

# Measurement of neutron production in atmospheric neutrino interactions at Super-Kamiokande

S. Han,<sup>2</sup> K. Abe,<sup>1,48</sup> C. Bronner,<sup>1</sup> Y. Hayato,<sup>1,48</sup> K. Hiraide,<sup>1,48</sup> K. Hosokawa,<sup>1</sup> K. Ieki,<sup>1,48</sup> M. Ikeda,<sup>1,48</sup> J. Kameda,<sup>1,48</sup> Y. Kanemura,<sup>1</sup> R. Kaneshima,<sup>1</sup> Y. Kashiwagi,<sup>1</sup> Y. Kataoka,<sup>1,48</sup> S. Miki,<sup>1</sup> S. Mine,<sup>1,6</sup> M. Miura,<sup>1,48</sup> S. Moriyama,<sup>1,48</sup> M. Nakahata,<sup>1,48</sup> Y. Nakano,<sup>1</sup> S. Nakayama,<sup>1,48</sup> Y. Noguchi,<sup>1</sup> K. Sato,<sup>1</sup> H. Sekiya,<sup>1,48</sup> H. Shiba,<sup>1</sup> K. Shimizu,<sup>1</sup> M. Shiozawa,<sup>1,48</sup> Y. Sonoda,<sup>1</sup> Y. Suzuki,<sup>1</sup> A. Takeda,<sup>1,48</sup> Y. Takemoto,<sup>1,48</sup> H. Tanaka,<sup>1,48</sup> T. Yano,<sup>1</sup> T. Kajita,<sup>2,48,22</sup> K. Okumura,<sup>2,48</sup> T. Tashiro,<sup>2</sup> T. Tomiya,<sup>2</sup> X. Wang,<sup>2</sup> S. Yoshida,<sup>2</sup> P. Fernandez,<sup>3</sup> L. Labarga,<sup>3</sup> N. Ospina,<sup>3</sup> B. Zaldivar,<sup>3</sup> B. W. Pointon,<sup>5,51</sup> E. Kearns,<sup>4,48</sup> J. L. Raaf,<sup>4</sup> L. Wan,<sup>4</sup> T. Wester,<sup>4</sup> J. Bian,<sup>6</sup> N. J. Grishevich,<sup>6</sup> M. B. Smy,<sup>6,48</sup> H. W. Sobel,<sup>6,48</sup> V. Takhistov,<sup>6,24</sup> A. Yankelevich,<sup>6</sup> J. Hill,<sup>7</sup> M. C. Jang,<sup>8</sup> S. H. Lee,<sup>8</sup> D. H. Moon,<sup>8</sup> R. G. Park,<sup>8</sup> B. Bodur,<sup>9</sup> K. Scholberg,<sup>9,48</sup> C. W. Walter,<sup>9,48</sup> A. Beauchêne,<sup>10</sup> O. Drapier,<sup>10</sup> A. Giampaolo,<sup>10</sup> Th. A. Mueller,<sup>10</sup> A. D. Santos,<sup>10</sup> P. Paganini,<sup>10</sup> B. Quilain,<sup>10</sup> R. Rogly,<sup>10</sup> T. Nakamura,<sup>11</sup> J. S. Jang,<sup>12</sup> L. N. Machado,<sup>13</sup> J. G. Learned,<sup>14</sup> K. Choi,<sup>15</sup> N. Iovine,<sup>15</sup> S. Cao,<sup>16</sup> L. H. V. Anthony,<sup>17</sup> D. Martin,<sup>17</sup> N. W. Prouse,<sup>17</sup> M. Scott,<sup>17</sup> Y. Uchida,<sup>17</sup> V. Berardi,<sup>18</sup> N. F. Calabria,<sup>18</sup> M. G. Catanesi,<sup>18</sup> E. Radicioni,<sup>18</sup> A. Langella,<sup>19</sup> G. De Rosa,<sup>19</sup> G. Collazuol,<sup>20</sup> F. Iacob,<sup>20</sup> M. Mattiazzi,<sup>20</sup> L. Ludovici,<sup>21</sup> M. Gonin,<sup>22</sup> G. Pronost,<sup>22</sup> C. Fujisawa,<sup>23</sup> Y. Maekawa,<sup>23</sup> Y. Nishimura,<sup>23</sup> R. Okazaki,<sup>23</sup> R. Akutsu,<sup>24</sup> M. Friend,<sup>24</sup> T. Hasegawa,<sup>24</sup> T. Ishida,<sup>24</sup> T. Kobayashi,<sup>24</sup> M. Jakkpu,<sup>24</sup> T. Matsubara,<sup>24</sup> T. Nakadaira,<sup>24</sup> K. Nakamura,<sup>24,48</sup> Y. Oyama,<sup>24</sup> K. Sakashita,<sup>24</sup> T. Sekiguchi,<sup>24</sup> T. Tsukamoto,<sup>24</sup> N. Bhuiyan,<sup>25</sup> G. T. Burton,<sup>25</sup> F. Di Lodovico,<sup>25</sup> J. Gao,<sup>25</sup> A. Goldsack,<sup>25</sup> T. Katori,<sup>25</sup> J. Migenda,<sup>25</sup> R. M. Ramsden,<sup>25</sup> Z. Xie,<sup>25</sup> S. Zsoldos,<sup>25,48</sup> A. T. Suzuki,<sup>26</sup> Y. Takagi,<sup>26</sup> Y. Takeuchi,<sup>26,48</sup> H. Zhong,<sup>26</sup> J. Feng,<sup>27</sup> L. Feng,<sup>27</sup> J. R. Hu,<sup>27</sup> Z. Hu,<sup>27</sup> M. Kawauae,<sup>27</sup> T. Kikawa,<sup>27</sup> M. Mori,<sup>27</sup> T. Nakaya,<sup>27,48</sup> R. A. Wendell,<sup>27,48</sup> K. Yasutome,<sup>27</sup> S. J. Jenkins,<sup>28</sup> N. McCauley,<sup>28</sup> P. Mehta,<sup>28</sup> A. Tarrant,<sup>28</sup> M. J. Wilking,<sup>29</sup> Y. Fukuda,<sup>30</sup> Y. Itow,<sup>31,32</sup> H. Menjo,<sup>31</sup> K. Ninomiya,<sup>31</sup> Y. Yoshioka,<sup>31</sup> J. Lagoda,<sup>33</sup> M. Mandal,<sup>33</sup> P. Mijakowski,<sup>33</sup> Y. S. Prabhu,<sup>33</sup> J. Zalipska,<sup>33</sup> M. Jia,<sup>34</sup> J. Jiang,<sup>34</sup> W. Shi,<sup>34</sup> C. Yanagisawa,<sup>34,\*</sup> M. Harada,<sup>35</sup> Y. Hino,<sup>35</sup> H. Ishino,<sup>35</sup> Y. Koshio,<sup>35,48</sup> F. Nakanishi,<sup>35</sup> S. Sakai,<sup>35</sup> T. Tada,<sup>35</sup> T. Tano,<sup>35</sup> T. Ishizuka,<sup>36</sup> G. Barr,<sup>37</sup> D. Barrow,<sup>37</sup> L. Cook,<sup>37,48</sup> S. Samani,<sup>37</sup> D. Wark,<sup>37,43</sup> A. Holin,<sup>38</sup> F. Nova,<sup>38</sup> S. Jung,<sup>39</sup> B. S. Yang,<sup>39</sup> J. Y. Yang,<sup>39</sup> J. Yoo,<sup>39</sup> J. E. P. Fannon,<sup>40</sup> L. Kneale,<sup>40</sup> M. Malek,<sup>40</sup> J. M. McElwee,<sup>40</sup> M. D. Thiesse,<sup>40</sup> L. F. Thompson,<sup>40</sup> S. T. Wilson,<sup>40</sup> H. Okazawa,<sup>41</sup> S. M. Lakshmi,<sup>42</sup> S. B. Kim,<sup>44</sup> E. Kwon,<sup>44</sup> J. W. Seo,<sup>44</sup> I. Yu,<sup>44</sup> A. K. Ichikawa,<sup>45</sup> K. D. Nakamura,<sup>45</sup> S. Tairafune,<sup>45</sup> K. Nishijima,<sup>46</sup> A. Eguchi,<sup>47</sup> K. Nakagiri,<sup>47</sup> Y. Nakajima,<sup>47,48</sup> S. Shima,<sup>47</sup> N. Taniuchi,<sup>47</sup> E. Watanabe,<sup>47</sup> M. Yokoyama,<sup>47,48</sup> P. de Perio,<sup>48</sup> S. Fujita,<sup>48</sup> C. Jesùs-Valls,<sup>48</sup> K. Martens,<sup>48</sup> K. M. Tsui,<sup>48</sup> M. R. Vagins,<sup>48,6</sup> J. Xia,<sup>48</sup> S. Izumiymaya,<sup>49</sup> M. Kuze,<sup>49</sup> R. Matsumoto,<sup>49</sup> M. Ishitsuka,<sup>50</sup> H. Ito,<sup>50</sup> Y. Ommura,<sup>50</sup> N. Shigeta,<sup>50</sup> M. Shinoki,<sup>50</sup> K. Yamauchi,<sup>50</sup> T. Yoshida,<sup>50</sup> R. Gaur,<sup>51</sup> V. Gousy-Leblanc,<sup>51,†</sup> M. Hartz,<sup>51</sup> A. Konaka,<sup>51</sup> X. Li,<sup>51</sup> S. Chen,<sup>52</sup> B. D. Xu,<sup>52</sup> B. Zhang,<sup>52</sup> M. Posiadala-Zezula,<sup>53</sup> S. B. Boyd,<sup>54</sup> R. Edwards,<sup>54</sup> D. Hadley,<sup>54</sup> M. Nicholson,<sup>54</sup> M. O'Flaherty,<sup>54</sup> B. Richards,<sup>54</sup> A. Ali,<sup>55,51</sup> B. Jamieson,<sup>55</sup> S. Amanai,<sup>56</sup> Ll. Martí,<sup>56</sup> A. Minamino,<sup>56</sup> and S. Suzuki<sup>56</sup>

(The Super-Kamiokande Collaboration)

<sup>1</sup>Kamioka Observatory, Institute for Cosmic Ray Research, University of Tokyo, Kamioka, Gifu 506-1205, Japan

<sup>2</sup>Research Center for Cosmic Neutrinos, Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba 277-8582, Japan

<sup>3</sup>Department of Theoretical Physics, University Autònoma Madrid, 28049 Madrid, Spain

<sup>4</sup>Department of Physics, Boston University, Boston, MA 02215, USA

<sup>5</sup>Department of Physics, British Columbia Institute of Technology, Burnaby, BC, V5G 3H2, Canada

<sup>6</sup>Department of Physics and Astronomy, University of California, Irvine, Irvine, CA 92697-4575, USA

<sup>7</sup>Department of Physics, California State University, Dominguez Hills, Carson, CA 90747, USA

<sup>8</sup>Institute for Universe and Elementary Particles, Chonnam National University, Gwangju 61186, Korea

<sup>9</sup>Department of Physics, Duke University, Durham NC 27708, USA

<sup>10</sup>Ecole Polytechnique, IN2P3-CNRS, Laboratoire Leprince-Ringuet, F-91120 Palaiseau, France

<sup>11</sup>Department of Physics, Gifu University, Gifu, Gifu 501-1193, Japan

<sup>12</sup>GIST College, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea

<sup>13</sup>School of Physics and Astronomy, University of Glasgow, Glasgow, Scotland, G12 8QQ, United Kingdom

<sup>14</sup>Department of Physics and Astronomy, University of Hawaii, Honolulu, HI 96822, USA

<sup>15</sup>Center for Underground Physics, Institute for Basic Science (IBS), Daejeon, 34126, Korea

<sup>16</sup>Institute For Interdisciplinary Research in Science and Education, ICISE, Quy Nhon, 55121, Vietnam

<sup>17</sup>Department of Physics, Imperial College London , London, SW7 2AZ, United Kingdom

<sup>18</sup>Dipartimento Interuniversitario di Fisica, INFN Sezione di Bari and Università e Politecnico di Bari, I-70125, Bari, Italy

<sup>19</sup>Dipartimento di Fisica, INFN Sezione di Napoli and Università di Napoli, I-80126, Napoli, Italy

<sup>20</sup>Dipartimento di Fisica, INFN Sezione di Padova and Università di Padova, I-35131, Padova, Italy

<sup>21</sup>INFN Sezione di Roma and Università di Roma "La Sapienza", I-00185, Roma, Italy

<sup>57</sup> <sup>22</sup> *ILANCE, CNRS - University of Tokyo International Research Laboratory, Kashiwa, Chiba 277-8582, Japan*

