

Data Management Plan

Experiment HAWC

Experiment description:

The High Altitude Water Cherenkov (HAWC) gamma-ray observatory is located at 4100 m above sea level near Pico de Orizaba in central Mexico. HAWC consists of 300 water tanks each 7.3 m in diameter and 4.5 m deep and containing 4 large area photomultiplier tubes. The tanks are close packed in order to detect the lowest energy extensive air showers whose directions are reconstructed to < 1 degree by timing the shower front. The deep tanks allow muons to be identified in order to distinguish the hadronic cosmic ray background from the gamma-ray signals. The wide field of view of ~ 2 sr and the continuous operation allow for $2/3$ of the sky to be surveyed each day. HAWC is the most sensitive gamma-ray observatory above ~ 10 TeV. The science goals of HAWC include the search for dark matter annihilation as well as the origin and acceleration mechanisms of the highest energy particles known. HAWC began full operations in March 2015 and plans to continue operations for 5 years.

DOE's roles in the experiment:

DOE HEP procured $1/3$ of the water Cherenkov tanks and the laser calibration system. DOE HEP also funded the Deputy Project Manager for Construction and the Operations Manager, both of which are Brenda Dingus at Los Alamos National Lab. DOE also funds science investigations in fundamental physics phenomena with HAWC, e.g. indirect dark matter searches, at Los Alamos National Lab, University of New Mexico, Rochester University, and Michigan Tech University.

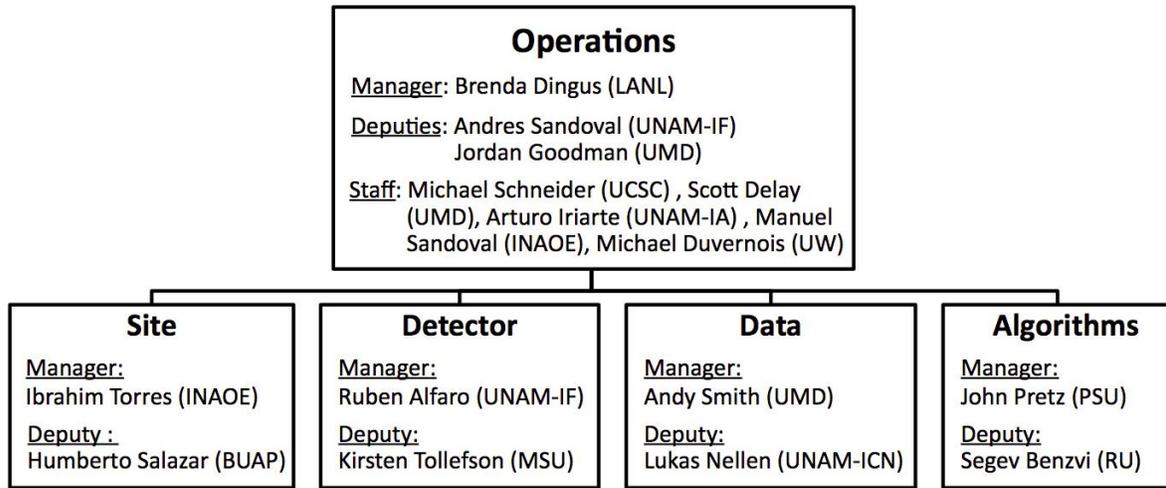
Partnerships:

NSF and CONACYT (Mexican science agency) are also partners in the funding of HAWC construction, operation, and science. There are no MOUs between these agencies. However, NSF is responsible for the data management and Andrew Smith at the University of Maryland is the HAWC data manager.

Organization – Agency/Lab level

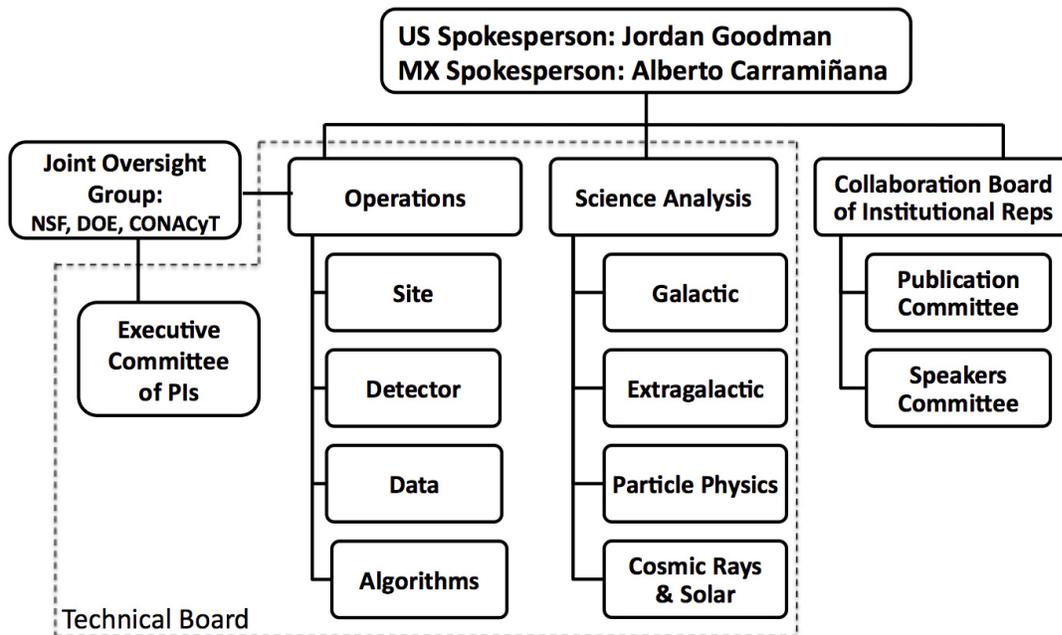
NSF funds the University of Maryland as the lead institution for the construction and operations of HAWC. Los Alamos National Lab is the lead institution that is funded by DOE HEP, and is the employer of the Operations Manager Brenda Dingus. Each institution is separately funded for science investigations and non-managerial operations tasks.

