CMS Physics Analysis Summary

Contact: cms-pog-conveners-egamma@cern.ch 03/08/2010

Electron reconstruction and identification at $\sqrt{s} = 7$ TeV

The CMS Collaboration

Abstract

We present results on the commissioning of CMS electron reconstruction and identification using the first LHC collisions at $\sqrt{}$ s = 7 TeV recorded by the CMS detector, corresponding to an integrated luminosity of $\sim$ 200 nb^{-1}. The electron and photon Level 1 and High Level Trigger efficiencies have been measured from data. The main observables for CMS electron reconstruction, identification, isolation and conversion rejection have been measured and compared to Monte Carlo simulation. The performance of the electron reconstruction and identification has been studied on data. Different selections proposed to be used in physics studies with electrons are studied and compared to the Monte Carlo expectations.

1 Introduction

In this paper, we present results on the performance of the electron reconstruction and identification in the CMS experiment. The analysis makes use of the first data collected during the 2010 run at the LHC at a centre-of-mass energy of $\sqrt{s}_{pp} = 7\,\text{TeV}$. The data collected correspond to an integrated luminosity of $\mathcal{L}_{\text{int}} \sim 200\,\text{nb}^{-1}$. The analysis presented here complements previous commissioning results obtained with 0.9 TeV data [1].

Different stages of the commissioning process are presented. As a preliminary check, we compare the main electron observables for low transverse momentum (p_T) candidates in minimum bias events with Monte Carlo expectations. Here the sample is mainly composed of "fakes" from charged hadrons, with a signal from "real" electrons mostly constituted by pairs from γ conversions. This sample is also used to study the performance of the Level-1 (L1) and High Level Trigger (HLT) triggers for electrons and photons. In a second step we concentrate on the p_T region beyond 20 GeV/c, using events selected either by electron/photon or jet triggers, to assess the performance of the electron reconstruction and identification on prompt electron signal and to measure the contamination from the various background sources.

Background studies are performed on electron candidates from QCD di-jet events tagged by the recoiling jet. For the signal, instead, we used tagged W electron candidates. With the integrated luminosity available at the time, electrons from Z decays could not be used to extract signal distributions but were enough to already provide a precise measurement of the reconstruction and selection efficiencies.

A brief description of the CMS detector is given in Section 2. In Section 3 we introduce the datasets which are used in this analysis. A short introduction to the algorithms used for electron reconstruction and the variables used for their identification is given in Section 4. In Section 5, the measurement of the performance of the CMS electron and photon L1 and HLT triggers is presented. A comparison between data and Monte Carlo for minimum bias triggered data is shown in Section 6. Section 7 presents the results obtained on the commissioning of electrons with $p_T > 20\,\text{GeV/c}$. Observables relevant for the seeding step of the electron reconstruction are compared to Monte Carlo for signal electrons and a first measurement of the electron reconstruction efficiency is presented. The principles for the selection of electron candidates are described and identification variables for signal and background candidates are compared with the Monte Carlo expectation. The measurement of the electron identification efficiency is also presented. Finally, distributions for background electron candidates are presented and the background fake rates for different level of tightness of the electron identification criteria are measured from data.

2 The CMS Detector

The CMS detector is a general-purpose detector built to explore the physics at the TeV scale and is described in detail in Ref. [2]. We provide here only a brief description of the main components that are relevant to this analysis.

The CMS electromagnetic calorimeter (ECAL) is a crystal calorimeter designed to have both excellent energy resolution and high granularity. These properties make it a powerful instrument to study photon production at the LHC. It is organized into two pseudorapidity (η) regions: the barrel ($|\eta| < 1.479$) and the endcap ($1.479 < |\eta| < 3.0$). The calorimeter is constructed with lead tungstate crystals that are arranged in a projective geometry. The crystals are 25.8 X_0 long in the barrel, and 24.7 X_0 in the endcap region. In the barrel region the front face of the crystal

is approximately 22×22 mm^2, corresponding to a granularity of $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. In the endcap region there is a lead-silicon-strip preshower detector (ES) consisting of two orthogonal silicon strip detectors that have a strip pitch of 1.9 mm. One plane of the detector is at a depth of 2 X_0 and the other is at 3 X_0. Avalanche Photodiodes (APD) are used to detect the scintillation light in the barrel region and Vacuum Photo-Triodes (VPT) are used in the endcap region.

The ECAL is surrounded by a brass/scintillator sampling hadron calorimeter (HCAL) with coverage up to $|\eta| = 3.0$. The scintillation light is converted by wavelength-shifting fibers, which are readout with hybrid photodiodes. The HCAL is sub-divided into towers with a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, corresponding to 5×5 ECAL crystals in the barrel region.

The CMS tracking system is composed of several layers of silicon pixel and silicon strip detectors and covers the region $|\eta| < 2.5$. The tracker and both calorimeters are embedded inside a 3.8 T solenoidal magnet. With the silicon tracker the transverse momentum resolution for high momentum tracks (100 GeV/c) is around 1 - 2% up to $|\eta|$ = 1.6, beyond which it degrades because of the reduced lever arm.

3 Data and Monte Carlo Samples

The data used in this analysis were collected with the CMS detector at the LHC collider operating at 7 TeV. They correspond to a total integrated luminosity of 198 nb^{-1}. Only data where the calorimeters and the tracker were fully operational were used in this analysis.

The data used for the trigger efficiency measurements are selected making use of minimum bias triggers. The first commissioning checks on minimum bias events described in Section 6 involve in addition a selection based on scintillator tiles placed upstream of both Forward Hadronic Calorimeters (Beam Scintillation Counters or BSC), in coincidence with the beam pick-up monitors (BPTX) and together with a requirement intended to eliminate beam background contamination. In order to eliminate losses caused by BSC inefficiencies, events with one reconstructed primary vertex with more than 4 tracks were also selected.

High p_T electron selections make use of events triggered by the CMS HLT electron or photon paths. The selection of high p_T electron candidates uses an HLT trigger which is seeded by the electromagnetic L1 trigger requiring a transverse energy of at least 5 GeV (L1_SingleEG5), and additionally only requires a reconstructed ECAL supercluster with uncorrected energy larger than 10 or 15 GeV (HLT_Photon10 and HLT_Photon15). HLT_Photon10 was used at lower luminosity and HLT_Photon15 at higher luminosity when the former was prescaled.

For the selection of fake electron candidates coming from background we instead use the single jet triggers with a threshold on the uncorrected jet energy of 15 GeV (HLT_Jet15U).

Electrons are reconstructed offline by applying the standard CMS electron reconstruction [3]. Anomalous events in the barrel part of the ECAL [4] were rejected using standard filters developed for this purpose. In the following, the missing transverse energy ($\not\!\!E_{\mathrm{T}}$) is computed using the particle flow algorithm [5].

The Monte Carlo samples used in this analysis were generated with PYTHIA version 6.4 [6]. Events were generated with a center of mass energy of 7 TeV and the CTEQ6L [7] parton distribution functions. The tune is the so-called D6T tune unless otherwise stated. The detailed CMS detector response is simulated using Geant4 [8]. The position and width of the beam spot in the

simulation were adjusted to the values determined from the data. The simulation was carried out using the most up-to-date description of the alignment, calibration and dead channel lists.

4 Electron reconstruction and identification: algorithms and tools

Electron reconstruction uses two complementary algorithms at the track seeding stage: 'tracker driven' seeding, more suitable for low p_T electrons as well as performing better for electrons inside jets and 'ECAL driven' seeding. The 'ECAL driven' algorithm starts by the reconstruction of ECAL "superclusters" of transverse energy $E_{\mathrm{T}} > 4$ GeV and is optimized for isolated electrons in the p_T range relevant for Z or W decays and down to $p_T \simeq 5$ GeV/c. "Supercluster"[9] is a group of one or more associated clusters of energy deposits in the ECAL constructed using an algorithm which takes account their characteristic narrow width in the η coordinate and their characteristic spread in ϕ due to the bending in the magnetic field of electrons radiating in the tracker material. As a first filtering step, superclusters are matched to track seeds (pairs or triplets of hits) in the inner tracker layers, and electron tracks are built from these track seeds. Trajectories are reconstructed using a dedicated modeling of the electron energy loss and fitted with a Gaussian Sum Filter (GSF) [10].