<sup>58</sup> <sup>23</sup> *Department of Physics, Keio University, Yokohama, Kanagawa, 223-8522, Japan*

<sup>59</sup> <sup>24</sup> *High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan*

<sup>60</sup> <sup>25</sup> *Department of Physics, King's College London, London, WC2R 2LS, UK*

<sup>61</sup> <sup>26</sup> *Department of Physics, Kobe University, Kobe, Hyogo 657-8501, Japan*

<sup>62</sup> <sup>27</sup> *Department of Physics, Kyoto University, Kyoto, Kyoto 606-8502, Japan*

<sup>63</sup> <sup>28</sup> *Department of Physics, University of Liverpool, Liverpool, L69 7ZE, United Kingdom*

<sup>64</sup> <sup>29</sup> *School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA*

<sup>65</sup> <sup>30</sup> *Department of Physics, Miyagi University of Education, Sendai, Miyagi 980-0845, Japan*

<sup>66</sup> <sup>31</sup> *Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Aichi 464-8602, Japan*

<sup>67</sup> <sup>32</sup> *Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Nagoya, Aichi 464-8602, Japan*

<sup>68</sup> <sup>33</sup> *National Centre For Nuclear Research, 02-093 Warsaw, Poland*

<sup>69</sup> <sup>34</sup> *Department of Physics and Astronomy, State University of New York at Stony Brook, NY 11794-3800, USA*

<sup>70</sup> <sup>35</sup> *Department of Physics, Okayama University, Okayama, Okayama 700-8530, Japan*

<sup>71</sup> <sup>36</sup> *Media Communication Center, Osaka Electro-Communication University, Neyagawa, Osaka, 572-8530, Japan*

<sup>72</sup> <sup>37</sup> *Department of Physics, Oxford University, Oxford, OX1 3PU, United Kingdom*

<sup>73</sup> <sup>38</sup> *Rutherford Appleton Laboratory, Harwell, Oxford, OX11 0QX, UK*

<sup>74</sup> <sup>39</sup> *Department of Physics, Seoul National University, Seoul 151-742, Korea*

<sup>75</sup> <sup>40</sup> *Department of Physics and Astronomy, University of Sheffield, S3 7RH, Sheffield, United Kingdom*

<sup>76</sup> <sup>41</sup> *Department of Informatics in Social Welfare, Shizuoka University of Welfare, Yaizu, Shizuoka, 425-8611, Japan*

<sup>77</sup> <sup>42</sup> *August Cichkowski Institute of Physics, University of Silesia in Katowice, 75 Putku Piechoty 1, 41-500 Chorzów, Poland*

<sup>78</sup> <sup>43</sup> *STFC, Rutherford Appleton Laboratory, Harwell Oxford, and*

<sup>79</sup> <sup>44</sup> *Daresbury Laboratory, Warrington, OX11 0QX, United Kingdom*

<sup>80</sup> <sup>45</sup> *Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea*

<sup>81</sup> <sup>46</sup> *Department of Physics, Faculty of Science, Tohoku University, Sendai, Miyagi, 980-8578, Japan*

<sup>82</sup> <sup>47</sup> *Department of Physics, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan*

<sup>83</sup> <sup>48</sup> *Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*

<sup>84</sup> <sup>49</sup> *Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of*

<sup>85</sup> <sup>50</sup> *Tokyo Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

<sup>86</sup> <sup>51</sup> *Department of Physics, Tokyo Institute of Technology, Meguro, Tokyo 152-8551, Japan*

<sup>87</sup> <sup>52</sup> *Department of Physics, Faculty of Science and Technology, Tokyo University of Science, Noda, Chiba 278-8510, Japan*

<sup>88</sup> <sup>53</sup> *TRIUMF, 4004 Wesbrook Mall, Vancouver, BC, V6T2A3, Canada*

<sup>89</sup> <sup>54</sup> *Department of Engineering Physics, Tsinghua University, Beijing, 100084, China*

<sup>90</sup> <sup>55</sup> *Faculty of Physics, University of Warsaw, Warsaw, 02-093, Poland*

<sup>91</sup> <sup>56</sup> *Department of Physics, University of Warwick, Coventry, CV4 7AL, UK*

<sup>92</sup> <sup>57</sup> *Department of Physics, University of Winnipeg, MB R3J 3L8, Canada*

<sup>93</sup> <sup>58</sup> *Department of Physics, Yokohama National University, Yokohama, Kanagawa, 240-8501, Japan*

(Dated: May 23, 2024)

We present measurements of neutron production following atmospheric neutrino interactions in water, as a function of electron-equivalent visible energy ranging from 30 MeV to 10 GeV. These measurements are based on 4,270 days of data collected by Super-Kamiokande, including 564.4 days with 0.011 wt% gadolinium added to the water to enhance neutron detection. The results were compared to various model predictions, which showed significant variability, particularly in the secondary spallation caused by energetic hadron propagation within the target water volume. Notably, the Liège intranuclear cascade model (INCL++) showed better agreement with our observations across all energy ranges compared to the Bertini cascade model and its variants. We also examine the differences in these models that contribute to the variations in their predictions.

## I. INTRODUCTION

Accurate reconstruction of incoming neutrinos interacting with target nuclei is crucial for precise measurements of neutrino oscillation parameters, including experimental tests of CP violation in neutrino oscillation.

There are two main limiting factors: inefficiency and uncertainty in reconstructing the visible final state, and invisible nuclear effects within the target nucleus, which cause the visible final states to differ from free nucleon target predictions. Low-energy hadrons are especially important because they probe nuclear effects, but until recently, they were poorly reconstructed in neutrino experiments due to higher energy thresholds. Recent and forthcoming precision measurements of final state hadrons with upgraded detectors provide benchmarks for neutrino-nucleus interaction models, which are essential for accurate characterization of incoming neutrinos.

\* also at BMCC/CUNY, Science Department, New York, New York, 1007, USA.

† also at University of Victoria, Department of Physics and Astronomy, PO Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada.

121 Neutrons have a unique status compared to charged  
 122 hadrons such as protons and pions. One notable feature  
 123 is their detectability via exothermic, radiative capture  
 124 reactions on nuclei:  ${}^A\text{X}(n, \gamma){}^{A+1}\text{X}$ . For most target ma-  
 125 terials, this is the dominant reaction channel for neutrons  
 126 below a few MeV, with the cross section peaking at ther-  
 127 mal energies. These reactions occur with characteristic  
 128 time scales and energies, allowing for clean signal selec-  
 129 tion. The absence of an energy threshold offers poten-  
 130 tial for hadron calorimetry and provides an effective way  
 131 to distinguish antineutrino charged current interactions  
 132 from neutrino interactions.

133 These advantages are particularly relevant for analyzing  
 134 GeV-scale atmospheric neutrino interactions, where  
 135 the neutrino and antineutrino fluxes are mixed, and neu-  
 136 trons are expected to carry significant energy. Neutron  
 137 tagging atmospheric neutrino interactions can enhance  
 138 sensitivity to the neutrino mass ordering and the CP vio-  
 139 lation effects. Additionally, it aids in searches for rare  
 140 events, such as proton decay and electron recoil due to  
 141 dark matter, which are not expected to produce neutrons,  
 142 or inverse beta decays with single neutron production in-  
 143 duced by  $O(1)$  MeV  $\bar{\nu}_e$  from supernovae.

144 Recently, several studies have measured neutron pro-  
 145 duction using both atmospheric and artificial neutrinos  
 146 with various target materials, such as water (T2K), heavy  
 147 water (SNO), and plastic (MINERvA), aiming to validate  
 148 neutrino-nucleus interaction models. The typical mod-  
 149eling approach for these studies, including downstream  
 150 particle transport, is illustrated in Fig. 1. 180

151 Differentiating the source of neutron production is 181  
 152 challenging as it is only inferred from secondary inter- 182  
 153 actions that produce charged particles. This makes neu- 183  
 154 trons produced through secondary spallation reactions 184  
 155 often indistinguishable from primary neutrons knocked 185  
 156 out directly by a neutrino. For these reasons, accurate 186  
 157 modeling of neutron interactions and spallation reactions 187  
 158 in the detector volume is essential for validating neutrino- 188  
 159 nucleus interaction models. Notably, observed deficits in 189  
 160 past measurements compared to predictions from various 190  
 161 neutrino event generators are often attributed to defici- 191  
 162 cies in modeling these spallation reactions. 192

163 Special attention must be given to modeling intranu- 193  
 164 clear hadron transport, as it significantly impacts the pre- 194  
 165 diction of observable final neutron multiplicity. This in- 195  
 166 volves controlling the intensity of the hadron shower and 196  
 167 the residual nuclear energy affecting subsequent nuclear 197  
 168 de-excitation. For hadron transport in the 0.1-1 GeV 198  
 169 range, semi-classical Intranuclear Cascade (INC) mod- 199  
 170 els are commonly used. These models approximate the 200  
 171 process as a cascade of binary collisions between hadron 201  
 172 projectiles and constituent nucleons. However, in many 202  
 173 cases, additional considerations beyond these assump- 203  
 174 tions are necessary. Different INC models handle such 204  
 175 corrections in various ways, often resulting in significant 205  
 176 disparities in predictions. It is worth noting that previ- 206  
 177 ous analyses have either used the Bertini cascade model 207  
 178 or its variants. 208

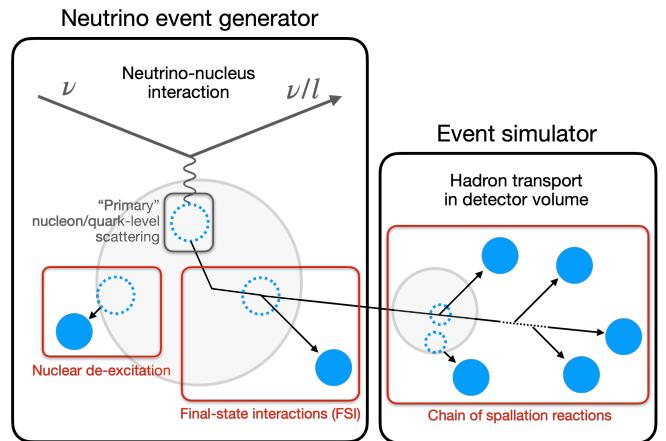


FIG. 1. Schematic of sources of nucleon production in a typical neutrino event simulation setup. Neutrino event generators sample outgoing hadrons from the initial neutrino interaction, occurring at either the nucleon or quark level, and handle the spallation of the target nucleus, including intranuclear hadron transport (also referred to as final-state interaction or FSI) followed by nuclear de-excitation. Subsequent transport of hadrons, potentially involving multiple spallation reactions, is managed by event simulators developed with particle transport codes. Dashed blue circles represent nucleon holes, black arrows represent nucleon trajectories, and solid blue circles represent detectable nucleons ejected from nuclei.

In this paper, we present our findings on the measurement of neutron production following atmospheric neutrino interactions, primarily with  ${}^{16}\text{O}$  in water. We analyzed events fully contained within the Super-Kamiokande detector from 2008 to 2022 and evaluated the average neutron capture multiplicity as a function of electron-equivalent visible energy, which is a calorimetric proxy for the momentum transfer. The observations were then compared with predictions generated by various combinations of spallation models, including the two widely used INC models: the Geant4 Bertini and Liège cascade models (INCL++). Evaluating multiple INC models with varying corrections can help estimate the uncertainty associated with secondary neutron production, which eventually affects neutrino reconstruction. This assessment is also relevant to recent studies of the FSI impact on multi-nucleon recoils and pion production, contributing to an improved estimation of their contamination in the analysis sample.

This paper is structured as follows. A brief overview of the SK detector is provided in Sec. II. The selection process of atmospheric neutrino events and neutron signals and the estimation of selection performance are described in Sec. III and IV. Sec. V discusses the methodology for obtaining average neutron capture multiplicity per visible energy bin and related systematic uncertainties. Sec. VI describes the various spallation models and how they were used to generate predictions. The comparison between observations and predictions, as well as the implications of the results, are discussed in Sec. VII and VIII.

TABLE I. SK operational phases and neutron-related characteristics. SK-IV, V, VI data were used in this analysis.

Phase	Dates	Livetime [days]	Gd concentration <sup>a</sup> [wt%]	Expected H( $n, \gamma$ ) [%]	( $n, \gamma$ ) ratio <sup>b</sup> H( $n, \gamma$ ) [%]	( $n, \gamma$ ) time constant <sup>c</sup> [μs]
SK-I-III	1996-2008	2805.9	-	>99.9	-	No data
SK-IV	2008-2018	3244.4	-	>99.9	-	$204.8 \pm 9.8$
SK-V	2019-2020	461.0	-	>99.9	-	$199.8 \pm 10.2$
SK-VI	2020-2022	564.4 <sup>d</sup>	$0.0111 \pm 0.0002$	49.6	49.8	$116.2 \pm 2.3$
SK-VII	2022-present	-	$0.0332 \pm 0.0001$	24.8	73.1	$62.0 \pm 1.1$

<sup>a</sup> Based on dissolved Gd mass.

