



Evaluation of potential biases in the detection efficiencies of the KM3NeT ARCA and ORCA detectors

THESIS

submitted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE

in

PHYSICS

Author :	Annet Konings
Student ID :	s2124645
Supervisor :	D.F.E. Samtleben
2 nd corrector :	T.H. Oosterkamp

Leiden, The Netherlands, June 17, 2022

Evaluation of potential biases in the detection efficiencies of the KM3NeT ARCA and ORCA detectors

Annet Konings

Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

June 17, 2022

Abstract

KM3NeT is a research infrastructure consisting of two neutrino detectors in the Mediterranean Sea. The ARCA detector searches for high-energy cosmic neutrinos, whereas with ORCA the fundamental properties of neutrinos are studied. Three possible biases in the detection efficiencies for ORCA and ARCA are investigated. Light from radioactive nearby K-40 decay compared to isotropic light from far away, and anisotropic light emitted by muons compared to simulated muon light are used to examine these biases for wavelength- and angular dependency. Firstly, certain DOMs in ARCA and ORCA show a lower efficiency due to a different gel transparency in the PMTs. Shadowing effects cause a bias in both ORCA (10%) and ARCA (14%). Finally, the sedimentation on top of ARCA DOMs affects the detection efficiencies in a different way than it does on ORCA DOMs.

Contents

1	Introduction	4
1.1	Neutrinos	4
1.2	KM3NeT	5
1.3	Muons	6
1.4	Cherenkov radiation	7
1.5	Background light	8
1.6	Research aims	10
2	Methods	13
2.1	Research instruments	13
2.2	Treatment of data	15
2.3	Calibration and simulation	16
2.4	Research data	17
2.4.1	Methods	19
3	Results	22
3.1	ARCA	22
3.1.1	Time residuals	22
3.1.2	Possible biases due to change in gel transparency . .	28
3.1.3	Shadowing and enhanced dark rate	40
3.1.4	Radioactivity and sedimentation	51
3.2	ORCA	59
3.2.1	Time residuals	60
3.2.2	Possible biases due to change in gel transparency . .	63
3.2.3	Shadowing	67
3.2.4	Radioactivity and sedimentation	69
4	Conclusion	76
A	Appendix	79
A.1	ARCA runs 9791-9796	79
A.2	ORCA	90

Introduction

1.1 Neutrinos

“How did the universe begin?” is one of the most fundamental questions of humanity. Particle physics and cosmology, two research fields engaged in questions such as this one, have known many great advances in the last 100 years. With research of elementary particles, new particles were discovered and more puzzles solved. However, the question of the start of the universe remains partly unanswered. It is also a difficult question to answer, since the universe has only existed for 13-14 billion years and not all particles of the early universe can reach us; some are even decayed due to interactions. Yet research is done that may shed light on this and other burning questions in particle physics and cosmology. Some of these researches focus on a tiny particle to answer big questions like these. This particle is called a neutrino.

Neutrinos are elementary particles of the lepton family. They have spin $\frac{1}{2}$ and are thus fermions. However, unlike electrons, they carry no electric charge and can only interact via two out of four fundamental forces: gravity and the weak force. The existence of these particles was hypothesised by Wolfgang Pauli in 1930. Physicists noticed that two conservation laws were violated during beta decay: the conservation of momentum and the conservation of energy. When the nucleus spontaneously emitted an electron, the final product had too little energy. The momentum conservation seemed to be violated since the proton was standing still instead of being pushed in the opposite direction of the electron. Because of the discrepancies in energy and momentum, physicists began looking for the particle that could solve this puzzle. Its solution was found a few decades later.

Clyde Cowan and Frederick Reines built a neutrino detector near a power plant. The fission reaction in this nuclear reactor produced the neutrinos they were looking for, confirming the existence of the neutrino in 1956 [1].

The Standard Model of particle physics gives us three ‘flavours’ of neutrinos and their anti-particle: the electron neutrino ν_e , muon neutrino ν_μ and tau neutrino ν_τ . A neutrino can change into the other flavours, which is called neutrino oscillation. This is only possible if the neutrinos have a mass and the masses of the flavours slightly differ. Instead of mass eigenstates the neutrinos ν_e , ν_μ and ν_τ have weak interaction eigenstates. Any neutrino weak interaction eigenstates can be expressed as a linear superposition of orthogonal mass eigenstates [2].

