

## Track Reconstruction for ISS-CREAM Resulting in Improved Energy and Charge Resolutions

J.P. Lundquist<sup>\*1,2,‡</sup>, Y. Amare<sup>2</sup>, D. Angelaszek<sup>2,3</sup>, N. Anthony<sup>3</sup>, G.H. Choi<sup>1</sup>, M. Chung<sup>2</sup>, M. Copley<sup>2</sup>, L. Derome<sup>4</sup>, L. Eraud<sup>4</sup>, C. Falana<sup>2</sup>, A. Gerrety<sup>2</sup>, L. Hagenau<sup>2</sup>, J.H. Han<sup>2</sup>, H.G. Huh<sup>2</sup>, Y.S. Hwang<sup>5</sup>, H.J. Hyun<sup>5</sup>, H.B. Jeon<sup>5</sup>, J.A. Jeon<sup>1</sup>, S. Jeong<sup>1</sup>, S.C. Kang<sup>5</sup>, H.J. Kim<sup>5</sup>, K.C. Kim<sup>2,3</sup>, M.H. Kim<sup>2</sup>, H.Y. Lee<sup>1</sup>, J. Lee<sup>1,5</sup>, M.H. Lee<sup>2</sup>, C. Lamb<sup>2</sup>, J. Liang<sup>2</sup>, L. Lu<sup>2</sup>, L. Lutz<sup>2</sup>, B. Mark<sup>2</sup>, A. Mechaca-Rocha<sup>6</sup>, T. Mernik<sup>2</sup>, M. Nester<sup>2</sup>, O. Ofoha<sup>2</sup>, H. Park<sup>5</sup>, I.H. Park<sup>1</sup>, J.M. Park<sup>5</sup>, N. Picot-Clemente<sup>2</sup>, S. Rostsky<sup>2</sup>, E.S. Seo<sup>2,3</sup>, J.R. Smith<sup>2</sup>, R. Takeishi<sup>1</sup>, T. Tatoli<sup>2</sup>, P. Walpole<sup>2</sup>, R.P. Weinmann<sup>2</sup>, J. Wu<sup>2</sup>, Z. Yin<sup>2,3</sup>, Y.S. Yoon<sup>2,3</sup>, H.G. Zhang<sup>2</sup>

<sup>1</sup> Dept. of Physics, Sungkyunkwan University, Republic of Korea

<sup>2</sup> Inst. for Phys. Sci. and Tech., University of Maryland, College Park, MD, USA

<sup>3</sup> Dept. of Physics, University of Maryland, College Park, MD, USA

<sup>4</sup> Laboratoire de Physique Subatomique et de Cosmologie, Grenoble, France

<sup>5</sup> Dept. of Physics, Kyungpook National University, Republic of Korea

<sup>6</sup> Instituto de Fisica, Universidad Nacional Autonoma de Mexico, Mexico

Cosmic Ray Energetics And Mass for the International Space Station (ISS-CREAM) has taken 1.5 years of direct measurements of high-energy cosmic ray (HECR) particles for energies from  $10^{12}$  to  $10^{15}$  eV. HECR particle identification is significantly improved by tracking particle-detector interactions from the calorimeter (CAL) back to the Silicon Charge Detector (SCD) for charge determination. A track finding algorithm resistant to such issues as particle multiplicity, backscatter, and electronic noise will be outlined. Also, shown is the energy resolution improvement, and the resulting all particle spectrum, provided by ensuring good particle tracks. This allows ISS-CREAM to investigate how the energy distributions evolve, for protons all the way to iron nuclei, and will provide important information for models of galactic sources and HECR propagation.

*36th International Cosmic Ray Conference -ICRC2019-*

*July 24th - August 1st, 2019*

*Madison, WI, U.S.A.*

\*Speaker.

†jplund@umd.edu

‡for the ISS-CREAM Collaboration list see PoS(ICRC2019)1177

## 1. Introduction

The ISS-CREAM experiment is built upon the successes of seven Antarctic balloon-borne CREAM missions that were flown between 2004 and 2016 ([1], [2]). The ISS-CREAM instrument is an external ISS payload successfully installed, and activated, on the Japanese Experiment Module External Facility 2 (JEM-EF 2) on August 22, 2017 and was deactivated in February 2019.

