PROCEEDINGS OF SCIENCE

Investigating high-energy neutrinos from blazars with a maximum-likelihood analysis of the IceCube Observatory data

Eleonora Barbano,[a,*] **Sara Buson,**[a] **Gaëtan Fichet de Clairfontaine,**[a] **Lenz Oswald,**[a] **Andrea Tramacere,**[b] **Giulia Illuminati,**[c] **Alessandra Azzollini,**[a] **Vardan Baghmanyan**[a] **and Leonard Pfeiffer**[a]

[a] *Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg,*
Emil-Fischer-St. 31, D-97074, Würzburg, Germany,

[b] *Department of Astronomy, University of Geneva,*
Chemin Pegasi 51, CH-1290 Versoix, Switzerland

[c] *INFN - Sezione di Bologna, Viale Berti-Pichat 6/2, 40127 Bologna, Italy*

E-mail: eleonora.barbano@physik.uni-wuerzburg.de

In the past decade, the IceCube observatory has established the presence of a diffuse flux of high energy neutrinos ($\geq$ 100 TeV to 10 PeV) that is consistent with an astrophysical origin [1], [2]. The population of sources responsible for this flux remains largely unknown.
Among the candidate sources of neutrinos, blazars have recently been suggested as promising emitters of the high-energy events detected by IceCube [3]. Our recent studies have provided evidence of a statistically significant spatial correlation between blazars from the 5th Roma-BZCat catalog (5BZCat) [7] and the IceCube southern [4], [5] and northern [6] celestial hemisphere data.
In this contribution, we present a Python-based tool: IceCubePy, that performs an extended unbinned maximum likelihood on the recently released public IceCube's 10-year neutrino point source event sample [8]. Upon its development and testing phase, IceCubePy will be released publicly as an open source and user-friendly code.

38th International Cosmic Ray Conference (ICRC2023)
26 July - 3 August, 2023
Nagoya, Japan

ICRC2023
38th International Cosmic Ray Conference
The Astroparticle Physics Conference

PoS(ICRC2023)1540

---

[*] Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

1. Introduction

The recent discovery of a diffuse high-energy astrophysical neutrino flux by the IceCube Collaboration [1], opens a new research field in astrophysics and astroparticle physics. Identifying the astrophysical origin of this high-energy component, represents one of the most compelling challenges. Among the plausible sources of high-energy neutrinos there are blazars, a class of active galactic nuclei (AGN). AGN are extremely bright objects in the universe and consist of a compact region that contains a supermassive black hole surrounded by an accretion disk and emit a relativistic jet of plasma. AGN, among which blazars, are thought to be good candidates as astrophysical neutrino emitters according to theoretical models (see [10, 11], and references therein). Relativistic jets could be capable to accelerate hadrons (and leptons) at very-high energy. The presence of external radiation fields (e.g. from the broad-line-region), can lead to the production of neutrinos trough photo-hadronic processes. A first potential association of the γ-ray bright blazar TXS 0506+056 with putative neutrino emission (chance probability at the 10^{-4} level [3]) has encouraged the scientific community to investigate in that direction. Other parallel works from our group are further exploring this path [22], [23].

The IceCube Neutrino Observatory, which has been operational for over 10 years, is currently the most sensitive detector for high-energy neutrinos. It is a large-scale experiment that extends nearly 3 kilometers into the Antarctic ice and consisting of 86 strings equipped with optical sensors. Composed of photo-multiplier tubes, these sensors can detect the Ĉerenkov light produced by charged particles resulting from the interaction of neutrinos with the ice nuclei through weak-force interactions. However, the number of astrophysical neutrinos tracked by the detector is only a small fraction, approximately 0.1%, of the sample. This latter is largely dominated by background noise, caused primarily by atmospheric neutrinos and muons. Therefore, it is crucial to develop a robust statistical framework that can pinpoint possible clusters of astrophysical neutrino events in the sky and associate them with their potential sources. The maximization of the extended unbinned likelihood is a widely used statistical method [21]. In the following sections, we present this method used by currently available codes, such as SkyLLH [15], as well as in the code under development IceCubePy, presented in this proceeding. IceCubePy is designed to be an open-source and user-friendly framework specifically tailored for analyzing the publicly available 10-year neutrino point source event sample [8] released by the IceCube collaboration in 2021.