On February 14, 2022, physicists published an article showing that a neutrino has a mass of no more than $0.8\text{eV}/c^2$ [3]. This allows them to travel at a speed close to the speed of light.

Apart from being created in nuclear reactors and experiments, neutrinos are created at several places in the universe. Electron neutrinos are created when stars use their ‘fuel’ in hydrogen burning. The collapse of a star into a supernova creates neutrinos that are released into the universe. Another example is that of cosmic ray nuclei interacting with particles in the Earth’s atmosphere. This starts a cascade where neutrinos are produced in decay reactions [2]. Several decay chains of isotopes such as U-238 and K-40 emit anti-neutrinos via beta decay. These are naturally present on Earth and provide information on radioactivity in the Earth [4].

Neutrino experiments may provide answers to many unanswered questions and unexplained phenomena. What did the universe look like in the first seconds after the Big Bang? Why is there much more matter than antimatter? And questions about neutrinos themselves: what are their masses? Do sterile neutrinos exist? To answer these questions and more, experiments are conducted by physicists all over the world. For example, IceCube uses detectors in the Antarctic ice. Another detector operates in the Mediterranean Sea: KM3NeT.

1.2 KM3NeT

KM3NeT, which stands for ‘cubic kilometre Neutrino Telescope’, is a state-of-the-art research infrastructure in the Mediterranean Sea. It is currently

under construction. The completed telescope will span a volume of several cubic kilometres. There are two detectors: ORCA and ARCA, each with a different aim. The ORCA (Oscillation Research with Cosmics in the Abyss) detector is located at 40 km off the coast of Toulon, France. With ORCA, fundamental properties of neutrinos are studied. Detecting neutrinos (few-GeV range) created by interactions of cosmic rays with the atmosphere of the Earth will give insights into the relative ordering of neutrino masses [5]. The aim of ARCA (Astroparticle Research with Cosmics in the Abyss) is to search for very high-energy (≥ 1 TeV) cosmic neutrinos. These neutrinos provide information about their origin in the universe, energy spectrum, and flavour composition. ARCA is located at a distance of 100 km from Portopalo di Capo Passero, Sicily, Italy [6].

The two detectors consist of several vertical strings. The strings consist of two ropes holding eighteen digital optical modules (DOMs). These DOMs are glass spheres of 42 cm in diameter, housing the light sensors [7]. The DOMs are designed to withstand the enormous pressure of the seawater: up to 350 times the normal atmospheric pressure. A total of 31 light sensors or PMTs (photomultiplier tubes) are distributed over a DOM. This distribution is such that the PMTs detect passing light coming from all directions [8].

1.3 Muons

Muons, discovered in 1936, are a member of the lepton family. Just as electrons and the tau particle they have a charge of $-e$, but their mass is 207 times that of the electron. Another important difference between the electron and the muon is their lifetimes: electrons are stable particles with a lifetime of at least $6.6 \cdot 10^{28}$ years, whereas the muon's lifetime averages $2.2 \mu\text{s}$ [9]. One of the sources of muons on Earth is a product of reactions involving cosmic rays. Cosmic rays consist mainly of atomic nuclei, including hydrogen nuclei, protons and helium nuclei. Although the exact sources of the cosmic rays are unknown, they are hypothesised to come from explosions from supernovae and AGN (Active Galactic Nuclei). When these relativistic particles hit the atmosphere they interact with atmospheric particles, creating pions and kaons. In turn, the pions and kaons react or decay into muons and muon neutrinos. Since the lifetime of a muon is short, it decays into an electron, muon neutrino, and electron antineutrino:

$$\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e.$$

The other possible decay is that of an anti-muon into a positron, muon antineutrino, and an electron neutrino via:

$$\mu^+ \rightarrow e^+ + \bar{\nu}_\mu + \nu_e.[2]$$

Despite its short lifetime a high-energy muon has a chance to reach the ground before decaying, because of length contraction. This makes it possible to detect them on Earth.

