MEASUREMENT OF THE FLAVOUR TAGGING EFFICIENCY ON A W+C-JET SAMPLE AT 7 TEV FOR THE ATLAS EXPERIMENT

The efficiency to mistakenly tag a c-jet as a b-jet is measured on a sample of c-jets produced in association with a W boson, with 2011 data collected by the ATLAS detector. The new b-tagging algorithms developed for Run 2 are designed to produce an improvement of the c-jet rejection, with respect to Run 1 performance. The same event selection strategy will be used to measure the W+c production cross section at 13 TeV, with 2015 and 2016 ATLAS datasets.

ATLAS experiment

ATLAS is a multi purpose experiment located at the CERN Large Hadron Collider. The detector is composed of three different sub-detectors: the Inner Detector, the Calorimeters and the Muon Spectrometer.

The correct identification of jets containing b-hadrons (hence b-tagging) is of capital importance for the hadron collider experiments. Measurements that benefit from b-tagging include Standard Model analyses, like Higgs and top quark physics, as well as searches for new physics phenomena beyond the Standard Model.
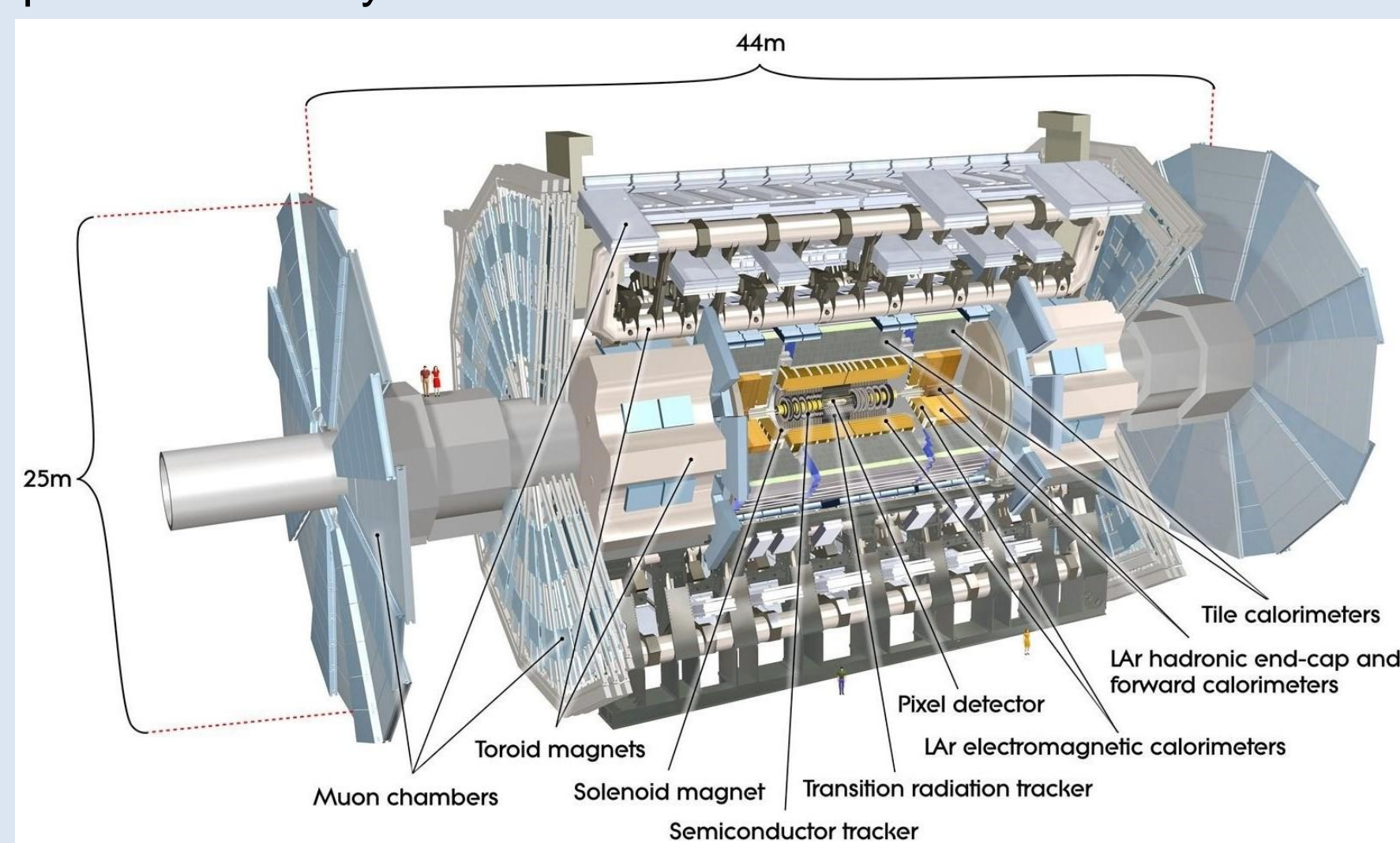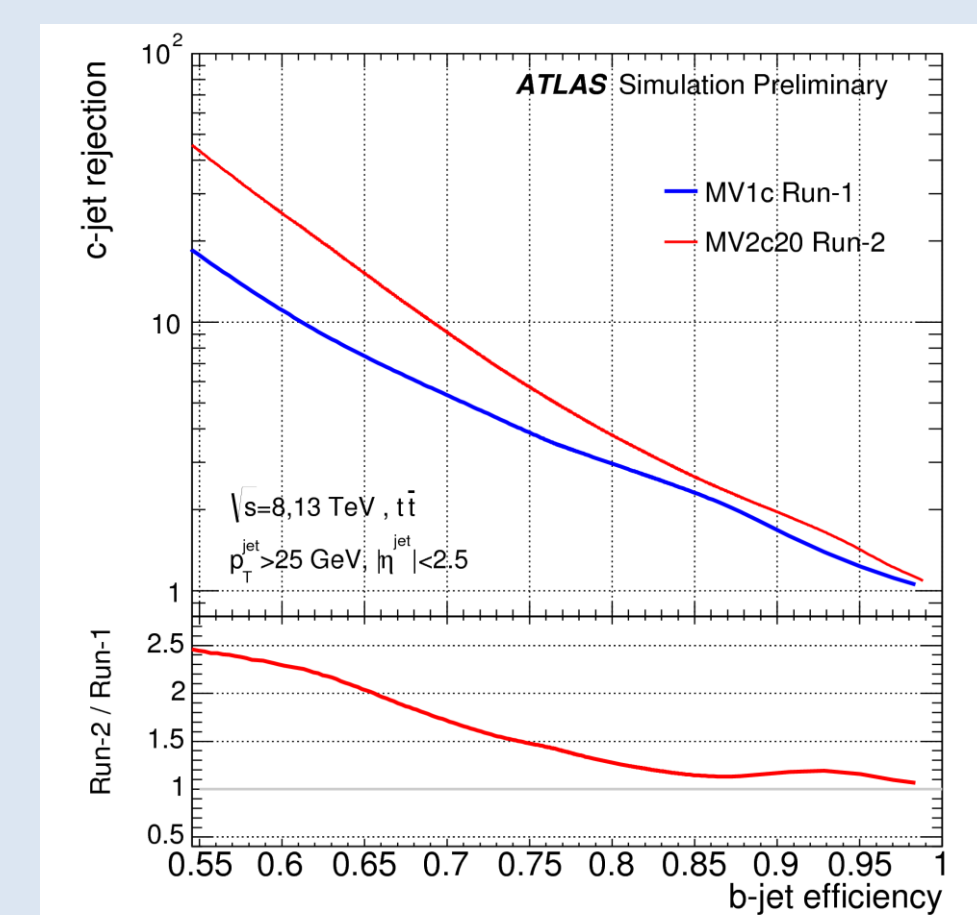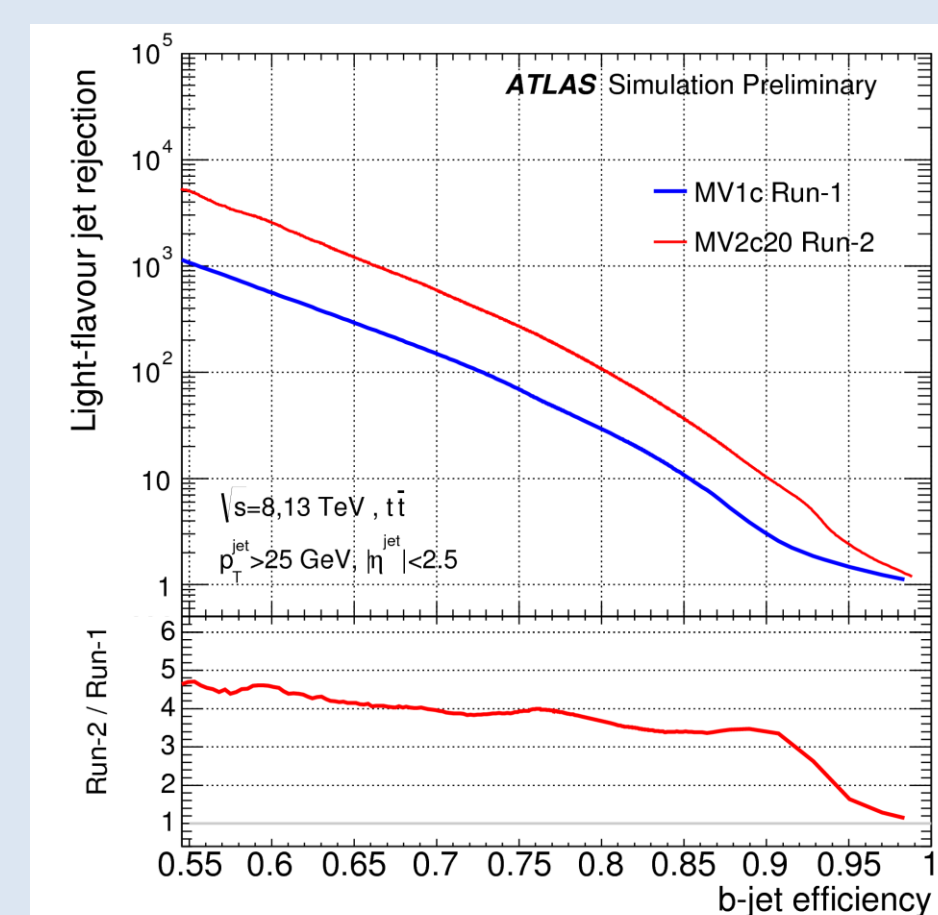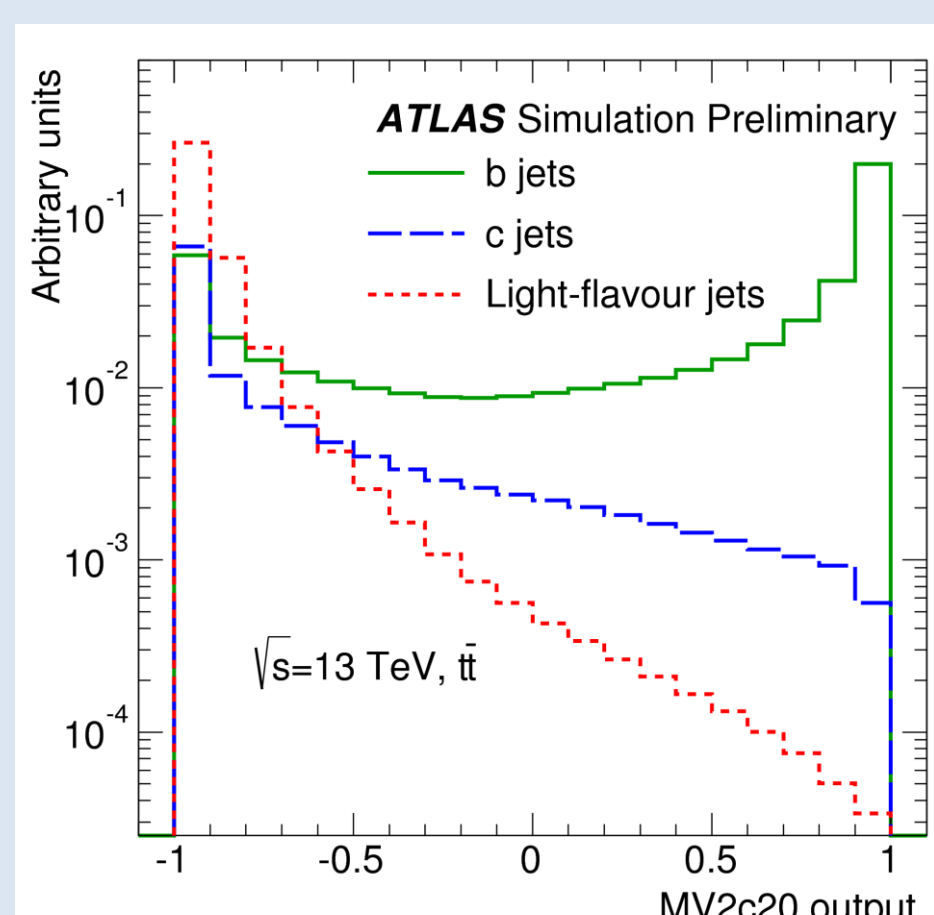

