
CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

March 27, 2007

Jet Performance in CMSSW_1_2_0

R. Harris

Fermilab, Batavia, IL, USA

S. Petrushanko, O. Kodolova

Moscow State University, Moscow, Russia

J. Widawsky, M. Zielinski

U. Rochester, Rochester, NY, USA

A. Bhatti

U. Rockefeller, New York, NY, USA

L. Apanasevich, N. Varelas

U. of Illinois at Chicago, Chicago, IL, USA

R. Cavanaugh, B. Scurlock

U. of Florida, Gainesville, FL, USA

Abstract

Jets in CMSSW_1_2_0 are blah blah blah.

1 Introduction

1.1 Jet and MET Defaults

1.2 CMSSW_1_2_0 Samples

DRAFT

2 MC Jet Corrections

Here we describe the MC Jet corrections which correct the calorimeter level jet (CaloJet) to have the same E_T as the particle level jet (GenJet). These corrections are based on QCD dijet Monte Carlo using Pythia and the CMS detector simulation. The methodology was developed previously [1] for OSCAR/ORCA and ported to CMSSW. We adopt this methodology in order to provide corrections as rapidly as possible to CMSSW data samples.

2.1 Jet Response for Jet Corrections

Measurements of jet response were made with the module SimJetResponseAnalysis in the MCJet package of the JetMETCorrections subsystem of CMSSW. Jets were reconstructed using the iterative cone and midpoint cone algorithms in CMSSW with the E scheme. We consider all GenJets in the event and match each GenJet with the closest CaloJet which minimizes

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}. \quad (1)$$

If the closest CaloJet is not within $\Delta R = 0.25$, the GenJet is discarded. For all passing GenJets we measure

$$\text{Jet Response} = \frac{\text{CaloJet } E_T}{\text{GenJet } E_T} \quad (2)$$

in the bins of GenJet E_T and CaloJet $|\eta|$ listed in table 1 and table 2:

10	12	15	20	27	35	45	57	72	90	120
150	200	300	400	550	750	1000	1400	2000	2900	4500

Table 1: Bin edges of GenJet E_T in GeV used for measurement of jet response.

0.0	0.226	0.441	0.751	0.991	1.260	1.496	1.757	2.046
2.295	2.487	2.690	2.916	3.284	4.0	4.4	4.8	

Table 2: Bin edges of CaloJet $|\eta|$ used for measurement of jet response.