1.4 Cherenkov radiation

Since neutrinos rarely interact with matter, neutrino detectors detect these particles indirectly. That is, neutrino telescopes detect the light emitted by electrically charged leptons and hadrons created by neutrino interactions. The second ingredient for neutrino detection is the medium. For a refractive index n the speed of light in the medium is c/n . When the charged particles move faster than this, they emit photons in a blueish light: Cherenkov radiation. This radiation is a coherent wavefront or shockwave comparable to the sonic boom from a supersonic aircraft. The emission angle θ_C of the wavefront relates to the refractive index and the speed of the particle in units of the speed of light ($v = \beta \cdot c$) as

$$\cos \theta_c = \frac{1}{\beta n}.$$

θ_c is also called the Cherenkov angle. Thus, the Cherenkov effect takes place for $\beta > 1/n$. These photons from the Cherenkov wavefront are exactly the photons that hit the detector. The direction the neutrino came from can be deduced from the direction of the charged lepton. Many neutrino detectors are either situated in ice or water. With a refractive index of around 1.3 for optical and UV light, and the need for very large detection areas, the (Ant-)Arctic ice and seas and lakes are the best choice for detectors. For a relativistic particle ($\beta \approx 1$) in seawater ($n \approx 1.35$) the corresponding Cherenkov angle is approximately 42° [2, 10]. The KM3NeT detectors use this Cherenkov principle as well. An artist's impression is given in figure 1.1.

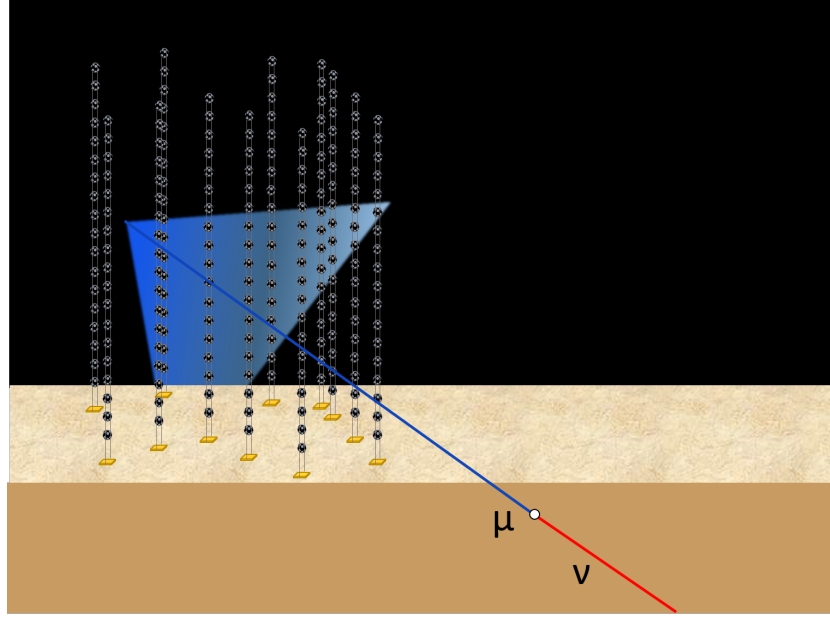


Figure 1.1: An artists impression of the KM3NeT detector. It shows the signal from a muon passing the strings. Its Cherenkov light is then detected by the DOMs [11].

1.5 Background light

Two kinds of background are detected by the telescope. The first one is random background: dark current and K-40 decay. The second one is non-random background, caused by bioluminescence and atmospheric muons. For ARCA, which detects cosmic neutrinos, the atmospheric neutrinos are also regarded as noise [12].

⁴⁰K decay

Potassium-40 is a naturally occurring radioactive isotope in seawater. The main decay reaction (89.3%) is via β -decay, resulting in ⁴⁰C, an electron, and an electron antineutrino:

$$^{40}\text{K} \rightarrow ^{40}\text{C} + e^- + \bar{\nu}_e$$

The produced electron sometimes surpasses the Cherenkov limit and emits Cherenkov radiation. However, a large fraction of the unwanted K-40 background can be filtered out to an acceptable level.

^{40}K is not only a nuisance: the PMTs are calibrated in dark rooms using this isotope, since it is also present in the glass of the DOMs. The difference between the time-offset of a PMT with respect to the DOM offset is determined with the use of K-40 in 'in-situ' calibration [10].

Dark current

Dark current is a result of (thermo-)electric noise in the PMTs. It is mostly due to spontaneous thermal emission from electrons by the photocathode [13].

Bioluminescence

Organisms like algae and other creatures in the sea emit *bioluminescent light*. This light is detected on a single DOM, resulting in an increased hit rate lasting several seconds. To filter out this bioluminescence, the rates of PMTs exceeding 20 kHz are not sent to shore. This is less present in ARCA than in ORCA, since ARCA is located at larger depth and a different location in the Mediterranean Sea.