The intent of the space-based ISS-CREAM mission is to help unravel the mysteries of high-energy cosmic ray (HECR) origins, acceleration, and propagation within the galaxy, by the direct measurement of incident particle composition at energies from  $\sim 10^{12}$  to  $\sim 10^{15}$  eV - and with greater statistics than a balloon experiment. HECR at the high end of this range have previously been measured from ground based arrays, that detect the secondary particle air showers induced by them in the Earth's atmosphere, which means that individual particle compositions are unknown due to air shower model uncertainties and statistical fluctuations.

Ground based all-particle HECR energy spectra measurements have shown a spectral steepening "knee," of the overall power law, at  $\sim 10^{15}$  eV that may be explained by a supernova remnant shock acceleration limit [3]. CREAM data showed spectrum hardening in the TeV range, compared to lower energy data from previous experiments, that may be due to acceleration hardening mechanisms and/or nearby source distribution [4]. Elemental spectra measurements in the ISS-CREAM energy range, where a composition change could be caused by the expected rigidity-dependent supernova acceleration limit, are crucial to understanding cosmic-ray acceleration and propagation.

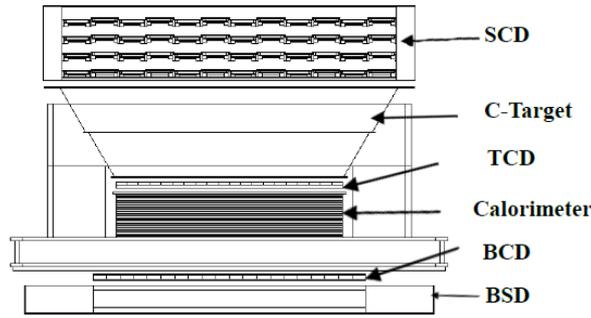
The particle tracking method within the ISS-CREAM detector described in the following sections improves the particle discrimination power, and energy resolution, as shown via Monte Carlo simulation. The pattern recognition noise filter further improves statistics and data quality. A preliminary all-particle distribution of cosmic rays with  $\sim 1$  year of exposure time will be shown.

## 2. ISS-CREAM Instrument

The ISS-CREAM instrument consists of complementary and redundant particle detectors as shown in Figure 1 [5]. An ionization calorimeter (CAL) determines HECR energy, provides particle tracking, and a high energy trigger. The four-layer Silicon Charge Detector (SCD) provides precise charge measurements to identify nuclei from  $1 \leq Z \leq 26$  at a resolution of 0.2e to 0.3e. Top and Bottom Counting Detectors (TCD/BCD) provide an independent lower energy trigger and shower profiles for electron/hadron separation. The Boronated Scintillator Detector (BSD) provides additional electron/hadron discrimination using thermal neutrons produced within the calorimeter.

The CAL is based on the CREAM CAL and is a stack of 20 tungsten plates each one radiation length ( $X_0$ ) thick with an area of  $50 \times 50$  cm<sup>2</sup>, interleaved with layers of scintillating-fiber ribbons 1 cm wide and 50 cm long that alternate directions in each layer. The alternating layers result in two 2d track images in the x and y directions, as the layers stretch across the CAL in a single direction.

The SCD at the top of the instrument, for the identification of cosmic rays, consists of four layers of silicon pixel sensors each with an active area of  $78 \times 74$  cm<sup>2</sup>. A cosmic ray passing through the sensor produces an ionization signal proportional to the square of the particle charge. The pixels  $1.55 \times 1.38$  cm<sup>2</sup> active area is optimized to reduce the effect of CAL backscatter.



**Figure 1:** ISS-CREAM configuration in the Monte Carlo simulations used in this analysis. Taken from [6]

The CAL trigger and TCD/BCD trigger initiate data collection for the entire detector. The high energy CAL hardware trigger requires at least six consecutive layer ribbons to have an energy deposit larger than  $\sim 40$  MeV, and at least one of them greater than  $\sim 76$  MeV. The TCD/BCD trigger is designed for high-charge nucleus events and is created when one channel in the TCD and two channels in the BCD show energy deposits greater than threshold values determined in each channel [7]. Further description of the ISS-CREAM instrument can be found in [8].