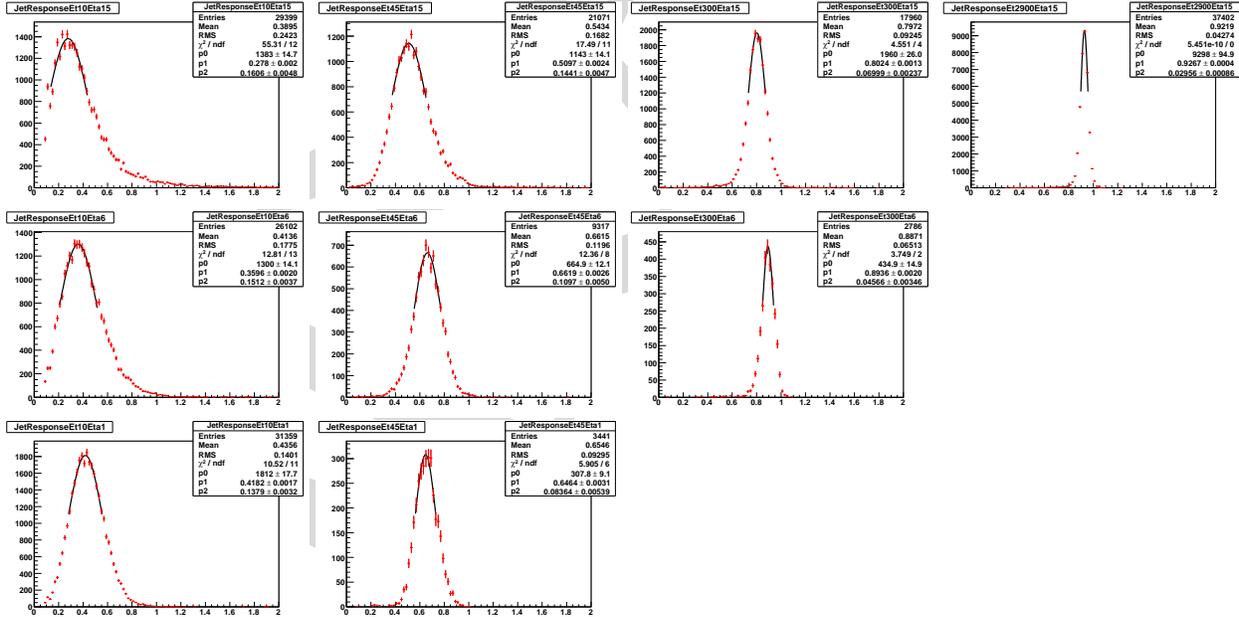


Figure 1: Jet response and fit for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T : $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

Example histograms of jet response are shown in Fig 1. Notice that for each η region (row) the response increases and the resolution improves with increasing E_T (column). To determine the peak jet response, the most probable value, we have fit each of the histograms in Fig 1 with Gaussians in the interval $\pm 1\sigma$ from the peak. If a full Gaussian fit is used instead of $\pm 1\sigma$, the mean value of the Gaussian increases by less than 3% depending on E_T . We use the mean value of the $\pm 1\sigma$ Gaussian fits to define the peak jet response at the average GenJet E_T in our bin. We use the peak jet response as input to determine the jet correction for the following reasons:

- The peak is easy to find and well defined.
- The peak is optimal for correcting resonances.
- The peak is less sensitive to thresholds and cuts applied to the calibration samples.

Example plots of peak jet response as a function of average GenJet E_T are shown in Fig. 2. The response increases smoothly with E_T . For each of the 16 bins of CaloJet $|\eta|$, we fit the response with the same parameterization used for ORCA [1]. The parameterization is compared to the response points in Fig. 2.