Atmospheric muons

The detectors are placed deep in the sea to minimise the number of muons produced in cosmic ray air showers reaching the detector. The atmospheric muon flux that remains is still several order of magnitudes larger than the expected muon rate induced by (cosmic) neutrinos. It is important to filter out this atmospheric muon flux, since they can produce hit patterns that mimic patterns of up-going muons and signals of neutrino-interactions. As an atmospheric muon traversing the Earth is almost impossible, only selecting muons with a reconstructed upwards trajectory filters out the atmospheric muons. However, the possibility exists that the reconstruction is not 100% accurate, resulting in atmospheric muons which are not excluded.

Since muons propagating downwards in the seawater lose energy, there is less Cherenkov light from atmospheric muons at larger depth. The PMT efficiencies can be determined from calibration, using this depth-dependence [14].

1.6 Research aims

Currently, KM3NeT has 10 strings for ORCA and 19 strings for ARCA in operation. This will be 115 and two times 115, respectively, when the detectors are fully operational. ORCA will take up a total volume of about 4 km^3 and ARCA of approximately 1 km^3 . Before and during the deployment of the strings, their performance must be evaluated. The first ones went into use in 2015. Phase 2.0 consists of the extension to implement all 115 ORCA respectively 230 ARCA strings [8].

It is no surprise that KM3NeT wants to achieve the best possible resolution and accuracy for the detectors. A better resolution means more precise measurements. Therefore, strings, DOMs, and PMTs must be tested and checked for performance biases.

Several factors that may affect the efficiency of the detectors have been and are being studied. One is a possible bias caused by the gel used in the PMTs. It is hypothesised that the gel used in some DOMs has different transmission properties than expected, whereby the differences could be wavelength-dependent. Two reports [15, 16] showed a lower efficiency for ORCA.0001 DOMs 1-6. The DOMs/PMTs that possibly have a lower efficiency due to different transmission properties of the gel for ARCA are: DOMs 7-15 of ARCA.0010, all DOMs of ARCA.0012, and the top half of ARCA.0013: E1-E6 and F1-F6 (PMTs 0-11).

Another example is the sedimentation on top of DOMs (see fig. 1.2). The longer the DOMs have been in the water, the more sedimentation has accumulated on the DOMs. The sedimentation for ORCA and ARCA are different: for ORCA the sedimentation comes mostly from organic sediments, while the sediments of the ARCA detector are expected to consist of volcanic ashes. Since the sedimentation lies on top of the (upward-facing) PMTs, this blocks the light from hitting the PMTs, resulting in a lower efficiency for these PMTs.

