

A new release of the KASCADE Cosmic Ray Data Centre (KCDC)

D. Kang^{*1}, J. Wochele¹, A. Haungs¹, S. Schoo¹, D. Wochele¹, W.D. Apel¹, J.C. Arteaga-Velázquez², K. Bekk¹, M. Bertaina³, J. Blümer^{1,4}, H. Bozdog¹, I.M. Brancus⁵, E. Cantoni^{3,6}, A. Chiavassa³, F. Cossavella⁴, K. Daumiller¹, V. de Souza⁷, F. Di Pierro³, P. Doll¹, R. Engel¹, D. Fuhrmann⁸, A. Gherghel-Lascu⁵, H.J. Gils¹, R. Glasstetter⁸, C. Grupen⁹, D. Heck¹, J.R. Hörandel¹⁰, D. Huber⁴, T. Huege¹, K.-H. Kampert⁸, H.O. Klages¹, K. Link⁴, P. Łuczak¹¹, H.J. Mathes¹, H.J. Mayer¹, J. Milke¹, B. Mitrica⁵, C. Morello⁶, J. Oehlschläger¹, S. Ostapchenko¹², N. Palmieri⁴, T. Pierog¹, H. Rebel¹, M. Roth¹, H. Schieler¹, F.G. Schröder¹, O. Sima¹³, G. Toma⁵, G.C. Trinchero⁶, H. Ulrich¹, A. Weindl¹, J. Zabierowski¹¹ -

KASCADE-Grande Collaboration

¹ Institut für Kernphysik, KIT - Karlsruhe Institute of Technology, Germany

² Universidad Michoacana, Inst. Física y Matemáticas, Morelia, Mexico

³ Dipartimento di Fisica, Università degli Studi di Torino, Italy

⁴ Institut für Experimentelle Kernphysik, KIT - Karlsruhe Institute of Technology, Germany

⁵ Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania

⁶ Osservatorio Astrofisico di Torino, INAF Torino, Italy

⁷ Universidade São Paulo, Instituto de Física de São Carlos, Brasil

⁸ Fachbereich Physik, Universität Wuppertal, Germany

⁹ Department of Physics, Siegen University, Germany

¹⁰ Dept. of Astrophysics, Radboud University Nijmegen, The Netherlands

¹¹ National Centre for Nuclear Research, Department of Astrophysics, Lodz, Poland

¹² Frankfurt Institute for Advanced Studies (FIAS), Frankfurt am Main, Germany

¹³ Department of Physics, University of Bucharest, Bucharest, Romania

E-mail: donghwa.kang@kit.edu

The KASCADE Cosmic Ray Data Centre (KCDC) is a web-based platform to provide scientific data for the general public. KASCADE and its extension KASCADE-Grande fully completed the data accumulation at the end of 2013 after more than 20 years of measurements. The main purpose of KCDC is to archive original data such as from KASCADE to offer long-term scientific data for the high-energy astroparticle physics community as well as for students and the interested public via a sophisticated web portal. Recently, a new version named NABOO was released, based on a completely new back-end database and software. In addition, the amount of data increased by factor 3 compared to the previous release. Data from the Grande detector array are now also available. In this contribution, the KCDC web portal and the data availability as well as future plans will be discussed.

35th International Cosmic Ray Conference

10-20 July, 2017

Bexco, Busan, Korea

*Speaker.

1. Introduction

Research in astroparticle physics, like cosmic rays physics, addresses some of the most fundamental questions in nature. There is an intimate connection between measurements and theoretical descriptions of astrophysical phenomena to provide the foundation for the sophisticated models of macroscopic astrophysical systems. In addition, scientists between experiments from ground-based and space-based devices have to share their incredibly detailed observations to study composition changes and energy release in astrophysical environments. For that, a diverse set of astrophysical data is required to be made public. The current trend, not only in astroparticle physics but also in particle physics is that people from all over the world can see images as soon as they are posted, and scientists can immediately download scientific data. This demonstrates from the initial stages the power of the Internet and the ability of the scientific community to share data quickly with other colleagues and with the general public. Some experiments in astroparticle physics have already adopted this fascinating idea and they have involved their scientific data in electronic publishing. The KASCADE Cosmic-ray Data Centre (KCDC) [1] is a pilot project in this respect.

