Automatic Monte Carlo Production for Imaging Air Cherenkov Telescopes

Daniela Dorner*$^{\|}$, Thomas Bretz and Marlene Doert^{x}**

ISDC Data Centre for Astrophysics, University of Geneva, 1290 Versoix, Switzerland
$^{\|}$*Institute for Particle Physics, ETH Zürich, 8093 Zürich, Switzerland*
**Universität Würzburg, 97074 Würzburg, Germany*
^{x}*TU Dortmund University, 44221 Dortmund, Germany*

Abstract. **Imaging Air Cherenkov Telescopes need Monte Carlo simulations to reconstruct the energy of the primary particle of the recorded shower event, as they lack an absolute calibration in energy. These Monte Carlo simulations have to take into account not only changes in the detector design and performance, but also variations in the atmospherical conditions.**

Especially for long-term monitoring, as it is planned with the emerging DWARF project, Monte Carlos well describing these effects on the data are mandatory. Reducing the systematics on the results, conclusions on the physics of the sources can be drawn from light curves spanning several years of data.

Using the flexible automation concept, developed for the analysis of MAGIC data, the automatic production of Monte Carlos can be set up. For the DWARF project, this is not only used for later production on demand, when the first data will arrive, but also for the design studies for which it allows the users to scan a large parameter space of different setups.

In addition, the automatic Monte Carlo production can be used for other current and future Imaging Air Cherenkov Telescopes like MAGIC or CTA.

Keywords: **automation, monte carlo simulation, Cherenkov telescopes**

I. INTRODUCTION

In the past years, Cherenkov telescopes of the current generation have discovered many new sources. With its special design, the **M**ajor **A**tmospheric **G**amma-ray **I**maging **C**herenkov telescope (MAGIC) has reached energies below 100 GeV [1]. From lower energies, Fermi is now closing the gap that has been left by EGRET and the past generation of Cherenkov telescopes. Future projects like the **C**herenkov **T**elescope **A**rray (CTA) aim for a better sensitivity and a larger energy range. Given the larger number of sources and therefore broader interest in astronomy, the future generation of Cherenkov telescopes is operated as observatories and opened to a wider community. An open data policy is also the aim of the **D**edicated multi**W**avelength **A**gn **R**esearch **F**acility (DWARF), a low-cost project for long-term monitoring of AGNs bright in TeV energies. The project starts with a small telescope on the Canary Island of La Palma [2], but aims at 24-hour coverage by establishing similar small telescopes around the globe [3].

For all current and future Cherenkov telescopes, Monte Carlo (MC) simulations are mandatory, to reconstruct the energy spectrum of the source.

II. IMAGING AIR CHERENKOV TECHNIQUE

In gamma-ray astronomy, there are two approaches: satellites for direct measurement from space and Imaging Air Cherenkov Telescopes (IACTs) for indirect measurements from ground. Due to the power law spectra of most sources, the first is confined by too low statistics at higher energies as the detector volume is limited in space. For the ground-based observations, the key issue is the interaction of gammas in the atmosphere which also constrains the energy range for the indirect measurement. Consequently, low energies are observed from space with satellites, while for higher energies ground-based indirect measurements are performed.

Particles interacting in the atmosphere generate a cascade of secondary particles which emit Cherenkov light when traveling faster than the speed of light in the medium. Reaching ground, the Cherenkov light is reflected into photon detectors. With a pixelized camera, an image of the shower is recorded. As not only gammas, but also protons, muon, electrons and heavy nuclei produce particle showers in the atmosphere, Cherenkov telescopes have to face a large background of about 1000 background events versus one gamma event from the source. Sensitivity and energy threshold of a Cherenkov telescope depend on several instrumental parameters, e.g. the light collection efficiency or the field of view, but also on the observation conditions, e.g. observation mode or zenith angle. A statistical analysis of the images provides various parameters describing the images [4], [5]. Based on this, the source direction and the particle type and its energy are reconstructed. Lacking an absolute calibration of the detector, the Imaging Air Cherenkov Technique (IACT) needs simulated data for this reconstruction. By comparing the real with the simulated events, the energy of the primary particle can be reconstructed. Either by parameterizing the energy with the calculated image parameters or by training a statistical learning method like Random Forest [6] with

the simulated data. Also for suppressing the background, MC events are used either to develop or train cuts or also using a statistical learning method.

III. SIMULATION

The simulation consists of several steps. First the development of the shower in the atmosphere and the emitted Cherenkov photons have to be simulated. There are different packages available. CORSIKA [7] is the one mainly used for Cherenkov telescopes. It provides several electro-magnetic and hadronic interaction models for different energy ranges. Within the CTA design study, other packages, like KASKADE or Geant4 [8], are tested and compared with CORSIKA.

