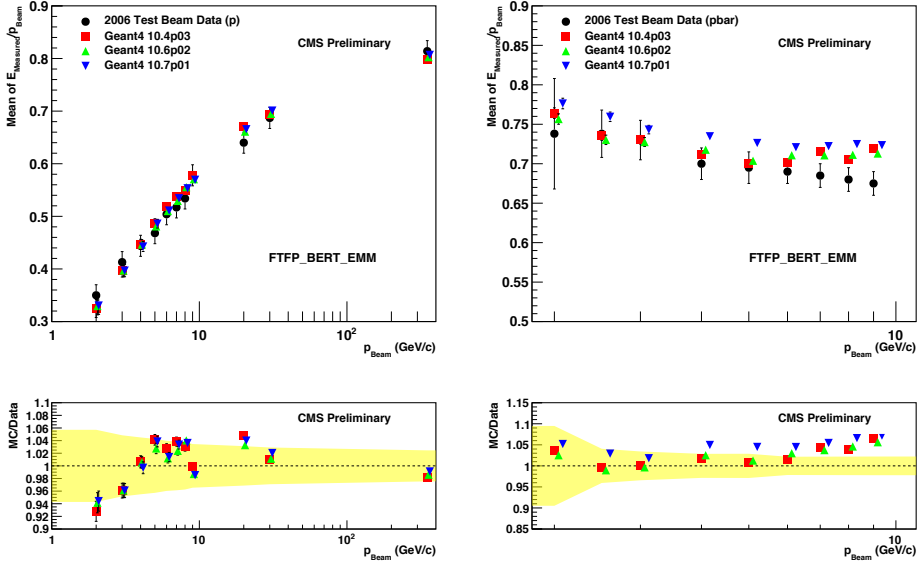
**Table 1.** Chisquare per data point between data and simulation in the measurement of energy response for pions, kaons, protons and anti-protons.

	Negative pions	Positive pions	Negative kaons	Positive kaons	Protons	Anti-protons
10.4.p03	0.45	0.73	26.2	26.8	0.80	1.78
10.6.p02	0.31	1.14	19.4	14.4	0.53	1.81
10.7.p01	0.28	0.93	20.8	14.9	0.63	3.84

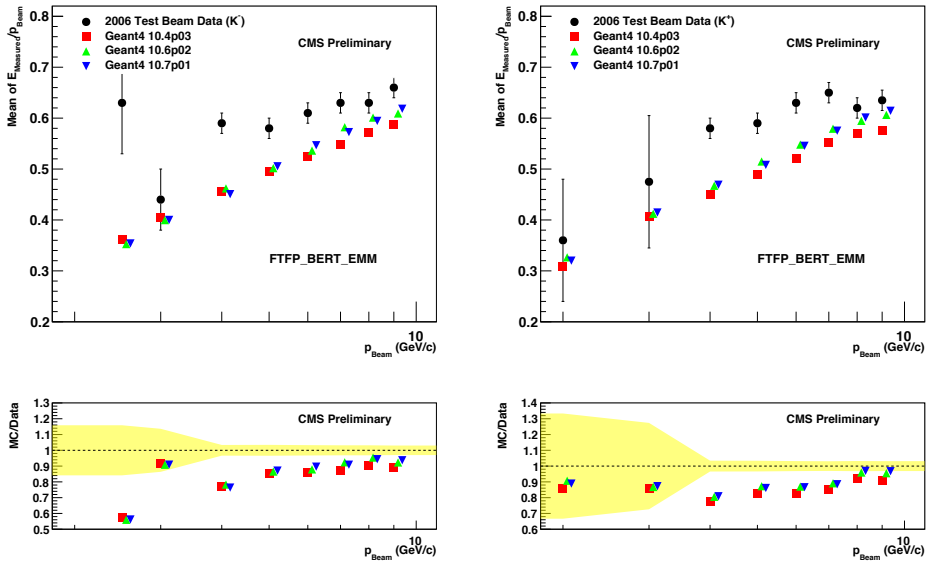


**Fig. 1.** Mean energy response of the combined calorimeter for negative pion (left) and positive pion (right) beams as a function of beam momenta.