In this proceedings, we describe an overview of the KCDC web portal as well as a new released version. A future plan will be also discussed.



Figure 1: Logo of the KASCADE Cosmic-ray Data Centre (KCDC)

2. KASCADE Cosmic Ray Data Centre (KCDC)

KCDC is a web portal, where data of KASCADE and the KASCADE-Grande experiments are made available for the interested public [2]. Due to low statistics, cosmic-ray particles above an energy of about 100 TeV have to be observed indirectly by measuring extensive air showers, i.e. the cascades of secondary particles generated in the atmosphere. Experiments for these measurements cover up to several square kilometers or even thousands of square kilometers by using large-area detector arrays. KASCADE [3] and KASCADE-Grande [4] are prominent experiments of such detectors. KASCADE stands for KARlsruhe Shower Core and Array Detector and -Grande denotes the tenfold larger version, covering an area of 0.5 km², with non-shielded and shielded detectors, a large-size hadron calorimeter and additional muon tracking devices. KASCADE and KASCADE-Grande located at the Campus North of the KIT have been operated for more than two decades. Operations for data taking were successfully completed at the end of 2013.

Since the first release in 2013, KCDC provides to the public the measured and reconstructed parameters of air showers. In addition, KCDC provides the conceptual design, how the data can be treated and processed so that they are also usable outside the community of experts in the research field. Detailed educational examples make a use also possible for high-school students and early stage researchers. The aim of the project KCDC is the installation and establishment of a public data centre for high-energy astroparticle physics based on the data of the KASCADE experiment. Moreover, with KCDC we provide to the public the selected data via a custom-made web page.

Figure 2 shows a schematic view of the KCDC project with keywords related to the various works: An extensive air shower of cosmic rays is measured by a large-area detector on ground, where its energy from 100 TeV up to 1 EeV is observed by KASCADE and KASCADE-Grande, respectively. The successfully measured raw data of air showers and accompanying simulated data are saved in the databank. Using the software tools these original shower data are made publicly available in the web portal, which is an interface between the data archive, the data centre software and the user. In the frame of the first release, KCDC already provides efforts to fulfill following basic requirements:

- To provide the experimentally reliable data for the scientific community to answer many unsolved scientific questions.
- To serve as a long term data archive as part of a multi-messenger data basis for the astrophysics community.
- To provide a basis for education and outreach to the general public.
- To develop, improve, optimize, and publish the software packages behind the data centre.

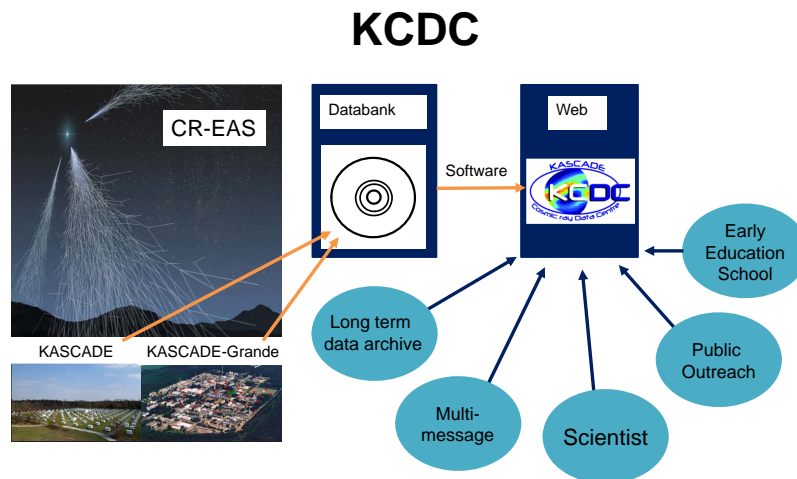


Figure 2: Schematic view of KCDC including various aims.

3. New release NABOO

In a first release in 2013, KCDC has published 160 Million air-shower events from the original KASCADE experiment providing 15 reconstructed parameters per event, including the reconstructed primary energy and direction of the impinging cosmic ray, but also information on the measurement itself, like the detected number of muons and electrons on the ground. The data storage was based on the MySQL database.

The second large release of KCDC was VULCAN about a year after the first release. The main change was the transition of the MySQL database to a mongoDB to store the data, since the usage of MongoDB is introduced to smoothen the way for adding more measurement parameters in the next release MERIDIAN. In addition, we added the new quantities of the number of reconstructed hadrons and the sum of their energies deposited in the KASCADE central calorimeter. The number of air-shower events provided to the public stayed the same. Documentations of the experiment and the KCDC manual were included as well.