<sup>b</sup> Based on simulated atmospheric neutrino events explained in Sec. III C

<sup>c</sup> Weighted mean of all Am/Be neutron source measurements, explained in Sec. IV B

<sup>d</sup> Excludes earlier runs which showed signs of non-uniform Gd concentration, i.e., varying time constant by position.

## 209 II. THE SUPER-KAMIOKANDE DETECTOR 251

210 Super-Kamiokande (SK) is an underground water 253  
 211 Cherenkov detector located in Gifu, Japan, with an over- 254  
 212 burden of 2,780 meters water equivalent. It consists of 255  
 213 two optically separated, concentric cylindrical volumes: 256  
 214 the inner detector (ID) with a diameter of 33.8 m and 257  
 215 a height of 36.2 m, containing 32.5 ktons of water and 258  
 216 equipped with 11,129 inward-facing 20-inch photomulti- 259  
 217 plier tubes (PMTs); and the outer detector (OD) with a 260  
 218 diameter of 39.3 m and a height of 41.4 m, containing 17.7 261  
 219 ktons of water and equipped with 1,885 outward-facing 8- 262  
 220 inch PMTs, serving as a cosmic-ray veto. A PMT signal 263  
 221 with over 0.25 photoelectron-equivalent charge is regis- 264  
 222 tered as a “hit”. If the number of ID or OD PMT hits 265  
 223 within a 200-ns sliding time window ( $N_{200\text{-ns}}$ ) exceeds a 266  
 224 given threshold, an event trigger is issued. 267

225 Charged particles, namely electrons and muons pro- 268  
 226 duced by charged-current neutrino interactions, are iden- 269  
 227 tified through Cherenkov radiation. The radiation is pro- 270  
 228 jected onto the PMTs as a characteristic ring pattern 271  
 229 with an opening angle of approximately 42 degrees from 272  
 230 the initial vertex. This ring pattern serves as the basis 273  
 231 for particle reconstruction. Neutrons are indirectly iden- 274  
 232 tified via Compton-scattered electrons resulting from ra- 275  
 233 diative neutron captures. In pure water, most occur on 276  
<sup>1</sup>H, emitting a single 2 MeV gamma-ray. With the recent 277  
 235 addition of gadolinium (Gd), the majority of neutrons are 278  
 236 expected to be captured by Gd isotopes, resulting in a 279  
 237 total gamma-radiated energy of around 8 MeV. 280

238 The performance of the  $O(1)$  MeV signal identifica- 277  
 239 tion algorithm is significantly influenced by variations in 278  
 240 detector characteristics. Parameters such as individual 279  
 241 PMT gain, timing characteristics, quantum efficiencies, 280  
 242 as well as optical absorption and scattering properties 281  
 243 in water, are continuously monitored using light sources 282  
 244 and cosmic-ray muons. Furthermore, the uncertainty in 283  
 245 Cherenkov ring energy reconstruction is assessed across 284  
 246 a broad energy spectrum using naturally occurring parti- 285  
 247 cles, including cosmic-ray muons, subsequent Michel elec- 286  
 248 trons, and decays of  $\pi^0$  produced in atmospheric neutrino 287  
 249 interactions. The evaluated energy scale uncertainty re- 288  
 250 mains within 2% across the  $O(10\text{-}10^4)$  MeV range. 289

The detector has operated through seven different phases. Neutron detection began with the fourth phase SK-IV (2008-2019), following the aforementioned electronics upgrade that allowed extended event recording up to 535 μs after certain ID triggers. This has enabled analysis of delayed neutron captures that occur with a time scale of  $O(10\text{-}100)$  μs following an atmospheric neutrino interaction. Between SK-IV and SK-V (2019-2020), the detector underwent refurbishment, during which malfunctioning PMTs were replaced. The last two phases, SK-VI (2020-2022) and SK-VII (2022-present), involved the dissolution of  $\text{Gd}_2(\text{SO}_4)_3$  into the water volume to enhance neutron detection efficiency. Table I summarizes the relevant operational conditions.

This study analyzes the atmospheric neutrino events fully contained in the ID during SK-IV, SK-V, and SK-VI, spanning livetime operation of 3705.4 days with pure water (SK-IV and SK-V) and 564.4 days with 0.011 wt% Gd-loaded water (SK-VI). Data from the ongoing SK-VII phase is not included in this analysis.

## III. ATMOSPHERIC NEUTRINO EVENTS

### A. Data reduction

We followed a typical selection process for atmospheric neutrino interactions that are fully contained within the ID, similar to previous studies conducted at SK. Additionally, all events were required to pass the ID trigger with the threshold  $N_{200\text{-ns}} \geq 58$  PMT hits, roughly corresponding to a 10 MeV electron, followed by the extended 535 μs event window for neutron detection.

Selected events were reconstructed as described in Sec. III B. To further reject low-energy backgrounds, we required that the reconstructed vertex be more than 1 m away from the ID tank wall (defining the fiducial volume with 23.2 kton of water) and that the reconstructed visible energy be larger than 30 MeV. These same reduction steps were also applied to simulated events. The remaining background contamination, mainly due to corner-clipping cosmic-ray muons and PMT discharges, was estimated to be below 0.2%, based on visual inspection.

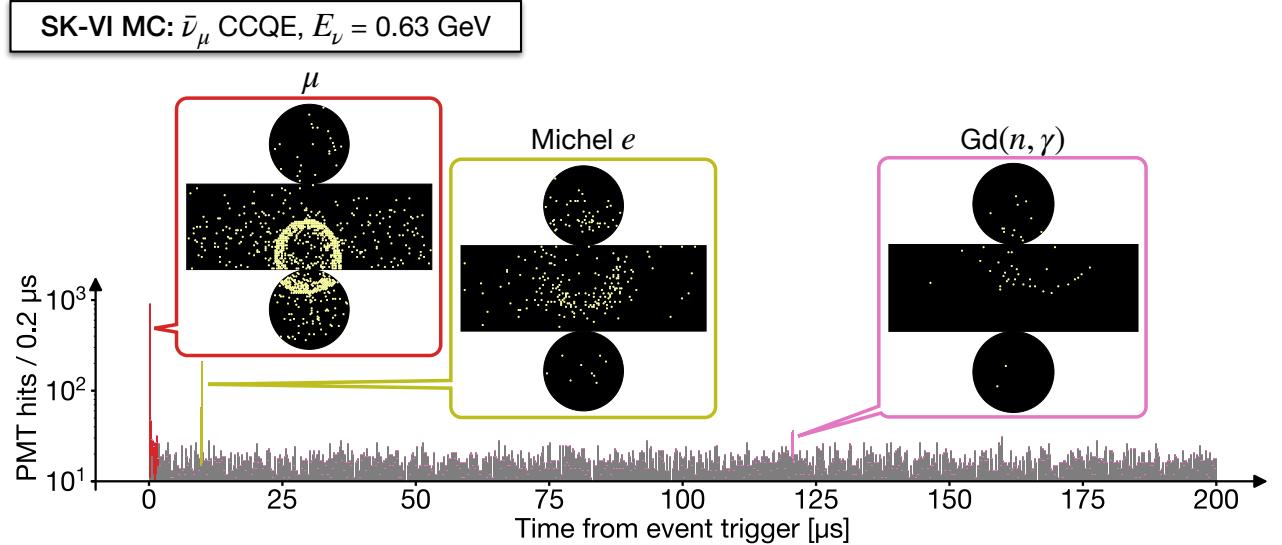


FIG. 2. A PMT hit time distribution for a typical  $\bar{\nu}_\mu$  charged-current quasi-elastic (CCQE) interaction with a muon and a neutron as final states. The zero is set at the event trigger. The event displays feature the “prompt” muon signal (red) and the two types of “delayed” coincident signals — Michel electrons from muon decay (olive) and neutron captures on Gd (pink). The gray bars represent randomly recorded background PMT hits. This event was simulated in the SK-VI configuration.

## 290 B. Reconstruction of prompt Cherenkov rings

291 Fig. 2 shows a typical neutrino event display at SK.  
 292 For the “prompt” (within 1-2  $\mu$ s from the event trigger)  
 293 radiation due to charged particles, we also followed the  
 294 typical Cherenkov ring reconstruction process as applied  
 295 in previous studies conducted at SK.

296 In this process, the neutrino interaction vertex is as-  
 297 sumed to be the vertex with minimal time spread in the  
 298 prompt PMT hits, after correcting for the photon time-  
 299 of-flight to each PMT. Each identified Cherenkov ring is  
 300 then classified into electron-like and muon-like, by eval-  
 301 uating likelihood ratios of observing each hit charge un-  
 302 der the assumptions of it being an electron or a muon.  
 303 This evaluation accounts for the fuzzy edges of the rings  
 304 caused by bremsstrahlung, which electrons are more  
 305 prone to than muons. The momentum of each ring is  
 306 determined based on the total observed charge associ-  
 307 ated with the ring, considering light attenuation in water,  
 308 scattered light, and the acceptance of each PMT.

309 The visible energy of an event is defined as the sum of  
 310 the reconstructed kinetic energies of each Cherenkov ring,  
 311 assuming each ring originates from an electron. While<sup>321</sup>  
 312 determining the momentum transfer in atmospheric neu-<sup>322</sup>  
 313 trino interactions can be challenging without knowledge<sup>323</sup>  
 314 of the neutrino direction and full reconstruction of final<sup>324</sup>  
 315 states, it is expected by simulation that both the aver-<sup>325</sup>  
 316 age momentum transfer squared ( $Q^2$ ) and the hadronic<sup>326</sup>  
 317 invariant mass ( $W$ ) will exhibit a linear increase with vis-<sup>327</sup>  
 318 ible energy, as shown in Fig. 3. In this study, we utilize<sup>328</sup>  
 319 visible energy as a proxy for interaction kinematics.<sup>329</sup>

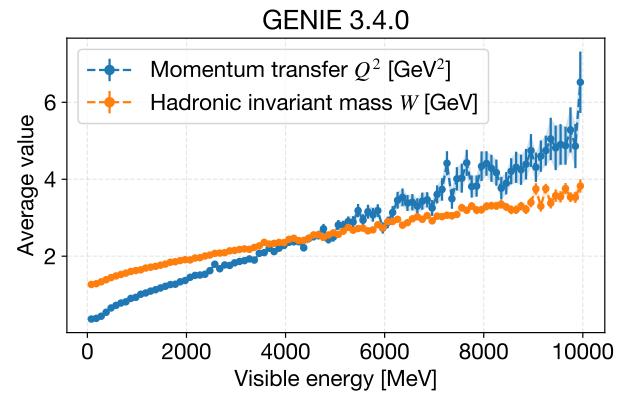


FIG. 3. Average values of squared momentum transfer  $Q^2$  (blue) and hadronic invariant mass  $W$  (orange) as a function of visible energy of SK atmospheric neutrino events simulated with GENIE 3.4.0.

## C. Default simulation setup

The simulation of atmospheric neutrino events involves a convolution of the atmospheric neutrino flux, neutrino event generator, and detector simulation. Here, we provide an overview of the default setup.

To determine the incoming neutrino kinematics and the event rate, we used the atmospheric  $\nu_e$ ,  $\nu_\mu$ ,  $\bar{\nu}_e$ , and  $\bar{\nu}_\mu$  flux calculations for the detector site without oscillation, as provided by Honda et al. (2006) for neutrino energy from 100 MeV to 10 TeV.

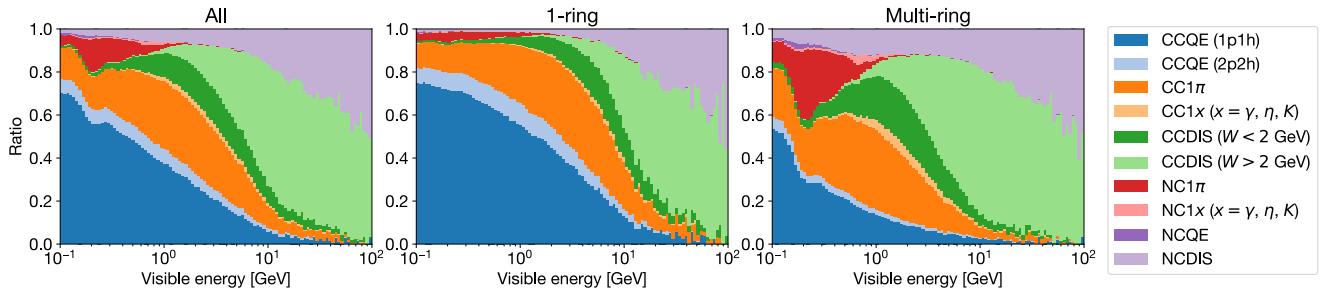


FIG. 4. Fraction of sampled interaction channels for all simulated events fully-contained in the fiducial volume (left), events with one reconstructed Cherenkov ring (middle), and events with two or more rings (right). CCQE events are dominant in the single-ring sample, while RES (labeled with  $1\pi$  or  $1x$ ) and DIS events are dominant in multi-ring samples.