### 3. Particle Tracking Method

#### 3.1 CAL Layer/Track Intersection

First, the calorimeter secondary particle energy deposit hits are cleaned of possible noise. Ribbon hits far from the shower core are also removed. Hardware sparsification of the hit ADC signals correspond to a cutoff of  $E < \sim 8$  MeV. Additional sparsification of CAL hits is done by removing those with  $E \leq 9.3$  MeV. If there are more than 25 hits remaining those with energies less than 0.6% of the total energy deposit (after sparsification) are removed. Then if there are still more than 25 hits remaining (in all layers) all but the 25 highest energy hits are removed. Finally, any layer with less than 1% of the total energy deposit is removed.

The first estimate for a particle track's intersection with a layer, the Center-of-Mass (COM), is the average ribbon hit position in the layer weighted by energy deposit. The hit energy sum in each layer is used for the linear regression weight.

The COM second estimate is found by removing hits  $1.41 \times \text{MAD}$  from the first COM and taking the weighted average of remaining hit positions, where MAD is the Mean Absolute Deviation  $\langle |COM1 - p_{hits}| \rangle$  and  $p_{hits}$  is the hit position in the x, or y, direction. MAD is more resistant to outliers than the standard deviation and has less uncertainty than the Median Absolute Deviation which is not appropriate for very low statistics.

#### 3.2 Weighted Repeated Mean Linear Regression

A somewhat novel linear regression method has been developed for this analysis that is robust against outliers and useful for very low statistics. It is a modification of Repeated Median Regression (RMR) which allows up to 50% of the data to be from a different model/process than what is of interest in the linear fit, i.e. noise mixed with a cosmic ray track [9]. The issue with RMR, in this

case, is that medians have an uncertainty generally at least  $\sim 25\%$  larger than the mean (for a large sample Gaussian distribution) - the result is poor resolution as tested via MC simulations [10].

The Weighted Repeated Mean Regression (WRMR) used here is done by first taking the trimmed mean of the slopes between a Center-of-Mass (COM) position and all other COMs. The trimmed mean used is the mean of 74% of the slopes, from each COM position, around the median - this allows for a balance between outlier removal and uncertainty. Then the track slope is the weighted mean of the trimmed means for each COM. The track intercept is the mean COM depth (from the CAL top),  $z$ , minus the slope times COM positions  $wmean(z_{COM} - slope \times P_{COM})$ .

The WRMR was tested against various alterations of the above, Theil-Sen estimation (and versions using weights and mean instead of median) [11], repeated median regression (and a weighted modification) [9], weighted least squares, least absolute residuals, bisquare iterated [12], and weighted RANSAC fitting methods (and STARSAC) [13], and was found to have the best track resolution via MC.

### 3.3 Outlier Removal and Fit Reiteration

An initial WRMR fit of the COMs is done by reweighting by COM energy, and CAL depth,  $w = E_{COM}^{1.11}/z_{COM}^{1/3}$ . This weighs higher energies slightly more and gives greater weight to layers further into the CAL where the shower axis is more stable.

Then the COM with the largest residual distance, from the fit, is removed if  $R > 0.58$  cm and there are more than seven COMs. This is repeated once. If there are still  $N > 7$  COMs, the highest energy remaining COM is set aside, and the mean residual and their mean absolute deviation (MAD) are calculated. The largest COMs are removed if they are larger than  $2.75 \times MAD$  until there's at least 6 COMs left. If COMs were removed the fit is redone with  $E_{COM}^2$  weights.

Finally, this algorithm is repeated once starting from the beginning after removing hits some distance from the current fit. If there are  $N_{hits} > 20$  ( $E > 9.3$  MeV) within 1 cm of the fit then hits outside this distance are cut. Else, if  $N_{hits} > 10$  within 2 cm then hits outside 2 cm are cut. Otherwise, hits greater than 3 cm from the fit are removed. Then the COM estimation and fitting are repeated.