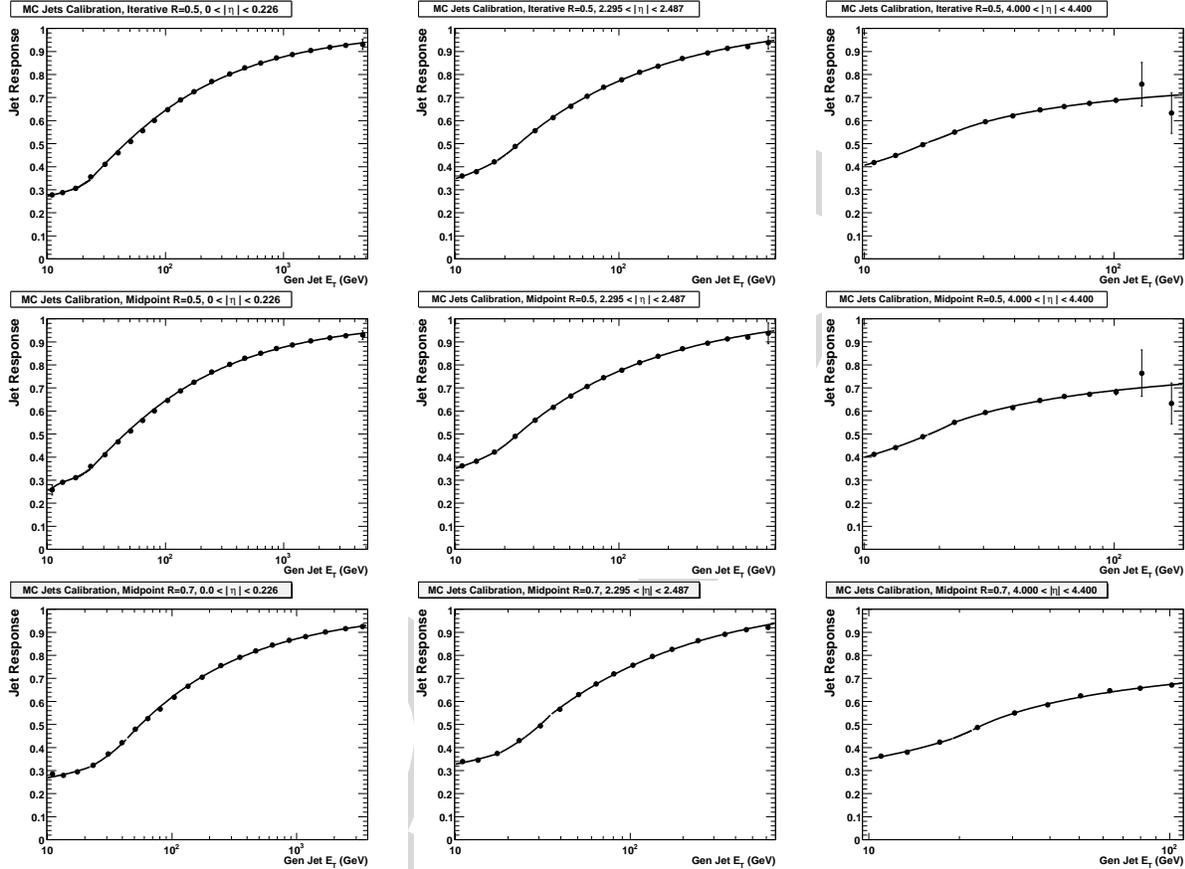


Figure 2: The jet response as a function of GenJet E_T (points) is compared to a parameterization of the response (curve). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of plot, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

Changes in jet response with different CMS detector simulations are illustrated in Fig. 3. The CMSSW_1_2_0 response is compared with both CSA06 [2] and ORCA [1]. CSA06 used CMSSW_0_8_3 simulation and CMSSW_1_0_3 reconstruction. We see that CMSSW_1_2_0 response is higher than CSA06 in the barrel and endcap. The following changes are known to have occurred and contribute to that difference

- The Hcal Endcap (HE) response in CSA06 was low by 35% because the reconstruction used the same calorimeter sampling fraction as the Hcal Barrel (HB), and this was fixed before CMSSW_1_2_0 was released.

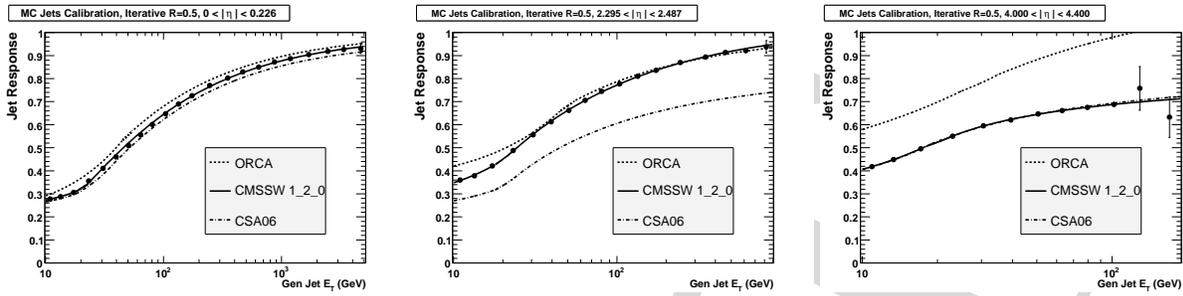


Figure 3: The jet response as a function of GenJet E_T from ORCA (dashed curve), CSA06 (dot-dashed curve), and CMSSW_1_2_0 (points and solid curve). The three plots, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

- The RecHit energy in HB and HE was increased by 3% in CMSSW_1_2_0 to account for signal lost outside of the 4 time slice window, and this correction was not present in CSA06.

The response in CMSSW_1_2_0 remains different than ORCA by a few percent in the the Barrel and Endcap, and by around 25% in the Forward.

The following are known differences between CMSSW_1_2_0 and ORCA.