330 Neutrino interactions in water were simulated using<sup>372</sup>  
 331 neutrino event generators, which compute the cross sec<sup>373</sup>  
 332 tions of each interaction channel and sample outgoing<sup>374</sup>  
 333 particle kinematics.<sup>375</sup>

334 NEUT 5.4.0 was used as the default neutrino event<sup>376</sup>  
 335 generator. It handles charged-current (CC) and neu-<sup>377</sup>  
 336 tral current (NC) exchanges, covering major interaction<sup>378</sup>  
 337 channels, including quasi-elastic (QE) scattering with ei-<sup>379</sup>  
 338 ther single (one-particle-one-hole, 1p1h) or double (two-<sup>380</sup>  
 339 particle-two-hole, 2p2h) nucleon knockout, single pion<sup>381</sup>  
 340 production due to  $\Delta$  resonance (RES), and multiple pion<sup>382</sup>  
 341 production due to deep inelastic scattering (DIS). In our<sup>383</sup>  
 342 default setup, QE interactions were modeled based on<sup>384</sup>  
 343 Nieves et al., which uses the Local Fermi Gas (LFG)<sup>385</sup>  
 344 model for initial target nucleons along with Random<sup>386</sup>  
 345 Phase Approximation (RPA) correction. Here, the axial<sup>387</sup>  
 346 mass value  $M_A$  was set to  $1.05 \text{ GeV}/c^2$ , and the BBBA05<sup>388</sup>  
 347 vector form factors were used. RES interactions were<sup>389</sup>  
 348 modeled following Rein and Sehgal. For DIS, the parton<sup>390</sup>  
 349 distribution function was taken from GRV98, with cor-<sup>391</sup>  
 350 rections to the low- $q^2$  range by Bodek and Yang. For low<sup>392</sup>  
 351 hadronic invariant masses ( $W < 2 \text{ GeV}$ ), a custom multi-<sup>393</sup>  
 352 pion production model was used to predict final states,<sup>394</sup>  
 353 while for higher  $W$ , Pythia 5.72 was used.<sup>395</sup>

354 Spallation of the target nucleus due to off-shell<sup>396</sup>  
 355 hadronic final states was handled separately. The<sup>397</sup>  
 356 transport of low-momentum ( $< 500 \text{ MeV}/c$ ) pions which<sup>398</sup>  
 357 are likely produced through  $\Delta$  resonances was modeled<sup>399</sup>  
 358 based on Oset et al. whose parameters were fine-tuned to<sup>400</sup>  
 359 fit external pion-nucleus scattering data. The FSI model-<sup>401</sup>  
 360 ing for other hadrons follows the INC approach, consider-<sup>402</sup>  
 361 ing elastic scattering and single/double pion production.<sup>403</sup>  
 362 Reaction cross sections with free target nucleons were<sup>404</sup>  
 363 sourced from GCALOR for nucleon projectiles and from<sup>405</sup>  
 364 external pion scattering data for higher momentum pion<sup>406</sup>  
 365 projectiles. For pion production, the formation length<sup>407</sup>  
 366 is considered as proportional to the incident hadron mo-<sup>408</sup>  
 367 mentum, with tuned coefficient. The de-excitation of an<sup>409</sup>  
 368 oxygen target following the knockout of a nucleon by a<sup>410</sup>  
 369 neutrino was modeled using tabulated occupation prob-<sup>411</sup>  
 370 abilities of nucleon energy states from [], and branching<sup>412</sup>  
 371 ratios for the knockout of each state from [].<sup>413</sup>

The fraction of sampled neutrino interaction channels in the final simulated event sample is shown in Fig. 4, divided by the number of reconstructed Cherenkov rings. Single-ring events tend to have a higher fraction of QE and a smaller fraction of DIS compared to multi-ring events. For the same visible energy, multi-ring events are expected to accompany more neutron signals than single-ring events, due to the larger fraction of DIS with larger energy transfer to the hadron system.

The simulations of subsequent particle transport in water and detector responses were conducted using Geant3.21. By default, hadron propagation above 10 GeV was simulated with FLUKA, while GCALOR, based on the Bertini cascade model as implemented in NMTC, was used for those below 10 GeV. For neutrons under 20 MeV particularly, GCALOR calls MICAP based on ENDF/B-V. Low-momentum pion transport was separately managed by invoking the NEUT pion FSI routine to ensure consistency between NEUT and Geant3.

For the simulation of the Gd-loaded SK-VI phase, we used the Geant4.10.03.p01 NeutronHP model based on ENDF/B-VII.1, to replace MICAP for neutron energies below 20 MeV. Additionally, we modeled the gamma-cascade resulting from neutron captures on the  $^{155}\text{Gd}$  and  $^{157}\text{Gd}$  targets using the ANNRI-Gd model.

The characteristics of individual PMTs and the optical parameters in water used in the detector simulator were finely adjusted to align with calibration data obtained from light sources and through-going cosmic-ray muon data. To realistically account for detector noise, randomly recorded PMT hits were appended as background, shown as gray bars in Fig. 2.

500 years' worth of atmospheric neutrino events were simulated for each SK phase and processed as described in Sec. III B and III A. The events in the final sample were weighted based on flavor oscillation within the standard three-flavor framework, calculated using Prob3++ with the current best-fit values of the oscillation parameters and the PREM model for Earth's matter density. Flux corrections accounting for solar activity were applied.

## IV. NEUTRON SIGNAL SELECTION

413 Neutrons, resulting from atmospheric neutrino inter-<sup>467</sup>  
 414 actions in the ID, are mainly captured by  $^1\text{H}$  in pure  
 415 water and  $^{155}/^{157}\text{Gd}$  in Gd-loaded water within  $O(100)$ <sup>468</sup>  
 416 ps. The resulting  $O(1)$  MeV gamma-rays scatter ener-<sup>469</sup>  
 417 getic electrons, producing Cherenkov photons that form<sup>470</sup>  
 418 faint rings on the ID tank wall, as illustrated in Fig. 2.<sup>471</sup>

419 Fast neutrons, along with subsequent gamma-rays and<sup>472</sup>  
 420 scattered electrons, typically travel only a few tens of  
 421 centimeters in water. In SK, we may approximate<sup>473</sup>  
 422 Cherenkov photons resulting from neutron captures as<sup>474</sup>  
 423 originating from a single vertex. Although reconstruct-<sup>475</sup>  
 424 ing the vertex is challenging due to the low number of<sup>476</sup>  
 425 PMT hits, *assuming* it to be close to another known ver-<sup>477</sup>  
 426 tex, such as the reconstructed neutrino interaction ver-<sup>478</sup>  
 427 tex, helps identify signals from random coincidences of<sup>479</sup>  
 428 PMT hits due to dark current. Remaining Michel elec-<sup>479</sup>  
 429 trons from muon decays can largely be mitigated through<sup>480</sup>  
 430 the implementation of time and energy cuts.<sup>481</sup>

### A. Signal selection algorithm

432 The signal selection algorithm is based on [], and con-<sup>485</sup>  
 433 sists of two stages.<sup>486</sup>

434 In the first candidate search stage, we initially subtract<sup>487</sup>  
 435 the expected photon time-of-flight (ToF) from the indi-<sup>488</sup>  
 436 vidual PMT hit times for a given *assumed* signal vertex,<sup>489</sup>  
 437 which was provided by the reconstructed neutrino event<sup>490</sup>  
 438 vertex. Then, we slide a time window of 14 ns width<sup>491</sup>  
 439 on the ToF-corrected PMT hit times to trigger on the  
 440 number of included PMT hits. The threshold was 5 for<sup>492</sup>  
 441 pure water phases SK-IV and V, and 7 for the Gd-loaded<sup>493</sup>  
 442 phase SK-VI. The search span for each event was [18, 534]<sup>494</sup>  
 443  $\mu\text{s}$  from the event trigger for SK-IV and V, and [3, 534]<sup>494</sup>  
 444  $\mu\text{s}$  for SK-VI with a shorter neutron capture time. For  
 445 overlapping candidates within 50 ns, only the candidate<sup>495</sup>  
 446 with the largest number of PMT hits is selected.<sup>495</sup>

447 In the second candidate classification stage, we extract<sup>496</sup>  
 448 features of each candidate and use a neural network to<sup>497</sup>  
 449 classify each candidate into signal and noise based on<sup>497</sup>  
 450 input features. These features characterize the signal en-<sup>498</sup>  
 451 ergy, the background hit level, timing spread assuming<sup>499</sup>  
 452 the vertex, correlation between the input vertex and the  
 453 hit PMT positions, correlation to the known properties<sup>500</sup>  
 454 of PMT noise, and angular correlation among hit PMTs<sup>501</sup>  
 455 relative to the Cherenkov cone opening angle.<sup>502</sup>

456 The major changes from the original algorithm include<sup>503</sup>  
 457 a simplified algorithm, a reduced set of features, and a  
 458 heuristically tuned neural network architecture. These  
 459 modifications aim to reduce performance bias between<sup>504</sup>  
 460 the data and the simulation that is used to train the  
 461 neural networks.

462 Here, we provide the definition and unit of each feature<sup>505</sup>  
 463 used for the classification of signal candidates, along with<sup>506</sup>  
 464 their expected distributions as shown in Fig. 5:<sup>507</sup>

- NHits

The number of selected PMT hits within the 14-ns sliding time window.

- NResHits

The number of PMT hits within [-100,+100] ns from the center of the 14-ns sliding time window, subtracted from NHits.

- TRMS [ns]

The root mean square (RMS) of the ToF-corrected time distribution of the selected PMT hits.

- FitGoodness

The normalized likelihood of the ToF-corrected time distribution of the selected PMT hits, given the *assumed* signal vertex and the Gaussian PMT timing resolution of 5 ns.

- DWall [cm]

The distance from the *assumed* signal vertex to the nearest tank wall.

- DWallMeanDir [cm]

The shorter of the radial and vertical distances from the *assumed* signal vertex to the tank wall, weighted by the mean of the unit vectors connecting the vertex to each hit PMT.

- BurstRatio

The ratio of the selected PMTs with a preceding hit within 10  $\mu\text{s}$ , which are likely caused by scintillation within the irradiated PMT glass.

- DarkLikelihood

The normalized log likelihood ratio based on measured individual PMT dark rates, given by:

$$\text{DarkLikelihood} = \sigma \left( \log \prod_{i=1}^{\text{NHits}} \frac{r_i}{\langle r \rangle} \right) \quad (1)$$

where  $\sigma$  represents the sigmoid function,  $r_i$  is the dark rate of the  $i^{\text{th}}$  PMT, and  $\langle r \rangle$  is the average dark rate of all ID PMTs.

- OpeningAngleStdev [deg]

The standard deviation of the opening angles of cones formed by every possible combination of three hit PMTs and the *assumed* signal vertex.

- Beta(k)

$$\text{Beta}(k) = \frac{2}{\text{NHits}(\text{NHits} - 1)} \sum_{i \neq j} P_k(\cos \theta_{ij}) \quad (2)$$

where  $P_k$  is the  $k^{\text{th}}$  Legendre polynomial and  $\theta_{ij}$  is the opening angle between the assumed signal vertex and the  $i^{\text{th}}$  and  $j^{\text{th}}$  hit PMTs.