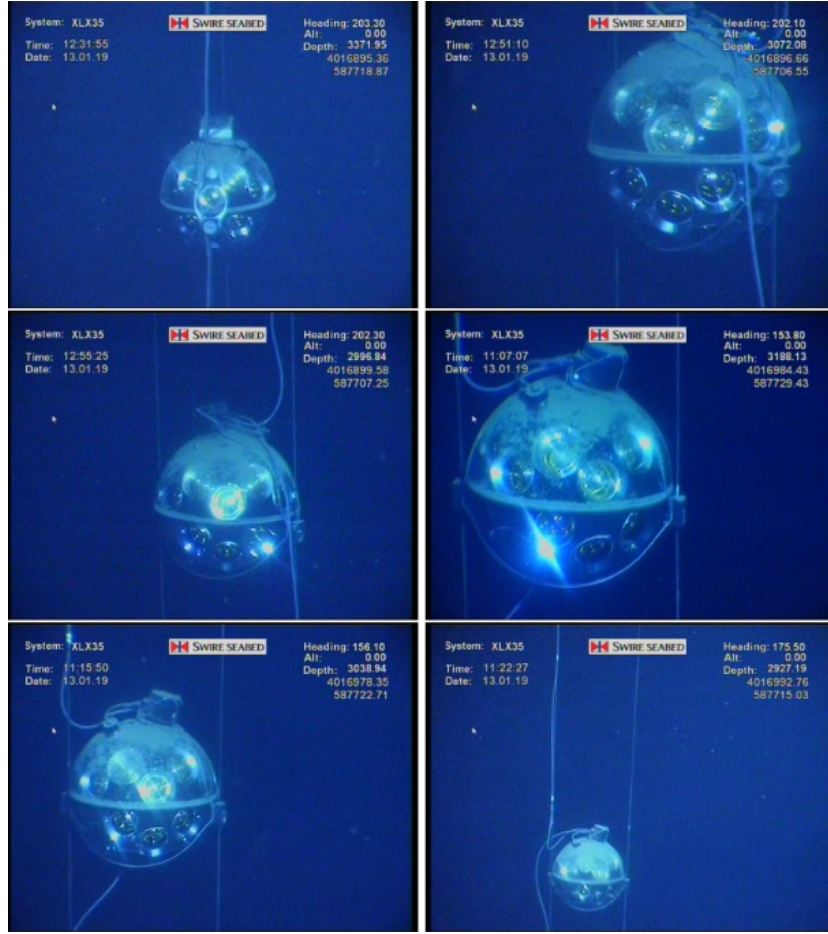


Figure 1.2: Picture of sedimentation on top of ARCA DOMs. These photos were taken during an inspection in 2019. Visible are (top left to bottom right): ARCA.0001 DOM 2, 10, 12, ARCA.0002 DOM 7, 11, 14 [17].

Studies [18] indicate that there is a decrease in efficiency caused by the structure of the DOMs, i.e. the equator tape and titanium collar (see figure 2.1). These structures are necessary, for example, to attach the DOMs to the cables of the string. A disadvantage is that they impede the photons from certain directions reaching the PMTs.

In this thesis the performance biases of the KM3NeT detectors ARCA and ORCA will be presented. To be more precise, three biases will be looked at. 1) The possible bias due to a difference in gel transparency. The efficiency is evaluated to be low (using nearby Cherenkov light from K40 decays), while the effective efficiency towards muon light might be different. The muon light has to travel a larger distance than the Cherenkov light from K40 decays. As a result the wavelength spectrum of this muon

light is different. 2) Shadowing of the equator tape and the titanium collar. A possibly enhanced dark rate is covered in this section as well. 3) The potential impact of radioactivity of the DOM glass in the efficiency evaluation for PMTs with large sedimentation. Normally, this signal is negligible and therefore neglected in the analysis of the data. However, when the top half of the DOM is covered due to sedimentation, the signal that the radioactivity sends to the PMT may be too substantial to neglect.

First, more information is presented in chapter 2 on how the detector works and the data acquisition. Then chapter 3.1 will address the three biases for the ARCA detector. The efficiencies of ORCA are covered in chapter 3.2. Finally, a conclusion is presented.

Chapter 2

Methods

This chapter will go more into depth into the data taking process and calibration. The working of the detector and its components is explained. With this knowledge, the data acquisition will be addressed.

2.1 Research instruments

This section gives some more information on characteristics of the several components of the ORCA and ARCA telescopes.

- Building blocks. ORCA consists of one building block with a radius of 106 meters; the two building blocks of ARCA have a radius of 538 meters each.
- Strings. ORCA strings are 20 meters apart and for ARCA the distance between strings is 90 meters. The top DOM for ORCA is at a depth of 2232 m, the bottom at 2385 m and the seabed at 2425 m. ARCA's top DOM, bottom DOM and the seabed are at depths of 2773 m, 3385 m and 3450 m, respectively. At the bottom of a string, a dead weight is attached to an anchor on the seabed. A buoy keeps the strings upright at the top. The DOMs, cables, anchor, and buoy together form a Detection Unit or DU for short. Each detection unit is assigned a number from 1 to 115 in order of deployment. For example, the name ARCA DU1 belongs to the first deployed DU for ARCA. 115 DUs together form a cylindrical building block. A similar naming scheme is used for naming the strings themselves. ORCA.0001 belongs to string 1 for ORCA.

- DOMs. The vertical spacing between DOMs is 9 meters for ORCA and 36 meters for ARCA. There are eighteen floors: a floor corresponds to all DOMs with the same vertical position. The lowest DOMs correspond to floor 1; the highest DOMs are grouped as floor 18.
- PMTs. The 31 PMTs in a single DOM are ordered through rings (figure 2.1). Ring A is the lowest ring and consists of one PMT that looks down. Rings B, C, D, E, F all consist of six PMTs. PMTs are numbered from one to six, except for the PMT of ring A. Each PMT in the DOM corresponds to a certain PMT channel. Channels 0-11 belong to the top half of the DOM (ring E and F). The ring convention is the old convention, so in this paper the channel numbers will be used. The effective area of a PMT is 14.8 cm^2 , which makes the total effective area for a DOM $31 \cdot 14.8 \text{ cm}^2 \approx 460 \text{ cm}^2$. Because the projected area of a DOM is 1348 cm^2 , this means that approximately one third of all photons encountering a DOM will hit a PMT.