- ORCA Hcal response was calibrated with Test Beam (TB) pions, but CMSSW Hcal response is not yet calibrated to TB pions. Improvements are expected in CMSSW_1_4_0.
- CMSSW_1_2_0 has very different response in HF due to problems in the calibration procedure and the use of an uncalibrated shower library. Improvements are expected in CMSSW_1_4_0

2.2 Jet Correction

The jet correction, k , is defined as

$$k = \frac{1}{\text{Jet Response}} \quad (3)$$

It is a multiplicative correction: the CaloJet Lorentz Vector, p , is multiplied by the jet correction to obtained a corrected CaloJet Lorentz Vector, p'

$$p' = kp \quad (4)$$

The parameterized jet response as a function of GenJet E_T in 16 slices of CaloJet $|\eta|$ are used as input to the MC Jet package in the JetMETCorrections subsystem of CMSSW.

The software applies a simple iteration procedure to derive the jet correction as a function of observed CaloJet E_T from the input Jet Response which is a function of true GenJet E_T . Let i be the iteration number, then k_i is the correction obtained in the i th iteration, and is equal to

$$k_i = \frac{1}{\text{Jet Response}(\text{CaloJet } E_T \times k_{i-1})} \quad (5)$$

where $k_0 = 1$. In equation 5 we are substituting an approximation for the GenJet E_T into the Jet Response function of GenJet E_T , and with each iteration the approximation becomes more precise. The software iterates ten times to obtain a value of the jet correction as a function of CaloJet E_T which has safely converged. In other words, for each reconstructed CaloJet the software solves the non-linear equation

$$\frac{\text{CaloJet } E_T}{\text{GenJet } E_T} = \text{Jet Response}(\text{GenJet } E_T) \quad (6)$$

using a simple iteration procedure.

The jet correction as a function of CaloJet E_T is obtained for the bins of CaloJet $|\eta|$ given. This means that for a fixed CaloJet E_T the correction is held constant within the bin of CaloJet $|\eta|$. The correction changes as a function of $|\eta|$ in discrete jumps, as one moves from one bin CaloJet $|\eta|$ to the next. We plan to replace this in the future with a correction that varies smoothly as a function of CaloJet η .

In Fig. 4 we show the jet correction as a function of CaloJet E_T in three bins of CaloJet $|\eta|$. Inverse to the jet response, the jet correction decreases with increasing jet E_T , asymptotically approaching 1 in the barrel and the endcap at high CaloJet E_T . We might expect the same behavior in the forward if the calibration of the HF was physically reasonable.

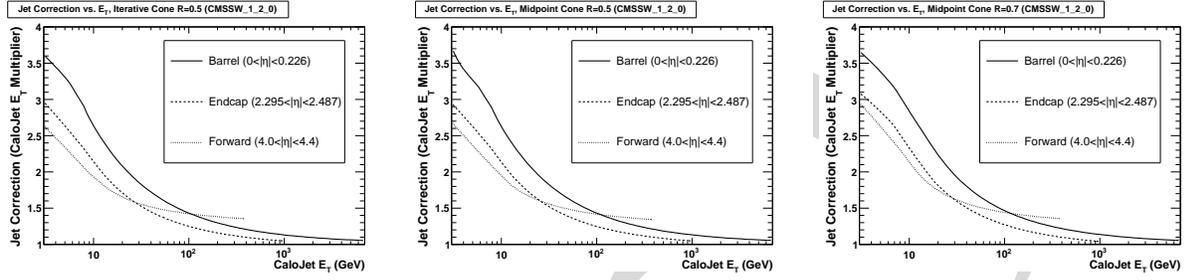


Figure 4: The jet correction as a function of CaloJet E_T for the iterative cone algorithm with a cone size of $R = 0.5$ (left plot), and the midpoint cone algorithm with a cone size of $R = 0.5$ (middle plot) and $R = 0.7$ (right plot).

2.3 Corrected Jet Response

The jet correction procedure above was used in the MC Jet package to produce collections of corrected CaloJets. The corrected jet collections were then input to the same SimJetResponseAnalysis module used to measure the raw jet response. The corrected jet response

$$\text{Corrected Jet Response} = \frac{\text{Corrected CaloJet } E_T}{\text{GenJet } E_T} \quad (7)$$

was histogrammed in the same bins of GenJet E_T and CaloJet $|\eta|$.

