



autoMAGIC Basics

autoMAGIC Workshop Padova

Jan Lukas Schubert

28th - 30th August 2024





Contents:

- Reminder: The MAGIC analysis chain
- Reminder: The Data Levels
- autoMAGIC's goal
- autoMAGIC's core-concept
- autoMAGIC's capabilities
- Results obtained with autoMAGIC





Data Levels







The MAGIC Analysis Pipeline

- Sorcerer: calibration
- Merpp: merge subsystem reports and data files
- Star: image cleaning and image parameterization
- Superstar: stereo reconstruction







The MAGIC Analysis Pipeline

- Coach: train models for the event reconstruction
- Melibea: apply models to the data
- DL3 Converter: Compute event lists and store them in DL3 format







AutoMAGIC Goals

- Main goal: automatize the MAGIC workflow at least for standard use cases
 - Save analyzer's time
- Reproducible Analyses
- Prevent storage and computation time waste
- Bigger picture: MAGIC legacy





AutoMAGIC Core-Concepts

- Wrap MARS executable in Python code
- Use a Database
- Use a huge cluster for computations
- Work on the PIC server (where the MAGIC data is stored)





AutoMAGIC – Wrapping MARS executables

- For each analysis step (i.e. each MARS executable) we have a wrapper
 - > Checks the input files
 - > Checks the input parameters and creates a suiting .rc file
 - Runs the corresponding MARS executable
 - Saves the output
 - > Writes information into the DB
- In principle, autoMAGIC is just an automated MARS!





AutoMAGIC – The Database

- Organizes and stores all analysis-related information
- Database: Tables with relationships
- Tables for different use cases:
 - > Organizing the Data itself
 - Organizing the MC data
 - > Organizing the Jobs for the MARS wrappers and keeping track of the job states
 - Organizing analysis-related information such as cleaning levels or moon conditions