In the new release MERIDIAN in 2016, we have added 3 quantities represented by arrays. This required a complete redesign of the Data Shop and of the data format offered. Thus the back-end as well as the front-end of the Data Shop have been rewritten. The back-end of KCDC is now based on a modular system, where newly added functionality can be implemented via plugins. This has been done in preparation of the planned software release.

A new version named NABOO was released on February 2017. NABOO is based on a completely new Data Shop back-end and offers more than twice as many events as former releases. Additionally, data from the KASCADE-Grande detector component have been added. The Data Shop is now a plugin system based on a newly written software package, which will also be made available for public usage.

A detailed description of the new and modified features follows below:

- We added about 290.000.000 events, so that now all 433.000.000 events recorded by the KASCADE and KASCADE-Grande experiment during 20 years of data taking are available for public usage.
- The newly added detector component is the KASCADE-Grande (called Grande below) detector array with 9 new quantities:
 - Xc: X-axis shower core position as reconstructed by Grande
 - Yc: Y-axis shower core position as reconstructed by Grande
 - Ze: Zenith angle as reconstructed by Grande
 - Az: Azimuth angle as reconstructed by Grande
 - Nch: Number of charged particles in Grande
 - Nmu: Number of muons in Grande
 - Age: Shower age parameter as reconstructed by Grande
 - GDeposit: Energy deposits of all charged particles in Grande per station
 - GArrival: Arrival time of the first particle in Grande per station

- The former published quantities e/γ -density of charged particles per m^2 and μ -density have been replaced by e/γ - and μ -energy deposit in MeV, respectively, which are directly measured and thus offer more flexible analysis.
- The preselections have been updated to the new data sets and extended.
- We changed our storage concept in mongoDB to speed up the processing of the requests.
- The KCDC-Manual has been extended.
- We cleared inactive accounts and user histories according to privacy protection guidelines.

4. Web portal

Figure 3 shows the index page of the current KCDC version called NABOO. It shows the main menu on the left side, which is fully customizable within the administrator interface. The first two pages cover the motivation and legal regulations concerning the usage of the data as well as the web portal. The information page contain documentation on KCDC, KASCADE, and the data format which are published. The documentations on the web page are still under development, however, there is a KCDC manual, which describes more details the KCDC, KASCADE and the data. There are pages for announcements and collections of frequently asked questions of common interest. Moreover, pages for the users to manage their account, for reporting a bug, and for publications related to KCDC.

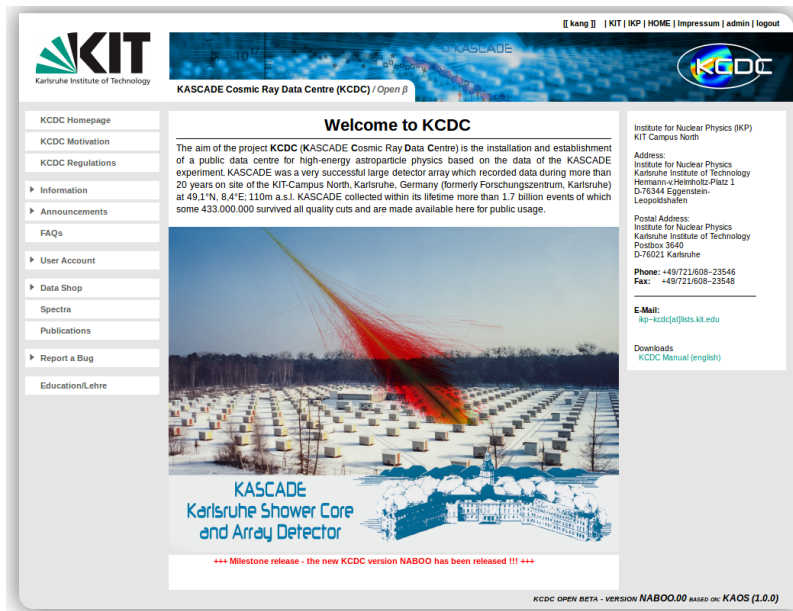


Figure 3: The index page of the website of KCDC, the KASCADE Cosmic-ray Data Centre (<https://kcdc.ikp.kit.edu>).