2. Maximization of the extended unbinned likelihood

The likelihood quantity $\mathcal{L}$ is build as the joint probability density function (PDF) f of observing N events x_i in the dataset with certain properties defined by the parameters $\vec{\theta}$. The closer the parameters associated $\vec{\theta}$ are to the real ones $\vec{\theta}^*$, the higher the likelihood value is. In experiments following a Poisson distribution, the extended likelihood formalism is used,

$$\mathcal{L}\left(\vec{\theta}\right) = \exp\left(-\lambda\right) \prod_{\mathrm{i}=1}^{N} \lambda \cdot f\left(x_\mathrm{i}|\vec{\theta}\right) , \tag{1}$$

where λ is the expectation number of detected events.

PoS(ICRC2023)1540

The IceCube Collaboration released a public 10-year neutrino point source event sample in 2021 [8]: a subsample of track-like events collected between 2008 and 2018. The expected number of astrophysical neutrinos is only a small fraction of the dataset, approximately 0.1%, given the large contamination due to the background, primarily caused by atmospheric neutrinos and muons. In order to find a possible link between astrophysical neutrinos and candidate cosmic sources, it aims testing two hypotheses in each region of interest in the sky; the background hypothesis $\mathcal{H}_0$ and the signal hypothesis $\mathcal{H}_1$. This latter is the assumption of having in the tested position $\vec{x}_s$ the presence of background events and also an astrophysical point source, emitting n_s neutrinos following a power-law with spectral index γ,

$$\frac{\mathrm{d}\phi}{\mathrm{dE}_\nu} = E^{-\gamma}\,. \tag{2}$$

The PDF $f\left(x_\mathrm{i}|\vec{\theta}\right)$ for each neutrino event x_i is a weighted sum of the signal term $\mathcal{S}_\mathrm{i}$ and the background term $\mathcal{B}_\mathrm{i}$,

$$f\left(x_\mathrm{i}|\vec{\theta}\right) = \frac{n_\mathrm{s}}{N}\mathcal{S}_\mathrm{i} + \frac{N - n_\mathrm{s}}{N}\mathcal{B}_\mathrm{i}\,. \tag{3}$$

Under the signal hypothesis, the extended likelihood function may be written as,

$$\mathcal{L}\left(\vec{\theta}\right) = \exp\left(-N\right)\prod_{\mathrm{i}=1}^{N}\frac{n_\mathrm{s}}{N}\mathcal{S}_\mathrm{i} + \left(1 - \frac{1}{N}\right)\mathcal{B}_\mathrm{i}\,. \tag{4}$$

The signal $\mathcal{S}_\mathrm{i}$ and the background $\mathcal{B}_\mathrm{i}$ PDFs can be factorized in two independent spatial (S_i, B_i) and energy ($\mathcal{E}_{\mathcal{S}\mathrm{i}}$, $\mathcal{E}_{\mathcal{B}\mathrm{i}}$) terms. The spatial term S_i of the signal PDF is, as in [8, 9, 12, 21], modeled as a bivariate Gaussian function. While the spatial term B_i of the background PDF is modeled as the density of neutrino events in stripes equally spaced in $\sin\left(\delta\right)$ in the celestial sky. We have,

$$S_\mathrm{i} = \frac{1}{2\pi\sigma_i^2}\exp\left(-\frac{|\vec{x}_\mathrm{i} - \vec{x}_\mathrm{s}|^2}{2\sigma_\mathrm{i}^2}\right), \tag{5}$$

$$B_\mathrm{i} = \frac{N_\nu \in \delta_\mathrm{i} \pm [\delta_\mathrm{min}; \delta_\mathrm{max}]}{2\pi\cdot\left(\sin(\delta_\mathrm{max}) - \sin\left(\delta_\mathrm{min}\right)\right)}, \tag{6}$$

where σ_i is the estimated angular error of the event $\vec{x}_\mathrm{i}$.