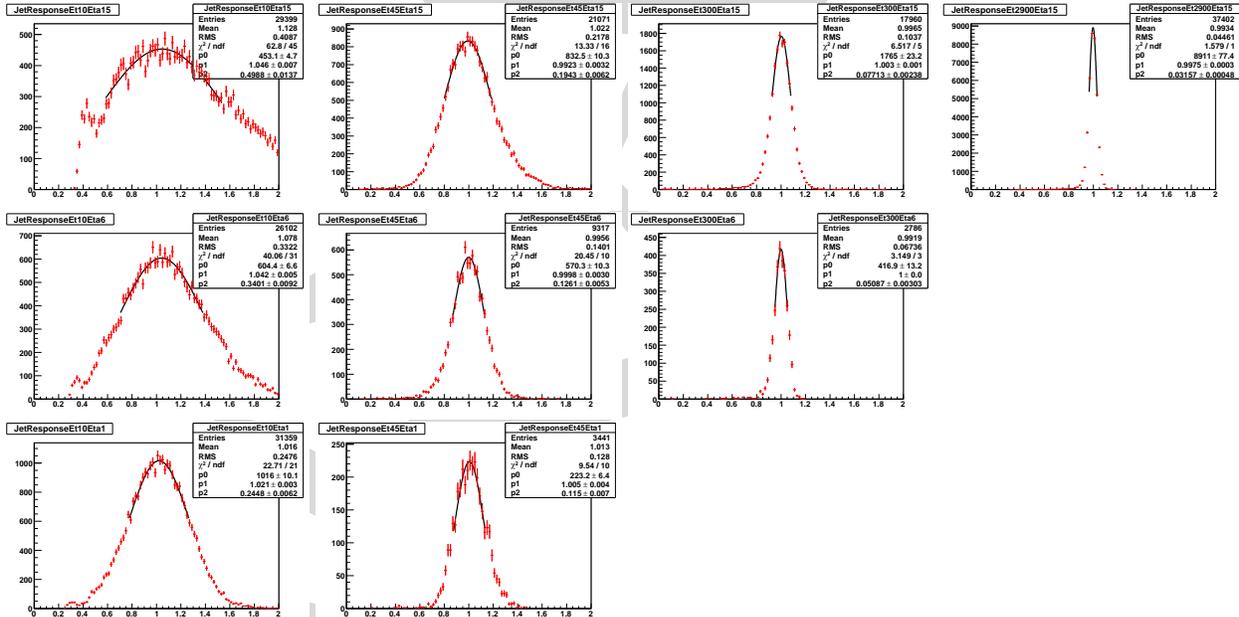


Figure 5: Corrected jet response and fit for the iterative cone algorithm with cone size $R = 0.5$. The number of jets as a function of corrected jet response (points) is compared to a Gaussian fit (curve) in the interval $\pm 1\sigma$ from the peak response. The three rows of plots, from top to bottom, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$. The four columns of plots, from left to right, are for the following regions of GenJet E_T : $10 < E_T < 12$, $45 < E_T < 57$, $300 < E_T < 400$, and $2900 < E_T < 4500$ GeV.

Example histograms of corrected jet response are shown in Fig 5. The corrected jet response peaks near 1, particularly at high GenJet E_T , indicating that the correction is working. At very low values of GenJet E_T , typically less than 30 GeV, the jet resolution is very poor, and the correction made to the jet varied significantly over the width of the jet response within the histogram. As a result the shape of the corrected jet response histogram at low GenJet

E_T in Fig 5 is slightly different from the shape of the uncorrected jet response histograms in Fig 1. As a result, for these very low energy jets, the correction does not work as well. To quantify how well the correction is working, we have determined the peak corrected jet response by fitting each histogram with Gaussians in the interval $\pm 1\sigma$ from the peak, as shown in the examples of Fig 5.

