Overview of simulation efforts in CTA

Michal Vraštil

Olomouc 2016

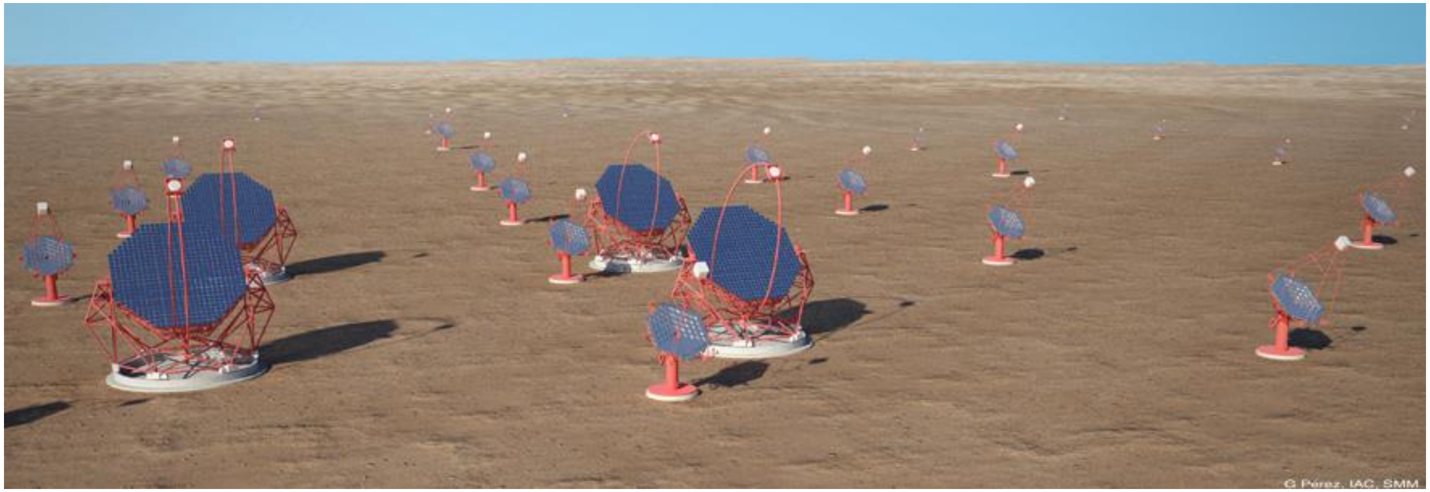
using materials of Maier, G., Arrabito L., Di Pierro F.

Outline

- **CTA Overview**
- **Monte Carlo Simulations**
  - use of MC simulations in CTA
  - MC pipeline
  - queries by MC
- **CTA computing resources**
  - current status of CTA computing Grid
  - organization process for official productions
- **Atmospheric Simulations**
  - goals of the simulation
  - studied extreme atmospheric profiles
  - results in ideal case and in extreme case

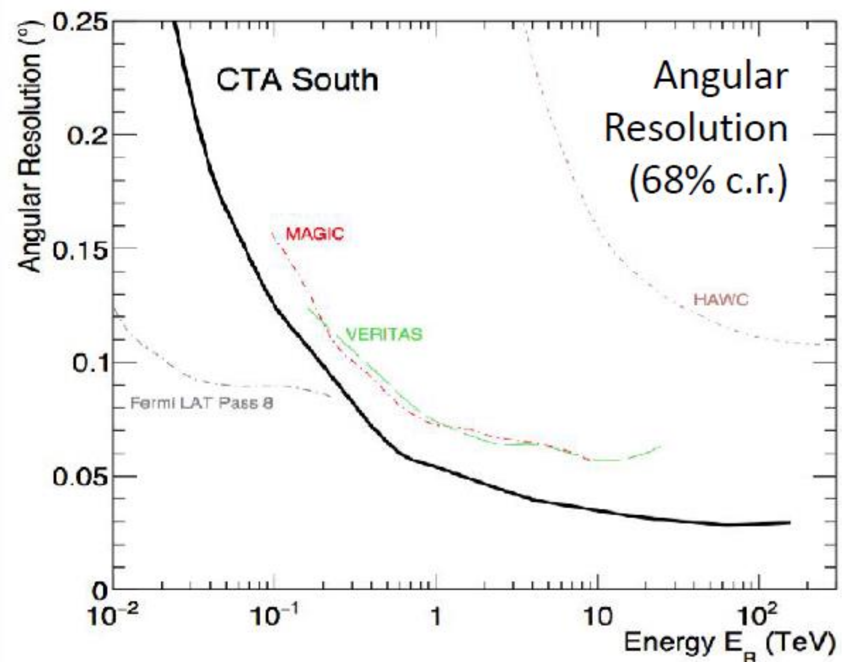
CTA Overview

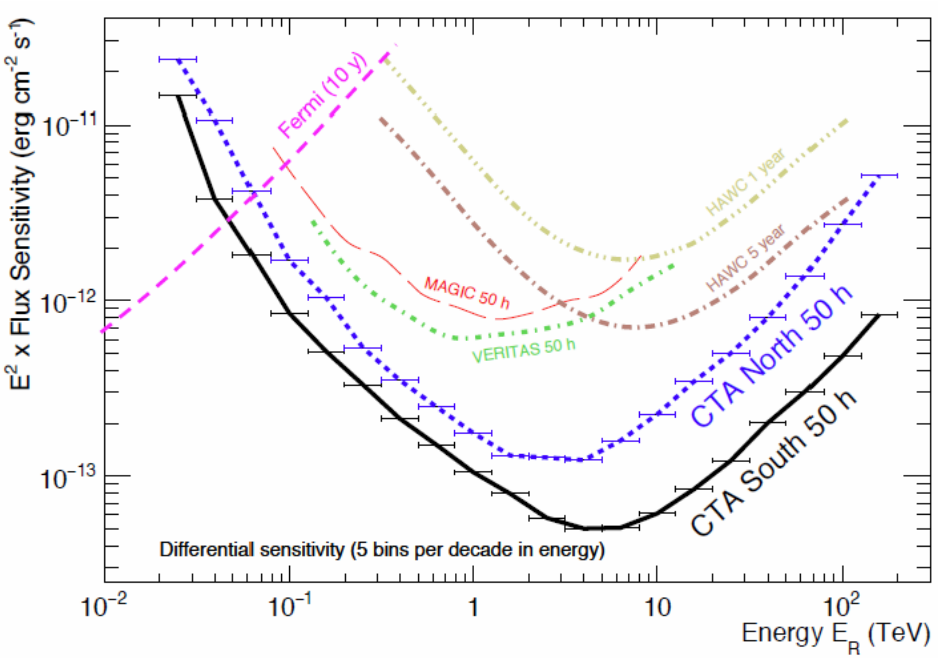
- Observatory for ground-based gamma-ray astronomy
  - imaging atmospheric Cherenkov telescope (IACT)
- Arrays in both hemispheres for full sky coverage
  - ESO, Paranal, Chile in the south; ORM, La Palma, Spain in the north
  - arrays consist of large-, medium- and small-size telesopes
- Key science drivers
  - Understanding the origin of cosmic rays and their role in the Universe
  - Understanding the nature and variety of particle acceleration around black holes
  - Searching for the ultimate nature of matter and physics beyond the Standard Model