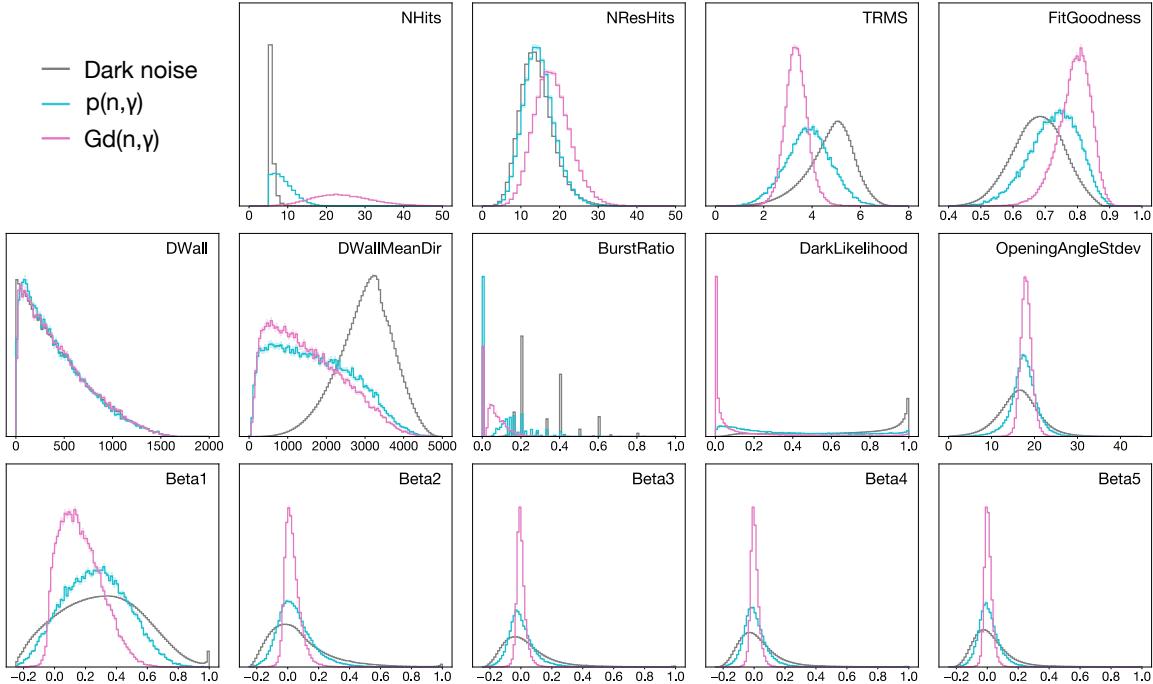


FIG. 5. The features (area-normalized) of neutron capture signals and noise from the thermal neutron MC used for training the neural network for SK-VI phase.

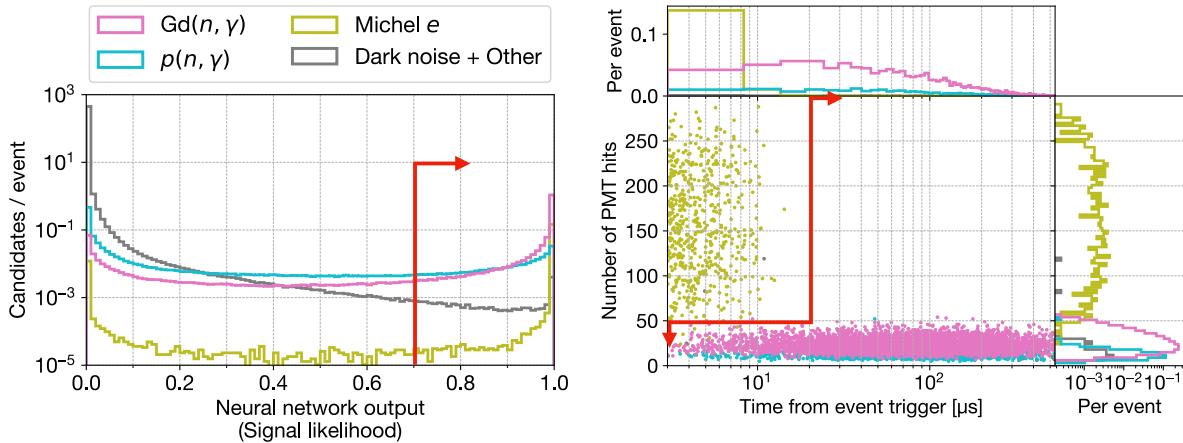


FIG. 6. The neural network response (left) to neutron captures, Michel electrons, and other backgrounds, including dark noise and gamma-rays from nuclear de-excitation following muon captures, along with the time distribution of the neural network-selected signal candidates (right). Corresponding cut points are indicated by red arrows.

508 A feed-forward fully connected neural network with an<sub>517</sub>  
 509 input layer of 14 features and three dense layers of 128<sub>518</sub>  
 510 ReLU-activated nodes, each with a 50% dropout rate,<sub>519</sub>  
 511 and a single sigmoid output node, was constructed using<sub>520</sub>  
 512 Keras 2.6.0. Input features were standardized. Weights<sub>521</sub>  
 513 and biases were initialized following He et al. and up-<sub>522</sub>  
 514 dated by iteratively minimizing the binary cross-entropy<sub>523</sub>  
 515 iteratively on each batch of candidates in the training<sub>524</sub>  
 516 dataset generated with thermal neutron simulation.<sub>525</sub>

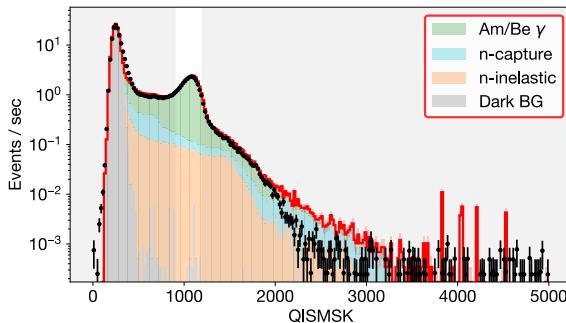
Candidates with a final neural network output greater than 0.7 were selected as signals. Candidates with more than the typical number of PMT hits ( $N_{\text{Hits}} > 50$ ) and occurring earlier than the typical timescale ( $< 20 \mu\text{s}$ ) for neutron captures were regarded as Michel electrons and thus rejected. The effectiveness of these selections is illustrated in Fig. 6. Applying this to cosmic-ray muons decaying within the ID, the estimated Michel electron efficiency was  $98.4 \pm 1.3\%$ , with a purity of  $98.7 \pm 0.5\%$ .

## 526 B. Signal selection performance on calibration data

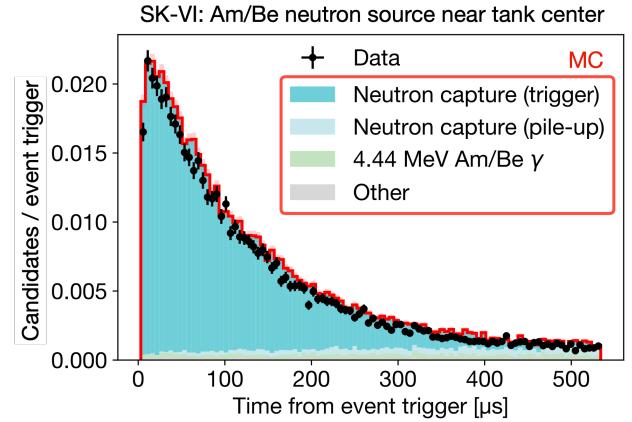
527 An Am/Be neutron source with a measured total intensity of  $236.8 \pm 5.0$  neutrons/s was used to obtain calibration data for estimating the neutron detection performance. The first-excited state of the alpha-absorbed  ${}^9\text{Be}$ , with a roughly 60% branching ratio, emits a fast neutron and a 4.44 MeV gamma-ray simultaneously. This source was encapsulated with Bismuth Germanate (BGO,  $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ ) crystals so that the 4.44 MeV gamma-rays can induce scintillation. The setup was deployed in various positions within the ID, and events were recorded for 30 minutes to 1 hour. Event triggers with a charge yield corresponding to the scintillation induced by the 4.44 MeV gamma-rays were regarded as the single neutron control sample.

541 The observed light yield distribution was compared with dedicated simulation, as shown in Fig. 7. The simulation accounts for continuous source activity and pile-up, by reorganizing the simulated detector response to 542 Am/Be neutron emission on a single global time axis, based on the measured total neutron intensity and the 543 estimated branching ratios to each excited state of alpha-absorbed  ${}^9\text{Be}$ . As shown in Fig. 7, this simulation accurately 544 models event triggers due to ambient neutron captures and 545 neutron inelastic interactions within scintillator crystal. The 546 contamination of such unwanted event triggers in the 4.44 MeV 547 gamma-ray event selection was estimated to be at a few percent level. 548

549 Within the selected events in the single neutron control 550 sample, signal candidates were obtained following the algorithm 551 described in Sec. IV A, with the *assumed* signal 552 vertex set at the source position. 553



578 FIG. 7. Distribution of total deposit charge,  $[-0.5, 1.0]$   $\mu\text{s}$  579 around event triggers obtained with Am/Be neutron source 580 within the ID. The black points correspond to SK-VI data 581 with the source near the ID tank center, while the red line 582 represent simulated prediction. The highlighted area near 1,000 583 photoelectrons represent event selection window for 4.44 MeV 584 gamma-ray induced scintillation. While  $\sim 95\%$  of the selected 585 events are expected to be actually due to the Am/Be gamma- 586 rays (green), the remaining  $\sim 5\%$  are expected to be triggered 587 by ambient neutron captures (blue) and neutrons inelastically 588 producing charged particles within the scintillator (orange). 589



590 FIG. 8. Exponential decrease of the selected neutron signal 591 candidates as a function of the time from the selected event 592 triggers with the Am/Be neutron source positioned near the 593 ID tank center, in the SK-VI phase. The label “Neutron 594 capture (trigger)” indicates captures of neutrons that are 595 produced within 350 ns ahead of the event trigger, while the 596 label “Neutron capture (pile-up)” indicates captures of piled-up 597 neutrons without such correlation to the event trigger. 598

Fig. 8 shows an example time distribution of the selected signal candidates. Such distributions of the time  $t$  were fitted with a function  $f$  of the form:

$$f(t) = A(1 - e^{-t/\tau_{\text{thermal}}})e^{-t/\tau_{\text{capture}}} + B \quad (3)$$

where the normalization constant  $A$ , the background constant  $B$ , the neutron thermalization time scale  $\tau_{\text{thermal}}$ , and the neutron capture time constant  $\tau_{\text{capture}}$  are free parameters. The signal efficiency was evaluated as the number of identified signals per selected event triggers, corrected by the constant background term  $B$ . Fig. 9 shows the estimated neutron detection efficiencies for various source positions in the ID.

The major sources of systematic uncertainty are summarized in Table II. In the pure water phases (SK-IV and V), the dominant source of uncertainty lies in the potential bias caused by the calibration setup, such as the unwanted event triggers or time correlation of false positives to event triggers, often leading to an overestimation of the background constant  $B$ . The size of this uncertainty was conservatively estimated by comparing the true and estimated signal efficiencies from the simulations and quantifying the fluctuation within each SK phase. In the Gd-loaded phase (SK-VI), the dominant source of uncertainty is in the fraction and the  $\gamma$  emission model of the  $\text{Gd}(n, \gamma)$  reaction. The size of these uncertainties are estimated based on the evaluated thermal neutron capture cross section uncertainties in ENDF/B-VII.1 and variations in the estimated signal efficiency with the modified ANNRI-Gd model where the continuum gamma cascade of the excited Gd was replaced by EGLO model instead of the default SLO model.

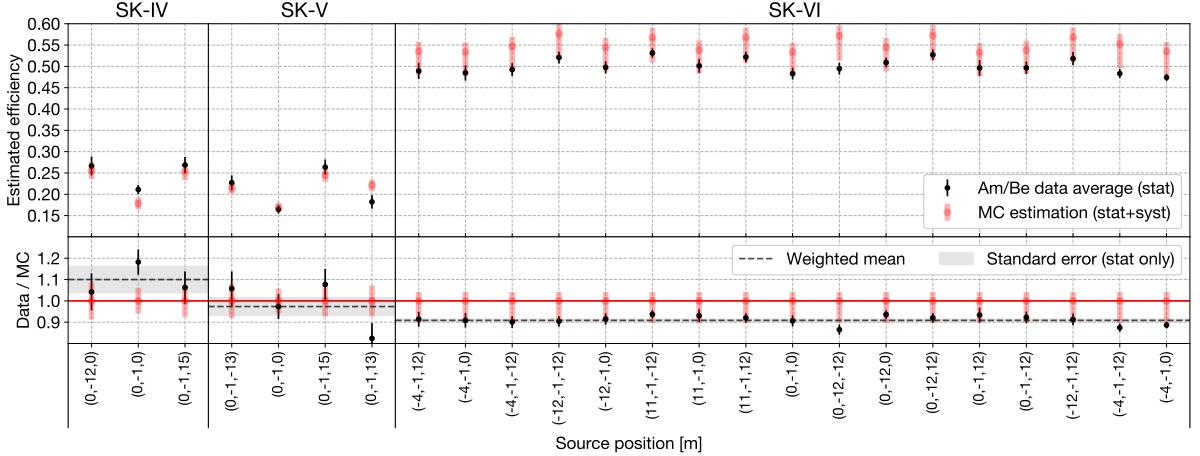


FIG. 9. Estimated neutron tagging efficiency for each calibration position within the tank. For positions in SK-VI where there are multiple measurements at different dates, the estimated efficiencies were averaged. The systematic uncertainties considered for MC simulation are summarized in Table II.