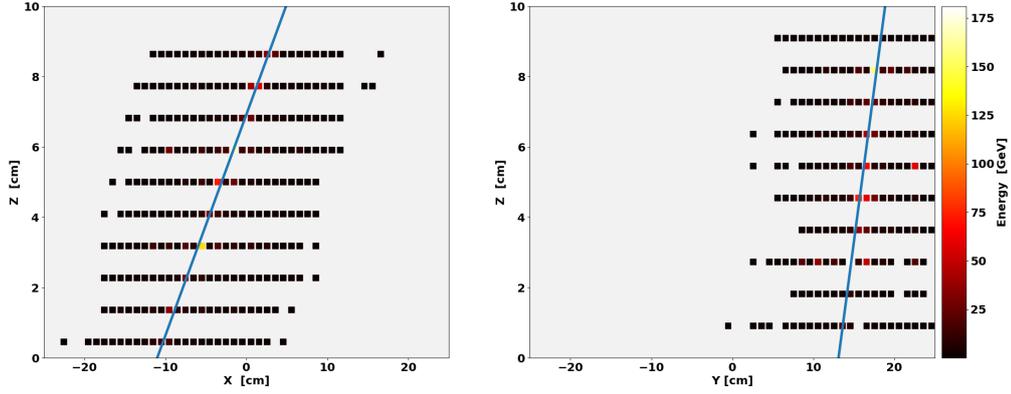
The fit is done for the  $x$ , and  $y$ , directions and the result is the 3d cosmic ray particle CAL track. The highest energy event ( $E \sim 222$  TeV) in this analysis is shown as an example in Figure 2.

All arbitrary parameters in this fitting method were chosen for the best track resolution in 1 TeV proton MC simulations and were found to result in even better resolutions at higher energies.

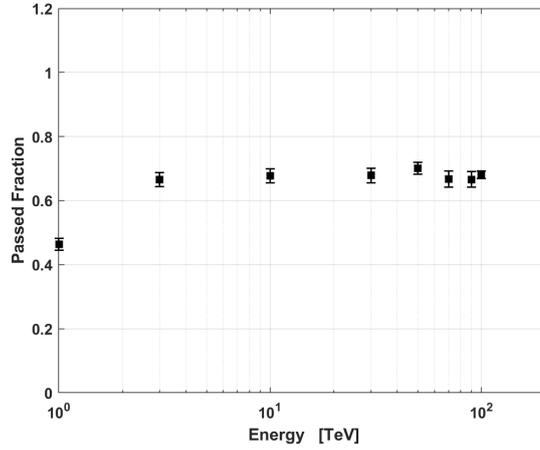
## 4. Event Selection

### 4.1 Software Trigger and Quality Cuts

To increase statistics all hardware triggers were used which necessitates the need for a CAL software trigger. If there are at least 8 COMs in the  $x$ , or  $y$  direction, each with an adjacent layer (they do not have to be consecutive) all with a hit  $E > 15$  MeV the event is "triggered." This is a looser cut than the CAL hardware trigger and was chosen for improved tracking resolution, noise removal and increased statistics. The particle track is also required to be within the CAL geometry. The acceptance of these two cuts is shown in Figure 3 to be stable above 10 TeV.



**Figure 2:** The highest energy event in this analysis ( $E \approx 222$  TeV) and the CAL particle track fit.



**Figure 3:** MC event fraction passed by the software trigger and inside CAL requirement (1 to 200 TeV).

The last cut is on the normalized residual sum (after outlier removal),  $\Sigma R/N_{COM} < 0.28$ . This removes a number of CAL noise events, that are introduced by the low energy T/BCD trigger, and improves energy and track resolutions. No MC simulation events are removed.

#### 4.2 Noise Filters

Further noise filtering is necessary to avoid removing data based on noisy time frames, again due to the use of the low energy T/BCD trigger for calorimeter events. None of the cuts in this section remove events from MC simulations as noise is not included in the MC used. Two straightforward cuts require at least 3 COM layers with an adjacent layer each with a hit  $E > 15$  MeV in and  $x$  and  $y$ , and at least 5 hits of any energy in the  $x$ , and  $y$ , layers each.

Another event filter used is based on entropy and is simply  $\Sigma \log(E_{hit}^2)$  for hit energies in GeV. Highly disordered events are removed by requiring Entropy  $\geq -1955$  in  $x$ , and  $y$ , layers. Highly ordered events from anomalous channel readouts are cut by requiring Entropy  $< 100$  in  $x$ , and  $y$ .