$\mathcal{E}_{\mathcal{S}\mathrm{i}}$ and $\mathcal{E}_{\mathcal{B}\mathrm{i}}$ are defined from bidimensional histograms of the energy proxy of the muon event E_i and $\sin\left(\delta_i\right)$ [18]. The background energy term $\mathcal{E}_{\mathcal{B}\mathrm{i}}$ is obtained from the real data, while the signal energy term $\mathcal{E}_{\mathcal{S}\mathrm{i}}$ can be obtained trough *Monte Carlo* simulations. The likelihood function for the signal hypothesis includes two free parameters, the spectral index γ and the emitted signal neutrinos n_s, while the background hypothesis assumes $n_\mathrm{s} = 0$,

$$\mathcal{L}\left(\vec{x}_\mathrm{i}, E_\mathrm{i}|n_\mathrm{s}, \gamma\right) = \exp\left(-N\right)\prod_{i=1}^{N}\left(\frac{n_\mathrm{s}}{N}\mathcal{S}_\mathrm{i}\left(|\vec{x}_\mathrm{i} - \vec{x}_\mathrm{s}|, \mathrm{i}; \gamma\right) + \left(1 - \frac{n_\mathrm{s}}{N}\right)\mathcal{B}_\mathrm{i}\left(\sin\delta_\mathrm{i}, E_\mathrm{i}\right)\right), \tag{7}$$

$$\mathcal{L}\left(\vec{x}_\mathrm{i}, E_\mathrm{i}|n_\mathrm{s} = 0\right) = \exp\left(-N\right)\prod_{\mathrm{i}=1}^{N}\mathcal{B}_\mathrm{i}\left(\sin\delta_\mathrm{i}, E_\mathrm{i}\right)\,. \tag{8}$$

PoS(ICRC2023)1540

The test statistic ($\mathcal{TS}$) is calculated to assess the validity of the hypothesis, for best parameters $\hat{n}_s$ and $\hat{\gamma}$. For the Wilk's theorem [14],

$$\begin{aligned}\mathcal{TS} &= -2\log\left(\frac{\mathcal{L}(n_s=0)}{\mathcal{L}(\hat{n}_s,\hat{\gamma})}\right) ,\\ &= 2\sum_{i=1}^{N}\log\left(1+\frac{\hat{n}_s}{N}\left(\frac{S_i(|\vec{x}_i-\vec{x}_s|)}{B_i(\sin\delta_i)}\times\frac{\mathcal{E}_{\mathcal{S}i}(\sin\delta_i,E_i;\hat{\gamma})}{\mathcal{E}_{\mathcal{B}i}(\sin\delta_i,E_i)}-1\right)\right) .\end{aligned} \tag{9}$$

The $\mathcal{TS}$ value can be converted to p-value, which represents the probability of observing a discrepancy from the background,

$$p\text{−value} = \int_{\mathcal{TS}_{\text{obs}}}^{\infty} g\left(\mathcal{TS}|\mathcal{H}_0\right) , \tag{10}$$

where g is the normalized distribution of test statistic founded under the assumption $\mathcal{H}_0$.

IceCubePy under development is based on this formalism and is written in Python and optimized to produce full sky maps or maps around regions of interest. As part of the development process of such framework, the open-source Python3-based tool SkyLLH [15], is used for cross-checking purposes.

3. SkyLLh

SkyLLH [15] is a Python-based tool for likelihood analyses. It is available as an open-source project on the IceCube Collaboration's GitHub repository [16]. The tool provides a modular framework for implementing the likelihood functions and performs the maximization of the unbinned extended likelihood. The available tutorial explains how to handle the public 10-year IceCube point-source event sample and includes an example applied to the position of TXS 0506+056. In our analysis, we used SkyLLH to perform a point-source likelihood analysis in a 20×20 matrix (where each pixel covers 0.2 squared degrees) at the position of the latter blazar.

The result is a $\mathcal{TS}$ map, shown in Fig. 1, covering $4° \times 4°$ centered in the vicinity of TXS 0506+056. The coordinates of the blazar are highlighted with a red cross. The green dot represents the position of the best-fit $\mathcal{TS}$, that corresponds to the value $\mathcal{TS}_{\text{max}}$. The map is realized with 2 degrees of freedom: n_s and γ. The contours for 68%, 90% and 95% of confidence, here in dashed, point-dashed and continuous line, are also shown. They are realized by filling a map calculating $\mathcal{TS}_{\text{max}} - \mathcal{TS}$ in each pixel. It is then evaluated the right end of integration for a probability density function of $\chi^2_{2\text{dof}}$ by taking the 68%, 90% and 95% of the distribution. Finally, it is highlighted with a line the correspondent position of pixels for which $\mathcal{TS}_{\text{max}} - \mathcal{TS}$ is equal to the evaluated quantities. The angular distance between the position of TXS 0506+056 and the best-fit $\mathcal{TS}$ value is of 0.21°, compatible within the error of the map resolution and with the median source offset of the best-fit position provided for the blazar (0.21°) [19].