TABLE II. Major sources of systematic uncertainty in the neutron detection efficiency estimated with the Am/Be neutron source.

Source	SK-IV	SK-V	SK-VI
Am/Be neutron characterization	0.5%	0.9%	0.5%
Detector response	2.2%	3.3%	1.2%
Bias due to calibration setup	6.9%	4.6%	1.1%
Gd( $n, \gamma$ ) fraction	-	-	2.1%
Gd( $n, \gamma$ ) $\gamma$ emission model	-	-	2.6%
Total	7.3%	5.7%	3.8%

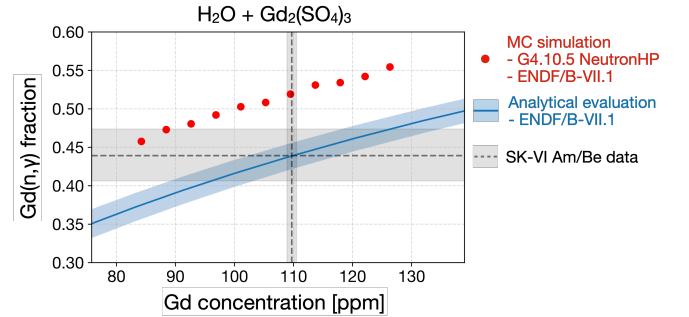


FIG. 10. Comparison between predicted and observed Gd( $n, \gamma$ ) fraction in Gd-loaded water as a function of Gd concentration. The red dots represent predictions by MC simulation using the NeutronHP model in Geant4.10.05.p01 with neutron cross sections from ENDF/B-VII.1, while the blue line and shades represent the analytically evaluated values using the same ENDF/B-VII.1 dataset and its associated uncertainties. Additionally, the gray line and shades show the estimated Gd( $n, \gamma$ ) fraction and its uncertainty.

The noticeable difference between the observed and predicted signal efficiencies in SK-VI is thought to stem from the overestimation of the Gd( $n, \gamma$ ) fraction in our MC simulation setup. Fig. 10 compares the MC-simulated Gd( $n, \gamma$ ) fraction  $r_{\text{Gd}}$  with the analytically evaluated fraction, assuming completely thermalized neutrons:

$$r_{\text{Gd}} \approx 1 - r_{\text{H}} \approx 1 - \frac{n_{\text{H}} g_{\text{H}}(T) \sigma_{\text{H}}(v_{\text{thermal}})}{\sum_i n_i g_i(T) \sigma_i(v_{\text{thermal}})} \quad (4)$$

Here, for the  $i^{\text{th}}$  isotope,  $n_i$  is the number density,  $g_i(T)$  is the Westcott  $g$ -factor for temperature  $T$ , and  $\sigma_i(v_{\text{thermal}})$  is the neutron capture cross section evaluated at thermal neutron speed  $v_{\text{thermal}} = 2200 \text{ m/s}$ .

The value estimated with SK-VI Am/Be data agrees well with the analytical evaluation based on ENDF/B-VII.1, while both differ from the MC simulation using Geant4.10.05.p01 NeutronHP and ENDF/B-VII.1. The reason seems to be that the NeutronHP model considers the thermal motion of hydrogen as free rather than bound in a water molecule, thus underestimating the fraction of  $^1\text{H}(n, \gamma)$  that competes with the Gd( $n, \gamma$ ) reaction.

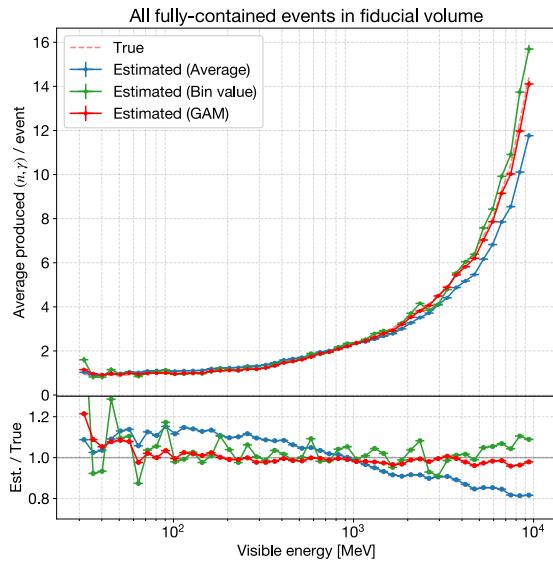
To account for such a difference between signal efficiencies evaluated with Am/Be data and MC simulation, we used the weighted mean of the ratios of the two evaluated at all Am/Be source positions as the correction factor. The obtained efficiency correction factor was  $1.10 \pm 0.11$  for SK-IV,  $0.97 \pm 0.07$  for SK-V, and  $0.91 \pm 0.04$  for SK-VI, including both statistical and systematic errors.

The average fitted neutron capture time constant for pure water phase data and Gd-loaded data were  $200.35 \pm 3.74 \mu\text{s}$  and  $116.9 \pm 0.3 \mu\text{s}$  respectively, aligning well with the predicted thermal neutron capture time constants  $204.7 \pm 5.3 \mu\text{s}$  and  $114.9 \pm 2.5 \mu\text{s}$  prediction based on ENDF/B-VII.1 cross sections and uncertainties.

622 C. GAM regression of signal selection performance<sup>647</sup>  
 623 on atmospheric neutrino simulation<sup>648</sup>

624 For atmospheric neutrino events, additional factors<sup>649</sup>  
 625 such as the resolution of the neutrino vertex reconstruc<sup>650</sup>  
 626 tion and the large outgoing neutron kinetic energy may<sup>651</sup>  
 627 affect the performance of signal selection. To improve the<sup>652</sup>  
 628 accuracy of signal detection performance regression, we  
 629 trained Generalized Additive Models (GAMs) on default  
 630 atmospheric neutrino event simulations. A total of six<sup>653</sup>  
 631 GAMs were constructed across the three SK phases, for  
 632 two output metrics: signal efficiency and false positive  
 633 rate, using LinearGAM in pyGAM 0.9.0.<sup>654</sup>

634 The input features considered were five reconstructed  
 635 neutrino event variables: visible energy, Cherenkov ring  
 636 multiplicity, the particle type of the most energetic ring,  
 637 and the radial and vertical displacements of the neutrino  
 638 interaction vertex. GAMs were fitted against the distri-  
 639 butions of the average simulated output metric per bin  
 640 in the five-dimensional feature space. No assumption of  
 641 feature correlation was made, and reasonable smoothing  
 642 was applied to prevent overfitting. The  $1\sigma$  prediction in-  
 643 terval was obtained based on the statistical errors in each  
 644 bin. Fig. 11 shows the performance of a trained GAM in<sup>664</sup>  
 645 estimating true average neutron capture multiplicity per<sup>665</sup>  
 646 visible energy bin, compared with alternative methods.<sup>666</sup>



675 FIG. 11. Comparison of three methods for estimating signal  
 676 detection performance in an unseen simulation dataset. The  
 677 average signal multiplicity is plotted against visible energy,  
 678 with signal efficiency and false positive rate estimated on an  
 679 event-by-event basis. The blue curve assumes a single aver-  
 680 age metric for all events. The green curve involves looking  
 681 up average metrics binned by reconstructed event variables.  
 682 The red curve uses regressed metrics with trained GAM. For  
 683 the red and green curves, binned performance metrics were  
 684 extracted from the default SK-IV MC simulation.

685 V. AVERAGE NEUTRON CAPTURE  
 686 MULTIPLICITY ESTIMATION

687 The average neutron capture signal multiplicity  $\langle N \rangle$   
 688 per visible energy bin is computed as the average of the  
 689 expected number of neutron captures estimated on an  
 690 event-by-event basis, as follows:

$$\langle N \rangle = \left\langle \frac{N_i^{\text{detected}} - N_i^{\text{BG}}}{\epsilon_i} \right\rangle \quad (5)$$

691 Here,  $N_i^{\text{detected}}$  is the count of detected signals,  $N_i^{\text{BG}}$  is  
 692 the estimated number of false positives, and  $\epsilon_i$  is the es-  
 693 timated signal detection efficiency of the  $i^{\text{th}}$  event.  $N_i^{\text{BG}}$   
 694 and  $\epsilon_i$  are estimated based on the GAM regression de-  
 695 scribed in Sec. IV C.  $\epsilon_i$  is further corrected by the factors  
 696 obtained by comparing Am/Be calibration data and sim-  
 697 ulation as described in Sec. IV B.

698 The following major systematic uncertainties affecting  
 699 signal counting were evaluated on a bin-by-bin basis:

700 (1) Signal efficiency scale

701 This includes uncertainties in calibrated efficiency cor-  
 702 rection factors, as well as neutron momentum and de-  
 703 tector modeling, which are accounted for by compar-  
 704 ing the results across different SK phases as described  
 705 in Sec. VII.

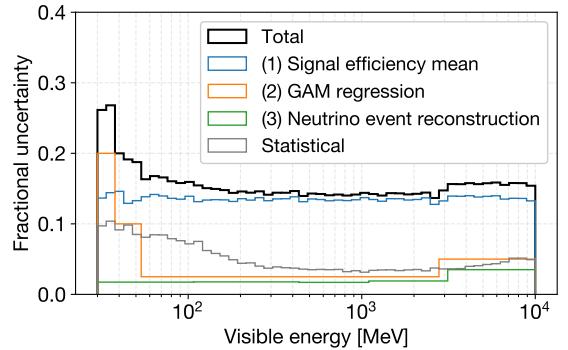
706 (2) GAM regression

707 This refers to the difference between the true signal  
 708 multiplicity in simulation and the signal multiplicity  
 709 estimated through the fitted GAM.

710 (3) Neutrino event reconstruction

711 This assumes a uniform 2% visible energy resolution.

712 Fig. 12 illustrates the distribution of fractional uncer-  
 713 tainties assigned to each visible energy bin. The most  
 714 dominant factor is the signal efficiency scale.



715 FIG. 12. Fractional uncertainties assigned to the average  
 716 signal multiplicity on a bin-by-bin basis.

## 678 VI. SPALLATION MODEL PREDICTIONS

679 To compare with the data, we prepared predictions  
 680 of the average neutron multiplicity versus neutrino  
 681 event visible energy using various combinations of neu-  
 682 trino event generators and secondary spallation mod-  
 683 els. We tested the following neutrino event genera-  
 684 tors: NEUT 5.4.0, NEUT 5.6.3, and GENIE 3.4.0 with  
 685 G\_18a\_10x\_02\_11b physics tunes, where  $x \in a,b,c,d$ . Here,  
 686 a, b, c, and d represent the use of INTRANUKE/hA,  
 687 hN, the Liège INC model (INCL++), and the Geant4  
 688 Bertini cascade model (G4Bertini) for FSI modeling, re-  
 689 spectively. NEUT 5.4.0 was used as described in Sec.  
 690 III C. NEUT 5.6.3 used the same interaction models as  
 691 NEUT 5.4.0 but with a different nuclear binding energy,  
 692 resulting in a slightly smaller fraction of CCQE inter-  
 693 actions. GENIE 3.4.0 with the specified physics tunes  
 694 uses similar models for QE and single-pion production as  
 695 the two NEUT versions but employs a different FSI and  
 696 final-state hadronization model. The ‘hA’ is an empirical  
 697 FSI model tuned to cross section data on composite nu-  
 698 clei, while the other three are full INC models based on  
 699 free nucleon cross-sections. INCL++ was coupled with  
 700 ABLA07 to describe nuclear de-excitation.

701 We also tested secondary spallation models using the  
 702 SK detector simulation software based on Geant3.21 and  
 703 Geant4.10.05.p01. The models tested were: SK-IV/V  
 704 default, SK-VI default, GCALOR, the Geant4 Bertini  
 705 cascade model (BERT\_HP), and the Liège INC model  
 706 (INCLXX\_HP). The first two are the default models for  
 707 the pure water phases and the Gd-loaded phase, differ-  
 708 ing in their use of MICAP (based on ENDF/B-V) or  
 709 NeutronHP (based on ENDF/B-VII.1) for low-energy  
 710 neutron propagation below 20 MeV. They also differ  
 711 from GCALOR, which uses its own hadron propaga-  
 712 tion for pion propagation, whereas the default options  
 713 use the NEUT pion FSI routine. The latter two mod-  
 714 els, BERT\_HP and INCLXX\_HP, were implemented in  
 715 Geant4.10.05.p01 and paired with NeutronHP for low-  
 716 energy neutron propagation. INCLXX\_HP is addition-  
 717 ally coupled with the Geant4 Precompound and Evapo-  
 718 ration models to describe nuclear de-excitation.