The filtering performed at the seeding step is complemented by a preselection. For candidates found only by the 'tracker driven' seeding algorithm, the preselection is based on a multivariate analysis as described in [5]. For candidates found by the 'ECAL driven' seeding algorithm, the preselection is based on the matching between the GSF track and the supercluster in η and ϕ [3]. The few 'ECAL driven' electron candidates ($\sim$1% for isolated electrons) not accepted by these matching cuts but passing the multivariate preselection are also kept.

5 L1 and High Level Trigger

During the CMS 7 TeV collision data taking period, the CMS L1 e/γ trigger was fully deployed and operational. Noisy or absent ECAL channels were masked (2.5% in the barrel and 5% in the endcaps).

The L1 trigger decision is based on electron/photon trigger candidates which uses local energy deposits called trigger primitives as inputs. Details on the ECAL trigger algorithm can be found in Ref. [2]. The transverse energy of each ECAL trigger tower is computed in 250 MeV steps, then adjacent trigger towers are summed to form pairs using a 3×3 sliding window around the trigger primitive with maximum energy. The total transverse energy of a pair is used to compute L1 trigger energy values. An isolation, based on the amount of energy deposited in the towers around the central one, is used to separate candidates into isolated and non-isolated. Only the four most energetic isolated and non-isolated L1 candidates are sent to the global trigger which generates the final decision . The L1 trigger requires a transverse energy above a configurable cut: L1_SingleEG5 is requiring $E_T > 5$ GeV. The L1_SingleEG5 ran unprescaled in the dataset which has been considered and is used as a seed for the High Level Trigger electron and photon paths used in this analysis.

The performance of the L1 electron/photon trigger was evaluated in terms of efficiency using selected minimum bias collisions as described in Section3. This sample is unbiased with respect to the trigger under study.

Figure 1 shows the L1Single_EG5 trigger efficiency in the ECAL barrel and ECAL endcaps. Efficiency is computed with respect to reconstructed electrons, selected using dedicated cuts suitable to enrich the minimum bias sample in electrons from conversions, with a purity higher

than 80%. The trigger efficiency is shown as a function of the electron supercluster E_T as determined from the offline reconstruction.

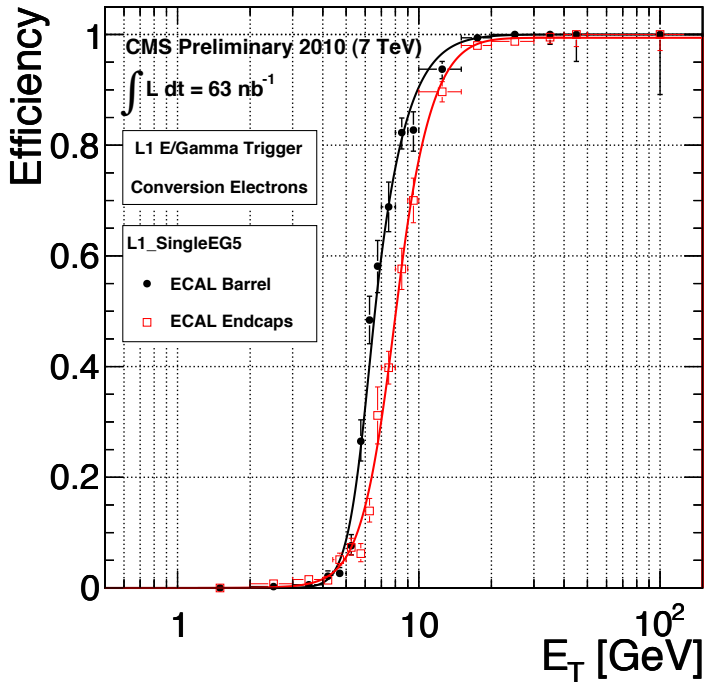


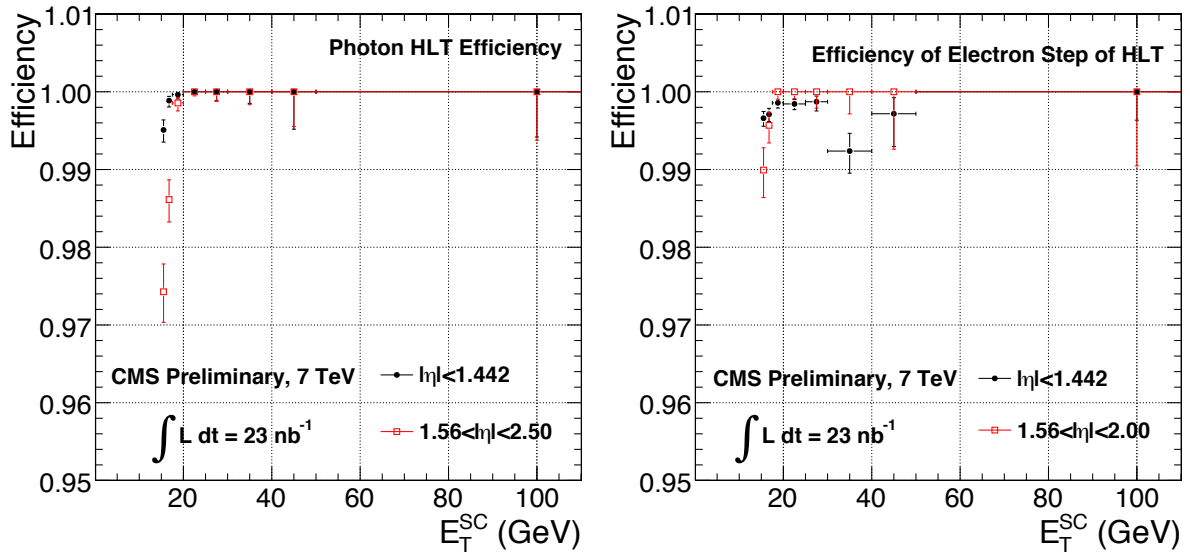
Figure 1: Level-1 EG5 trigger efficiency for electron candidates from minimum bias data as a function of the electron supercluster transverse energy for candidates in the ECAL barrel (black dots) and in the ECAL endcaps (red empty squares). The horizontal bars indicate the bin size.

At HLT, electron and photon selection proceeds requiring a supercluster with E_T above a given threshold matching an electromagnetic L1 candidate. The HLT runs the standard ECAL superclustering algorithm with almost identical settings to the offline reconstruction. The electron paths additionally require a hit in the pixel layers of the CMS detector compatible with an electron trajectory, with matching requirements currently looser than the offline requirements in most regions of the detector. Therefore, electron paths are a complete subset of the photon paths when the E_T threshold is identical. More details on the electron HLT can be found in Ref. [2].

The sample used to evaluate the HLT efficiency is selected with the same criteria used for L1 efficiency estimation. The primary photon trigger used for this analysis is the HLT_Photon15 trigger, which requires a matching L1 candidate with $E_T > 5$ GeV and an online supercluster with $E_T > 15$ GeV. The corresponding electron trigger is called HLT_Ele15_LW.

Efficiency to pass HLT_Photon15 is computed with respect to an offline electron which has already produced a L1_SingleEG5 candidate: it is shown in Figure 2. Also in this case it is evaluated as a function of the reconstructed electron supercluster E_T. The efficiency vs E_T for an offline ECAL seeded electron to pass the HLT_Ele15LW path, having already passed the HLT_Photon15 path is also shown in Figure 2.