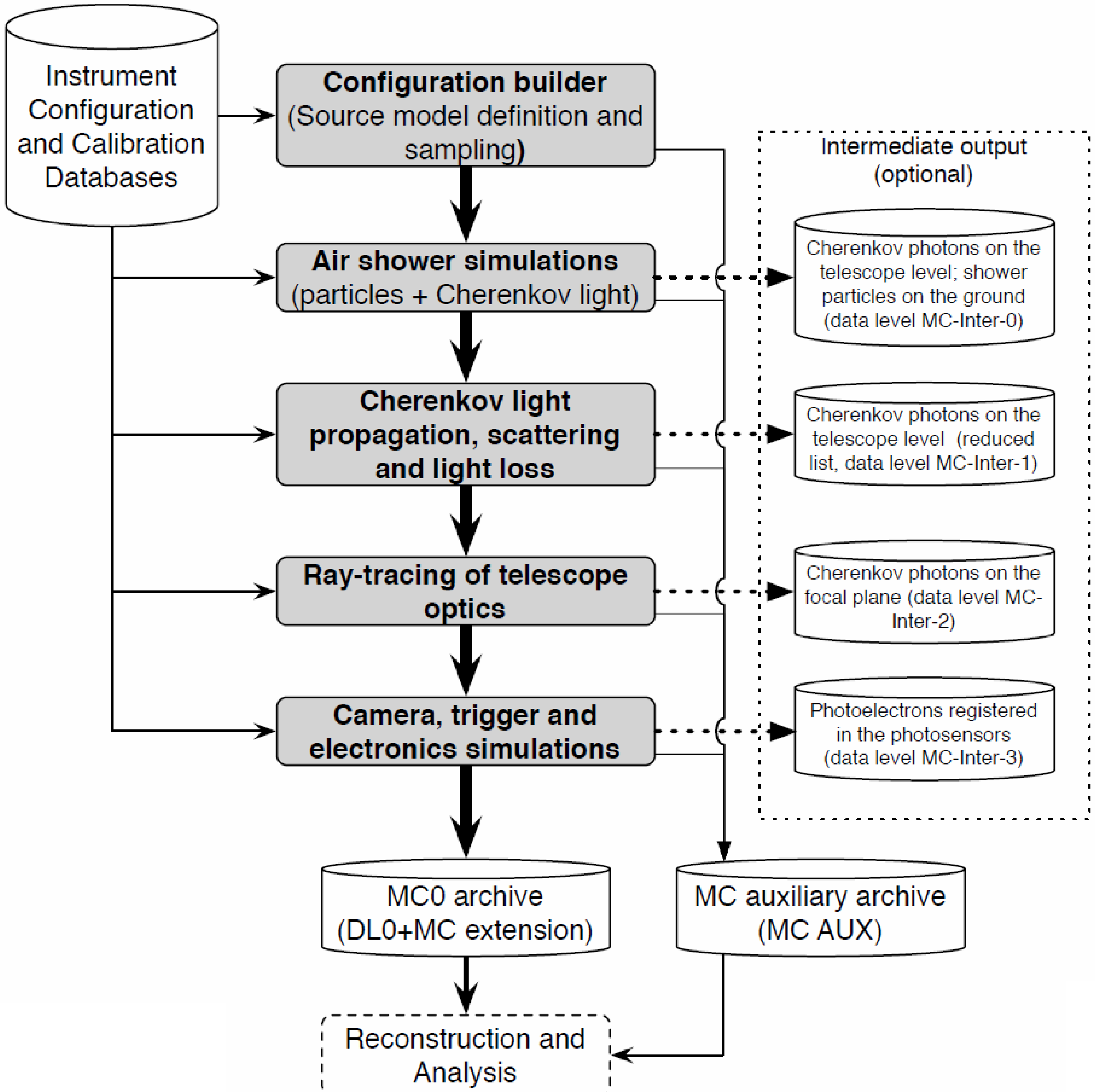
Use of MC simulations in CTA

- performance characterization of CTA
- almost any science analysis in CTA (instrument response functions)
- optimisation and validation of system parameters during prototyping and (pre-) construction phase
- determination of upgrade scenarios
- development of reconstruction algorithms
- evaluation of systematic uncertainties





MC Pipeline



Instrument Configuration and Calibration Databases
Configuration builder (Source model definition and sampling)
Air shower simulations (particles + Cherenkov light)
Cherenkov light propagation, scattering and light loss
Ray-tracing of telescope optics
Camera, trigger and electronics simulations
Intermediate output (optional)
Cherenkov photons on the telescope level; shower particles on the ground (data level MC-Inter-0)
Cherenkov photons on the telescope level (reduced list, data level MC-Inter-1)
Cherenkov photons on the focal plane (data level MC-Inter-2)
Photoelectrons registered in the photosensors (data level MC-Inter-3)
MC0 archive (DL0+MC extension)
MC auxiliary archive (MC AUX)
Reconstruction and Analysis

Analysis Types

- **Level A**
  - on-site, real-time analysis
  - no requirements on systematic uncertainty
  - simplified analysis using time-averaged instrument response functions (IRFs)
- **Level B**
  - on-site, next-day analysis
  - no requirements on systematic uncertainty
  - sophisticated analysis using time-averaged IRFs
- **Level C**
  - off-site analysis
  - 10% requirement on systematic uncertainty of the energy scale
  - most sophisticated analysis using period-wise IRFs

Time-averaged vs period-wise IRFs

Time-averaged ('classical') IRFs

- generate full MC sets for each point in a large phase space
- any significant change in the instrument (e.g. trigger settings, HV changes) requires a new MC set
- any significant change in the atmospheric conditions requires a new MC set
- required for Level A, Level B, basic analysis of all CTA observations

Period-wise IRFs

- period-wise: produce MC sets for certain observation periods
- requires a reasonable quick turnaround of all calibration / validation steps
- only done for regions of interests where the systematic uncertainty requirement is applicable
- needs to be a fully (semi?) automatic process
- by definition much closer to reality than full phase-space approach – no need for data correction!