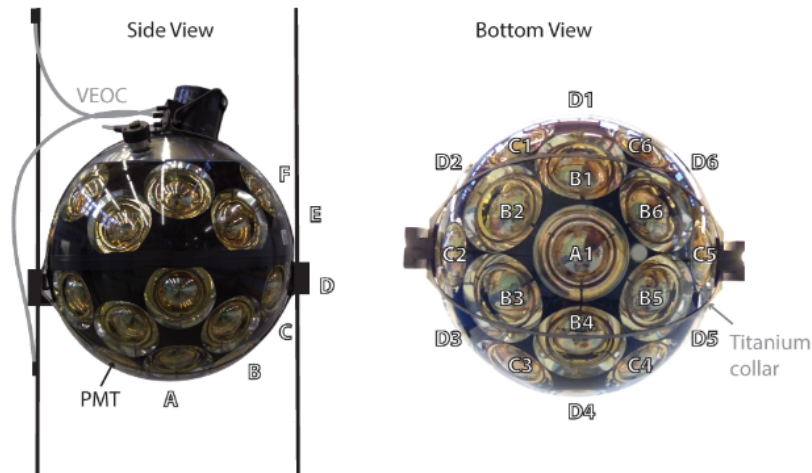


Figure 2.1: Side view (left) and bottom view (right) of a DOM. The PMTs are ordered in rings. The ring facing the sea floor is ring A, which consists of only one PMT. Rings B-F consist of six PMTs each indicated with a number of 1-6.

Both ORCA and ARCA use the same technology. The difference in scientific goals determines the difference in volume and geometry. This results in a much denser distribution of the sensor module in ORCA compared to ARCA in order to have the sensitivity needed to detect the lower energy atmospheric muons.

Data from the PMTs is collected in the Central Logic Board of a DOM and goes via optical fibres to onshore stations. The exact process will be described in section 2.3.

In order to monitor the position of the DOMs due to the sea current, the DOMs are equipped with an acoustic receiver, a digital compass and a tilt- and accelerometer [10].

2.2 Treatment of data

When one or more photons hit the PMT, this hit is characterised by three things: first, the PMT ID. Secondly, the hit time, which is the time at which the signal exceeds the threshold. Finally, the time it stays above the threshold. This is called the *ToT*: Time over Threshold. Both the hit time and the ToT are rounded to an integer number of nanoseconds. Together, the hit time and the ToT form a *level zero hit* (L0 hit). All L0 hits are bundled in slices of 0.1 s. The all-data-to-shore approach is used, meaning that all the L0 hits are sent to shore without any processing taking place.

A signal hitting the PMT triggers electrons via the photoelectric effect. The time between the real signal hitting the cathode and the electron hitting the anode is called the *transit time* (TT). It takes about 30 ns, with a spread of 2 ns.

L0 hits are sorted and ordered in L0 timeslices that contain all L0 hits of the entire detector recorded during 0.1 s.

An L0 detected by a single PMT is called a *single rate*. If two or more L0 hits are recorded on the same DOM within a period of 25 ns, these hits are grouped together as a *level one* (L1) hit. The *multiplicity* is the number of different PMTs contributing to an L1 hit.

Trigger algorithms are used to detect possible neutrino interactions among the data. A triggered event is then a snapshot of the L0s recorded within a time window of the trigger. The triggered events can be shower- or track-like events.

A *run* usually lasts for several hours, in which the data are taken without changing parameters controlling the detector and the readout [10, 19].

The storing and processing of the data is performed at three computer centres.

2.3 Calibration and simulation

Various simulation packages are available for the simulation of particles, such as atmospheric muons. Several assumptions and simplifications are made in the simulation software packages to reduce the computational load.

MUPAGE is used in the simulation of atmospheric muons and events they cause. Both single muons and muon bundles are simulated. A muon bundle consists of multiple muons generated in the same cosmic ray air shower. Several properties of atmospheric muons are parameterised, such as their energies and zenith angles.