Since there are no other HLT requirements except for an on-line supercluster with $E_T > 15$ GeV matched to a L1 candidate, the HLT component of the trigger efficiency is close to 100% for



(a) (b)

Figure 2: The electron/photon HLT efficiencies: (a) HLT_Photon15 efficiency for an offline reconstructed electron matched to a L1_SingleEG5 candidate as a function of E_T, and (b) the HLT_Ele15LW efficiency for an offline reconstructed electron which passes HLT_Photon15 as a function of E_T.

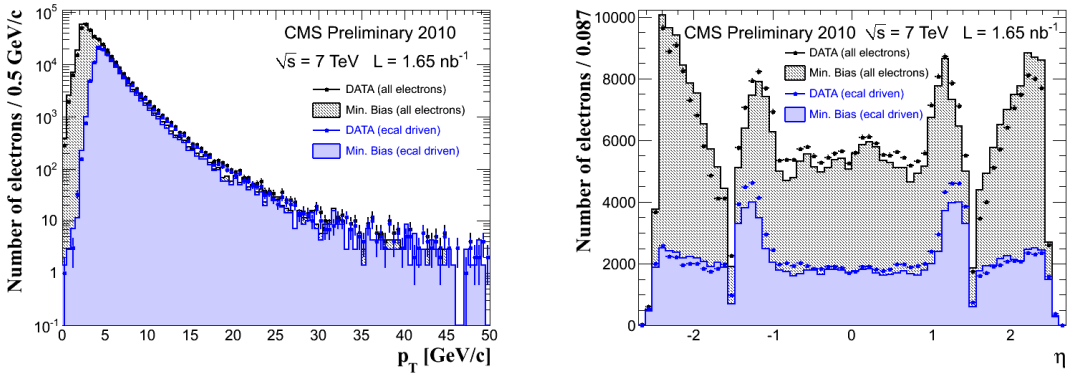
photon triggers. The online pixel-matching requirement of the electron triggers is close to 100% efficient w.r.t to the offline pixel-matching requirement up to an η of 2.0.

6 Commissioning with minimum bias data

We present here the results obtained on an inclusive sample of electrons selected from minimum bias events.

Figure 3 presents the kinematical p_T and η distributions of the reconstructed electron candidates. The data corresponds to an integrated luminosity of $\mathcal{L}_{\text{int}} \sim 3$ nb^{-1}. The Monte Carlo distribution is normalized to the total number of electron candidates in the data. No other selection than the reconstruction preselection is applied to electron candidates. The composition of this inclusive sample, as estimated from the Monte Carlo, is that $\sim$ 4.6% of the electron candidates are "real" electrons (34.9% from D decays and 46.7% from B decays, 13.5 % from Dalitz decays and few from J/Ψ decays), 33.9% come from γ conversions and 61.5 % are fakes from hadrons. As expected, the 'ECAL driven' algorithm contribution dominates above 4 GeV while the contribution of pure 'tracker driven' electrons allows coverage of the low p_T range. The relative fraction of electron candidates found by the 'ECAL' or the tracker driven seeding algorithm is found to be well reproduced. The differences in the shape of the η distribution for 'ECAL driven' electron candidates in particular in the endcaps mainly arise from the 4 GeV E_T cut and the preselection cuts applied in the 'ECAL driven' reconstruction. A reasonable agreement is found with the simulation. Discrepancies in the η distribution are essentially due to an imperfect calibration of the ECAL endcaps.

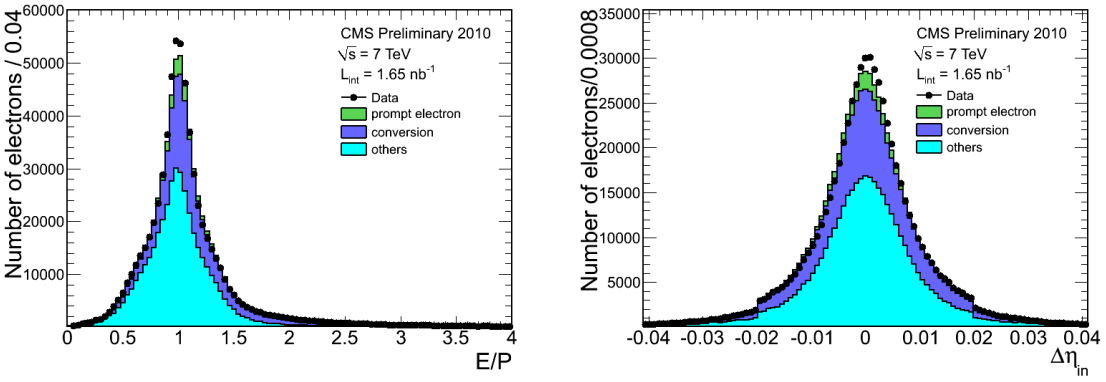
Figure 4 presents the comparison between data and Monte Carlo of the distributions of the ratio E/p of the electron SC energy over the track momentum and the position matching $\Delta\eta_{in}$.



(a) (b)

Figure 3: Kinematical distributions of electron candidates in minimum bias events (dots) compared with Monte Carlo (histograms) for all reconstructed candidates and for reconstructed candidates from the ECAL driven seeding algorithm: (a) transverse momentum distribution and (b) pseudorapidity.

A good agreement is found with the simulation, apart from a small discrepancy in $\Delta\eta_{in}$ being understood as due to a displaced position of the ECAL with respect to the tracker [4]. In position matching variables, some biases are also expected from the position measurement in the ECAL being performed on hadrons and conversion electrons.



(a) (b)

Figure 4: Distributions of electron candidate track-cluster matching variables in minimum bias events in data (dots) compared with Monte Carlo (histograms): (a) ratio E/p between the supercluster energy and the GSF track momentum and (b) difference $\Delta\eta_{in}$ between the supercluster position and the track extrapolation from the innermost measurement.

7 Commissioning of electrons with $p_T > 20$ GeV/c

We present in the section the results obtained for the commissioning of the electrons with p_T >20 GeV/c. As a prompt electron signal source, electron candidates tagged from the W decay and triggered from electron/photon triggers have been used. To study the background contamination a sample of reconstructed electron candidates tagged from di-jet events and trig-

gered from jet triggers has been exploited. In Section 7.1 we first show results on the electron reconstruction for signal electron candidates , and then present in Section 7.2 the performances of the electron selection.

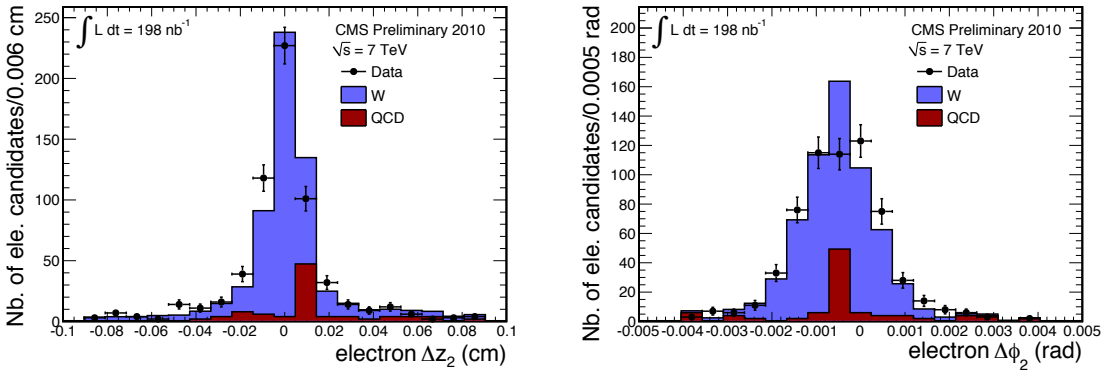
7.1 Electron reconstruction

7.1.1 Seeding Distributions from W $\longrightarrow e\nu$ events

The ECAL-driven seeding starts by the reconstruction of superclusters in the ECAL and takes advantage of the fact that the supercluster energy weighted position is on the helix corresponding to the initial electron trajectory. Back-propagating the measured supercluster energy and position through the magnetic field, a prediction on the innermost tracker layers is inferred and used to match standard tracker seeds with the electron supercluster.