5. Data Shop

After an user account has been created, the user gets an access to enter the page of the data shop. The registration is required to ensure that the 'End User License Agreement' ¹ is read, i.e. the legal aspects of public data are understood. In the data shop, the user can either download the preselected data, or create individual selection. When a request for a subset of the available data has been submitted, it is processed asynchronously in the background. The user will get a notification when the processing of the job has finished, and can download the data which is kept available for two weeks.

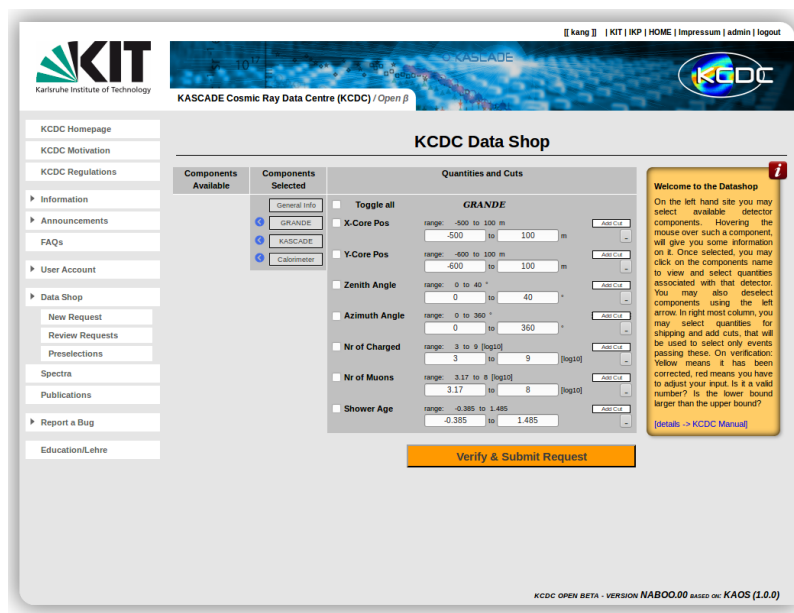


Figure 4: The page of the Data Shop of KCDC, showing the parameters per air shower to be selected.

The page of the data shop in its current version is shown in Fig. 4. In the left column, the available detector components are listed and the selected components are listed in the second column. The General Info in the second column contains the information of time, run and event number. The third column contains a list of quantities belonging to the currently selected detector component (e.g. Grande in Fig. 4). In addition, the user can add the selection criteria that are to be applied during the process of the request. The available range for each parameter is also displayed. After selecting quantities and defining selection criteria, it is possible to choose the file format which will be used to store the selected data. Three formats are currently supported: ASCII, ROOT [5] and HDF5 [6]. ASCII files do not require the usage of special libraries. ROOT and HDF5 files are commonly used in high-energy physics and often in astronomy communities, respectively. A detailed description of the software and the individual selection can be found in Ref. [7].

¹There is no standard procedure yet available for open data publication. In cooperation with KIT and its law department, we developed an own license based on the EULA (End User License Agreement) model adapted from that one often used for software.

6. Education

An important goal of the public data release is to be conveyed to a broad audience including all levels of education to fulfill the requests from the "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities" [8].

The goal of having detailed tutorials, i.e. an 'education portal', is also connected with having a visible outreach of astroparticle physics. The target groups for the tutorials are teachers and pupils in high schools or students in early stages. This means a tutorial has to provide at least:

- i a basic knowledge on cosmic ray physics, astrophysics and related topics;
- ii the required software and KCDC data (preferably as a preselection);
- iii a gradual explanation of a simple data analysis;
- iv a modern programming language code example;
- v the interpretation and discussion of the result.