Typical queries by MC group

- what is the atmospheric density profile for 2019, May 5th, 10:00 UTC at CTA South?
- what is the average (mean + RMS) atmospheric density profile for 2019, May at CTA South?
- what is the throughput correction for each individual telescope for 2019, May 5th relative to 2019, April 1st?
- expect to get one single value for each quantity, even if there are several measurements exists: CCF has the expertise to determine best values
- **expect to get for every quantity:**
  mean value + statistical uncertainty + systematic uncertainty

MC simulation pipeline software

- CTA MC simulation pipeline software is in development
  - technical development of data model / configuration reader and writer
  - compilation of all knowledge about CTA from the instrument teams and CCFs
- data production
  - CORSIKA (shower simulation) => sim_telarray (telescopes response)
  - CORSIKA => GrOPTICS (ray-tracing) => CARE (camera and readout)
- reconstruction of events
  - analysis of data (calibration, FADC trace integration, image and stereo parameter analysis, response function calculation)
- productions run on GRID

CTA Computing Grid (CTACG)

- **CTA Computing Grid**
  - use of EGI grid throughout the CTA Virtual Organization (since 2008)
  - use of DIRAC to access grid resources (since 2011)

- **CTA Virtual Organization**
  - open to any CTA member having a grid certificate
  - supported by 20 EGI sites in 7 countries
    + 1 ARC site in Sweden
  - eventual new OSG (US) resources in future
- **DIRAC for CTA**
  - Workload and Data Management System
  - dedicated server instance at CC-IN2P3, PIC and DESY
  - CTA-DIRAC software extension

CTACG resources

- CPU: 6000 – 8000 cores available on average
- about 2.2 PB (+ 0.7 PB of tape) distributed among 6 main sites
- additional 50 TB at Frascati and Torino for specific studies
- disk fills rapidly during MC massive productions

Site	Available Disk (TB)	Used Disk (TB)	Total Disk (TB)
CYFRONET-LCG2	16	627	643
DESY-ZEUTHEN	7	648	655
IN2P3-CC	105	249	354
GRIF (LPNHE+CEA)	17 (7+10)	182 (112+70)	200 (120+80)
IN2P3-LAPP	29	89	118
INFN-T1	114	172	286
Total	288	1967 (87%)	2255

by Luisa Arrabito

Organization process

- **Computing Model**
  - MC productions are run everywhere
  - output data are stored at 6 main Storage Elements
  - MC analysis is run at restricted number of sites (for efficiency reasons)
- **MC group**
  - provides software for production and analysis
  - runs production and analysis on local clusters and on the grid
  - produce final results
- **Production team**
  - ports MC software into DIRAC workflows
  - runs massive productions on the grid and low-level stages of analysis

Atmospheric Simulations

by Federico Di Pierro

Goals of Atmospheric Simulations

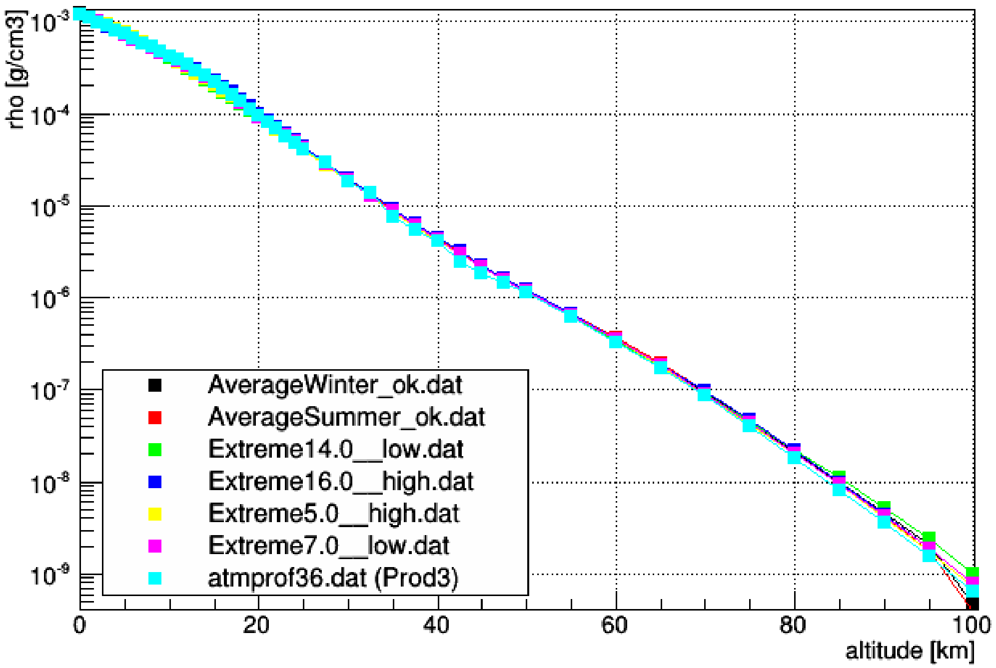
- atmosphere has a strong influence on the data
  - the air-shower development, the variation of the Cherenkov angle with altitude, the loss of photons due to scattering and absorption of Cherenkov light out of the camera field-of-view, resulting in dimmer images and the (multiple-)scattering of photons back into the camera, resulting in blurred images
- we need to estimate:
  - the effect of different atmospheric profiles on CTA performance
  - the effect of aerosols (dust, clouds) under different conditions
  - the effect of atmospheric calibration uncertainties on reconstructed energy, flux and pointing uncertainties
  - the possible strategy for generating observation-wise MC simulations

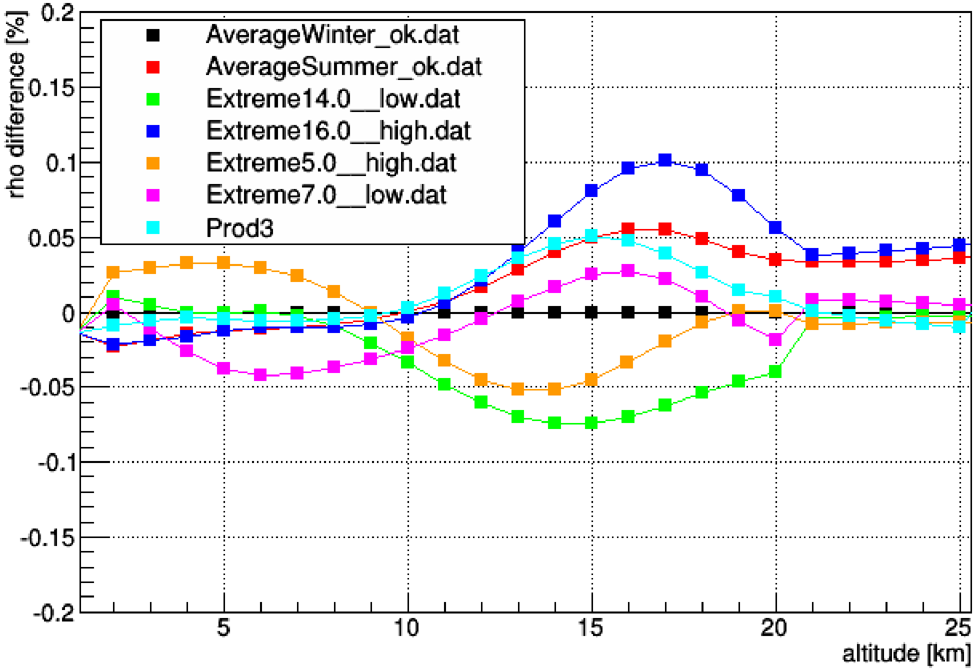
Extreme Atmospheric Profiles

- data from the La Palma site
  - very well known atmosphere, lots of measurements already available (thanks to Markus Gaug)
  - profiles from GDAS (< 25 km) and NRLMSISE-00 (25 - 100 km)
  - other data from GDAS (e.g. ECMWF) are available and currently being explored