We are developing an open-source and user-friendly framework, IceCubePy, with the aim of allowing full reproducibility of the results. An initial example will be presented in the next section, with comparison with SkyLLH to cross-check our preliminary results.

PoS(ICRC2023)1540

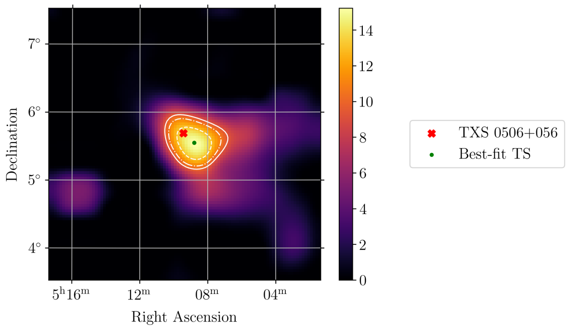


Figure 1: $\mathcal{TS}$ map centered in the vicinity of the position of TXS 0506+056 realized with SkyLLH. The red cross represents the position of the blazar, the green dot represents the position of best-fit for the $\mathcal{TS}$. The contour lines are shown respectively by a white dashed line for 68%, a dashed-dot line for 90% and by a continuous line for the 95% of containment levels.

4. IceCubePy

The framework presented here is a Python-based, object-oriented code that aims to implement the maximization of the extended unbinned likelihood detailed in the previous sections. It is designed to analyze the publicly available IceCube's 10-year neutrino point source data. IceCubePy computes the maximization of the likelihood taking into account either the spatial term or the spatial term along with the energy term [21]. To add the latter term, we simulated the emission of neutrino events [20] from point-like sources incorporating the Poisson distribution to emulate the IceCube detector response. The instrument response function (IRFs) of the instrument, provided by the IceCube collaboration along with the 10-year neutrino point source dataset, is also taken into account. To date, the framework provides test statistic maps, as future development we plan to include the possibility to also generate p-value maps.

5. Results

In this section, we will present the first preliminary results of our framework, IceCubePy, along with a comparison with SkyLLH.

Fig. 2 illustrates a background test statistic sky map created using our framework IceCubePy. The code utilizes the Python library Healpy to generate and visualize the celestial coordinates of the analysis. In this case, the sky map has a resolution of NSIDE = 256, resulting in a pixel coverage of approximately 0.05 square degrees. The data used to generate this map are from periods between 2011 and 2018, representing events detected in the final configuration of the IceCube experiment with all its 86 strings deployed. To construct the background maps, the data were randomized in right ascension while preserving the distinctive distribution in declination. This randomization effectively disrupts the presence of any astrophysical signals while retaining the inherent structure of the dataset. At the current stage of development, it takes approximately 4 hours to generate such a map assuming a fixed value of $\gamma = -2.0$.

The code provides the option to apply a close-up view in the test statistic map, allowing for a more detailed analysis on the region of interest.