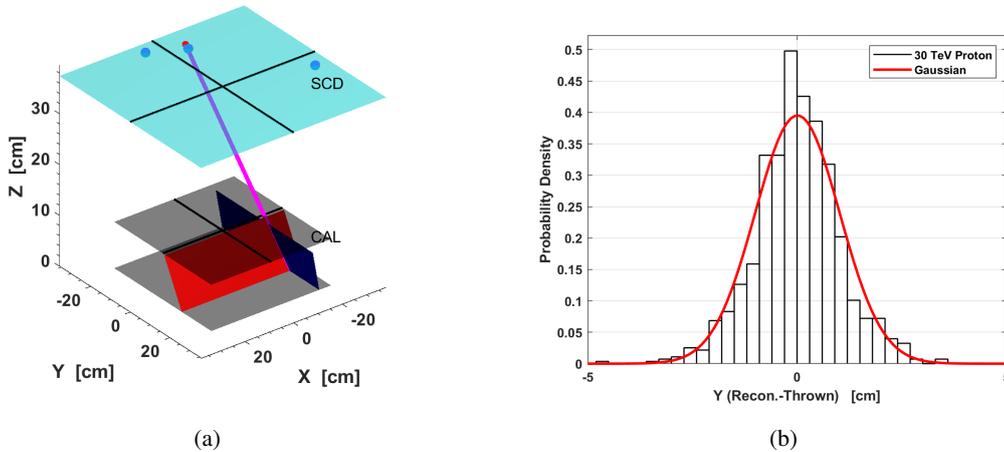
The final noise filters are done via simple pattern recognition. All hit positions  $E \geq 9.8$  MeV (Figure 2) are set to one and all others zero. Adjacent hits in a layer are subtracted, which results in a matrix with one less column, then another adjacent subtraction is done, resulting in two less columns, and the absolute value is taken. The number of two's in this matrix are counted and if  $N > 100$  the event is considered noise. This removes triggers with many adjacent hits. A third adjacent subtraction is done and the cut threshold of the number of one's ( $N > 90$ ) removes sparse noise triggers, and the number of three's ( $N > 93$ ) removes triggers with many adjacent hits.

## 5. Results

### 5.1 SCD Track Resolution

Backtracking the CAL track to the SCD at the top of the detector allows the identification of incident particle species using SCD hits within a  $3\sigma$  “circle of confusion” around the tracks SCD intersection. An MC example of a CAL track extrapolated to the SCD is shown in Figure 4(a).

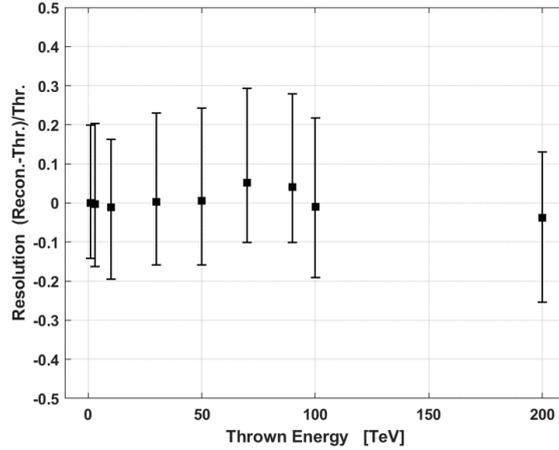
The average  $1\sigma$  resolution, found by fitting Gaussian distributions to the MC reconstructed track intersections minus the thrown/known intersections with the topmost SCD1, is  $\sim 1$  cm (which improves the closer the SCD is to the CAL) over the entire energy range. The 30 TeV proton MC resolution in SCD1 for the y-direction is shown in Figure 5. This corresponds to a  $5 \times 5$  SCD pixel  $3\sigma$  diameter circle of confusion for particle identification. Resolutions of higher Z primary particles should improve even further due to the production of more secondary particles in the CAL.



**Figure 4:** (a) A MC example of a CAL track extrapolated to the SCD. (b) Track reconstruction resolution using 30 TeV proton MC. The resolution is the Gaussian  $\sigma$  of the difference between the track reconstruction SCD intersection subtracted by the thrown/known particle intersection and is 1 cm.

### 5.2 Energy Resolution

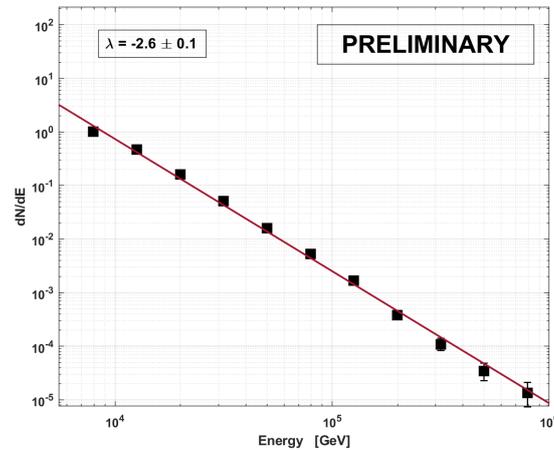
The energy resolution and bias found by using this analysis to reconstruct proton MC ([6]) from 1 TeV to 200 TeV is shown in Figure 5. This shows very small bias and an energy resolution average of  $\pm 18\%$  and does not include the effects of electronics noise or CAL calibration uncertainty which are estimated to add an additional  $\sim 10\%$  to the resolution [14].