719 Predictions of the average neutron capture multiplicity  
 720 per visible energy bin for each model combination  
 721 were made as follows. First, we obtained the momen-  
 722 tum distributions of all particles resulting from neutrino-  
 723 nucleus interactions within each visible energy bin. For  
 724 hadrons that contribute most to neutron production —  
 725 namely, neutrons, protons, and  $\pi^\pm$  — we calculated the  
 726 average neutron capture multiplicity using a momentum-  
 727 to-multiplicity table generated with hadron particle gun  
 728 simulations in water. Finally, we determined the ratios  
 729 between each model combination and the default setup  
 730 (NEUT 5.4.0 with SK-VI default). These ratios were  
 731 then used to scale the average signal multiplicity obtained  
 732 from the full MC simulation in the default setup, provid-  
 733 ing a more accurate prediction by including unaccounted  
 734 neutron production from other hadrons.

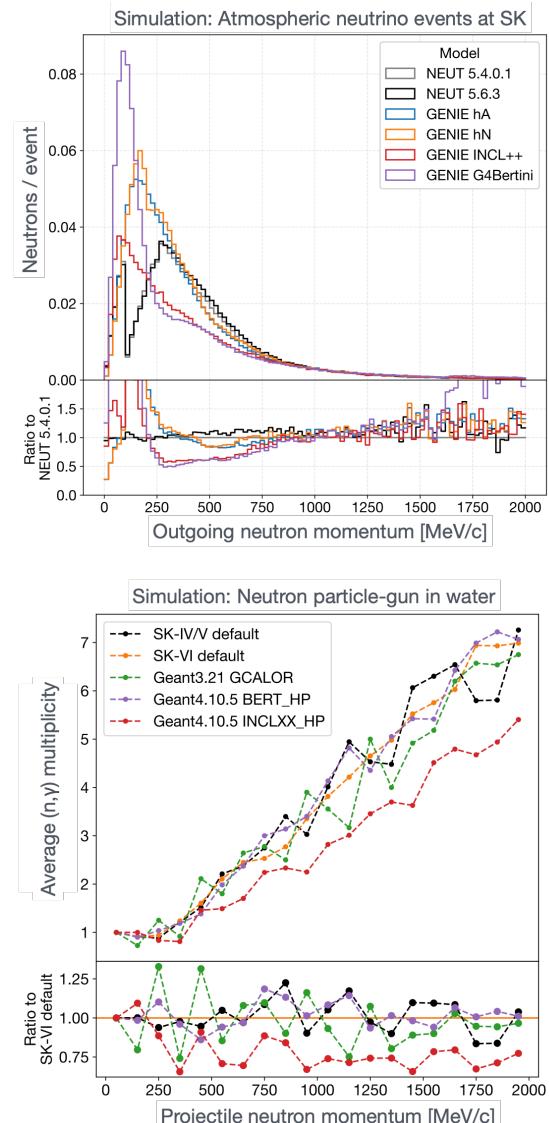


FIG. 13. Model comparison: the outgoing neutron momentum distribution per atmospheric neutrino event (top) and the average neutron capture multiplicity per projectile neutron momentum bin (bottom). In the bottom figure, each momentum bin includes roughly 100 hadron projectiles.

Figure 13 shows the outgoing hadron momentum distribution and the average signal multiplicity per projectile hadron momentum bin for neutrons, which are expected to account for more than 60% of secondary neutron production. The outgoing neutron and proton momentum distributions near or below the oxygen Fermi momentum (225 MeV/c) varied significantly among the different neutrino event generators. INCLXX\_HP predicted a 20-30% lower average neutron capture multiplicity across all projectile neutron momentum bins. The specific reasons for these differences in model predictions, in terms of physics modeling, are discussed in Sec. VIII.

## VII. RESULTS

Table III summarizes the number of atmospheric neutrino events in the final data sample and the number of detected neutron capture signals for each SK phase.

TABLE III. Summary of atmospheric neutrino events in the final data sample for analysis and detected neutron capture signals. The shown errors are purely statistical.

	SK-IV	SK-V	SK-VI
$\nu$ events	29,942	4,231	5,203
Events/day	$9.23 \pm 0.05$	$9.18 \pm 0.14$	$9.22 \pm 0.13$
Signals	15,705	2,035	5,752
Signals/event	$0.525 \pm 0.004$	$0.481 \pm 0.011$	$1.106 \pm 0.015$
$\langle N \rangle$	$2.04 \pm 0.01$	$2.33 \pm 0.05$	$2.36 \pm 0.02$

Figure 14 illustrates the consistency of the estimated average neutron capture multiplicity per visible energy bin across different SK phases. To assess the impact of uncertainty in the displacement of the neutron capture vertex from the neutrino vertex, an additional algorithm was applied to SK-VI data. This algorithm reconstructs the  $Gd(n, \gamma)$  vertex independently of the neutrino reconstruction, providing a reference for the results. The relative scale difference, which may be due to not fully accounting for all time variations in detector characteristics — especially during the extended SK-IV phase — was included as a systematic uncertainty in the signal efficiency mean, as described in Sec. V.

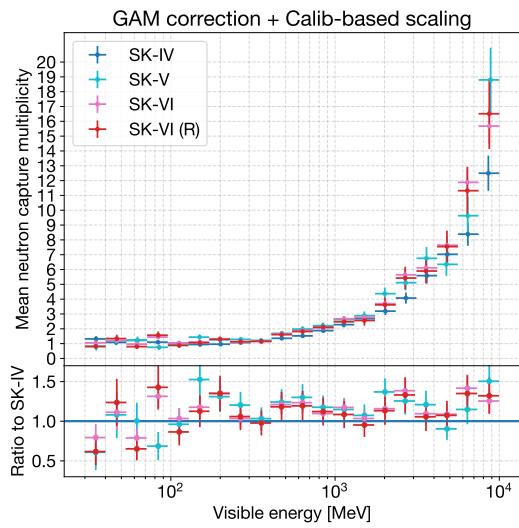


FIG. 14. Comparison of average neutron capture multiplicity across SK phases, using GAMs for signal selection performance regression, with corrections based on the Am/Be calibration. Only statistical errors are shown. The label SK-VI (R) uses an independently reconstructed  $Gd(n, \gamma)$  vertex as the assumed signal vertex, serving as a reference.

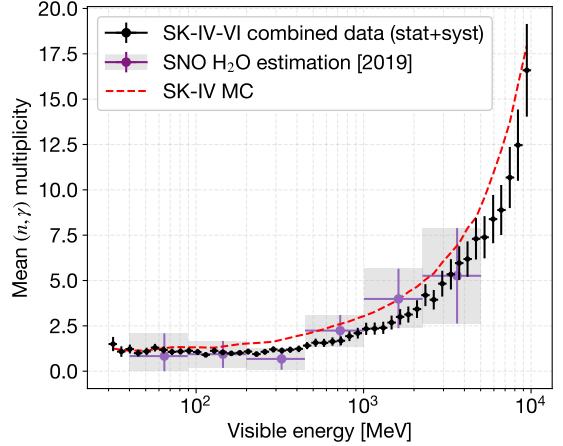


FIG. 15. Average neutron capture multiplicity estimated for atmospheric neutrino interactions versus visible energy. The black dots and error bars represent combined SK data, including statistical and systematic uncertainties. Purple crosses with shading indicate estimates of neutron production in water derived from SNO measurements using a D<sub>2</sub>O target volume. The red dashed line represents the true average neutron capture multiplicity per visible energy bin obtained from SK-IV MC simulation.

The combined data estimate was compared with the SK-IV simulation and previous neutron production estimates in water from SNO, as shown in Fig. 15. While the combined data estimate aligned well with the SNO estimate, it was 20-40% lower than the SK-IV simulation results.

### A. Linearity

We observed a linear relationship between the average neutron capture multiplicity and the visible energy of each event. This relationship was modeled using a function of the form  $y = ax + b$ , where  $x$  represents the average visible energy in each bin,  $y$  stands for the average neutron capture multiplicity, and  $a$  and  $b$  denote the slope and intercept, respectively, as free parameters. The dataset was categorized by SK phases as well as single-ring and multi-ring events.

The data exhibit consistency across different SK phases. Furthermore, both the slopes and intercepts were greater in multi-ring events compared to single-ring events, as anticipated from the SK-IV MC simulation. This aligns with the notion that multi-ring events are more frequently caused by DIS events, resulting in higher energy transfer to the hadron system, as depicted in Fig. 4. However, both the observed slopes and intercepts were notably smaller than those expected from the MC simulation.

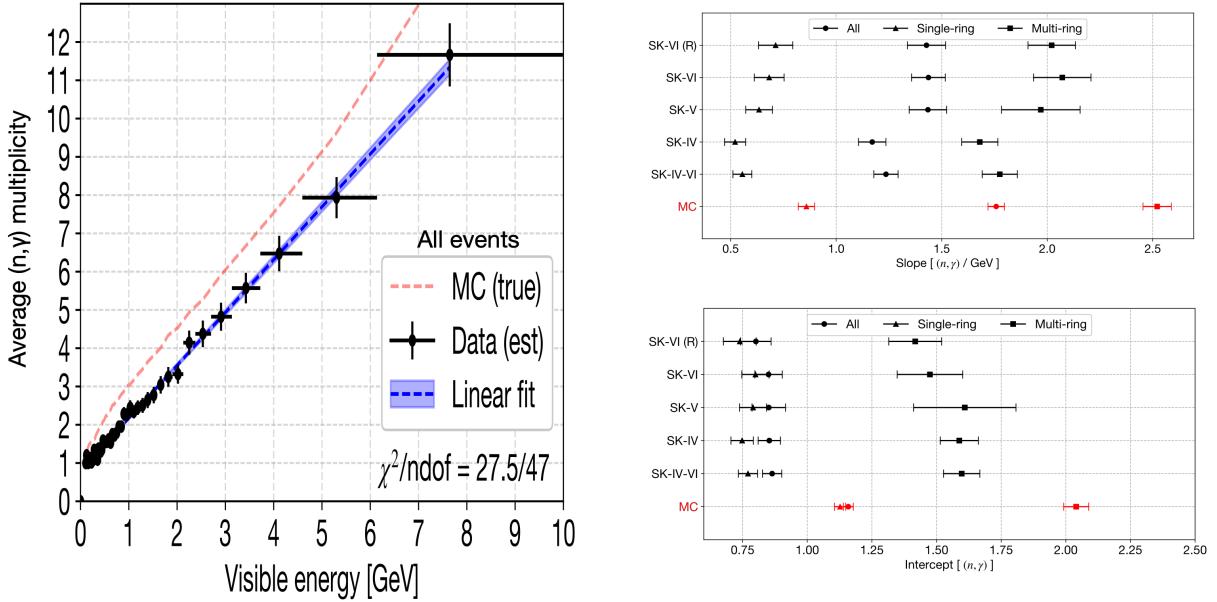


FIG. 16. A linear fit for the overall final data sample (left) and the fitted slopes and intercepts for each subcategory (right). The error bars on the left plot include both statistical and systematic errors. The shaded area on the left plot and the error bars on the right plot represent the standard errors associated with the linear fit. The label ‘MC’ denotes the SK-IV simulation.

## 791 B. Comparison with spallation model predictions

792 Predictions from various model combinations, as de-  
 793 scribed in Sec. VI, were compared with the combined  
 794 data results, as shown in Fig. 17. Compared to the  
 795 default SK-VI simulation setup, represented by the red  
 796 dashed line, the data showed a 20-40% deficit, partic-  
 797 ularly in the  $O(1)$  GeV visible energy range. Varying  
 798 neutrino event generator options, which mainly affect the  
 799 outgoing hadron kinematics from the target nucleus as  
 800 illustrated in Fig. 13, resulted in a 5-15% change in the  
 801 overall prediction. However, the combined data estima-  
 802 tion was clearly inconsistent with the prediction range of  
 803 the various neutrino event generator options. Changing  
 804 the secondary spallation model option had a more signif-  
 805 icant impact, causing a -20% to +20% difference in the  
 806 overall prediction compared to the default setup.

807 The secondary spallation model with the smallest pre-  
 808 diction was INCLXX\\_HP, and its combination with the  
 809 default neutrino event generator option NEUT 5.4.0  
 810 showed the best agreement with our data. This is ev-  
 811 ident when comparing the fitted slopes and intercepts  
 812 of each model combination’s predictions with those of  
 813 the combined data, as illustrated in Fig. 18. There  
 814 was a clear distinction in the fitted slopes between IN-  
 815 CLXX\\_HP and the other secondary spallation options,  
 816 especially in multi-ring events with larger hadron pro-  
 817 duction. There was also a slight variation in the fitted  
 818 intercepts among different event generator options, with  
 819 NEUT and GENIE 3.4.0 combined with the INCL++  
 820 option aligning better with the data.