We present here a comparison of the electron track seeding distributions as obtained from tagged electrons from W events. Events are tagged requiring the missing transverse energy $\not{E}_{\mathrm{T}}$ to be above 40 GeV. In addition, the matched electrons are required to be isolated and to pass shower shape requirements. No cut is applied on the track-cluster matching to produce unbiased residuals in the seeding windows.

Figure 5 presents the distributions of the differences between the expected position and the hit position in the second layer of pixels for electrons with ECAL-driven seeds. The data correspond to an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 198\,\mathrm{nb}^{-1}$. The distributions are presented for the barrel pixel detector and for both z and ϕ coordinates. The Monte Carlo distributions are normalized to the total number of events observed in the data. The background considered is from QCD di-jet events. A good agreement is found with the simulation. A bias is observed for the ϕ coordinate in both data and Monte Carlo which is mainly due to the offset of the beam spot with respect to the origin of the CMS reference frame.



(a) (b)

Figure 5: Differences between measured and predicted hit position on the second barrel pixel layer for the ECAL-driven seeding for tagged signal electron candidates from W events in data (dots) and compared with Monte Carlo (histograms): (a) z coordinate and (b) ϕ coordinate.

7.1.2 Measurement of the reconstruction efficiency

The baseline method for the measurement of the electron reconstruction efficiency relies on ″tag-and-probe″ from Z decays [11]. A well defined electron is used as the ″tag″. We require

the electron supercluster to have E_T >20 GeV and the electron has to pass standard identification and isolation cuts of the Z→ee selection. Probes are defined as superclusters with E_T >20 GeV which match, together with the tag, the nominal Z mass. The probe superclusters are then used to measure the efficiency by counting the fraction of reconstructed electrons corresponding to the probe superclusters (passing probes) over the total number of probe superclusters. Approximately 70 Z→ee events have been used to measure the reconstruction efficiency. An efficiency value of 99.3% ± 1.4% (resp. 96.8% ± 3.4%) is obtained for electrons in the ECAL barrel (resp. in the ECAL endcaps), in good agreement with the expected efficiency from the Monte Carlo simulation of 98.5% (resp. 96.1%).

As a complementary method, W events are used to determined the reconstruction efficiency. Superclusters are selected requiring $E_T > 20$ GeV in the ECAL region relevant for the electron reconstruction ($|\eta| < 2.5$). In order to reduce the di-jet background, a requirement on shower shape of the seed cluster of the supercluster is applied, together with tracker, ECAL and HCAL isolation requirements. To further reduced the QCD di-jet background, events are also required to have no jet with $E_T > 25$ GeV (corrected transverse energy) in the region $|\eta|$ <3.0 and to satisfy $\not{E}_T / E_T^{SC} > 0.3$. An unbinned maximum likelihood (ML) fit is applied to the transverse mass, fixing the W signal shape from Monte Carlo and floating all background shape parameters on data. The fit returns simultaneously the signal and background yields together with their reconstruction efficiency.

The results are summarized in Figure 6 where the ratio of the measured efficiency to the expected efficiency from Monte Carlo is presented for both the "tag-and-probe" method using Z→ee events and for the ML fit based on W→eν events. Results are shown separately for electrons in the ECAL barrel and in the ECAL endcaps. A good agreement is found between the two methods and the ratio values are found well compatible with unity. The combined efficiency values and their errors are also indicated.

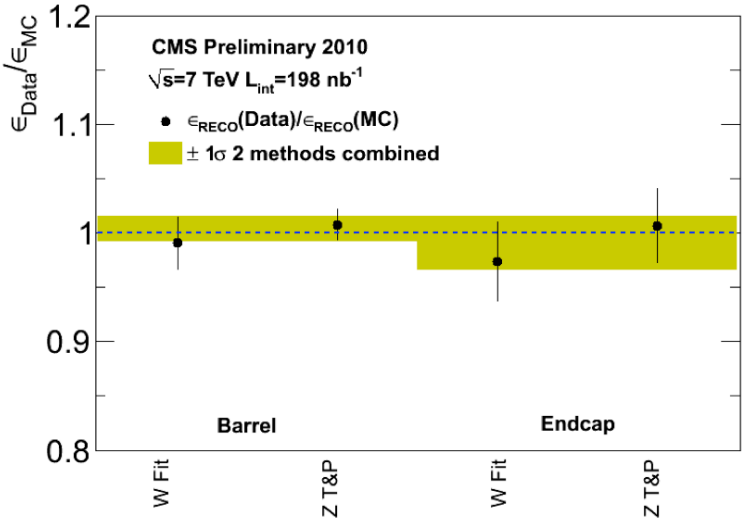


Figure 6: Ratio of the measured electron reconstruction efficiency over the expected efficiency from Monte Carlo. The values are shown for both the "tag-and-probe" method using Z→ee events and for the ML fit based on W events, as well as separately for the ECAL barrel and ECAL endcaps. Combined values of the efficiency and their errors are also indicated (yellow bands).

7.2 Electron selection

Electron selection variables may be categorized in 3 groups: identification (ID) variables, isolation variables and conversion rejection variables.

The electron identification variables that have been found to be the most powerful, and are used in the selection, are: the energy-momentum match between the seed cluster and the track E_{seed}/p_{in}, the variables measuring spatial matching between the track and the supercluster, $\Delta\eta_{in}$ and $\Delta\phi_{in}$, the supercluster η width, $\sigma_{i\eta i\eta}$ (as taken from the covariance matrix using logarithmic weights), and the hadronic leakage variable H/E. The supercluster η width is to a very good approximation unaffected by the spreading due to the magnetic field of the showering in the tracker material.

Isolation variables are computed in three sub-detectors: the tracker, the ECAL, and the HCAL. Transverse energy/momentum sums are evaluated in regions of $\Delta R < 0.3$. As electrons undergo bremsstrahlung energy loss in the tracker material, care is taken to remove from the isolation sums the contributions from bremsstrahlung photons and possible resulting conversion electrons.

Three sets of variables are used to reject electrons from conversions. The transverse impact parameter is used to discriminate electrons from conversions as they will have, on average, a greater distance to the beam position. We also expect an electron from a conversion to have missing hits in the innermost tracker layers. Finally, electrons from conversions are identified by finding the partner conversion track. This is done by looking for an oppositely charged track whose polar angle is close to the one of the electron track and whose distance at the point where the two tracks are parallel is small.

7.2.1 Electron selection algorithms

For early analyses an electron selection using simple cuts, rather than a multi-variate approach, provides a useful tool to understand the data and make comparison with Monte Carlo. At the moment two different approaches are being commissioned.

One applies simple cuts on the variables described above. Different cuts are used for electron candidates found in the ECAL barrel and the ECAL endcaps. Simple cuts give simplicity and transparency for early analyses.

A series of reference selections have been produced using Monte Carlo samples of graded severity with efficiency for prompt electrons of E_T >20 GeV having nominal values of 95%,90%, 85%, 80% and 70%. We consider for this PAS only the 95% and the 80% working points, and we will refer to them as "WP95" and "WP80". The effect of the application of the selection on signal and background can be seen in an inclusive sample of electron with $E_T > 25$ GeV as shown in Figure 7. Here the transverse mass (M_T) as reconstructed from the electron candidate p_T and the $\not{E}_T$ is used to better separate the signal (from $W \rightarrow e\nu$) from the background coming from di-jet events where one of the jets provides a fake (charged hadron/π^0 overlap, or early showering charged hadron), or real (heavy flavor decay, or photon conversion) electron. The Monte Carlo well reproduces the relative reduction of signal and background yields as the selection tightens. This result provides a first level of validation of the selection.