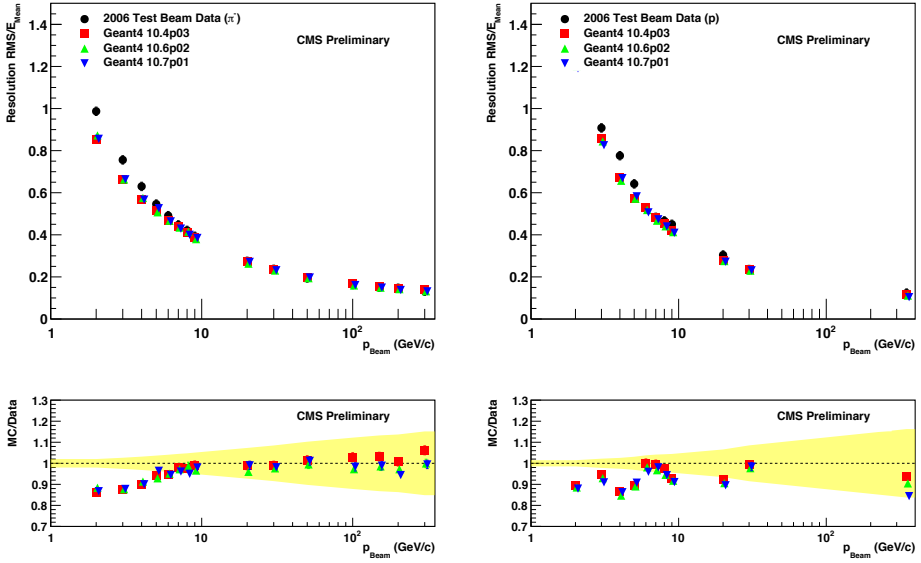
The energy resolution for pion and protons is also compared for negative pions and protons (Fig. 4). There is the statistical agreement between data and simulation at momentum above 8 GeV/c. At low momentum, simulation underestimates energy resolution, which could be due to simplified simulation of the detector response.



**Fig. 2.** Mean energy response of the combined calorimeter for proton (left) and anti-proton (right) beams as a function of beam momenta.



**Fig 3.** Mean energy response of the combined calorimeter for negative kaon (left) and positive kaon (right) beams as a function of beam momenta.



**Fig. 4.** Energy resolution of combined calorimeter for  $\pi^-$  (left) and proton (right) beams as a function of beam momenta.

## 4 Validation versus Run 2 data

The analysis method [15] for comparing data and Monte Carlo simulation of calorimeters is applied to the three set of Geant4 simulation using Geant4 versions 10.4.p03, 10.6.p02 and 10.7.p01. The ratio of calorimeter energy measurement to track momentum for isolated charged hadrons is studied. For that, well measured charged tracks reaching the ECAL surface are selected by imposing isolation criteria based on the NxN matrix analysis on the calorimeter surface. Final cuts are following:

- No extra tracks in the isolation region.
- Energy cut of 2 GeV for neutral isolation.
- No additional good primary vertex in the event in order to avoid PU effect.

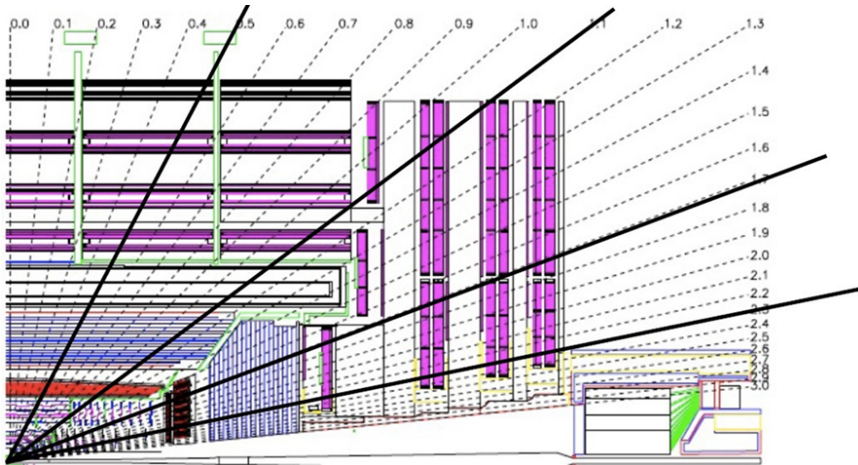
Calorimeter energy is measured in a NxN matrix of calorimeter cells around the impact point of the track. Two versions of NxN matrix are defined:

- 7x7 or 11x11 matrix for the ECAL.
- 3x3 or 5x5 matrix for the HCAL.

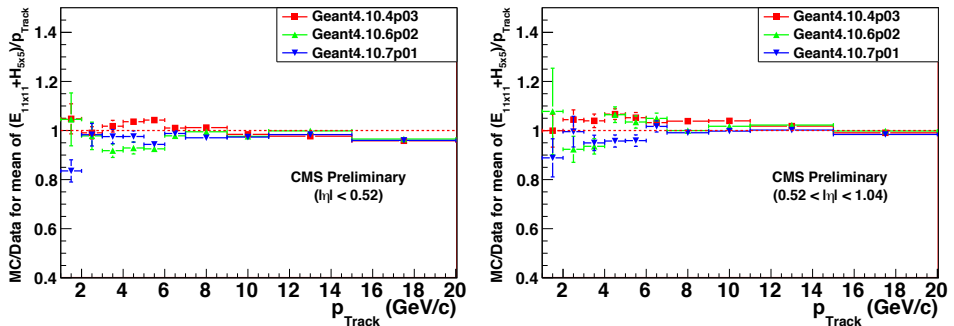
Two low luminosity data sets from the 2016B runs are used for this analysis. Results of comparisons are shown as a function of particle momentum (Figs. 6-7) for three different eta regions (Fig. 5) and three versions of Geant4 10.4.p03 (used during 2018), 10.6.p02 (used during 2020) and 10.7.p01 (to be used in the future).