Organization – Experiment level



The chart above shows the experiment level organization and the responsible managers with their institutions. The acronyms are Los Alamos National Lab (LANL), Universidad Nacional Autónoma de México (UNAM), Instituto de Física (IF), Instituto de Ciencias Nucleares (ICN), Instituto de Astronomía (IA), University of California Santa Cruz (UCSC), University of Maryland (UMD), University of Wisconsin (UW), Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Benemérita Universidad Autónoma de Puebla (BUAP), Michigan State University (MSU), Pennsylvania State University (PSU), University of Rochester (RU).

Collaboration:



The chart above shows the overall organization of the which consists of 15 institutions in the US and 15 institutions in Mexico with a total of ~100 scientists. The spokespeople are elected for 2

year terms and are limited to 2 consecutive terms. There is one spokesperson from the US and one from Mexico. A charter that has been approved by the collaboration describes the governance of the collaboration.

Data policy management:

The data management policies are set by the collaboration board, which is composed of a representative of each institution on the collaboration. The data manager is responsible for the archiving and public release of the data.

Data Description & Processing: [SC Req. 3]

The raw data produced by HAWC consists of timing and time-over-threshold information for pulses generated by any of 1200 photomultiplier tubes (PMTs) that are hit by Cherenkov light during an atmospheric shower and the rates of these hits recorded by scalars. The raw data rate is 20MB/sec and is continuously collected. This data is transported by vehicle weekly to UNAM-ICN and transferred to UMD over the internet. A copy of the raw data is kept at both UNAM-ICN and UMD.

Intermediate-level data products include reconstructed air shower directions, core locations, energy, and gamma/hadron likelihood. It also includes simulated data of PMT hits from Monte Carlo simulations. New intermediate-level data products will be produced by re-processing raw data as algorithms and calibrations are improved. Again two copies of this data will be kept with one at UMD and one at UNAM-ICN. These products require approximately 20-30% of the disk space as the raw data depending on how many times the data is reprocessed.

High-level data products will bin the intermediate-level data into sky maps with cuts on different parameters such as energy and gamma/hadron parameter as well as time. Source catalogs will be produced from these maps. The highest-level products will be scholarly papers published in the scientific literature.

Data Products and Releases: [SC Req. 4]

HAWC will provide prompt alerts of transient sources to the Gamma-ray Coordinates Network (GCN), the Astrophysical Multi-messenger Observatory Network (AMON) and IAU Astronomer's Telegram (ATel) in order to enable multiwavelength and multimessenger observations of these transients.

After a proprietary period of 2 years, high-level data products will begin to be released. The first release will be a catalog of detected source followed by gamma-ray and cosmic-ray sky maps. These maps will be updated yearly.

Access to raw and intermediate data will likely require that a research collaboration of some sort be formed or worked out on request through consultation with the collaboration. Note that the size of the anticipated 5-year raw data set for the entire HAWC detector will be approximately

3.5 Petabytes. There is no funding currently available to make other copies of such a large data set.

Plan for Serving Data to the Collaboration and Community: [SC Req. 1]

Skymaps will be provided in FITS format (an astronomical standard format maintained by NASA) on a web server after a period of vetting and analysis by collaboration members which will not exceed 2 years. We will also provide a web-based tool to examine the sky maps and extract HAWC source fluxes or flux upper limits for all areas of the sky.

Plan for Archiving Data: [SC Req. 1]

The HAWC data will be archived at both UMD and UNAM-ICN for at least 3 years after the operations funding from DOE HEP for HAWC ends.

Plan for Making Data Used in Publications Available: [SC Req. 2]

All data points shown in the published graph will also be available in a machine-readable form that can be obtained upon request from any co-author.

Responsiveness to SC Statement on Digital Data Management

This data management plan fully follows SC Statement on Digital Data Management.