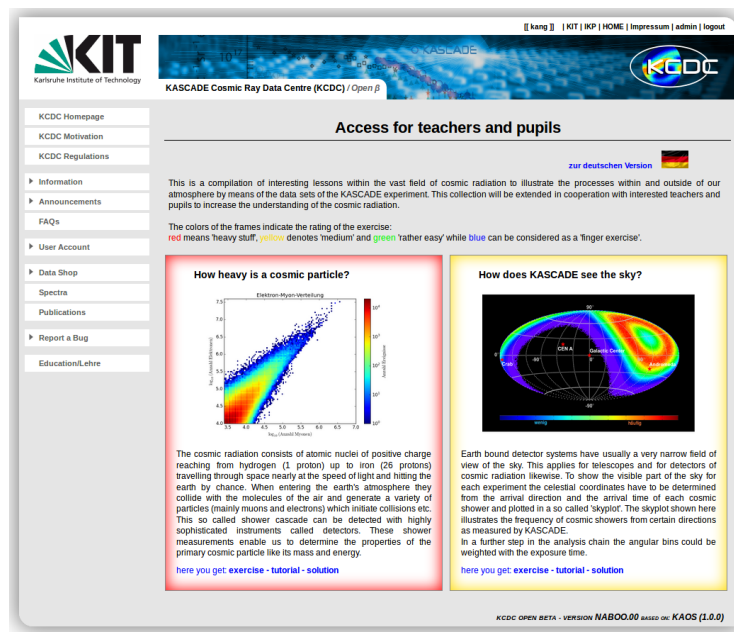


Figure 5: The page of the educational section of KCDC. The German and English versions are presently available and also other languages will be added in the near future.

Two examples of a simple data analysis for school students already exist in the web portal KCDC and they are shown in Fig. 5. The first example (left in Fig. 5) is the determination of the characteristic of cosmic rays, such as a mass or energy. To determine the mass and energy of the original cosmic rays particles, the measured number of electrons and muons on ground has to be compared, since these numbers depend on the shower development of the original cosmic ray particles. The second one (right in Fig. 5) is to generate a skymap from the measured directions

of cosmic rays. In order to represent the visible area in the sky, the sky coordinate has to be determined from the measured incidence direction and from the arrival time of the particles. The distribution of the directions can then be displayed in a skymap corresponding to the sky visible by KASCADE. Tutorials for bachelor or master students will be more sophisticated and complicated in comparison to those for pupils, e.g. the reconstruction of the air showers, anisotropy studies, etc. From the education portal, pupils or undergraduate students can do a simple data analysis, i.e. download the measured data, apply selection criteria, select reasonable events and obtain the result, so that they can be earlier adopted to the field of astroparticle physics.

7. Future plan

KCDC currently has more than 120 registered users demonstrating the acceptance of the portal among researchers as well as the general public. News are distributed via an Email to the users and among other activities, by a KCDC Twitter account (#KCDC_KIT). The basic features for a data providing web portal are already implemented in KCDC, however, there are several tasks and ideas for further development and improvement of KCDC. In particular, the educational portal has to be extended with more examples and in various languages.

In the new release of NABOO, data from the KASCADE-Grande detector component have been included to cover a larger part of the energy spectrum. For deeper investigations of the air-shower parameters, e.g. for composition analyses, full simulations of individual events are necessary. Therefore, as the first step, a new release with the inclusion of the full air-shower simulations will be evaluated and planned for the future. In addition, the publication of the accompanying software tools for open access will be achieved.

Another plan for the future is to open KCDC for another type of shower data. Radio data from the LOPES experiment [9], which was co-located with KASCADE, will be included. Due to the different observation technique, the data structure from the LOPES antennas as well as calibration procedures are different from the ground-based KASCADE experiment, as well as the entire data analysis. Hence, an adoption of the data platform is required in direction of further generalization of KCDC.

References

- [1] <https://kcdc.ikp.kit.edu>
- [2] A. Haungs et al., KASCADE-Grande collaboration, J. Phys. Conf. Ser. 632 (2015) 012011
- [3] T. Antoni et al., KASCADE collaboration, Nucl. Instr. Meth. A513 (2003) 490
- [4] W.D. Apel et al., KASCADE-Grande collaboration, Nucl. Instr. and Meth. A620 (2010) 202
- [5] R. Brun and F. Rademakers, Nucl. Instr. Meth. A389 (1997) 81; <https://root.cern.ch>
- [6] The HDF Group, Hierarchical Data Format; <https://www.hdfgroup.org/hdf5/>
- [7] S. Schoo, Doctoral Dissertation, KIT Karlsruhe (2016); DOI:10.5445/IR/1000055797
- [8] https://en.wikipedia.org/wiki/Berlin_Declaration_on_Open_Access_to_Knowledge_in_the_Sciences_and_Humanities
- [9] H. Falcke et al., Nature 435 (2005) 313