The database is the absolute core of autoMAGIC





AutoMAGIC - The Database

| alembic_version | | coach_types | | corsika_versions | | exit_codes | | hv_settings | | H_tables | | 13_tables | 13_tables | | | jobs_flute_mataju_exclude | | jobs_sorcerer | | jobs_selectmc_mataju_exclude_pixels | | |
|---------------------|------------|--|--------------------|-------------------------------|-------------------|--------------------------------|----------|---------------------------|-------------------|-----------------------------------|----------|-----------------------|---------------------------|----------------------|---------------|--------------------------------|----------|------------------------------|----------------------|-------------------------------------|-------------------|--|
| version_num V | ARCHAR(32) | id INTEGER | | M INTEGER | | id INTEGI | ER | id INTEGER | | M INTEGER | | id INTE | GER | ы | INTEGER | id | INTEGER | id | INTEGER | н | INTEGER | |
| | | name VARCHAR | | name VARCHAR | | mars_executable VARCHU | AR | name VARCHAR | | name VARCHAR | | name VARO | HAR | name | VARCHAR | start_date | DATETIME | run_id | INTEGER | monte_carlo_set_id | INTEGER | |
| | | | | | | code INTEGI | ER | | | | | | | release_date | DATETIME | stop_date | DATETIME | output_M1 | WARCHAR | output_file | VARCHAR | |
| | | | | | | description VARCHI | AR . | | | | | | | tarball | VARCHAR | coach_job_id | INTEGER | output_M2 | WARCHAR | mars_version_id | INTEGER | |
| | | | | | | | | | | | | | | | | pointing_id | INTEGER | mara_version_id | INTEGER | state_id | INTEGER | |
| monte carlo product | lona | monte carlo triggera | | sources | | states | | telescopes | | zenith ranges | | enelvala perioda | | Jobs add data | | hadronness_cut | FLOAT | state_id | INTEGER | exit_code_id | INTEGER | |
| M | INTEGER | id INTEGER | | M INTEGER | | id INTEGER | | M INTEG | 8 | M INTEGER | | id | INTEGER | id | INTEGER | hadisoners of from efficiency | POORT | exit_code_sorcerer_M1_id | INTEGER | timestamp | DATETIME | |
| 00000 | ARCHAR | name VARCHAR | | name VARCHAR | | name VARCHAR | | name VARCHA | | name VAROAR | | Dame | VARCHAR | source id | INTEGER | thate3 and from efficiency | BOOLEAN | eur_code_sorcerer_w2_id | DATETIME | pointing_id | INTEGER | |
| path | ARCHAR | | | ra FLOAT | | | | nom by min FLOA | a. | ALCORA VAROAR | | description | VARCHAR | date | DATETIME | quantia badroperr or | R OAT | uneump | | | | |
| identifier 1 | ARCHAR | | | dec FLOAT | | | | nom_hv_max FLOA | π | zd_max VARDIAR | | mc_production_id | INTEGER | max_subruns_numb | INTEGER | quantile theta2 cut | FLOAT | dD obs mellbes matslu e | volude associations | | | |
| | | | | | | | | red_hv_min FLOA | ar . | | | start_date | DATETIME | mars_version_id | INTEGER - | binning id | INTEGER | ish dit commiter id | INTEGE | jobs_melibea_ | mataju | |
| | | | | | | | | red_hv_max FLOP | π | | _ | stop_date | DATETIME | state_id | INTEGER | run_fold | BOOLEAN | ich ohr melihes id | INTEGE | id id | INTEGER | |
| moon_conditions | | observations | | jobs_selectmc | | monte_carlo_runs | | | | coach_train_files | | | | timestamp | DATETIME | mars_version_id | INTEGER | , | | coach_job_id | INTEGER | |
| м п | TEGER | id i | INTEGER | id IN | TEGER | id INTEG | SER | | | job_coach_id INTEGE | R | | | | | state_id | INTEGER | | | superstar, job, a | d INTEGER | |
| name V2 | ACHAR | source_id I | INTEGER | monte_carlo_set_id IN | ITEGER | calibrated_M1 VARCH | AR | | | job_selectmc_id INTEGE | R | | | | | exit_code_id | INTEGER | | | mars_version_ic | d INTEGER | |
| dc_min I | NTEGER | date D/ | ATETIME | moon_condition_id IN | TEGER | calibrated_M2 VARCH | sAR | | | | | | | | | timestamp | DATETIME | dl3_obs_melibea_mataju_si | andard_associations | state_id | INTEGER | |
| dc,max 1 | NTEGER | duration_in_s | FLOAT | split VA | RCHAR | monte_carlo_set_id INTEG | SER | | | | | | | | | | | job_dt3_converter_id | INTEGE | ext_code_d | INTEGER | |
| cUv1 | FLOAT | mola_duration_in_s | FLOAT | output_file_1 VA | RCHAR | number VARCI | GAR | | | | | | | | | jobs_flute_mataju_standard | | job_obs_melibea_id | INTEGE | e unestanto | | |
| 0,042 | PLOAT | mola_significance_se | FLOAT | output(ne,2 vA | ACHAR | first_calibrated whitch | SAR | coach coralica associatio | | view coper | | jobs mr mellbes | | John star superst | - | id . | INTEGER | | | | | |
| noise_iv_mean | PLOAT | moia_significance_ne | FLOAT | mars_version_id in | ILCOR | Mast_callerated VARCH | 50X | inh canch id | INTEGER | M INTEGER | | 14 | INTEGER | 14 | INTEGER | start_date | DATETIME | dia ma malikas standar | | | | |
| hose, w, ms | PLOAT | | | Scate_10 IN | TEGER | | | coraika version id | INTEGER | name VARCHAR | | reach inh id | INTEGER | an id | INTEGER | stop_date | DATETIME | ish dib sometra id | and a contraction | jobs_mc_supersta | r_mataju_standard | |