A second approach to electron selection is to separate electron candidates into categories according to observables that are sensitive to the amount of bremsstrahlung [3]. These variables are the fraction $fbrem$ of radiated energy as measured from the innermost and outermost state of the electron track and the ratio E/p between the supercluster energy and the measured track

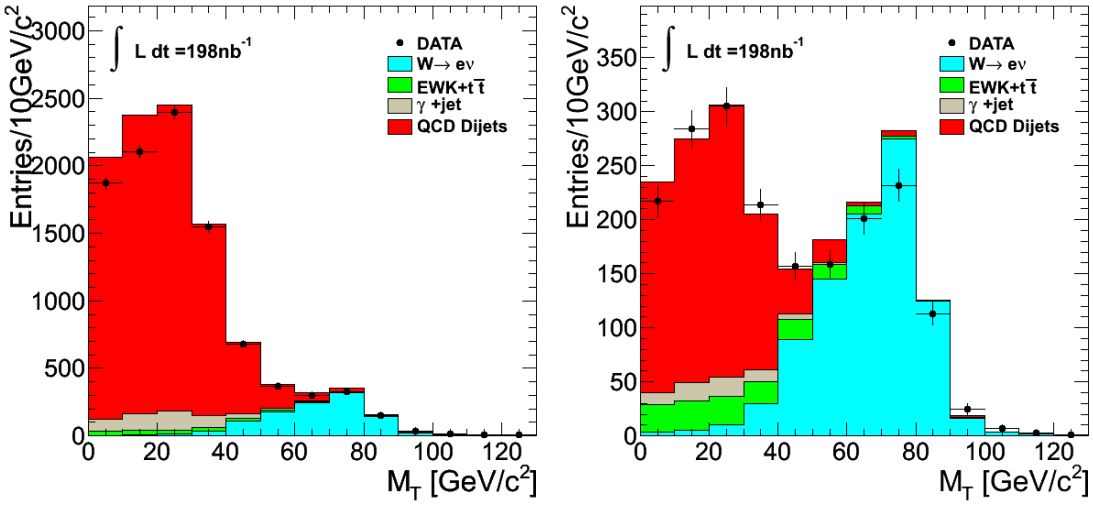


Figure 7: Transverse mass distribution for candidates with supercluster E_T >25 GeV in data (dots) compared with Monte Carlo (histograms) after (a) WP95 and (b) WP80 selections.

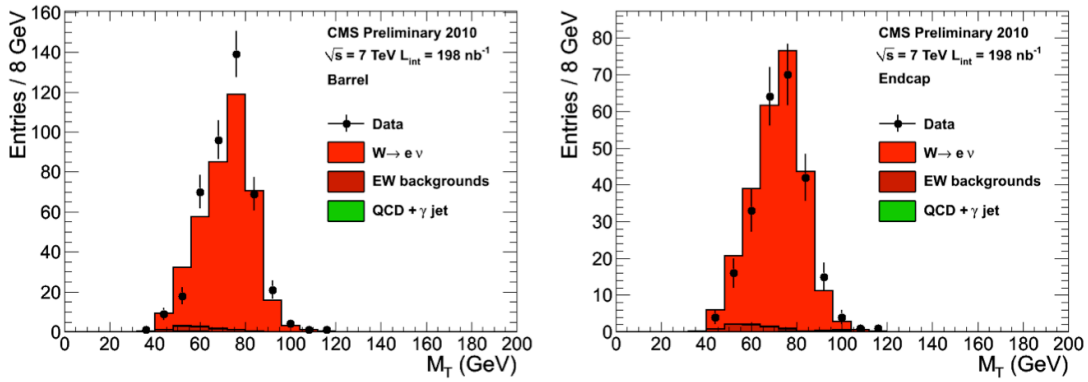
momentum at the vertex. Three categories are defined so to separate electrons with quite different measurement characteristics and purity, applying cuts in the ($fbrem$, E/p) plane.

The cuts are optimized to give the best signal to background ratio (s/b) for single electrons. This is done optimizing the cuts in each of the three categories as well as separating electron candidates in the ECAL barrel and ECAL endcaps and across different E_T regions. For a given set of cuts, the method ensures that individual electrons all have a s/b larger than some given value. Cuts have been determined for nine different s/b values, each level of cuts reducing the background rate by about a factor of two. The selection therefore ranges from very high efficiency that might be useful in multilepton events to very large background rejection that might be useful to detect single electrons. For the purpose of this PAS we only consider two working points which we will refer as "Cic Loose" and "Cic SuperTight", respectively targeted for efficiency of about 95% and 85% on prompt signal electrons from W and Z decays.

7.2.2 Signal Electrons Studies from W $\longrightarrow$ eν

A clean sample of electrons candidates is obtained from W $\rightarrow$ eν events. The selection for tagged signal electrons requires a cut in the electron transverse energy ($20 < E_T < 60$ GeV), a cut in the $(\not\!\!E_T, \sum E_T)$ plane as well as a cut on the azimuthal angle between the electron and the $\not\!\!E_T$ ($\Delta\phi_{e,\not\!\!E_T} > 0.75$). For many studies we also require that the electron candidate pass the very loose electron selection criteria to better define the tagged sample (this requirement is not included for the electron ID efficiency measurement). Figure 8 shows the transverse mass distribution for the selected events. The distributions are normalized to the integrated luminosity. The data agree well with the expectation from simulation both in the ECAL barrel and in ECAL the endcaps. The background is expected to be very small with a main contribution from $W \rightarrow \tau\nu \rightarrow e\nu\nu\nu$.

Figure 9 presents the distributions for tagged signal electron candidates of some electron ID and isolation variables. The distributions shown are: the ratio E/p of the supercluster energy in the ECAL over the track momentum at the innermost state, the ϕ difference $\Delta\phi_{in}$ between the supercluster position and the track extrapolation from the innermost state, the supercluster



(a) (b)

Figure 8: Transverse mass distribution for tagged signal electron candidates from W events in data (dots) compared with Monte Carlo (histograms) for candidates in (a) ECAL barrel and (b) ECAL endcaps.

η width $\sigma_{i\eta i\eta}$ and the combined isolation sums using the ECAL, HCAL and the tracker. The distributions are shown separately for electron candidates in the ECAL barrel and in the ECAL endcaps. A good agreement is found with the simulation.