Extreme Atmospheric Profiles

- data from the La Palma site
  - very well known atmosphere, lots of measurements already available (thanks to Markus Gaug)
  - profiles from GDAS (< 25 km) and NRLMSISE-00 (25 - 100 km)
  - other data from GDAS (e.g. ECMWF) are available and currently being explored

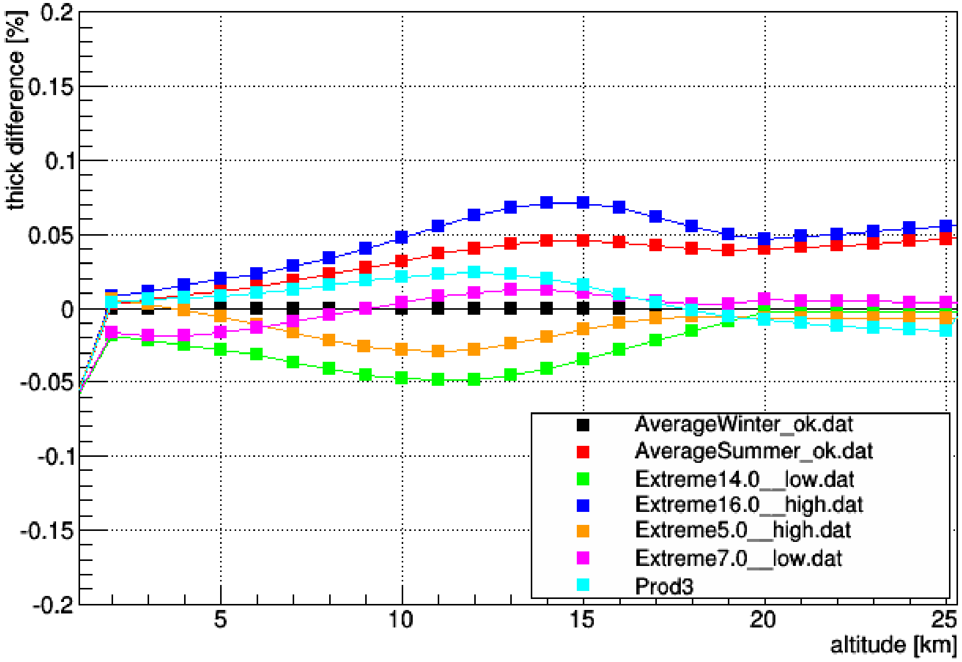


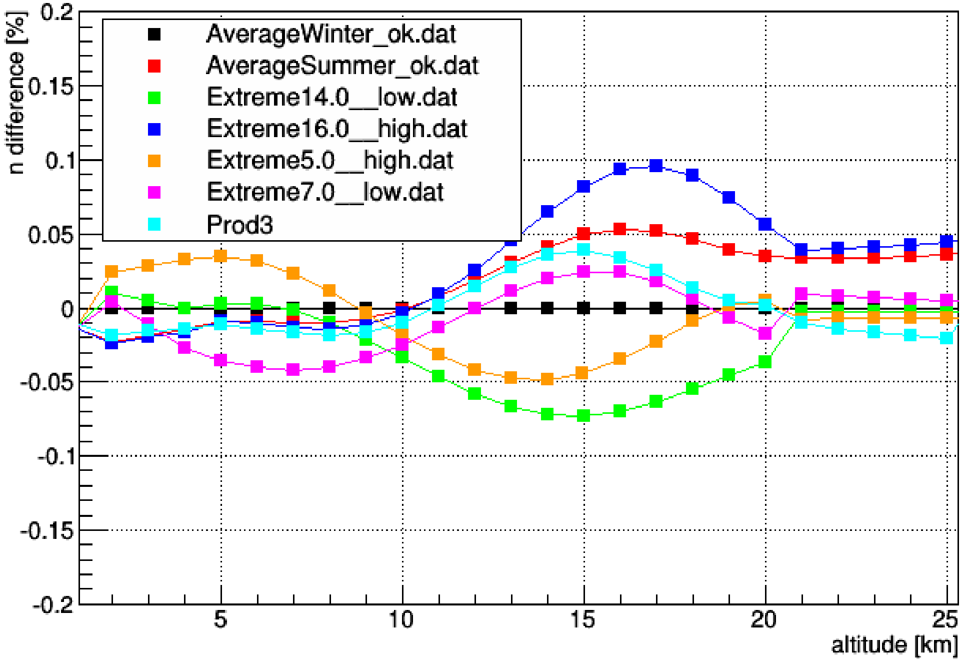


  - the different "extreme" altitudes have been chosen to maximize the effects on the high-energy or on the low-energy events

Extreme Atmospheric Profiles

- data from the La Palma site
  - very well known atmosphere, lots of measurements already available (thanks to Markus Gaug)
  - profiles from GDAS (< 25 km) and NRLMSISE-00 (25 - 100 km)
  - other data from GDAS (e.g. ECMWF) are available and currently being explored