The level of disagreement between data and MC is estimated from the deviation of the ratio (Data/MC) from 1. The mean level of disagreement is 2.1% in the barrel, 5.0% in the endcap and 3.6% in the transition region for the version 10.4.p03. The corresponding numbers for 10.6.p02 are 2.1%, 2.3%, 2.1% for Geant4 10.6.p02, and 2.6%, 1.9%, 1.0% for Geant4 version 10.7.p01. This level is computed as an average absolute deviation at all momentums excluding the 1 GeV/c point, which has a large uncertainty. Geant4 version

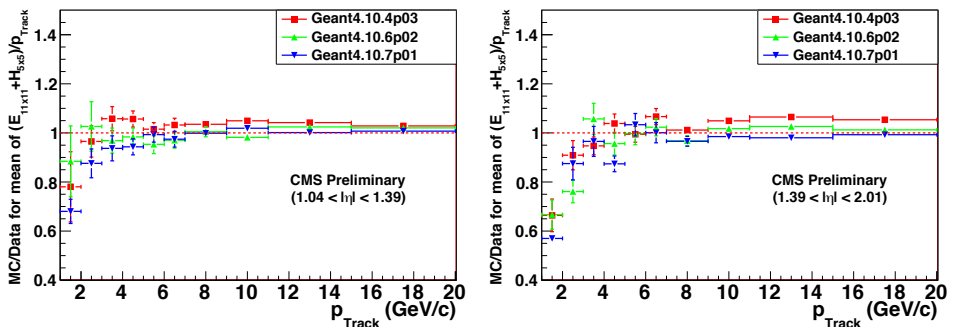
10.7.p01 with this level of agreement between data and MC is a suitable version for the Run 3 MC production.



**Fig. 5.** CMS detector layout with four calorimeter regions for isolated hadron analysis.



**Fig. 6.** MC/data ratio for calorimetric response of isolated charged hadrons in the central part (left) and side part of the barrel Calorimeter.



**Fig. 7.** MC/data ratio for calorimetric response of isolated charged hadrons in the transition region between barrel and endcap (left) and in the endcap (calorimeter) (right).

## 5 Summary

Predictions of the chosen physics list, FTFP\_BERT\_EMM from three different Geant4 versions (10.4.p03, 10.6.p02, 10.7.p01) are compared with the data. Test beam data with identified particle type are used as one source of validation, while isolated charged particles from collision data are used as the second source. The new version shows improved agreement with the data for kaons, some deterioration for positive pions, and acceptable agreement for negative pions, protons and anti-protons. The good overall agreement between data and MC leads CMS experiment to change the Geant4 version to 10.7.p01 for the future Monte Carlo production campaign.

## References

1. S. Chatrchyan et al., JINST **3**, S08004 (2008)
2. The Geant4 Collaboration (S. Agostinelli et al.), Nucl. Instr. Meth. A **506**, 250-303 (2003)
3. J. Allison et al., IEEE Trans. Nucl. Sci. **53**, 270-278 (2006)
4. J. Allison et al., Nucl. Instr. Meth. A **835**, 186-225 (2016)
5. D.J. Lange et al., J. Phys.: Conf. Ser. **608**, 012056 (2015)
6. M. Hildreth et al., J. Phys.: Conf. Ser. **664**, 072022 (2015)
7. M. Hildreth et al., J. Phys.: Conf. Series **898**, 042040 (2017)
8. S. Abdulline et al., European Journal of Physics **C60**, 359 (2009)
9. S. Abdouline et al., “Calorimetry Task Force Report”, CMS-NOTE-2008/25
10. S. Abdouline et al., “Calorimetry Task Force Report”, CMS-NOTE-2010/007
11. CMS Experiment, CMS-PAS-JME-10-008 (2010) (<http://cds.cern.ch/record/1279141>)
12. A. Ribon et al., IEEE NSS Conference, Dresden (2008) NSS **N02-89**
13. J. B. Birks, Proc. Phys. Soc. **A64**, 874 (1951)
14. A. Howard et al., JINST **15**, C05073 (2020)
15. S. Banerjee, J. Phys.: Conf. Series **898**, 042005 (2017)