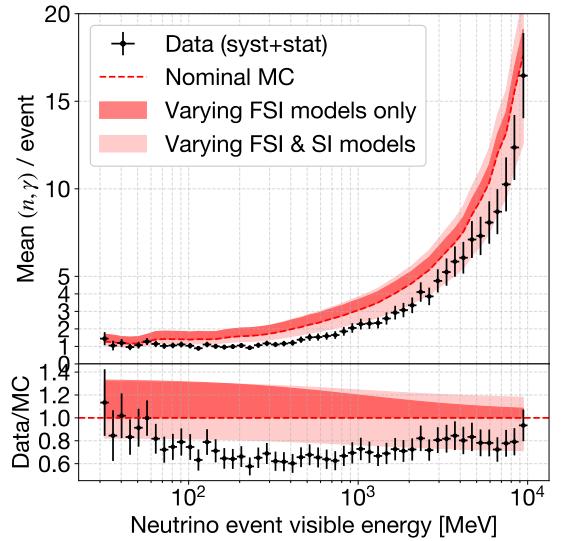


FIG. 17. Comparison of various model predictions with the combined data, showing average neutron capture multiplicity as a function of atmospheric neutrino event visible energy. The black dots with error bars represent the combined data results, including statistical and systematic uncertainties. The dashed red line represents the prediction from the full SK-VI MC simulation using the default setup. The thick red shaded area shows the range of predictions when varying the neutrino event generator options, while keeping the secondary spallation model fixed to the SK-VI default. The light red shaded area shows the range of predictions from all tested model combinations.

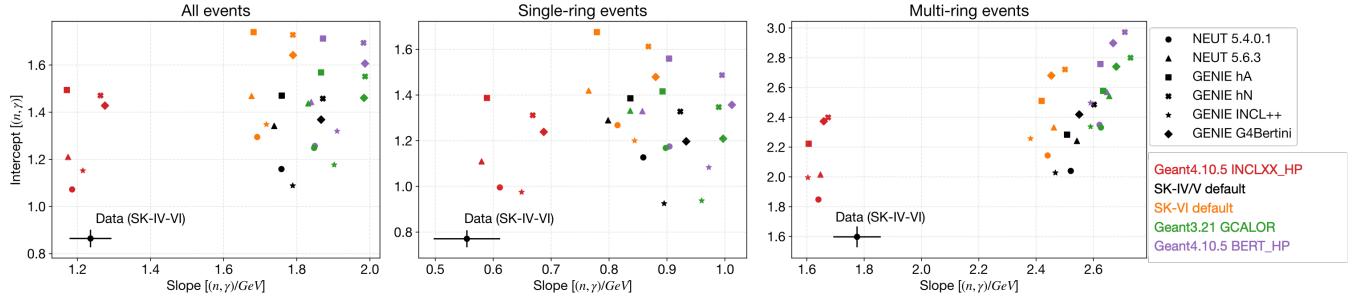


FIG. 18. Scatter plots of fitted slopes and intercepts in model predictions and data (black circles with error bars), shown for all events (left), single-ring events (middle), and multi-ring events (right). Different shapes represent various neutrino event generator options, while different colors indicate the secondary spallation model options used.

821

### VIII. DISCUSSION

822 Figure 19 shows that the linear slope is directly related  
 823 to the secondary spallation model, while the fitted intercept  
 824 is largely independent of the secondary spallation  
 825 model but more dependent on the neutrino event gener-  
 826 ator option. For low visible energy, most neutrons origi-  
 827 nate from the neutrino interaction on the target nucleus,  
 828 but for higher visible energy events, neutrons produced  
 829 through secondary spallation reactions in water predom-  
 830 inate.

831 The observed relationship between the neutrino event's  
 832 visible energy and the average neutron capture multi-  
 833 plicity provides discriminant power for secondary spal-  
 834 lation models, as indicated by the separation between  
 835 INCLXX.HP and other models in Figure 18. The other  
 836 secondary spallation models tested are based on the orig-  
 837 inal Bertini cascade model with a nuclear de-excitation  
 838 model by Griffin and Weisskopf et al., hence showing sim-  
 839 ilar slopes.

840 Figure 13 offers insights into the potential reasons.  
 841 For example, comparing the outgoing neutron momen-  
 842 tum distribution between the two GENIE FSI options,  
 843 INCL++ and G4Bertini, suggests that the difference in  
 844 slopes between INCLXX.HP and the other models is  
 845 largely due to neutrons below Fermi momentum pro-  
 846 duced through the de-excitation process. A significant  
 847 reduction in neutron and gamma-ray multiplicity has  
 848 been reported when the nuclear de-excitation module  
 849 is changed from the Bertini native to the Geant4 Pre-  
 850 compound model. Additionally, different cascade stop-  
 851 ping criteria can result in different residual energies left  
 852 for de-excitation, leading to varying nucleon emission  
 853 rates. The Liège model's consideration of cluster forma-  
 854 tion among outgoing nucleons near the nuclear surface  
 855 is also expected to reduce the number of outgoing neu-  
 856 trons. Our results highlight the importance of accurately  
 857 modeling secondary spallation to predict neutron multi-  
 858 plicity in neutrino events, suggesting that overpredictions  
 859 reported in previous studies, especially at lower neutron  
 860 energies, could be due to using the Bertini cascade model  
 861 or its variants.

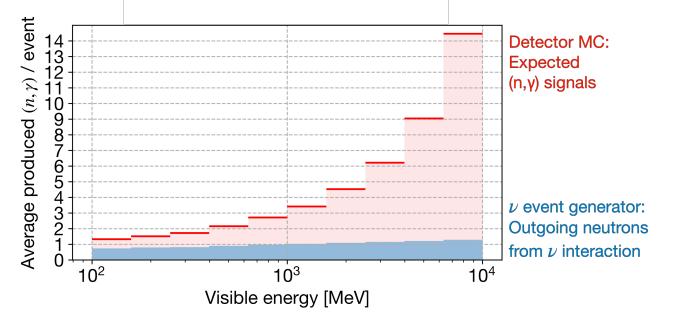


FIG. 19. Average neutron capture multiplicity as a function of visible energy, as predicted by the default SK-IV simulation setup. The blue histogram shows the number of neutrons pre-  
 dicted by the neutrino event generator (NEUT 5.4.0), while the red bars represent the average total neutron captures per bin, accounting for the effect of secondary spallation modeled by Geant 3.21 in the SK-IV default setup.

Regarding differences among neutrino event generator options, Pauli blocking implementation appears to be a significant factor. NEUT strictly blocks all nucleon ejection below oxygen Fermi momentum, resulting in a large dip in the outgoing neutron momentum distribution near Fermi momentum, while GENIE native hN and hA do not consider Pauli blocking. The Liège model uses probabilistic Pauli blocking, considering nucleon holes. Additionally, INCL++ and G4Bertini, with medium corrections such as nucleon repulsion and correlations of nucleon interactions, predict much lower neutron production at medium energies compared to other options.

The largest difference among neutrino event generator predictions below 100 MeV visible energy is about 30% between GENIE hA/hN and NEUT 5.4.0 predictions with the same secondary spallation model. These model differences are comparable to our current systematic uncertainty estimate at the lower energy end, and there is additional shape uncertainty that needs to be considered in the linearity fit. However, FSI models with lower neutron production are preferred.

883

## IX. SUMMARY

940

## ACKNOWLEDGMENTS

We gratefully acknowledge the cooperation of the Kamioka Mining and Smelting Company. The Super-Kamiokande experiment has been built and operated from funding by the Japanese Ministry of Education, Culture, Sports, Science and Technology, the U.S. Department of Energy, and the U.S. National Science Foundation. Some of us have been supported by funds from the National Research Foundation of Korea NRF-2009-0083526 (KNRC) funded by the Ministry of Science, ICT, and Future Planning and the Ministry of Education (2018R1D1A1B07049158, 2021R1I1A1A01059559), the Japan Society for the Promotion of Science, the National Natural Science Foundation of China under Grants No.11620101004, the Spanish Ministry of Science, Universities and Innovation (grant PGC2018-099388-B-I00), the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Scinet and Westgrid consortia of Compute Canada, the National Science Centre (UMO-2018/30/E/ST2/00441) and the Ministry of Education and Science (DIR/WK/2017/05), Poland, the Science and Technology Facilities Council (STFC) and GridPPP, UK, the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no.754496, H2020-MSCA-RISE-2018 JENNIFER2 grant agreement no.822070, and H2020-MSCA-RISE-2019 SK2HK grant agreement no. 872549.

884 Accurately modeling neutron production in neutrino<sup>941</sup>  
 885 interactions is essential for characterizing incoming neu<sup>942</sup>  
 886 trinos, which is crucial for advancing precision mea<sup>943</sup>  
 887 surements of neutrino oscillation parameters and rare<sup>944</sup>  
 888 event searches involving neutron tagging. Recent studies<sup>945</sup>  
 889 have suggested potential inaccuracies in the widely used<sup>946</sup>  
 890 Bertini cascade model for describing secondary spallation<sup>947</sup>  
 891 caused by energetic hadrons.<sup>948</sup>

892 In this paper, we present the measurement of total<sup>949</sup>  
 893 neutron production following atmospheric neutrino in<sup>950</sup>  
 894 teractions within the Super-Kamiokande (SK) detector's<sup>951</sup>  
 895 water volume. In SK, neutrons with kinetic energies be<sup>952</sup>  
 896 low a few MeV are typically captured by surrounding<sup>953</sup>  
 897 nuclei, such as  $^1\text{H}$  or  $^{155}/^{157}\text{Gd}$ , radiating 2-8 MeV of<sup>954</sup>  
 898 energy through gamma-rays. We used a neutron cap<sup>955</sup>  
 899 ture detection algorithm based on a simple low-energy<sup>956</sup>  
 900 trigger and a neural network binary classification, cali<sup>957</sup>  
 901 brated with an Am/Be neutron source. Neutron capture<sup>958</sup>  
 902 multiplicity was estimated on an event-by-event basis by<sup>959</sup>  
 903 predicting detection efficiency using a multivariate non<sup>960</sup>  
 904 linear regression technique. Atmospheric neutrino events<sup>961</sup>  
 905 were binned by their electron-equivalent "visible energy,"<sup>962</sup>  
 906 which is a semi-calorimetric measure of the neutrino mo<sup>963</sup>  
 907 mentum transfer. The average neutron capture multi<sup>964</sup>  
 908 plicity per visible energy bin was then compared with<sup>965</sup>  
 909 predictions from various combinations of neutrino event<sup>966</sup>  
 910 generator options and secondary spallation models.<sup>967</sup>

911 Our study stands out from previous neutron produc<sup>968</sup>  
 912 tion measurements due to several key aspects. We uti<sup>969</sup>  
 913 lized the largest dataset of neutrino events in water, cov<sup>970</sup>  
 914 ering the widest visible energy range from 30 MeV to<sup>971</sup>  
 915 10 GeV. This extensive dataset provides significant dis<sup>972</sup>  
 916 criminative power for spallation models. With reduced<sup>973</sup>  
 917 uncertainties, we quantified the linearity in observed av<sup>974</sup>  
 918 erage neutron capture multiplicity versus visible energy.<sup>975</sup>  
 919 The observed differential increase of the average signal<sup>976</sup>  
 920 multiplicity per visible energy was larger with multi-ring<sup>977</sup>  
 921 events, characterized by a higher fraction of inelastic in<sup>978</sup>  
 922 teractions, compared to single-ring events.

923 We also evaluated the performance of the Liège model<sup>979</sup>  
 924 (INCL++), which includes several low-energy correc<sup>980</sup>  
 925 tions and has shown good agreement with measurements<sup>981</sup>  
 926 within INC energy limits, as an alternative to the Bertini<sup>982</sup>  
 927 model. While a neutron deficit was observed compared<sup>983</sup>  
 928 to predictions using the Bertini cascade model and its<sup>984</sup>  
 929 variants for secondary spallation, better agreement was<sup>985</sup>  
 930 achieved with the Liège cascade model for fast intranu<sup>986</sup>  
 931 clear hadron transport, coupled with the Geant4 Precom<sup>987</sup>  
 932 pound and Evaporation models for nuclear de-excitation.<sup>988</sup>  
 933 This combination predicted significantly fewer neutrons<sup>989</sup>  
 934 below the oxygen Fermi momentum compared to the<sup>990</sup>  
 935 Bertini model. This highlights the importance of the ac<sup>991</sup>  
 936 curacy of secondary spallation models in predicting neu<sup>992</sup>  
 937 tron multiplicity in neutrino events and explains the pre<sup>993</sup>  
 938 viously reported deficit in low-momentum neutrons com<sup>994</sup>  
 939 pared to predictions using the Bertini cascade model.