PoS(ICRC2023)1540

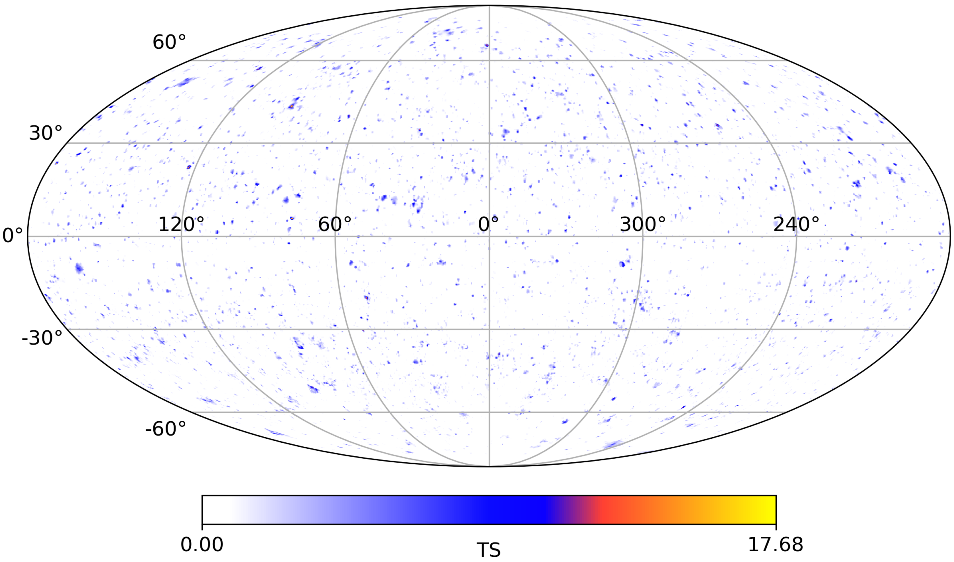
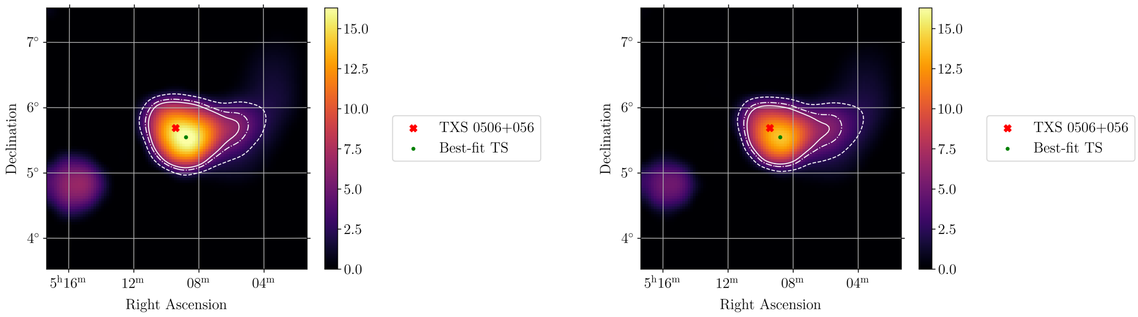


Figure 2: Background sky map realized with our framework, selecting the events collected between 2011 and 2018 from the IceCube 10-year dataset. The $\mathcal{TS}$ map is done with the Healpy library, choosing a NSIDE resolution of 256 and fixing spectral index to $\gamma = -2.0$.

As preliminary investigation, we analyzed the region of TXS 0506+056, fixing the spectrum index to −2.0. Figs. 3a shows the $\mathcal{TS}$ map computed with IceCubePy, Fig. 3b shows the same region of interest analyzed with SkyLLH, fixing coherently the spectrum index to −2.0. Both maps are computed performing a point-source likelihood analysis in a 20×20 matrix (where each pixel covers 0.2 squared degrees). The livetime considered for this analysis is the same as that of the background map. However, unlike the background map, the events used for the likelihood analysis in those cases were not randomized in right ascension. Both codes, reconstruct the best-fit $\mathcal{TS}$ in the same sky position (green dot). The blazar and the best-fit $\mathcal{TS}$ position are 0.21° apart. The offset is of the order of the map resolution, i.e. 0.2°. At the best-fit $\mathcal{TS}$ position, with IceCubePy we observe a number of neutrino events ($n_s = 12$), compatible with the one founded with SkyLLH ($n_s = 11$).



(a) Same as Fig. 1, with IceCubePy.

(b) Same as Fig. 1, with the SkyLLH code.

Figure 3: $\mathcal{TS}$ map around the position of TXS 0506+056, assuming $\gamma = -2.0$.

PoS(ICRC2023)1540

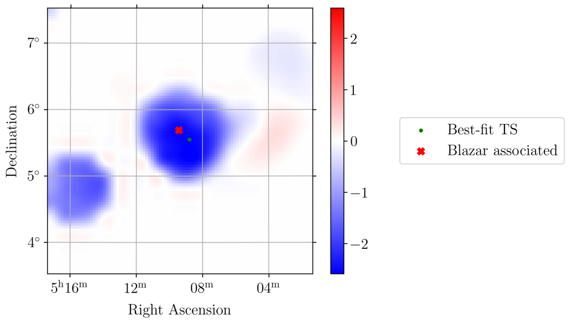


Figure 4: $\Delta\mathcal{TS}$ map, showing the pixel per pixel difference between the $\mathcal{TS}$ maps showed in Fig. 3b and in Fig. 3a.