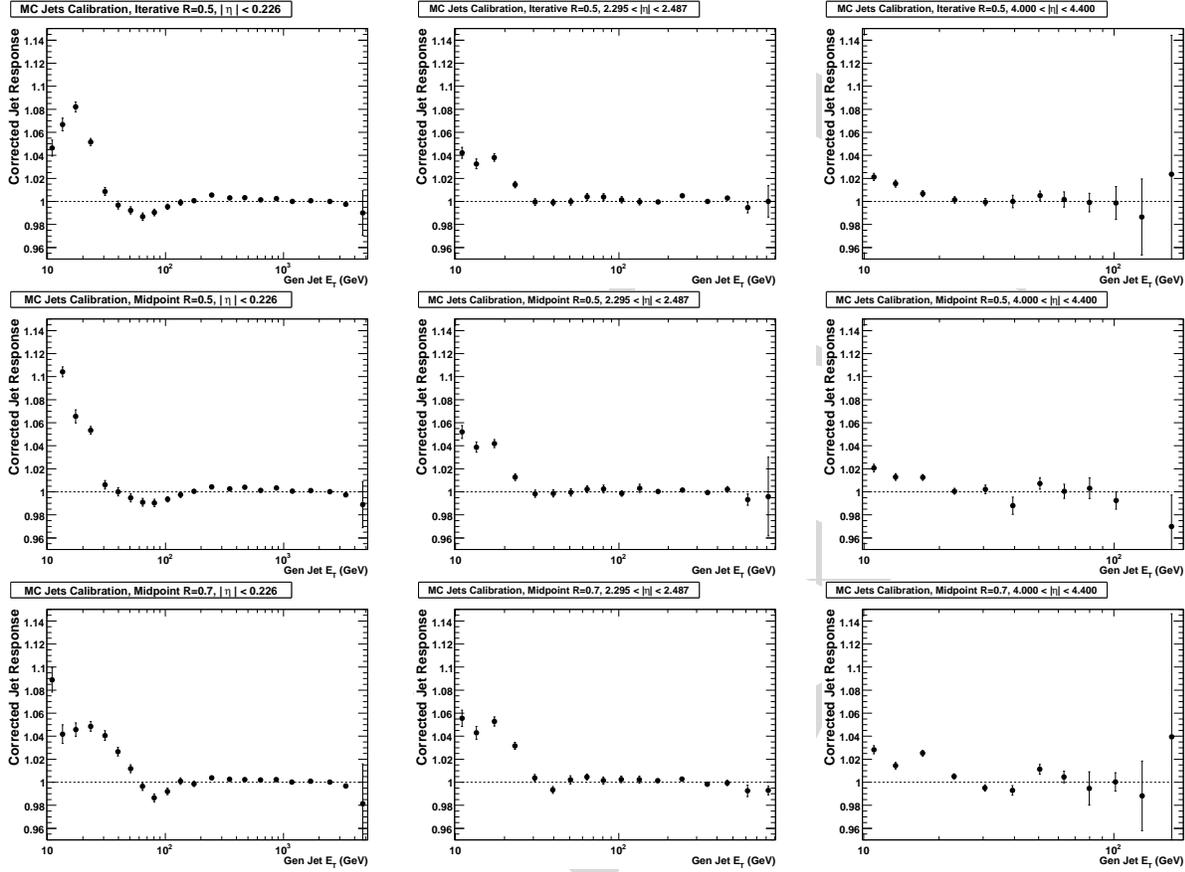


Figure 6: The corrected jet response as a function of GenJet E_T (points). The three rows of plots, from top to bottom, are for the iterative cone algorithm with a cone size $R = 0.5$, and for the midpoint cone algorithm with a cone size of $R = 0.5$ and $R = 0.7$. The three columns of plot, from left to right, are for the following regions of CaloJet η : $0 < |\eta| < 0.226$, $2.295 < |\eta| < 2.487$, and $4 < |\eta| < 4.4$.

Example plots of peak corrected jet response as a function of average GenJet E_T are shown in Fig. 6. The correction is working to within roughly 1% accuracy for GenJet $E_T > 30$ GeV. The accuracy is significantly worse at lower E_T , particularly in the barrel where the jet resolution is worse than the endcap or forward for fixed E_T . At high E_T where the jet resolution is pretty good, and the correction does not vary much within a single histogram of jet response, the corrected jet response peaks at 1 to within roughly 1%. However, at low jet E_T , where the jet resolution is poor, the corrected jet response peaks 2% to 10% above the expected value 1. We note that GenJets with E_T less than 30 GeV, corresponding to CaloJets with uncorrected E_T less than 10 GeV in the barrel, are very difficult to understand for many reasons, including the one just mentioned.