A crucial part for the shower simulation is the atmosphere, as it is part of the detector for IACTs. Weather conditions or meteorological effects, like e.g. calima [9], can affect the data and therefore have to be simulated properly. If the effect on the data is known like in the case of calima, the data can be corrected as described in [10]. However, the method is limited at very low energies where only proper simulations can help.

Next part of the simulation is the detector, where the mirror, the photon detectors, and the complete electronic chain have to be simulated. For this, the different experiments are using different software packages.

For DWARF, a flexible simulation framework, called ceres, has been implemented in the software package *Mars - CheObs ed.* [11]. Currently the design study of the project is on-going and several instrumental configurations are set up in ceres and compared.

In MAGIC, the detector simulation is split in two programs, one simulating the mirror, the other the camera and the electronics behind. In ceres, also the MAGIC setup is implemented to provide an independent cross-check of the program. In CTA, the H.E.S.S. simulation packages [12] are used for the design study. For the analysis, the MC are treated like the real data and the various software packages are used.

IV. REQUIREMENTS

Taking into account the different projects and the programs used therein, an automation concept has to handle not only different programs, setups and production and analysis chains, but also different computing resources at the various sites.

Apart from that, there are different purposes of the MC production. While MAGIC is operational and needs the MCs mainly for the analysis, DWARF and CTA are in the design and building phase. For a design study, events have to be simulated for various telescope configurations. The output should be easy manageable and comparable. In the case of DWARF, not only production of MC for the design study are needed. As soon as the operation of the telescope starts, proper MC have to be produced on demand to ensure a consistent analysis of the long-term monitoring. Studying bright AGNs, which are variable sources, MC taking into account all

technical components and observational conditions, are vital.

Like real data, also MC events have to be analyzed and the output has to be processed in the further analysis or the design study.

Taking care of all these different requirements and setups cannot be done manually. An automatic production and processing of the MC data is mandatory. In addition, it leaves the physicists the time to concentrate on the interpretation of the results.

V. AUTOMATION CONCEPT

For the analysis of MAGIC data, an automation concept has been developed [13], [14]. Slightly enhanced, it is now also used for the production and processing of simulated data.

The core part is a database which stores the processing status of the data. The interaction with the software is done with scripts. This setup is completed by a cron job and serval root macros interacting with the database. Apart from the status, the database also stores information on the data, the analysis and the results, and in case of the MC, in addition the simulation setup. For this, the structure of the database is adapted for the various experiments. While for real data the initial information comes from the telescope, this is different for the MC. Here the desired parameters have to be inserted into the database to start the production. For this initial insert, two possibilities are considered in case of DWARF: On the one hand, the user wants to simulate a certain setup for a study. On the other hand, the data recorded with the telescope have certain parameters, conditions and setup. This information has to be extracted from the data and the output of the auxiliary systems. Evaluating this information, the input of the MC database has to be created. As soon as the database is filled, the production is started automatically.

A. Database

The key part of the database, for steering production and processing, are the 'status tables'. Each step of the analysis has one column in one of the status tables. By inserting a time stamp, the successful completion of the step is marked. In case of failure, not only start and stop time, but also return and failure codes are inserted. Like this, the status of each step can be monitored.

B. Dependencies

The dependencies between the steps are stored in a setup file. This ensures that the steps are performed in the correct order.

C. Scripts

Using the dependencies setup file and the status tables, various scripts are running the production and analysis. They query from the database for which run or data set a certain step has to be done, execute the step and insert the new status into the database.

VI. FEATURES

Thanks to its setup with status tables, re-running steps can be easily done by simply resetting the according columns in the database. This can be useful in case of changing hardware or technical problems which have to be fixed in the analysis. When the software is ready to cope with the new or problematic data, these runs are simply reseted in the database. The same applies in case of a new software version, e.g. with improved analysis methods.

The setup can also be copied easily, allowing for both testing a new method on all or a subset of data with a second setup and running two setups in parallel. For example when a new method for background suppression is developed, its suppression power can be tested on all data. With high statistics, a comparison with the standard method can be provided. In case a good method is found, there is also the possibility to run two methods in parallel to have both results available. Doing the systematic study for the implementation of the timing information in the image cleaning and the background suppression [16], such a test setup has been used.

Due to its flexibility, adding a new step in the analysis chain is simple: the according column in the database has to be created, the dependencies have to be added in the setup file, and a script to execute the step has to be written, which is mainly copy and paste in most cases. Like this, not only enhancing the analysis with new steps or methods is easy, but also a setup for automatic processing for a new experiment can be done with few time and effort.