b-tagging performance

The identification of b-jets is performed with several algorithms, exploiting the long lifetime, high mass and high decay multiplicity of b-hadrons. Due to similar properties, c-jets are easily misidentified as b-jets.

Run 2 b-tagging discriminants (MV2) are produced combining a set of sensitive variables with multivariate techniques. The performance for several background mixtures of c-jets and light jets in the training sample has been evaluated in order to improve light and c-jet rejection [1, 2].

Performance is first measured on ttbar simulated samples, then the efficiencies are calibrated on data, measuring data-to-MC scale factors for different b-tagging working points.

MV2c20 discriminant variable, light and c-jet rejections versus b-jet efficiency are compared to the performance of the Run 1 b-tagging algorithm (MV1). The improvements are mainly due to the introduction of an extra pixel layer in the detector (IBL) and to enhancements in the b-tagging and tracking algorithms [1]. The c-jet rejection is expected to improve even more in 2016 (a factor of 40% for 77% b-tagging efficiency), thanks to further enhancements achieved in the b-tagging performance of the new variable MV2c10 [2].

Analysis Strategy

To perform the measurement of the b-tagging efficiency on c-jets, a high purity c-jet sample is needed. The analysis is performed on a sample of c-jets produced in association with a W boson.

The W boson is reconstructed via its decay into an electron and a neutrino, and the c-jet is identified via a soft muon stemming from the semileptonic c-hadron decay.

The c-jet sample is extracted exploiting the electron-muon charge correlation, as the difference between events with opposite sign (OS) and same sign (SS) charged leptons.

$$N^{OS-SS} = N^{OS} - N^{SS}$$

Event Selection:

- single electron triggers
- 1 electron: $p_T > 25$ GeV, $|\eta| < 2.47$
- neutrino: missing $E_T > 25$ GeV
- W boson: $m_T > 40$ GeV
- 1 jet : $p_T > 25$ GeV, $|\eta| < 2.5$
- 1 muon: $p_T > 4$ GeV, $|\eta| < 2.5$
- ΔR(muon, jet) < 0.5

Sample composition

Due to the associated production, the W+c signal has mainly OS leptons, while backgrounds almost evenly populate the OS and SS lepton samples.

The main backgrounds are W+light jet and multijet, both estimated with data-driven methods.