**Figure 5:** The proton MC energy resolution and bias  $(E_{recon} - E_{thrown})/E_{thrown}$  over the 1 TeV to 200 TeV energy range. This shows very small bias and an energy resolution average of  $\pm 18\%$ .

### 5.3 All-particle Energy Distribution

The resulting all-particle energy distribution  $dN/dE$  is shown in Figure 6. The power law spectral fit index,  $\lambda = -2.6 \pm 0.1$  (95% uncertainty), is in good agreement with previous experiments.



**Figure 6:** The resulting all-particle energy distribution  $dN/dE$  of this analysis. The power law spectral fit index (first low energy bin excluded),  $\lambda = -2.6 \pm 0.1$  (95% fit uncertainty), is in good agreement with previous experiments.

The absolute flux energy spectrum is not shown as the low energy efficiency is not fully known and the CAL ribbon by ribbon energy calibration corrections are not yet fully completed [15]. Gain stitching combining the low, mid, and high ADC readout ranges is not used here which will improve the energy resolution. Also, deconvolution of the energy bins remain to be done, that is a correction for the  $N$  events in each bin taking into account the Lorentzian tail of the energy resolutions that cause bin spillover.

## 6. Summary

Incident cosmic ray particle tracking in the ISS-CREAM calorimeter using a new Weighted Repeated Mean Regression (WRMR) method, in addition to noise filters, has resulted in an SCD backtracking resolution of  $\sim 1$  cm, for improved particle identification, and an improved energy resolution of  $\pm 18\%$ . The use of all hardware triggers, a software trigger, and pattern recognition noise filters also improves statistics.

Future updates should include power-law MC distributions, other MC compositions, the all-particle absolute spectrum, and the absolute spectra of individual particles. Additionally, the tracking resolution can be further improved by including particle hits in the T/BCD and the SCD.

## Acknowledgements

This work was supported in the U.S. by NASA grant NNX17AB41G, in Korea by National Research Foundation grants 2018R1A2A1A05022685 and 2018R1A6A1A06024970, and their predecessor grants. It was also supported in France by IN2P3/CNRS and CNES and in Mexico by DGAPA-UNAM project IN109617. The authors thank NASA GSFC WFF and its contractors for engineering support and project management, JSC ISS Program Office for launch support and ISS accommodation, MSFC for the operational support, and KSC and SpaceX for the launch support.

## References

- [1] Seo, E. et al., *Advances in Space Research* **33** (2004) 1777.
- [2] Kim, K. C. et al., *PoS ICRC2017* (2018) 182.
- [3] Gaisser, T. K., *J. Phys. Conf. Ser.* **47** (2006) 15.
- [4] Ahn, H. S. et al., *Astrophys. J.* **714** (2010) L89.
- [5] Seo, E. et al., *Advances in Space Research* **53** (2014) 1451.
- [6] Wu, J. et al., Monte carlo simulations of the ISS-CREAM instrument, in *Proceedings, 36th ICRC, Madison, WI July 24th - August 1st, 2019*.
- [7] Hyun, H. et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **787** (2015) 134, New Developments in Photodetection NDIP14.
- [8] Smith, J. R. et al., *PoS ICRC2017* (2018) 199.
- [9] Siegel, A. F., *Biometrika* **69** (1982) 242.
- [10] Rider, P. R., *Journal of the American Statistical Association* **55** (1960) 148.
- [11] Sen, P. K., *Journal of the American Statistical Association* **63** (1968) 1379.
- [12] Street, J. O., Carroll, R. J., and Ruppert, D., *The American Statistician* **42** (1988) 152.
- [13] Jongmoo Choi and Medioni, G., Starsac: Stable random sample consensus for parameter estimation, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 675–682, 2009.
- [14] Kim, K. C. et al., On-orbit performance of the ISS-CREAM calorimeter, in *Proceedings, 36th ICRC, Madison, WI July 24th - August 1st, 2019*.
- [15] Picot-Clément, N. et al., *PoS ICRC2017* (2018) 247.