| warries Vi | Incluse. | | | timestame (A) | TETIME | | | 21-1V | | r min FLOAT | | selectric job id | INTEGER | d M | FLOAT | coach_job_jd | INTEGER | ish mi melihes standard | A INTEGER | id | INTEGER | |
| Per anni | | | | uncounty or | | | | | | r max FLOAT | | mars version id | INTEGER | d h2 | FLOAT | hadronness_cut | FLOAT | projection of the second | | IS_SESS_SEC | BOOLEAN | |
| | | | | | | | | | | | | state_id | INTEGER | noise.lv.mean | FLOAT | theta2_cut | FLOAT | | | monoe_carlo_run_le | INTEGER | |
| | | | | | | | | | | | | exit_code_id | INTEGER | noise_lv_rms | FLOAT | hadronness_cut_from_efficiency | BOOLEAN | | | outpoche | INTEGER. | |
| coach_off_files | | jobs_melibea | | selectmc_inputs | | flute_mc_melibea_association | 15 | flute_obs_mellbea_assoc | lations | subruns | | timestamp | DATETIME | mars_version_id | INTEGER - | theta2_cut_from_efficiency | BOOLEAN | dl3_mc_melibea_exclude | associations | state id | INTEGER | |
| job_coach_id | INTEGER | id INT | TEGER | job_selectmc_id | INTEGER | job_flute_id INT | TEGER | job_flute_id | INTEGER | id | INTEGER | | | output_file | VARCHAR | quantile_hadronness_cut | FLOAT | job_dl3_converter_id | INTEGER | exit code id | INTEGER | |
| job_superstar_id | INTEGER | ceach_job_id INI | TEGER | jab_mc_superstar_id | INTEGER | job_mc_melibea_id INT | TEGER | job_obs_melibea_id | INTEGER | calibrated_file_M1 | VARCHAR | | | state_id | INTEGER | quantile_theta2_cut | FLOAT | job_mc_melibea_exclude_id | INTEGER | timestamp | DATETIME | |
| | | superstar_job_id INT | TEGER | | | | | | | calibrated_file_M2 | VARCHAR | | | exit_code_star_M1_ic | INTEGER | binning_id | INTEGER | | | | | |
| | | mars_version_id INT | TEGER | | | | | | | run_id | INTEGER | | | exit_code_star_M2_ic | INTEGER | run_fold | BOOLEAN | | | | | |
| | | state_id INT | TEGER | | | | | | | subrun_number | VARCHAR | | | exit_code_superstar, | id INTEGER | mars_version_id | INTEGER | flute_obs_melibea_mataju | exclude_associations | | | |
| lef errore | | exit_code_id IN1 | TEGER | dia aka malikan anandaria | | | | | | 2d_M1 | FLOAT | John much | | timestamp | DATETIME | 10200_04 | INTEGER | job_flute_id | INTEG | ER | | |
| in cypes | | timestamp DATI | ETIME | un jour mendes association | J'IS | coach_boor ce_associations | | runs | | 2d_M2 | FLOAT | Jouriceach | | | | exe_code_id | INTEGER | job_obs_melibea_id | INTEG | ER | | |
| Name VAROIAS | | | | ish ohs melihes id | INTEGER | source id INTEGO | | observation id | INTEGER | az_M1 | PLOAT | normalization id | INTEGER | | | unietarip | | | | | | |
| | | | | , | | | | run number | VARCHAR | 1 In case M1 | ROM | zenith id | INTEGER | | | | | | | | | |
| | | | | | | | | wobble offset | RIDAT | 1 In case MY | ROM | trigger id | INTEGER | | | jobs dl3 converter mataju excl | ude | coach_off_files_mataju | | flute_mc_melibea_excl | ude_associations | |
| | | | | | | | | wobble angle | FLOAT | cinations | ROAT | meen condition id | INTEGER | | | м | INTEGER | job_ceach_id | TEGER | job_flute_mataju_id | INTEGE | |
| | | | | | | | | L1_table_id | INTEGER | transmission3km, mean | FLOAT | mars_version_id | INTEGER | | | coach_job_jd | INTEGER | job_superstar_id | TEGER | job,mc,melibea,mataju, | exclude_id INTEGE | |
| jobs_mc_superstar | | dl3_mc_melibea_associati | ions | selectmc_inputs_mataju_st | landard | pointing | | L3_table_id | INTEGER | transmission6km_mean | FLOAT | state_id | INTEGER | monte_carlo_sets | | irf_type_id | INTEGER | | | | | |
| ы | INTEGER | job_dl3_converter_id | INTEGER | job_selectmc_mataju_standa | rd_id INTEGER | id INTEGER | | zd | FLOAT | transmission5km_mean | FLOAT | exit_code_RFgh_id | INTEGER | id | INTEGER | hadronness_cut | FLOAT | jobs_mc_melibea_mataja | _exclude_pixels | | | |
| monte_carlo_run_id | INTEGER | job_mc_melibea_id | INTEGER | job_mc_superstar_mataju_sta | andard_id INTEGER | source_id INTEGER | | az | FLOAT | transmission12km_mean | FLOAT | exit_code_LUTs_id | INTEGER | path | VARCHAR | theta2_cut | FLOAT | id | INTEGER | | | |
| moon_condition_id | INTEGER | | | | | wabble_offset INTEGER | | L3t_rate_M1 | RLOAT | transmission/fikm_std | FLOAT | exit_code_RFdisp_id | INTEGER | production_id | INTEGER | quantile_hadronness_cut | FLOAT | coach,job,jd | INTEGER | | | |
| output_file | VARCHAR | | | | | wobble_angle INTEGER | | L3t_rate_M2 | FLOAT | domean_M1 | FLOAT | timestamp | DATETIME | zenith_id | INTEGER | quantile_theta2_cut | FLOAT | selectric_job_exclude_pixels | id INTEGER | | | |
| mars_version_id | INTEGER | | | | | | | cloudiness | FLOAT | dcmean_M2 | FLOAT | view_cone_id | INTEGER | corsika_version_id | INTEGER | hadronness_cut_from_efficiency | BOOLEAN | pointing_id | INTEGER | | | |