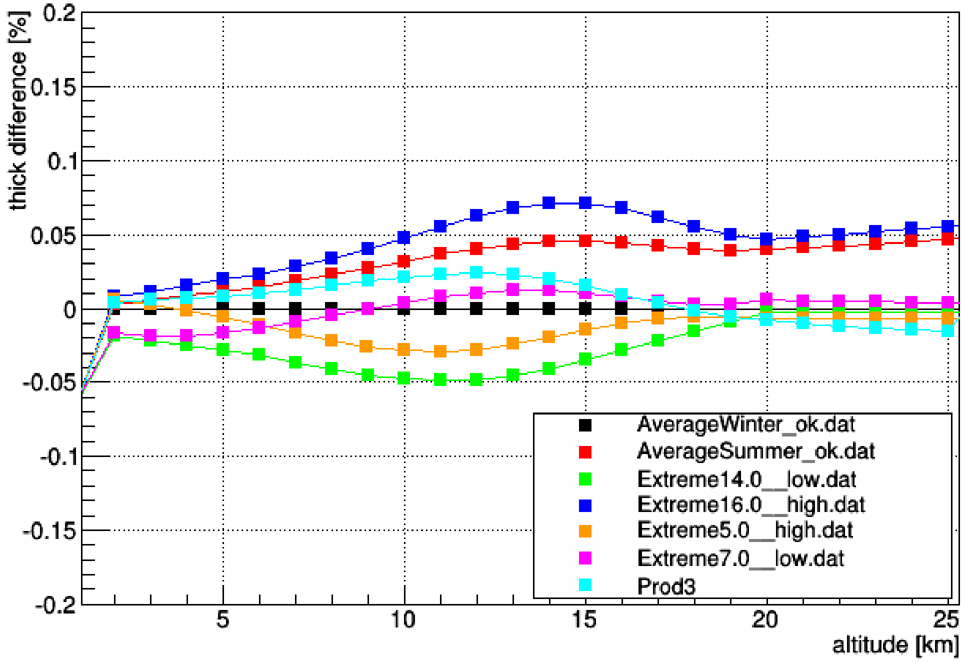


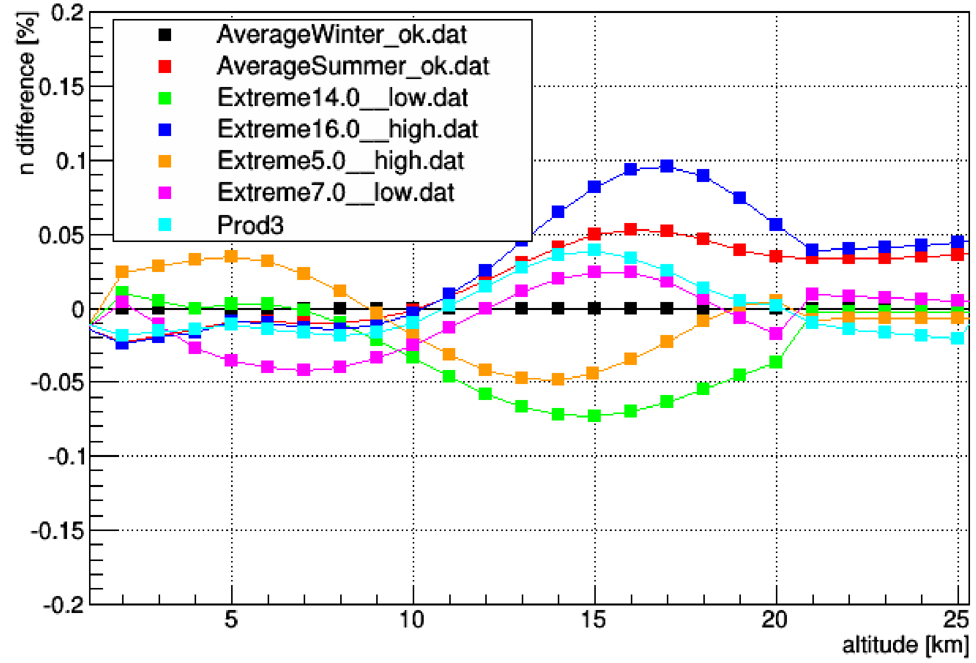


  - the different "extreme" altitudes have been chosen to maximize the effects on the high-energy or on the low-energy events

Extreme Atmospheric Profiles

- data from the La Palma site
  - very well known atmosphere, lots of measurements already available (thanks to Markus Gaug)
  - profiles from GDAS (< 25 km) and NRLMSISE-00 (25 - 100 km)
  - other data from GDAS (e.g. ECMWF) are available and currently being explored





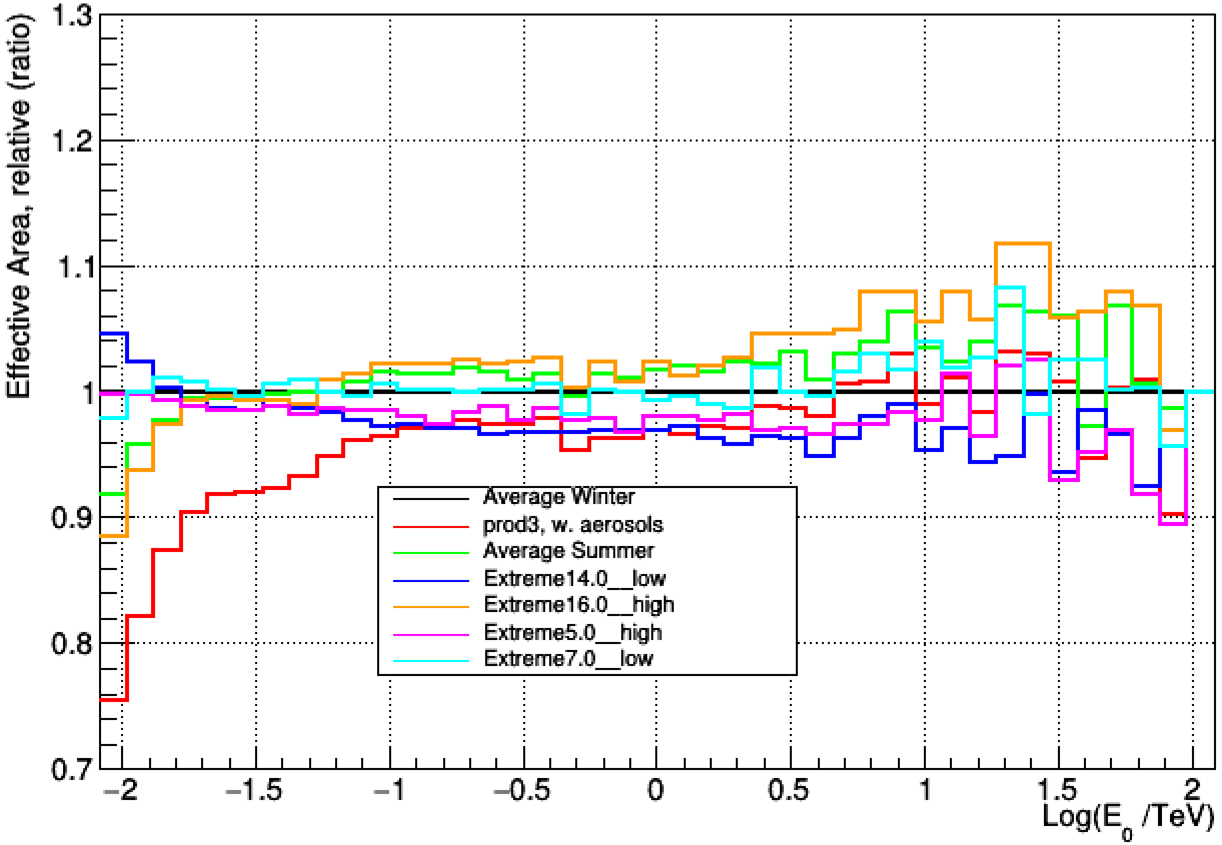
  - the different "extreme" altitudes have been chosen to maximize the effects on the high-energy or on the low-energy events
- atmospheric transmission obtained using MODTRAN, with these profiles as input