To simulate the propagation of muons through seawater, the software package JSirene is used. It is assumed that all muons follow a straight path. Along its trajectory, the muon emits photoelectrons until the energy of the muon gets below the Cherenkov limit. The KM3 package simulates the emission, propagation, and detection of photons. It also serves as a cross-check of JSirene.

The JPP software package is used for detector response. The specific characteristics of the telescope, such as the performance of the electronic circuit, are implemented in the JPP packages. The PMT characteristics, which may vary over time and from PMT to PMT, are tuned as well [10, 20].

A note must be made that the simulations do not fully reflect the proper angle and energy distribution of the atmospheric muons. As a result, the Monte Carlo simulations ('MC', used to simulate events with for systems with many coupled degrees of freedom) will have a discrepancy when compared with the real data. Tuning efforts are currently ongoing to resolve this caveat. In addition, the simulation of the light creation in the water is under scrutiny to adequately include various effects, such as delta-electrons. Another factor that may not be fully accurately described is the angular acceptance of the PMTs.

Track reconstruction

The trajectories of highly relativistic muons can be determined by event reconstruction. The first step is using a prefit, which gives an estimate for the muons position, direction, and time at which the muon passes through its position. These first estimates are found by looping over approximately 2000 muon directions. In steps of approximately one degree the full solid angle is covered. The most likely position and time are determined for all sampled directions. The twelve best fitted combinations of position and

time form a set, for which a maximum likelihood fit is performed with the Levenberg-Marquandt algorithm. All five independent trajectory parameters are fitted simultaneously. These results are then used to find the start position of the muon, defined as the position at which the first detected light was emitted and thus gives the position of the neutrino interaction [14].

A *time residual* or t_{res} for short is defined as the difference between the measured L0 time and the expected, reconstructed hit time:

$$t_{res} = t_{measured} - t_{predicted}.$$

This value is positive if the measured hit time is larger than the reconstructed hit time [10].

⁴⁰K calibration

Detector performance is monitored over time with the isotropic ⁴⁰K signal from the sea. Correlated hits from ⁴⁰K decays in PMT pairs are selected via a small time-window. The distribution of the time differences in L0 hits is used to, among other things, determine the detection efficiency for pairs of PMTs within a DOM. For 31 DOMs this gives 465 pairs. The ratio between the measured rate and the expected rate gives the relative detection efficiency of that PMT. A Gaussian distribution with area $f(\theta_{i,j}) \cdot \epsilon_i \cdot \epsilon_j$ is fitted simultaneously to all PMT pairs within a DOM. The expected coincidence rate $f(\theta_{i,j}) \cdot \epsilon_i \cdot \epsilon_j$ has an opening angle between PMTs i and j of $\theta_{i,j}$. ϵ_i and ϵ_j are the detection efficiencies of the two PMTs. Solving this equation for all PMT-pairs gives the relative detection efficiencies for the 31 PMTs.

The relative detection efficiency is one of the inputs for (Monte Carlo) simulations and will be used in this thesis for the analysis [14].

2.4 Research data

The data used for section 3.1 is from ARCA6 and ARCA8, meaning that there are six and eight strings used in the analysis, respectively.

Data from ARCA8 were used to investigate the relation between single rates and efficiency. Run 10672 has been processed to provide single rate measurements (jra_94_10672.root) and efficiencies

(KM3NeT_00000094_00010672_JFIT_K40.root). This run was taken on 9 and 10 November 2021, lasting six hours.

These two files contain data of eight functional strings: ARCA.0009, ARCA.0010, ARCA.0011, ARCA.0012, ARCA.0013, ARCA.0014, ARCA.0015, and ARCA.0018. The efficiencies of the PMTs are measured as detection efficiencies.

The rates are stored in histograms with the PMT channel on the x-axis and the single rate in kHz on the y-axis. The efficiencies are stored in 2D histograms with the PMT channel on the x-axis and the efficiency on the y-axis. The z-axis describes the number of times a certain efficiency occurs per PMT channel. The two histograms were combined via the PMT channels.

For some strings, one or multiple DOMs or PMTs are not functional for various reasons. They may have been broken during transport, for example. These were excluded from the results.

To investigate how well the simulated muons match the measured muons, ARCA6 was used, because this was the most recent version for which a Monte Carlo simulation existed. The real data were obtained from

KM3NeT_00000075/v6.2/reco/datav6.2.jchain.aanet.00009843.root and KM3NeT_00000075/v6.2/reco/datav6.2.jchain.aanet.00009844.root. This is data for run 9843 (six hours of data taking, 19-20 June 2021) and for run 9844 (six hours of data taking, 20 June 2021) from version 6.2. The two runs 9843 and 9844 were combined in the analysis.