| state_id | INTEGER | and a stand of the stand of the stand of the | analuda almala | | | and the first matrix | | transmission3km_mean | FLOAT | meanhy_M1 | FLOAT | source_set | INTEGER | num_files | INTEGER | theta2_cut_from_efficiency | BOOLEAN | mars_version_id | INTEGER | | | |
| exit_code_id | INTEGER | selecting inputs inataju | exclose pixels | coach_source_associations | Unacayo | coach_train_mes_macapo | | transmission6km_mean | FLOAT | meanhy_M2 | FLOAT | off_mc_events_ratio | FLOAT | crigger_id | INTEGER - | mars_version_id | INTEGER | state_id | INTEGER | | | |
| timestamp | DATETIME | job_selectmc_mataju_exclu | ale poxes is in in | rock jab,coach,ia | INTEGER | Job coace is invited | | transmission9km_mean | FLOAT | ebs_start_M1 | DATETIME | underpopulated_bins | INTEGER | num_events_zenith_ | sin INTEGER | state_id | INTEGER | ext_code_id | INTEGER | | | |
| | | joe inclupertar inataje e | enous presion | | | jus selecting is in the | | transmission12km_mean | FLOAT | obs_start_M2 | DATETIME | | | view_cone_id | REGETVE | exit_code_id | INTEGER | smestamp | DATETIME | | | |
| | | | | | | | | domean_M1 | FLOAT | obs_stop_M1 | DATETIME | | | | | timestamp | DATETIME | | | | | |
| | | | | | | | | domean_M2 | PLOAT | ebs_stop_M2 | DATETIME | | | | | pointing_id | INTEGER | jobs_mc_melibea_mat | ju_standard | | | |
| | | | | | | | | automotors | INTEGER | duration_in_s_M1 | PLOAT | | | | | | | id | INTEGER | | | |
| jobs_coach_mataju | | coach_corsika_association | ns_mataju | jobs_flute | | jobs_dl3_converter | | c a id | INTEGER | duradon_in_s_M2 | FLOAT | c.p.files | | binning_flute | | jobs_dl3_converter_mataju_stan | dard | coach_job_id | INTEGER | | | |
| ы | INTEGER | job_coach_id | INTEGER | ы | INTEGER | id | INTEGER | | | number storgader reports M2 | INTEGER | id | INTEGER | id INT | GER | м | INTEGER | selectric_job_standard_ie | INTEGER | | | |
| production_id | INTEGER - | corsika_version_id | INTEGER | start_date | DATETIME | coach_job_id | INTEGER | | | number identified stars M1 | ROAT | p_run_number | VARCHAR | name VAR | HAR | coach(job,id | INTEGER | mars_version_id | INTEGER | | | |
| zenith_id | INTEGER | | | stop_date | DATETIME | inf_type_id | INTEGER | | | number identified stars M2 | FLOAT | c_run_number | VARCHAR | | | irf_type_id | INTEGER | state_id | INTEGER | | | |
| view_cone_id | INTEGER | | | coach_job_id | INTEGER | mans_version_id | INTEGER | | | rumber, events | INTEGER | c_file_M1 | VARCHAR | | | nadronness_cut | FLOAT | exit_code_id | INTEGER | | | |
| trigger_id | INTEGER | | | mars_version_id | INTEGER | state_id | INTEGER | | | raw_file_M1 | VARCHAR | c_file_M2 | VARCHAR | | | thetaz_cut | ROAT | timestamp | DATETIME | | | |
| source_set | INTEGER | | | state_id | INTEGER | exit_code_id | INTEGER | | to ended a shade | raw_file_M2 | VARCHAR | p_file_M1 | VARCHAR | 100 miles | | quartite there? or | ROAT | | | | | |
| underpopulated_bins | INTEGER | jobs_star_superstar_mata | 4u | exit_code_id | INTEGER | timestamp | DATETIME | jobs_mc_superstar_mata | gu_exclude_pikels | | | p_file_M2 | VARCHAR | jobs_selectmc_ma | teju_standard | badronnes cut from affiliance | BOOLEAN | | | | | |
| off_mc_events_ratio | FLOAT | Id | INTEGER | timestamp | DATETIME | hadronness_cut | FLOAT | M | INTEGER | | | observation_id | INTEGER | Nd . | INTEGER | thera? cut from efficiency | BOOLEAN | | | | | |
| mars_version_id | INTEGER | mars_version_id | NIGOLO | binning_id | INTEGER | hadronness_out_from_efficiency | BOOLEAN | monte_cano_run_id | INTEGER | flute_mc_mellbea_standard_assoc | clations | | | monte_carlo_set_id | INTEGER | mars version id | INTEGER | | | | | |
| state_id | INTÉGER | date id | INTEGER | run_fold | BOOLEAN | quantile_hadronness_cut | FLOAT | output_ne | WARLINAR | job_flute_mataju_id | INTEGER | | | output, the | INTEGER | state_id | INTEGER | | | | | |
| exit_code_RFgh_id | WIEGER | exit code star M1 M | INTEGER | hadronness_cut | FLOAT | quantile_theta2_cut | FLOAT | mars_version_io | INTEGER | job_mc_melibea_mataja_standard_id | INTEGER | | | mars_version_H | INTEGER | exit_code_id | INTEGER | | | | | |
| excluse_COTS.Id | INTEGER | exit code star M2 id | INTEGER | mauromiest_out_from_efficienc | y BLOCEAN | that's out from officiency | ROOLEAN | exit code id | INTEGER | | | flute_obs_melibea_mat | aju_standard_associations | exit code id | INTEGER | timestamp | DATETIME | | | | | |
| timestamp | DATETIME | exit_code_superstar.id | INTEGER | quantile theta2 cut | FLOAT | o waa _con_rom_enciency | | timestamp | DATETIME | | | job_flute_id | INTEG | timestamp | DATETIME | | | | | | | |
| | | timestamp | DATETIME | theta2_cut | FLOAT | | | pointing_id | INTEGER | | | job_obs_melibea_id | INTEG | is_test_set | BOOLEAN | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | |





AutoMAGIC – The Cluster

- Cluster on the PIC server is equipped with HTCondor
- Allows for parallel processing of hundreds of jobs
- Jobs are submitted to the cluster, then HTCondor will manage to which working node the jobs are sent





AutoMAGIC – Working on the PIC Server

- Possibilities:
 - Use data stored on PIC no copies necessary
 - > Make use of the PIC Cluster
 - Immense amount of storage necessary for huge autoMAGIC productions
- Caveats:
 - > Data availability for calibrated and raw data can be problematic
 - > PIC is not the most stable server I know...





AutoMAGIC – Working Principle Overview





What autoMAGIC can do...

- Standard Analysis
- Moon Analysis
- MaTaJu analysis of standard trigger data
- MaTaJu analysis of Sum-Trigger-II data
- Diffuse analysis





What autoMAGIC cannot do (yet?)...

- Non-standard analyses "just changing one MARS parameter"
 - In autoMAGIC you cannot "just change" a parameter
 - > If the parameter is changed often, it can be implemented without huge effort
- The DL3 + Gammapy doesn't support the corrections for hazy atmosphere conditions (as far as I know
 - Stricter cuts on weather necessary (and hence less data available...)





Limitations and Caveats of autoMAGIC

- Usage of Database with many users simultaneously can be problematic in the current setup
 - > Need to design a workflow/concept to allow for this
- Main Bottleneck: Data availability at PIC
 - > Find a smoother way for accessing the calibrated data together with the PIC admins
- There still might be bugs!





AutoMAGIC results: Crab Nebula in DL3 paper







AutoMAGIC results: Crab Nebula Long-Term







Jan Lukas Schubert



AutoMAGIC results: Crab Pulsar Long-Term







AutoMAGIC results - Sum-Trigger-II Crab Pulsar Analysis







AutoMAGIC results – Mrk421 SED









AutoMAGIC List of Recent Productions

- Crab DL3 validation Paper
- Crab Longterm ST.03.06-ST.03.20
- Crab Sum-Trigger-II ST.03.06-ST.03.12
- Mrk421
- Multiple analysis of dim sources for the EHBL catalog paper





AutoMAGIC Future Plans

- I'm going to quit science at the end of October :'(
- Cyrus Walther (cyrus.walther@tu-dortmund.de) is going to take over the project
- David Venker's (Master student in Dortmund) work:
 - Survey work on FSRQs based on dark data
- Cyrus' work:
 - Modeling of AGN spectra, focusing on BL Lacs