Production, reconstruction and analysis

- CORSIKA
  - only gammas (point-like) at zenith angle of 20 deg
  - energy range 3 GeV – 100 TeV with spectral index -2
- sim_telarray
  - 4 large-size telescopes and 15 medium-size telescopes (Flash-cam)
- Eventdisplay
  - standard production analysis

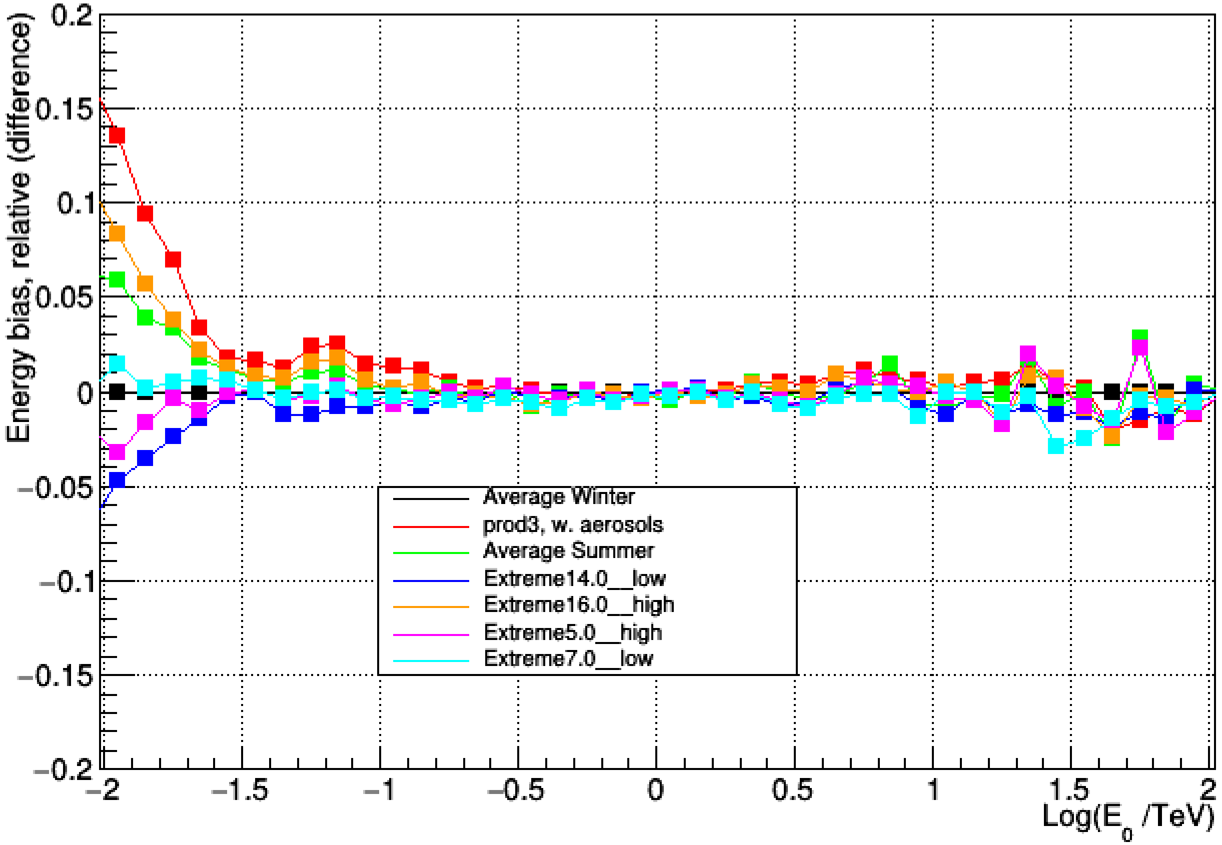
Effective area

- maximum difference for the atmosphere including aerosols
- the goal uncertainty on the knowledge of the effective area of CTA is 8% (including all uncertainties on the telescope hardware)



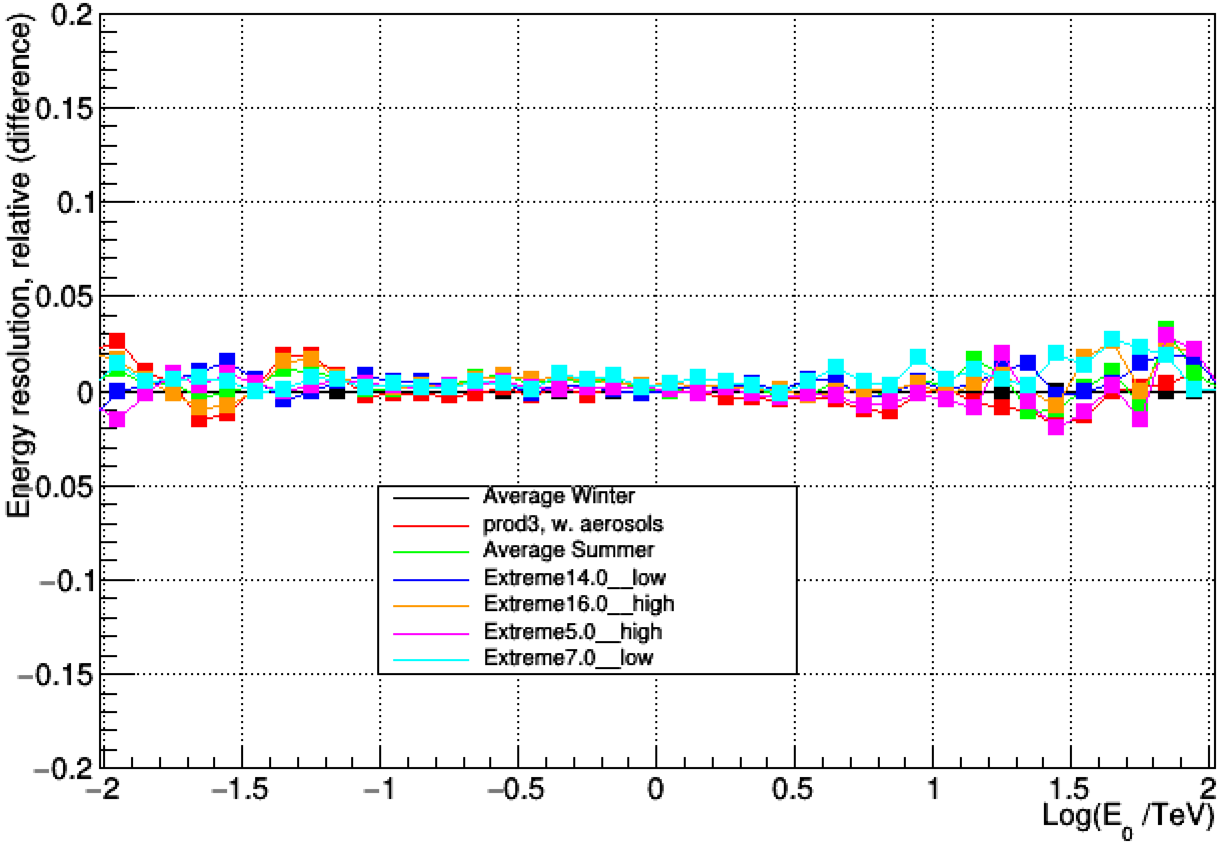
Energy bias

- not expected a large effect, we are using the same atmosphere for filling the LUTs and for the data production
- large effects < 20 GeV are due to a selection of the events which suffered large fluctuations, above 20 GeV only ±2% effect



