On Hidden Uncertainties (in CORSIKA).

and other tools

Johannes Knapp
DESY Zeuthen

HELMHOLTZ ASSOCIATION

DESY

CORSIKA and other simulation/analysis tools tend to be used with blind confidence

and often with minimal knowledge of the relevant literature (manuals, limitations, critical comparisons, known deficiencies, …).

and with insufficient MC statistics

This can have consequences on the claimed results which are usually

neither captured by the quoted systematic errors,
nor discussed in the resulting publications.

“unknown unknowns”

CORSIKA is a multi-purpose simulation for
4-d (space & time) particle showers in the atmosphere.

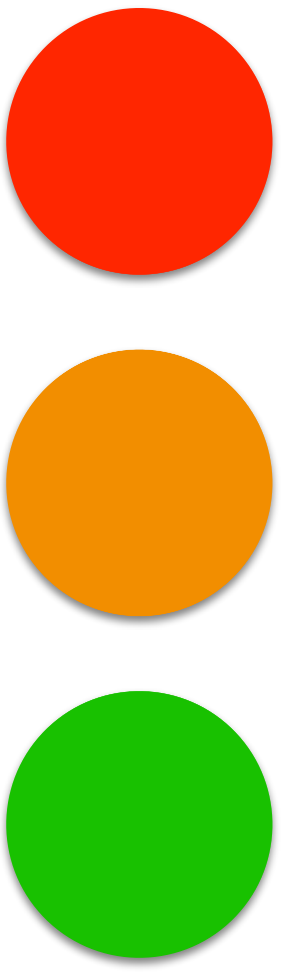
type, energy, momentum, direction, arrival time,
location of each secondary particle
Cherenkov, fluorescence, radio emission

particle numbers,
lateral, longitudinal and time distribution
of charged particles in showers

CORSIKA handles: hadronic interactions
electromagnetic interactions
particle decays
particle tracking (deflection, losses, …)

Shower simulation require also a detailed and adequate simulation
of the detector response, triggering, readout, event reconstruction …

a few examples …



thanks for input to D Heck

Particles:

masses: relative uncertainties are mostly $<10^{-4}$,
(some short-lived charmed/bottom particles are somewhat less well known)

Mass and lifetime parameters are updated typically all 5 years.

lifetimes: relative uncertainties $<4 \times 10^{-3}$ for most particles
(except charmed/bottom particles and those with lifetimes $<10^{-12}$ sec), which seems sufficient in the shower development also at highest energies.

resonances: are neglected (due to very short lifetime).
Only Omega, Rho, Kaon*, and Delta resonances are considered explicitly without transport, but kinematically correct decay.

Decays:

branching ratios: all down to 1% level.

In special cases even much lower:

i) K^0_s ⟶ $pi^\pm + e^\pm +$ (anti) ν_e B.R. = $7x10^{-4}$

large contribution to high-energy nu(e) flux.

ii) η ⟶ $\mu^+ + \mu^- + \gamma$ B.R. = $3x10^{-4}$

contribution to high-energy muon flux.

kinematics: exact kinematics is used everywhere,

also in 3-body decays.

Especially decay μ ⟶ $e + \nu_e + \nu_\mu$

includes muon polarisation and spin correlations

decays of charmed/bottom particles are performed in PYTHIA

Electromagnetic Interactions:

electrons and gammas treated by EGS4
(ionisation, bremsstrahlung, photo, Compton, pairs…)
production of $\mu^+ \mu^-$ pairs by gammas
LPM effect at high energies

photo-nuclear interactions of pi, rho(0) and omega mesons
but phi are neglected.
Recoil of nucleon is respected.

bremsstrahlung, pair, nuclear reactions of muons and tau-leptons
mu/tau multiple scattering with energy loss.

deflection in Earth magnetic. adaptive tracking.

e+e- pair creation of primary gammas in local Earth magnetic field
is considered only above the atmosphere (preshower),
but not in the atmosphere.
(becomes important at highest gamma energies $>10^{18}$ eV
in thin atmosphere in concurrence with ordinary e+e- pair creation.)

Particle Transport:

interaction cross section (large uncertainty):

cross sections unknown for STRANGE baryons and CHARMED/BOTTOM particles Eta, Rho mesons

(use proton and pion cross sections instead)

Resonances are not transported (very short lifetime $<10^{-22}$ sec).

range calculations: include

i) lifetime of unstable particles

ii) ionisation energy loss during transport in medium with increasing density

Deflection in Earth magnetic field

Low energy cut-off: reasonable cuts to be chosen by users (depending on their detector; not trivial)

Atmosphere:

atmospheric overburden & layering:
depends on pressure & temperature
varies with latitude (polar ... tropic),

optical atmospheric properties:
absorption, scattering, aerosols, refraction, …
Cherenkov & fluorescence yield

Both have seasonal & daytime variations.
Averages and variations are needed.

Users are responsible for a good atmospheric model:
regular local measurements and global modelling,
proper reflection in the air shower model throughout the year
(or proof that effects are negligible).