To the user, the concept provides web interfaces to monitor the status of the production and analysis, including information about failed or crashed processes. In addition, various tools make the other content of the database available, i.e. information about the data, the MC and their analyses. For data quality assurance, a lot of useful web interfaces and plotting tools are available enabling the user not only to check the quality of the analysis results, but also to control the long-term stability of the detector.

VII. ACHIEVEMENTS

At ISDC in Switzerland, the automation concept is used for the production of CTA MC. For the design study of the project, simulated data for different telescopes and configurations are needed. With an application of the concept, MC are produced and analyzed automatically.

In Dortmund, it is used to produce MC for MAGIC. For this, not only the single telescope and the stereo system, but also many other different parameters are taken into account. As MAGIC has been upgraded continuously, different telescope hardware is simulated for the different sets of data. For the different data sets, different observation modes, different atmospheric conditions and detector performance are included in the simulations. For tailored MC sets, a trajectory option has been implemented in CORSIKA to produce simulated

data with the fitting zenith angle and azimuth distribution.

In MAGIC, the concept is used to automatically analyst data. In the data center in Würzburg, the setup has been running stable since about five years. Since last year, it is not only used to process data, but also to produce and process MC for the design study of DWARF, for which CORSIKA and ceres are used. For MAGIC, also the processing of the simulated data is included in the chain. The MC are analyzed using the same programs and settings like for the data with the software package *Mars - CheObs ed.* [15].

Being used in three different sites, the package supports not only condor, but also sun grid engine and pbs as queuing systems.

The layout of the database is adapted to the special needs of each experiment and the purpose of the MC production. For the different projects, web sites to check the production and analysis status and results are provided. Within the MAGIC project, this includes also many tools for data quality assurance helping the user to analyze the data.

VIII. CONCLUSION AND OUTLOOK

The automation concept is in operation successfully since about five years. Its application has been enhanced: Apart from data analysis, it is now also spanning MC production and processing. In addition, new experiments and purposes are implemented, and the system is applied in different places.

For the future, further enhancements of the features for MC users are planned. Main issue here is the *production on demand*. This includes on the one hand design or special studies for which users have a special request for MCs to be produced. Such needs may be fulfilled by a web interface from which the simulation parameters can be selected. On the other hand, production on demand for future DWARF data will be implemented. In this case the requirements for the simulation setup comes from the observations themselves. Data quality, telescope performance and observation conditions are evaluated and translated into simulation parameters. Like this proper MC can be produced for each set of data.

REFERENCES

[1] E. Aliu et al., *Observation of Pulsed γ-Rays Above 25 GeV from the Crab Pulsar with MAGIC*", Science, 2008
[2] T. Bretz et al., *Status of the DWARF project for long-term monitoring of bright blazars*, these proceedings
[3] M. Backes et al., *Long-term monitoring of blazars - the DWARF network*, these proceedings
[4] A.M. Hillas, *Cerenkov light images of EAS produced by primary gamma*, ICRC1985
[5] B. Riegel, *A systematic study of the interdependence of IACT image parameters*, ICRC2005
[6] J. Albert et al., *Implementation of the Random Forest method for the Imaging Atmospheric Cherenkov Telescope MAGIC*, Nuclear Instruments and Methods in Physics Research A, April 2008
[7] D. Heck et al., *EAS Simulation with CORSIKA: A Users Manual (Version 6.023)*, 2004
[8] http://www.cern.ch/geant4

[9] J. P. Dunion and Velden C.S.,*The Impact of the Saharan Air Layer on Atlantic Tropical Cyclone Activity*, 2004

[10] D. Dorner et al., *A method to correct IACT data for atmospheric absorption due to the Saharan Air Layer*, A&A 2009

[11] T. Bretz and D. Dorner, *MARS-The Cherenkov Observatory edition*, American Institute of Physics Conference Series, 2008

[12] K. Bernlöhr, *Simulation of imaging atmospheric Cherenkov telescopes with CORSIKA and sim_telarray*, Astroparticle Physics October 2008

[13] T. Bretz and D. Dorner, *A concept for data managment and processing*, Proceedings of Towards a network of atmospheric Cherenkov detectors VII, 2005

[14] D. Dorner and T. Bretz, *Data Managment and Processing for the MAGIC Telescope*, Proceedings of Towards a network of atmospheric Cherenkov detectors VII, 2005

[15] T. Bretz and D. Dorner, *MARS - CheObs goes Monte Carlo*, these proceedings

[16] E. Aliu et al.,*Improving the performance of the single-dish Cherenkov telescope MAGIC through the use of signal timing*, Astroparticle Physics, 2009