Fig. 4 represents a preliminary comparison of results obtained with the two codes. In each pixel of the map, it shows the difference between the $\mathcal{TS}$ value estimated with SkyLLH and the one estimated with IceCubePy, for the maps shown in Fig. 3. We observe overeall higher $\mathcal{TS}$ values with IceCubePy than SkyLLH, at the postion of the central $\mathcal{TS}$-cluster (blue regions), suggesting that IceCubePy seems to systematically slightly over-estimate the $\mathcal{TS}$. While this mild discrepancy will require further investigation, the two analyses are overall in agreement. The $\mathcal{TS}$-values are within $\sim 15\%$, as well as their estimates best-fit positions of the central $\mathcal{TS}$-cluster are consistent.

6. Conclusion

In this study, we present IceCubePy, a consistent framework applied to the IceCube's 10-year dataset. Similarly to other frameworks developed before, we utilize the extended unbinned likelihood analysis, which is a robust statistical framework to identify possible clusters of astrophysical neutrino events in the sky and associate them with their potential sources. The code under development is optimized to allow extensive $\mathcal{TS}$ sky map computation, in full or zoom mode. In order to cross-check this code, we use the publicly available SkyLLH code. In this contribution, we present the initial stage in a sequence of cross-validation procedures for the framework currently being developed. Concerning our preliminary comparison, we were able to reproduce the $\mathcal{TS}$ map around the blazar TXS 0506+056, that is a candidate blazar-neutrinos association. We also obtained a number of expected neutrino events compatible with the SkyLLH tool. We reconstructed with both codes the actual position of the blazar within the 95% containment region. Although, discrepancies are seen between the two codes in the $\mathcal{TS}$ values obtained. Those differences are currently studied, and forthcoming analysis will be done to characterize them.

Acknowledgments

The authors are thankful to Dr. Marcos Santander, Dr. James Delaunay and Federico Visintini for the fruitful discussion. This work was supported by the European Research Council, ERC Starting grant *MessMapp*, S.B. Principal Investigator, under contract no. 949555.

PoS(ICRC2023)1540

References

[1] IceCube Collaboration 2013, Science, 342, 1242856.

[2] Aartsen, M. G., Abraham, K., Ackermann, M., et al. 2016, APJ, 833, 3.

[3] IceCube Collaboration, Aartsen, M. G., Ackermann, M., et al. 2018, Science, 361, eaat1378.

[4] Buson, S., Tramacere, A., Pfeiffer, L., et al. 2022, APJL, 933, L43.

[5] Buson, S., Tramacere, A., Pfeiffer, L., et al. 2022, ApJL, 934, L38.

[6] Buson, S., Tramacere, A., Oswald, L., et al. 2023, arXiv:2305.11263.

[7] Massaro, E., Maselli, A., Leto, C., et al. 2015, APSS, 357, 75.

[8] Aartsen, M. G., Ackermann, M., Adams, J., et al. 2020, PRL, 124, 051103.

[9] Abbasi, R., Abdou, Y., Abu-Zayyad, T., et al. 2011, APJ, 732, 18.

[10] Böttcher, M., Reimer, A., Sweeney, K., et al. 2013, ApJ, 768, 54.

[11] Dermer, C. D., Murase, K., & Inoue, Y. 2014, Journal of High Energy Astrophysics, 3, 29.

[12] Aartsen, M. G., Abraham, K., Ackermann, M., et al. 2017, APJ, 835, 151.

[13] IceCube Collaboration, Abbasi, R., Ackermann, M., et al. 2022, Science, 378, 538.

[14] S. S. Wilks 1938, The Annals of Mathematical Statistics, 9(1), 60-62.

[15] Wolf, M. 2019, 36th International Cosmic Ray Conference (ICRC2019), 36, 1035.

[16] IceCube Collaboration 2017, JINST 12, P03012.

[17] IceCube, Abbasi, R., Ackermann, M., et al. 2022, 37th International Cosmic Ray Conference, 1073.

[18] R. Stein, Search for Multi-Messenger Transients with IceCube and ZTF, p.69.

[19] IceCube, Abbasi, R., Ackermann, M., et al. 2022, 37th International Cosmic Ray Conference, 1138.

[20] Negro, M., Crnogorčević, M., Burns, E., et al. 2023, APJ, 951, 83.

[21] Braun, J., Dumm, J., De Palma, F., et al. 2008, Astroparticle Physics, 29, 299.

[22] Azzollini et al. in prep.

[23] Fichet de Clairfointaine et al. in prep.

PoS(ICRC2023)1540