After the OS-SS subtraction, the remaining sample is mainly composed of W+c-jet events (~90%), with a small contamination of W+light and multijet events [3].
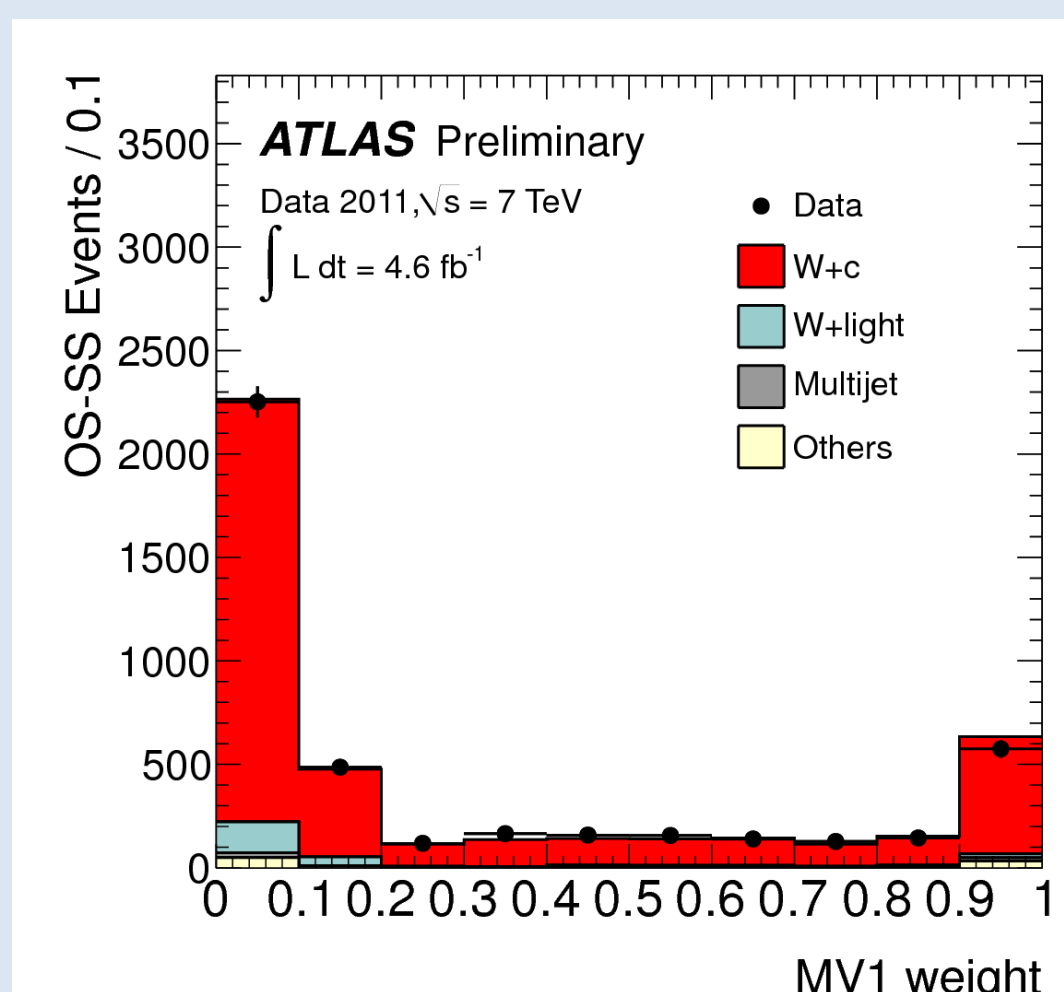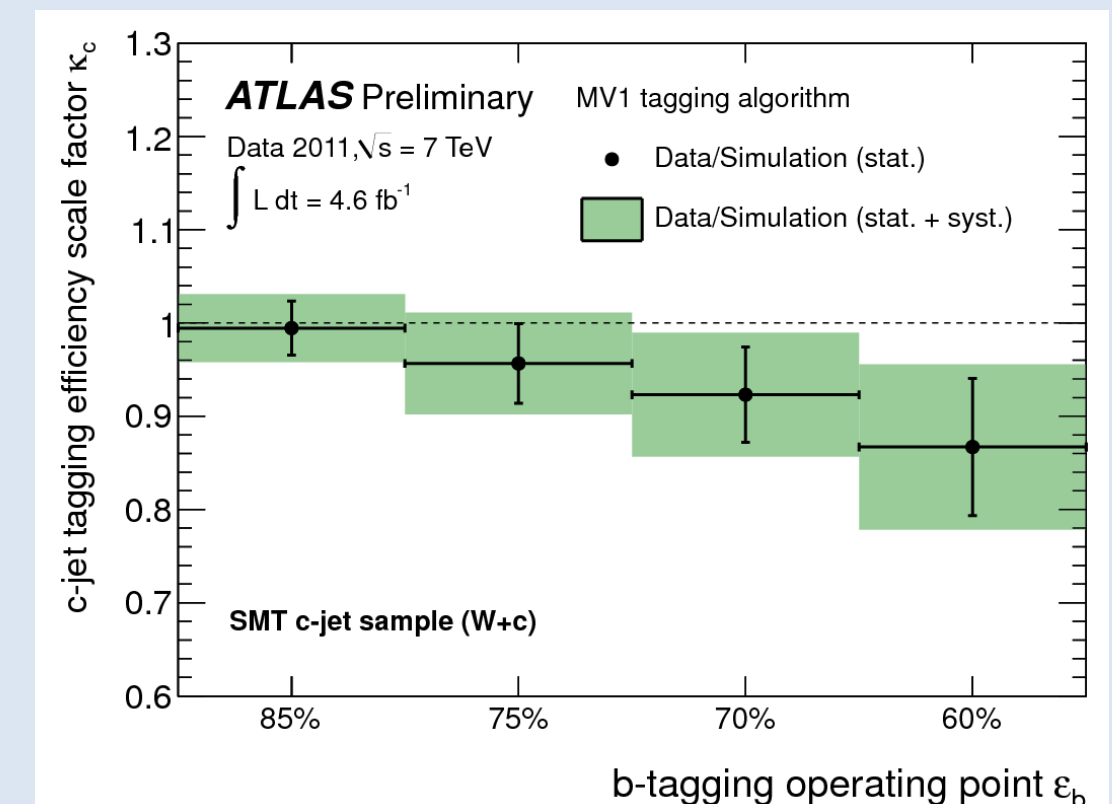