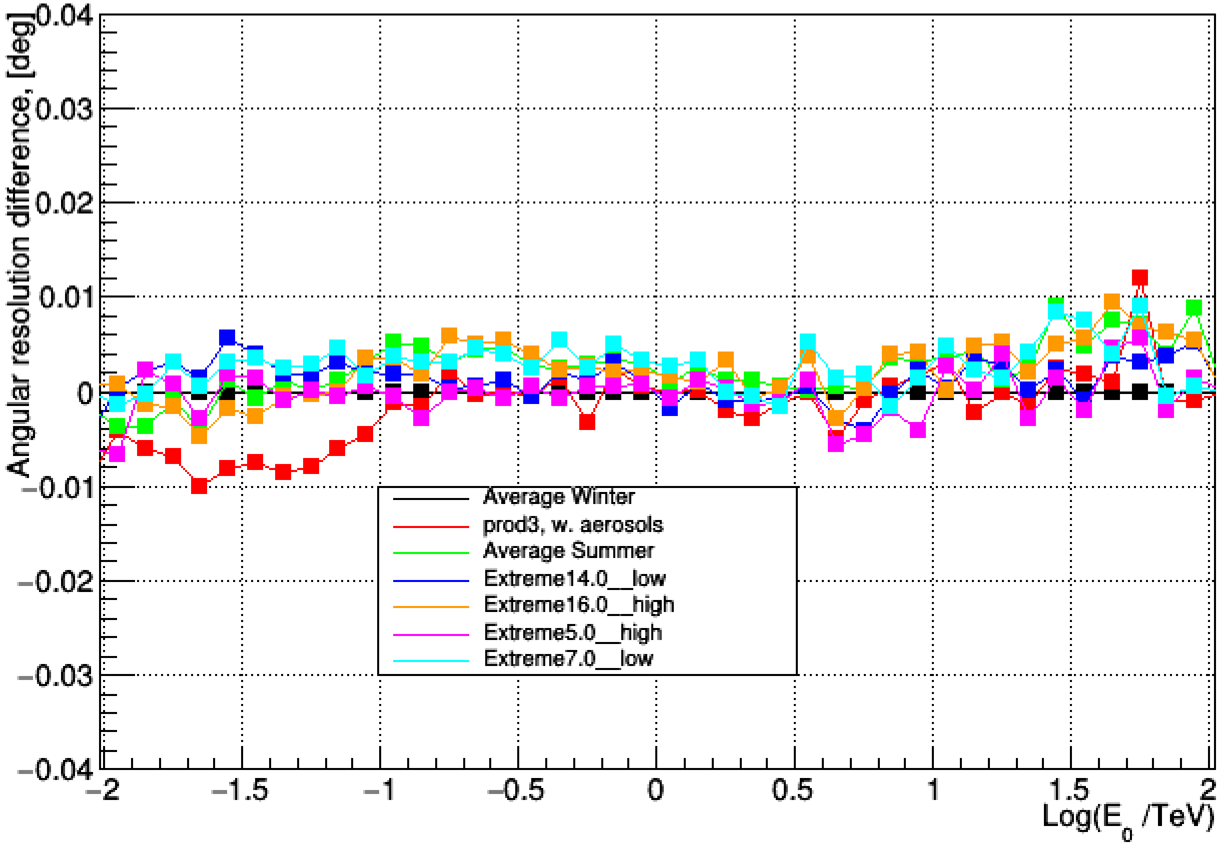
Energy resolutions

- ΔE interval including 68 % of the events
- very small ±2% effect



Angular resolutions

- we expect a second order effect on the angular resolution due to the different atmospheric profiles