An example of the jet response as a function of η before and after jet corrections is shown in Fig. 7. Before jet corrections the plot shows the response variations of the CMS detector simulation as a function of η . After the jet corrections the response is reasonably flat around 1. The vertical dotted lines show the edges of the $|\eta|$ bins in Table 2, while the points are the jet response measured in bins equal to the CaloTower η segmentation. Comparing the response variations after corrections as a function of η with the coarse binning of the jet correction into 16 $|\eta|$ bins, we get an idea as to the maximum level of improvement we can expect from a smoothly varying jet correction as a function of eta over our current binned correction as a function of $|\eta|$.

For completeness we note that Fig. 7 was made in bins of GenJet P_T (not E_T) and in this figure we are plotting the mean of the jet response distribution, not the peak of a $\pm 1\sigma$ Gaussian fit. We don't expect this to change any of our conclusions. Also note that the distribution is asymmetric in η for $|\eta| > 3$ due to a known problem with the eta values returned by the HF simulation and fixed after release CMSSW_1_2_0.

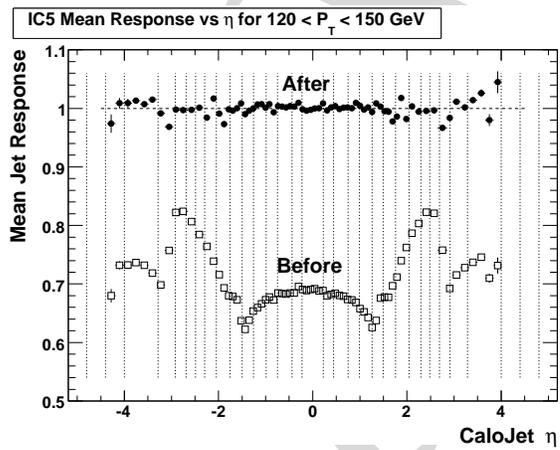


Figure 7: The jet response as a function of CaloJet η before jet corrections (boxes) and after jet corrections (closed circles), for GenJet P_T in the range $120 < P_T < 150$ GeV. Iterative cone algorithm with cone size $R = 0.5$.

3 Effect of Pileup

At nominal design luminosity ($10^{34} \text{ cm}^{-2} \text{ s}^{-1}$), LHC is expected to deliver on average about 17 proton-proton interactions per beam crossing. There are contributions both from particles produced in a trigger (in-time pileup) and from particles produced in the adjacent crossings (out-of-time pileup). Pileup (PU) of particles from different interactions will produce energy clusters in the calorimeter which can be misidentified as jets. The effect of in-time and full pileup have been studied on jet pseudorapidity.

3.1 Event Simulation

The particle-level events were generated with PYTHIA 6.227 [3]. The CMS detector simulation tool CMSSW_1_2_0 based on the GEANT4 package was used to simulate passage of particles through the detector and energy deposits in the sensitive volume. To simulate additional proton-proton interactions in a beam crossing, the signal events were mixed with a random number of minimum bias events in one crossing. The minimum bias events were generated with PYTHIA as inclusive di-jet events. The Poisson distribution with the average 5 was used to simulate in-time and full (in-time + out-time) pileup. The full pileup sample was simulated in the five crossings preceding and in the three crossings following the signal crossing.

Di-jet event samples are used in this study. To generate a sufficient number of high energy jets which are rare in an unconstrained di-jet sample, these data samples were produced with cuts on the hard process transverse momentum \hat{P}_T . The sample in \hat{P}_T bin 50-80 GeV is used for pileup observations.