7.2.3 Electron selection efficiency measurement

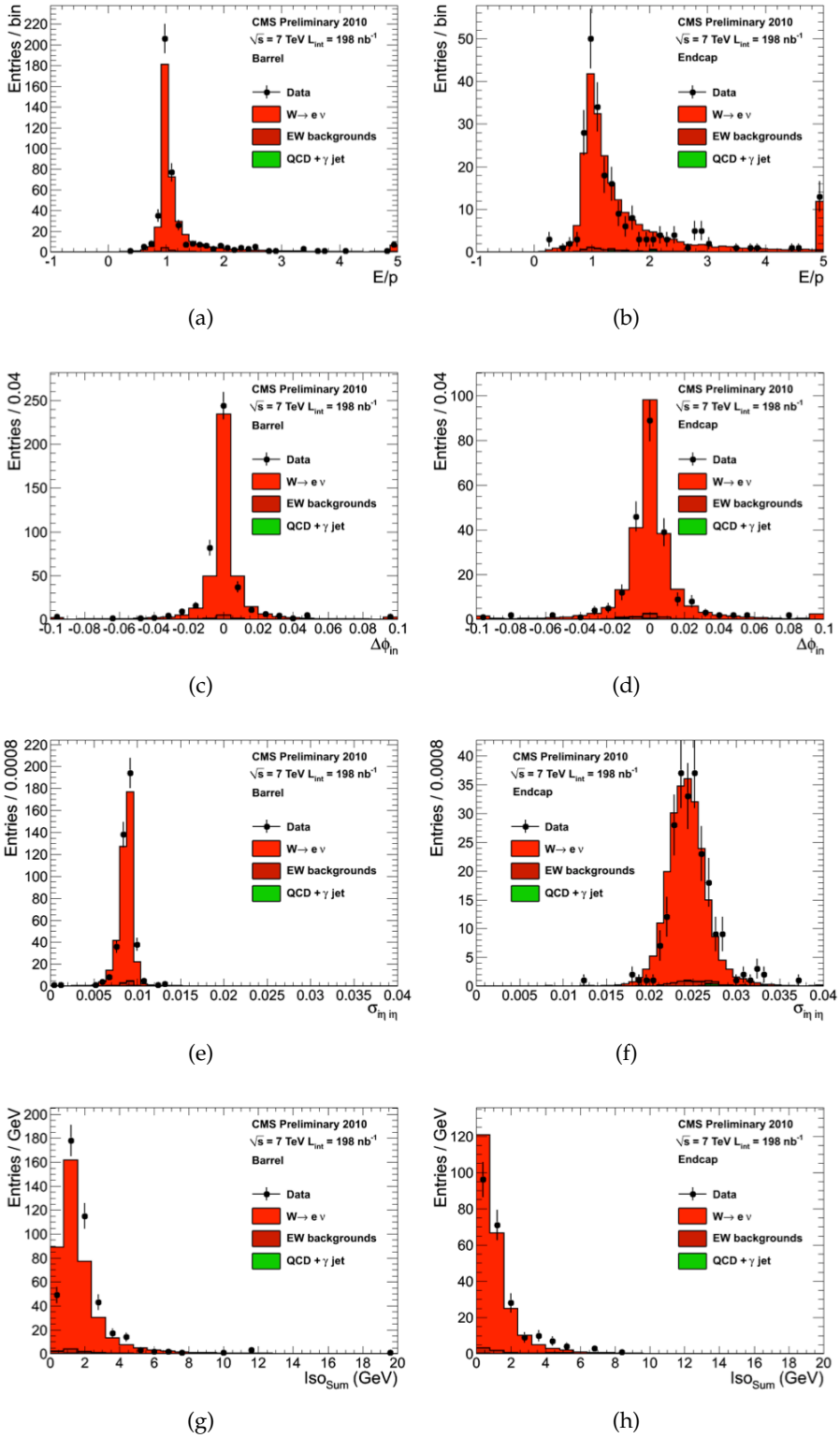
The baseline method for the measurement of the electron selection efficiency is the "tag-and-probe" from Z→ee as described in Section 7.1.2. Two alternative methods based on W events have also been used to determine the selection efficiency: a "tag-and-probe" method using the W tag described above and the ML fit described in Section 7.1.2. In the latter case, the treatment of the systematics errors is the same as described previously in Section 7.1.2.

The results are summarized in Figure 10 which presents the ratio of the measured efficiency to the expected efficiency from Monte Carlo for the three methods and for the two cut based selections corresponding to the 95% and the 80% working points and the two category based selections referred as "Cic Loose" and "Cic SuperTight". Results are shown separately for electron candidates in the ECAL barrel and in the ECAL endcaps. A good agreement is found between the three methods and within the errors the measured efficiencies are found to agree well with the Monte Carlo expectation. The combined efficiency values and their errors are also indicated.

The results as obtained from the "tag-and-probe" method using Z→ee events are finally presented in Table 1 and compared to Monte Carlo expectations for electron candidates in the ECAL barrel and in the ECAL endcaps and for the cut based and category based selections.

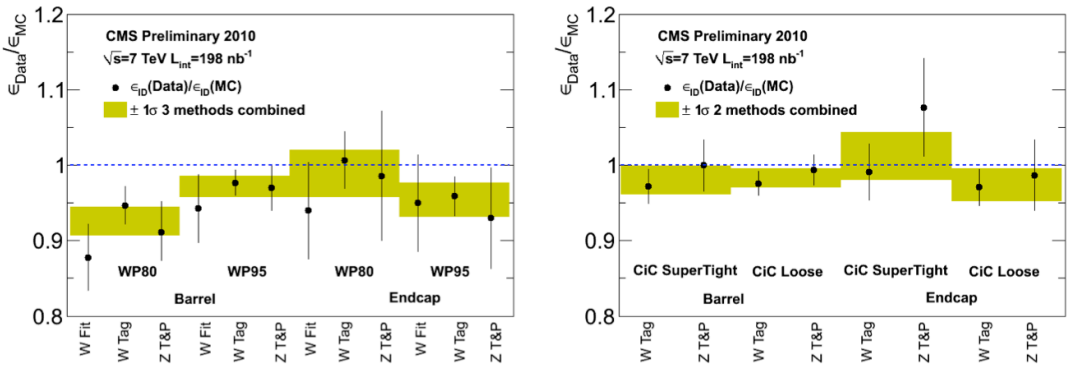
7.2.4 Background electron studies

Several requirements are applied to tag potential background electron candidates. We first require a jet in the event with (uncorrected) E_{T} greater than 20 GeV and satisfying one or more of the High Level Jet Triggers with $E_T > 15$ GeV. In order to eliminate tagging due to real electrons, the electromagnetic fraction of the jet is required to be less than 90%. The event must contain a reconstructed electron candidate with $E_T > 10$ GeV which is separated from the tag jet by $\Delta R > 0.4$. Finally, potential signal contamination from $W^{\pm}$ is suppressed by requiring $\not{E}_{\mathrm{T}}$ to be less than 30 GeV.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 9: Distribution of electron ID variables for tagged electron candidates from W events in data (dots) compared with Monte Carlo (histograms): (a,b) E/p_{in}, (c,d) $\Delta\phi_{in}$, (e,f) $\sigma_{i\eta i\eta}$, and (g,h) isolation sum. Distributions are presented separately for (a,c,e,g) the ECAL barrel and (b,d,f,h) the ECAL endcaps. The distributions are normalized to the integrated luminosity.



(a) (b)

Figure 10: Ratio of the measured electron selection efficiency over the expected efficiency from Monte Carlo for (a) the simple cuts selections and (b) the category based selections. The values are shown for both the "tag-and-probe" method using Z→ee events and for the "tag-and-probe" and ML fit methods based on W events, as well as separately for the ECAL barrel and ECAL endcaps. Combined values of the efficiency and their errors are also indicated (yellow bands).

	ECAL barrel			ECAL endcaps		
Selection	Efficiency data	Error (stat.+syst.)	Efficiency MC	Efficiency data	Error (stat.+syst.)	Efficiency MC
WP95%	92.5%	3.2%	95.4%	86.4%	6.7%	92.9%
WP80%	77.5%	4.7%	85.1%	75.1%	8.6%	76.2%
Cic Loose	96.4%	2.1%	97.0%	94.1%	4.7%	95.3%
Cic SuperTight	89.3%	3.4%	89.3%	85.5%	6.5%	79.4%

Table 1: Electron selection efficiency from "tag-and-probe" using Z→ee events for electron candidates in the ECAL barrel and in the ECAL endcaps and for the two cut based and category based selections.