b-tagging efficiency measurement

The efficiency, defined as the fraction of W+c signal events passing the b-tagging requirement, is measured separately for data and MC. The data efficiency is measured taking into account the possibility to tag background events, while the MC efficiency is measured directly on W+c signal events.

The data-to-MC scale factors are calculated as the ratio of data to MC efficiencies, for each of the b-tagging variable working points.

$$\varepsilon(\text{Data}) = \frac{N^{OS-SS}_{tag} - bkg^{OS-SS}_{tag}}{N^{OS-SS}_{tot} - bkg^{OS-SS}_{tot}} \qquad \varepsilon(\text{MC}) = \frac{sig^{OS-SS}_{tag}}{sig^{OS-SS}_{tot}} \qquad SF = \frac{\varepsilon(\text{Data})}{\varepsilon(\text{MC})}$$

Extrapolation and results

The tagging efficiencies are measured on a c-jet sample with the c-jet decaying semi-leptonically. Since the selection of semi-leptonic decaying c-jet has an effect on the probability to select different types of c-hadrons, the c-hadron composition of this sample differs slightly with respect to an inclusive c-jet sample. Extrapolation coefficients are determined for data and MC, in order to extrapolate the results to be valid for a general sample of c-jets.

First results from this method applied to 13 TeV data are expected soon!

References

1) Expected Performance of ATLAS b-Tagging Algorithms in Run II
https://cds.cern.ch/record/2032459 ATL-PHYS-PUB-2015-022
2) Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run
https://cds.cern.ch/record/2160731 ATL-PHYS-PUB-2016-012
3) Calibration of the b-tagging efficiency for c jets with the ATLAS detector using events with a W boson produced in association with a single c quark
https://cds.cern.ch/record/1640162 ATLAS-CONF-2013-109