3.2 Jets Pseudorapidity

The dependence of pseudorapidity distribution of the calo jets with $10 < P_T < 20$ GeV in the dijet sample with the signal event \hat{P}_T in the range $50 < \hat{P}_T < 80$ GeV for no pileup, in-time pileup and full pileup is illustrated in Fig. 8. The low number of calo jets being reconstructed in the calorimeter is due to non-linear nature of average calorimeter response as a function of jet rapidity. The forward calorimeter doesn't suffer from out-of-time pileup due to its very short response time. This can be seen in the Fig. 8(b) and Fig. 8(c). The endcap towers near the HE/HF boundary have the largest transverse size in the calorimeter and due to the largest energy flow from pileup events in the region, these towers tends to have the maximum transverse energy. It is due to this effect that the horns are visible in full pileup sample (Fig. 8(c)) around $\sim |\eta| = 3$.

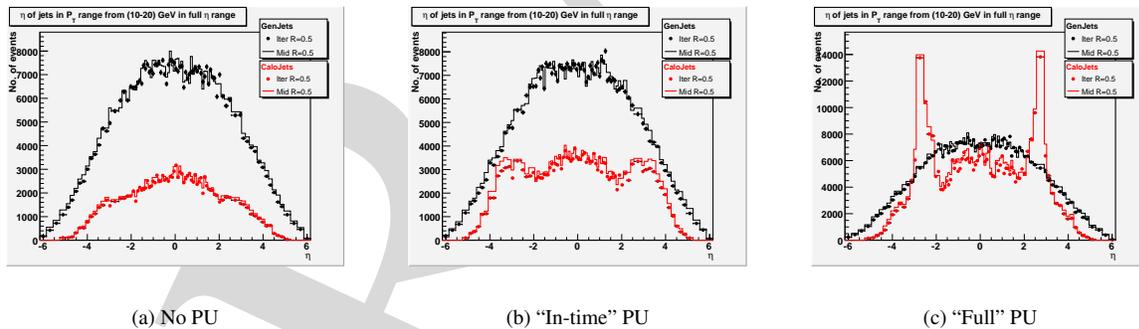


Figure 8: The pseudorapidity(η) distribution of jets with $10 < P_T < 20$ GeV in $50 < \hat{P}_T < 80$ GeV samples for different cases of pileup.

Fig. 9 shows the jets η distribution for $40 < P_T < 60$ GeV in $50 < \hat{P}_T < 80$ GeV samples. It is evident from the figures that there is little impact of pileup details on calo jet distributions at higher P_T as the minbias events contain soft jets.

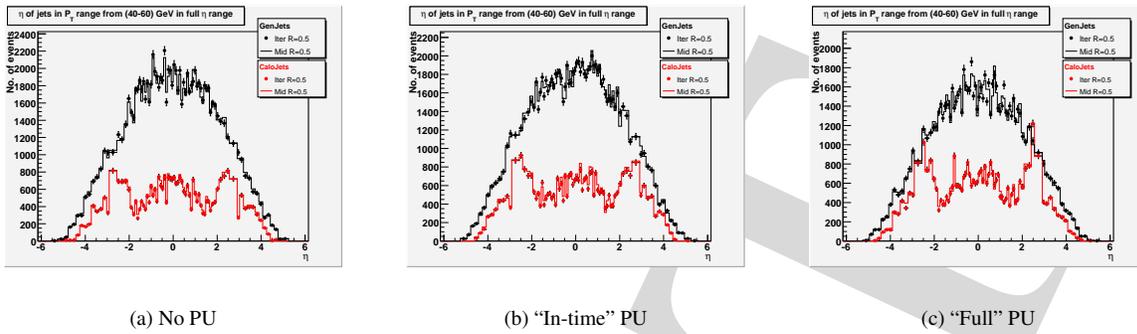


Figure 9: The pseudorapidity distribution of jets with $40 < P_T < 60$ GeV in $50 < \hat{P}_T < 80$ GeV samples for different cases of pileup.

4 Jet Response and Resolutions

5 Jet Efficiencies

6 Dijet Balance

7 MET Performance

8 Conclusions

References

- [1] CMS Note 2006/036, A. Heister et al., "Measurement of Jets with the CMS Detector at LHC".
- [2] CMS Note 2007/006, The CMS Collaboration, "CMS Computing, Software and Analysis Challenge in 2006 (CSA06) Summary".
- [3] Comput. Phys. Commun. 82(1994) 74.