Figure 11 presents the distributions for tagged background electron candidates of some electron ID and isolation variables. The distributions shown are: the ϕ difference $\Delta\phi_{in}$ between the supercluster position and the track extrapolation from the innermost state, the supercluster η width $\sigma_{i\eta i\eta}$, the fraction $fbrem$ of radiated energy as measured by the tracker and the combined isolation sums using the ECAL, HCAL and the tracker. The distributions are shown separately for electron candidates in the ECAL barrel and in the ECAL endcaps. The Monte Carlo distributions are normalized to the number of background electron candidates in the data. We generally observe a good agreement between data and simulation, however, there is some shift to lower $\sigma_{i\eta i\eta}$ for data in the ECAL endcaps as well as some excess in the tail of $fbrem$ also in the ECAL endcaps.

7.2.5 Electron Fake Rate Measurement

Beyond the simple data-versus-simulation agreement between the rate of reconstructed electron candidates and the distributions of the variables used for electron ID, the determination of fake rates after the application of electron ID is a key performance measure for the detector.

The fake rate is defined as the fraction of reconstructed electron candidates that pass the back-

ground selection and a given electron ID over the total number of electron candidates passing the background selection.

Figure 12 shows the fake rate as a function of p_{T} and η for the simple selection and the WP95 and WP80 working point as well as for the categorized selection and the "Cic Loose" and "Cic SuperTight" working points. The fake rate is found to be in reasonable agreement between data and simulation for both the simple and the categorized electron selections.

8 Conclusion

Progress on the commissioning and performance of electron reconstruction and identification has been made using $\sim$200 nb^{-1} of 7 TeV data recorded during the spring 2010. The Level 1 and HLT efficiencies of electron/photon triggers have been measured from data. The main electron distributions from candidates in the minimum bias, from background electron candidates using jet triggers and from signal electron candidates from W events have been measured and compared to Monte Carlo simulation. The simulation is shown to be overall in good agreement with the data. The electron reconstruction, selection efficiencies and fake rate have been measured from data and found in good agreement with the expectation. The commissioning of electron reconstruction and selection will proceed making use of the large sample of tagged candidates from Z that will be available with integrated luminosities of order $\sim$1 pb^{-1}.

References

[1] CMS Collaboration, "Electromagnetic physics objects commissioning with first LHC data", *CMS PAS* **EGM-2010/001** (2010).

[2] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **0803** (2008). doi:10.1088/1748-0221/3/08/S08004.

[3] S. Baffioni et al., "Electron reconstruction in CMS", *Eur. Phys. J. C* **49** (2007), no. 3, 1099.

[4] CMS Collaboration, "Electromagnetic calorimeter commissioning and performance with 7 TeV data", *CMS PAS* **EGM-2010/002** (2010).

[5] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", *CMS PAS* **PFT-2010/001** (2010).

[6] T. Sjostrand, P. Eden, C. Friberg et al., "High-Energy-Physics Event Generation with PYTHIA 6.1", *Comput. Phys. Commun.* **135** (2001) 238.

[7] J. P. et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", *Journal of High Energy Physics* **2002** (2002). doi:10.1088/1126-6708/2002/07/012.

[8] S. Agostinelli, J. Allison, K. Amako et al., "G4–a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303. doi:10.1016/S0168-9002(03)01368-8.

[9] CMS Collaboration, "CMS physics : Technical Design Report", *CERN/LHCC* **CERN-LHCC-2006-001** (2006).

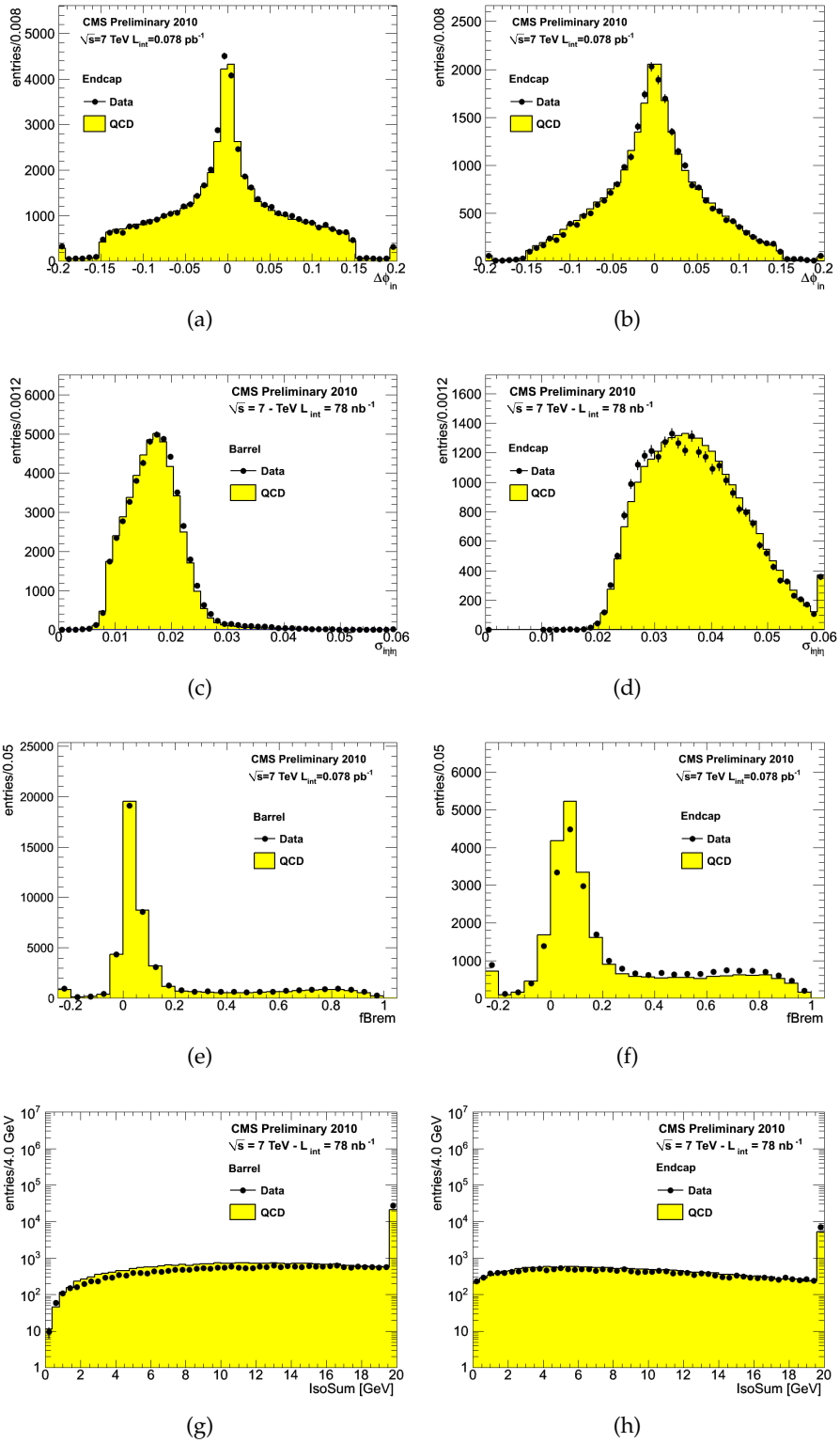


Figure 11: Distribution of electron ID variables for tagged background electron candidates from di-jet events in data (dots) compared with Monte Carlo (histograms): (a,b) $\Delta\phi_{in}$, (c,d) $\sigma_{i\eta i\eta}$, (e,f) fbrem and (g,h) isolation sum. Distributions are presented separately for (a,c,e,g) the ECAL barrel and (b,d,f,h) the ECAL endcaps.