Magnetic field: specified by user (horizontal & vertical components)
Dependent on latitude, longitude and year.
Introduces non-trivial shower shapes, rarely reflected in the analysis

Hadronic Interactions:

by far largest uncertainty for composition studies
phenomenological: no fundamental theory
considerable freedom for model builders (see talks of models)

energy conservation:
for most models by few (<5) %

nuclei projectiles/targets:
Glauber approach (which variant?), multiple interactions.

strange/charmed secondary particles can be produced, but
not initiate another interaction.
(treated as standard baryons/mesons)

extrapolations to UHE largely uncertain

should be correct for all energies, primaries, targets

use of outdated / fast but bad models

Documentation

good users manual,
materials from 4 CORSIKA Schools

Physics description (partly outdated)

>1800 citations (~100-150 / yr) to CORSIKA Physics Description
but no consistent collection of results on CORSIKA performance

KfK 4998
November 1992

The Karlsruhe Extensive Air Shower Simulation Code CORSIKA

J. N. Capdevielle, P. Gabriel, H. J. Gils, P. Grieder,
D. Heck, J. Knapp, H. J. Mayer, J. Oehlschläger,
H. Rebel, G. Schatz, T. Thouw
Institut für Kernphysik

Kernforschungszentrum Karlsruhe

Forschungszentrum Karlsruhe
Technik und Umwelt
Wissenschaftliche Berichte
FZKA 6019

CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers

D. Heck, J. Knapp, J. N. Capdevielle,
G. Schatz, T. Thouw
Institut für Kernphysik

Februar 1998

Preface to KfK 4998 (1992)

Analyzing experimental data on Extensive Air Showers (EAS) or planning corresponding experiments requires a detailed theoretical modeling of the cascade which develops when a high energy primary particle enters the atmosphere. This can only be achieved by detailed Monte Carlo calculations taking into account all knowledge of high energy strong and electromagnetic interactions. Therefore, a number of computer programs has been written to simulate the development of EAS in the atmosphere and a considerable number of publications exists discussing the results of such calculations. A common feature of all these publications is that it is difficult, if not impossible, to ascertain in detail which assumptions have been made in the programs for the interaction models, which approximations have been employed to reduce computer time, how experimental data have been converted into the unmeasured quantities required in the calculations (such as nucleus-nucleus cross sections, e.g.) etc.
This is the more embarrassing, since our knowledge of high energy interactions - though much better today than ten years ago - is still incomplete in important features. This makes results from different groups difficult to compare, to say the least. In addition, the relevant programs are of a considerable size which - as experience shows - makes programming errors almost unavoidable, in spite of all undoubted efforts of the authors. We therefore feel that further progress in the field of EAS simulation will only be achieved, if the groups engaged in this work make their programs available to (and, hence, checkable by) other colleagues. **This procedure has been adopted in high energy physics and has proved to be very successful.** It is in the spirit of these remarks that we describe in this report the physics underlying the CORSIKA program developed during the last years by a combined Bern-Bordeaux-Karlsruhe effort. We also plan to publish a listing of the program as soon as some more checks of computational and programming details have been performed. We invite all colleagues interested in EAS simulation to propose improvements, point out errors or bring forward reservations concerning assumptions or approximations which we have made. We feel that this is a necessary next step to improve our understanding of EAS.

… could be said almost unchanged
for any CR composition analysis.

CORSIKA & GEANT are de-facto standard tools

Why does every experiment
does have its own
calib / reconstruction / analysis method / stat tests ??

can we define here standards, too?
get some convergence …
(deconvolution for spectra, BDTs for classification, template fitting,
blind analyses for source searches, …)

Ideally:

A thorough and detailed assessment of all relevant systematic errors **must** be demonstrated by the teams publishing a result (expt., theor., method.).

One should be very careful with results if this has not been done.

Good examples will define “good practice” for all.

e.g.:

data used for tuning models
optimising cuts
calibration
comparing with collider data

quote version numbers of the models used
quote CORSIKA version
describe the atmospheric details
quote the precise versions of the analysis methods

Trend: open data, open methods, open software, …
for checking & re-analysis

Each experiment should have a **living archive** of plots, comparisons, simulations, methods, up-to-date references

to motivate **all** their calibration, data selection cuts, choice of algorithms, fit ranges, error estimates …

and

to update their results whenever new tools, models, input data … become available.

Such material **should be made public** (and hence checkable) and continuously updated.

PhD theses,
internal reports

Much more needed than just a few-page paper.

For model builders and users of Shower simulations:

Collect systematically a common set of reference data

— from collider experiments to tune models
— from CR experiments (low to high energies,)
to compare with CR data.

— apply models within CORSIKA for comparison
(e.g. using the “interaction test” feature of CORSIKA

Document the comparison of sims with reference plots
for each model variant.
Have it openly accessible.

Let’s push the trend towards open data in cosmic ray physics.

Demand openness from others and provide it yourself.

A “rating” for experiments / models / analysis tools / papers ?

How well understood is an
experiment / model / simulation & analysis tool?
How much background info is publicly available ?

+++, ++, +, 0, -, - -, - - -

e.g. Fermi LAT: +++

+++, ++, +, 0, -, - -, - - -

we should
increase syst errors by: 20 50 100 200 500 1000%

Don’t tolerate/accept **- -** and **- - -** results.

Homework for you:

How to rate the experiments / models / methods that are presented at this workshop?

Homework for you:

How to rate the experiments / models / methods that are presented at this workshop?

My ranking of CORSIKA: ++