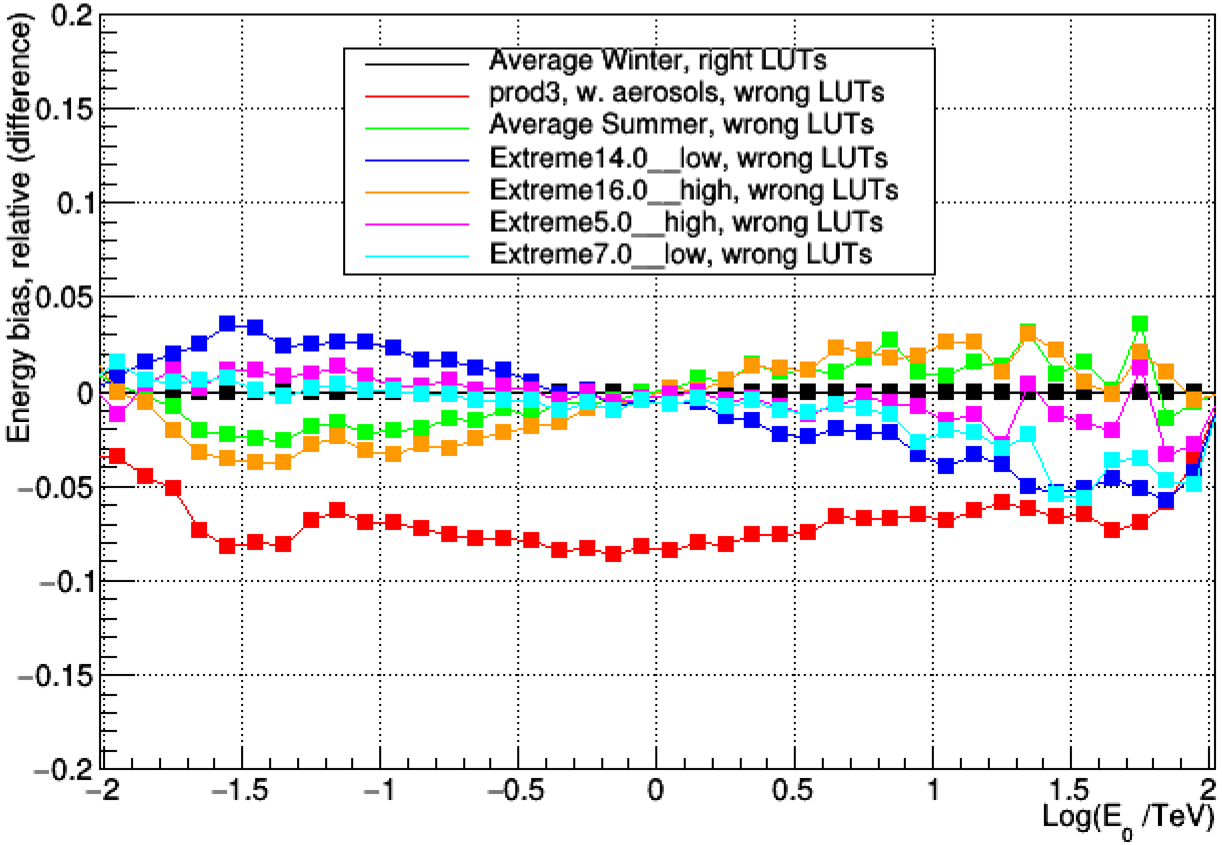


No atmospheric calibrations (preliminary)

- using wrong profile to reconstruct data produced extreme profiles
  - expected a large effect on the energy bias and only minor effects on the resolution

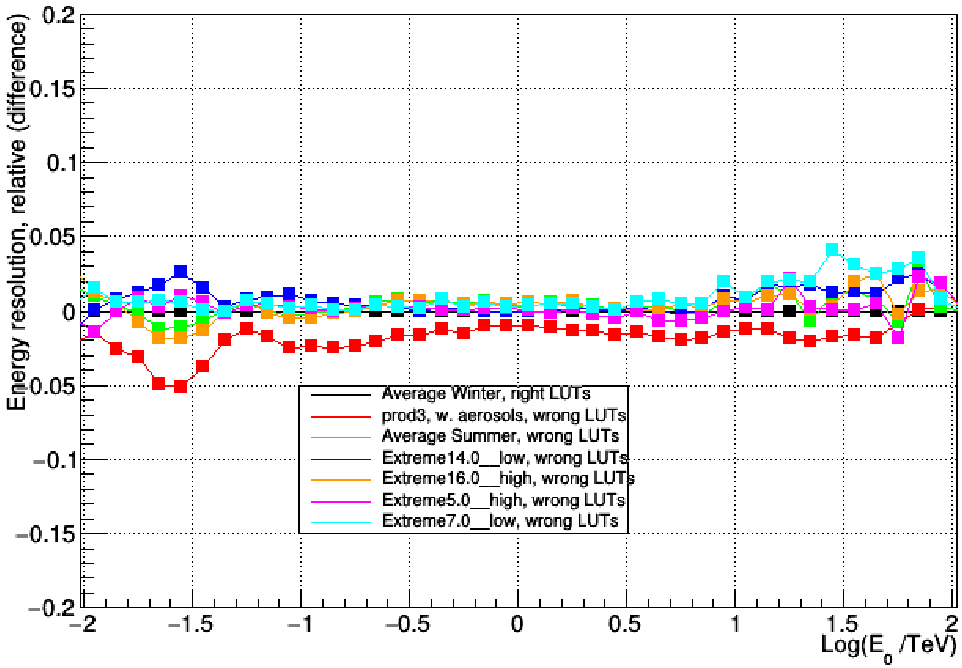
No atmospheric calibrations (preliminary)

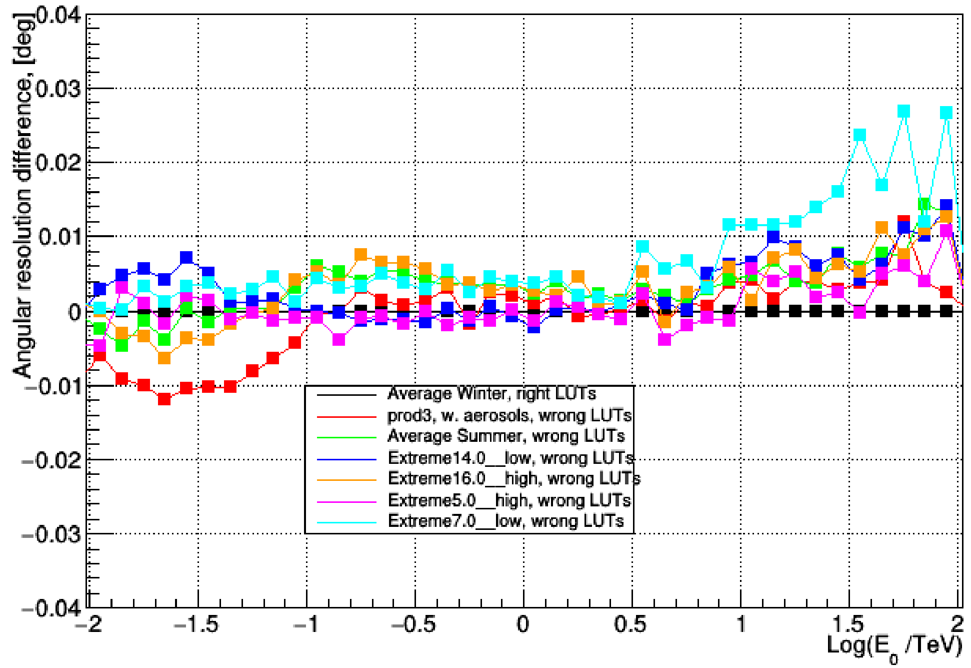
- using wrong profile to reconstruct data produced extreme profiles
  - expected a large effect on the energy bias and only minor effects on the resolution
- purely molecular profiles show an energy dependency (±4%) in energy bias
  - dangerous for energy spectra (only 2% systematic shift at all energies is allowed)



No atmospheric calibrations (preliminary)

- using wrong profile to reconstruct data produced extreme profiles
  - expected a large effect on the energy bias and only minor effects on the resolution
- purely molecular profiles show an energy dependency (±4%) in energy bias
  - dangerous for energy spectra





- for energy and angular resolutions these effects are rather small as expected

Future plans

- to study with MODTRAN the dependencies of the Optical Depth on the variations of the atmospheric profiles (density, thickness, refractive index, RH, water vapour fraction...)
- to introduce the aerosols in the atmospheric transmission
  - different kind of aerosols
  - different altitude and thickness
- same studies at different zenith angle
- to study the effects of the uncertainties of the currently foreseen CTA atmospheric measurements

Summary

- **Monte Carlo Simulations**
  - used almost in everywhere (characterization, analyses, optimization, …)
  - time-averaged and period-wise IRFs
  - input from atmospheric measurements
- **CTA computing resources**
  - GRID – Europe-wide computing network
  - open to any member of CTA
- **Atmospheric Simulations**
  - study of different atmospheric conditions on the La Palma site
  - need precise atmospheric measurements to reconstruct data