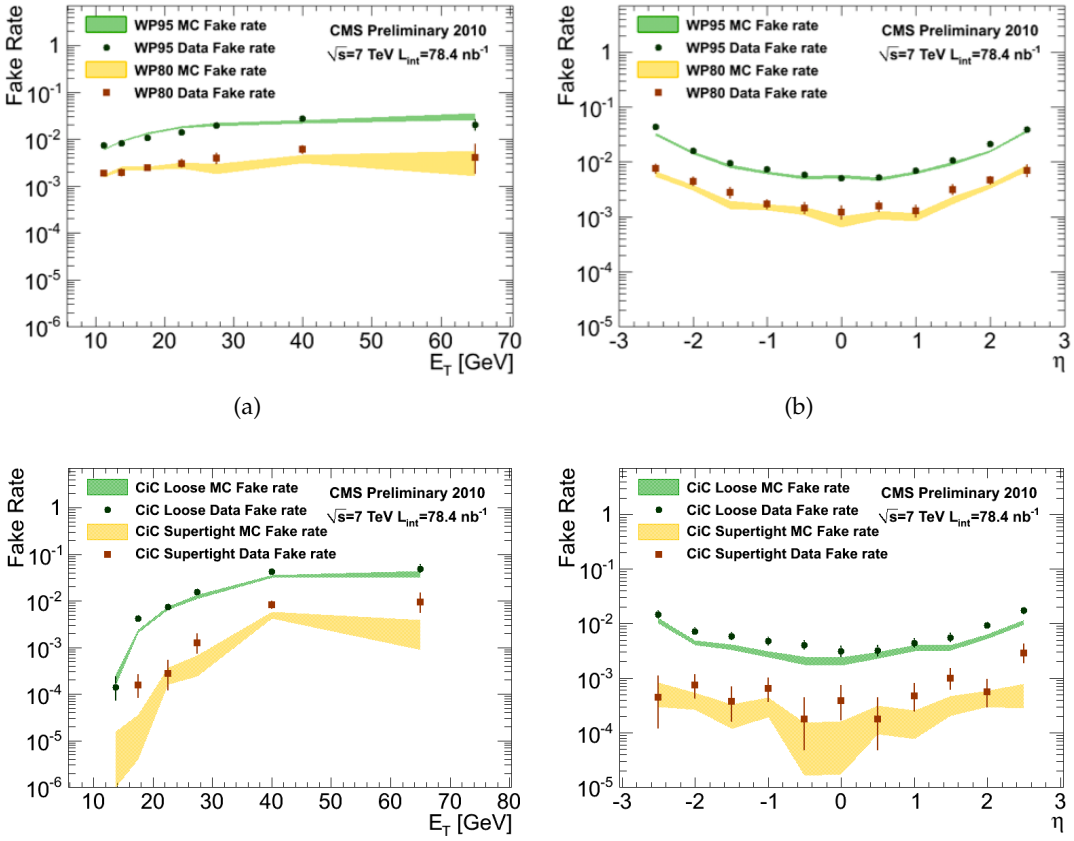


Figure 12: Electron fake rate per reconstructed electron candidate as a function of (a) E_{T} and (b) η for the WP95 and WP80 electron selections in data and Monte Carlo. The same quantity is shown for the "Cic Loose" and "Cic SuperTight" as a function of (c) E_{T} and (d) η.

[10] W. Adam, R. Fruhwirth, A. Strandlie, T. Todorov, "Reconstructions of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC", *J. Phys. G: Nucl. Part. Phys.* **31** (2005) N9–N20.

[11] CMS Collaboration, "Measuring Electron Efficiencies at CMS with Early Data", *CMS PAS EGM-07-001* **EGM-2007/001** (2008).

[12] F. R. L. D. M. Pivk, "sPlot: a statistical tool to unfold data distributions", *Nucl.Instrum.Meth.A* **555** (2005) 356–369. doi:10.1016/j.nima.2005.08.106.

A Electron-ID Variables from sPlot Technique

For the W ML fit technique we also checked the distribution of the identification variables using the "sPlots" technique [12] and compared them to the Monte Carlo expectations as shown in Fig. 13.

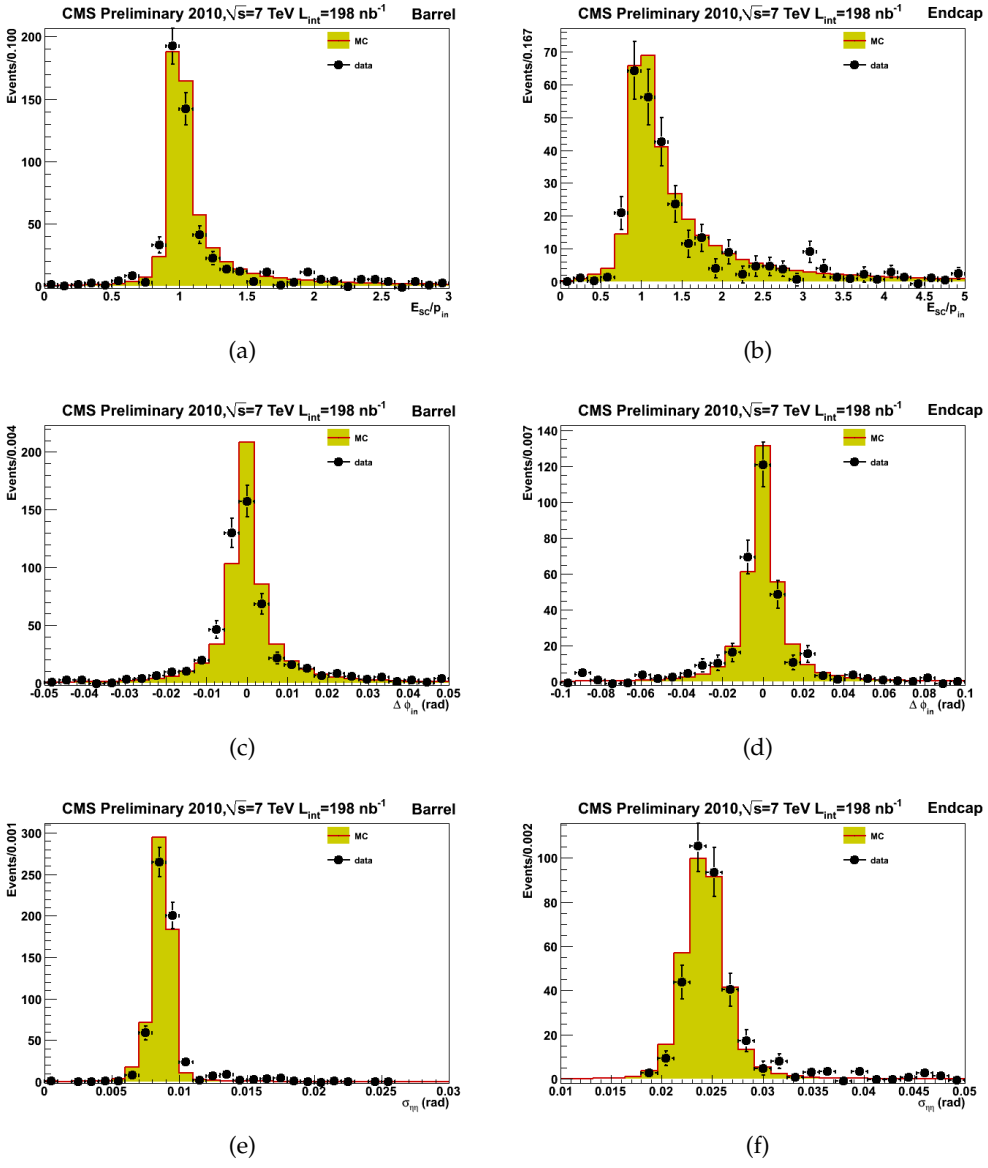


Figure 13: Distribution of electron ID variables for signal electron candidates in data as extracted using the sPlot technique (dots) compared with Monte Carlo (histograms): (a,b) $\Delta\phi_{in}$, (c,d) $\sigma_{i\eta i\eta}$, (e,f) fbrem and (g,h) isolation sum. Distributions are presented separately for (a,c,e,g) the ECAL barrel and (b,d,f,h) the ECAL endcaps.