The efficiencies are evaluated for run 9844 and assumed they have not changed significantly with respect to run 9843.

In addition to runs 9843 and 9844, runs 9791-9796 were analysed in the exact same way. Each run took 6 hours, with run 9791 starting on 13 June 2021 and run 9796 ending on 15 June 2021. The results of these runs are shown in the appendix. Since data taking is a continuous process, the choice of runs is arbitrary. The analysis of runs 9791-9796 serves as a consistency check.

The ORCA detector is analysed in section 3.2. Data taking for runs 10838-10842 began 11 September 2021 and ended 12 September 2021, with each of these runs lasting six hours. Three MC files for every run compare the simulated muons to the measured muons. Software version 7.1 is applied to both the data and MC. The analysed strings are: ORCA.0001, ORCA.0002, ORCA.0003, ORCA.0009, ORCA.0010, and ORCA.0011. The efficiencies are for run 10840. This file was chosen because it is in the middle of the used data runs. It can therefore be assumed that the efficiencies have not changed for the other runs.

2.4.1 Methods

For both ARCA and ORCA the data from muon light was processed in order to analyse it properly. First, events with no reconstructed track were removed from the data. Then the number of hits for the real data was divided by the number of hits for the MC. This ratio is a measure for the efficiency of the detector. The efficiencies were plotted as well to see whether they can explain some results of the ratio data/MC.

With the ratio of data/MC and the relationship between single rates and efficiency, three possible biases were investigated:

Possible biases due to change in gel transparency

The first bias that has been investigated is the potential bias due to a change in gel transparency, which is assumed to affect string ARCA.0010 DOMs 7-15, ARCA.0012 and ARCA.0013 ring E and F. It is not yet clear what the cause of the change in transmission properties is. For the single rates, a fit was made through the data sorted by string. The possibly affected DOMs and PMTs were fitted separately from those which are probably behaving normally. The comparison of these fits gave the first clue whether the DOMs are indeed affected. The hypothesis was that the transparency of the gel is not the same for all wavelengths, thus causing a lower efficiency.

To detect a possible pattern in the potentially affected PMTs, in section 3.1.1 the corrected rate has been calculated. The corrected rate was obtained by subtracting the offset from the measured rate. The offset was taken from the fit of all PMTs that are not affected. This was divided by the efficiency. Thus:

$$rate_{corrected} = \frac{rate_{measured} - offset}{efficiency}. \quad (2.1)$$

With the PMTs on the y-axis and the DOMs on the x-axis, the corrected rates are shown per string. If the efficiency is low, the corrected rate will be higher than the measured rate. The last figure in this section shows the distributions of corrected rates for the affected and normal PMTs. The means of the distributions were calculated and compared.

Shadowing and enhanced dark rate

The second possible bias is caused by the equator tape and the titanium belt. They shadow some parts of the detector, resulting in less photons hitting the affected PMTs.

At three dedicated positions an enhanced dark rate has been observed in recent DOM productions. It has been investigated if this effect is also observable in-situ. A comparison in fits of these PMTs and the average fit of all PMTs gave the extent of this potential bias.

The shadowing has been examined in two ways. With the single rates, the data from PMTs (possibly) affected by a difference in gel transparency were left out. Then the fits of all data were compared to the fits of only the shadowed PMTs. The second investigation used the ratio data/MC of the time residuals, for which rings C and E were plotted. This gave two plots per ring: one for the shadowed PMTs (C2, C5 and E2, E5) and one for the PMTs which are not shadowed (C1, C3, C4, C6 and E1, E3, E4, E6.) The ratios were plotted as distributions per ring, as well. Using the data from the single rates, the fit of the dark rate for PMTs A1, B5, and B6 was compared to that of all data. In the case of an enhanced dark rate, the fit of PMTs A1, B5, and B6 would have a fit with a steeper slope.

Radioactivity and sedimentation

The third possible bias emerges when the DOMs are covered with sedimentation. The current efficiency evaluation uses correlated hits from K40 decays in PMT pairs to estimate the efficiency. It does not consider that a part of this light actually is a result of the radioactivity in the glass of the DOMs. In a DOM covered with sedimentation there may still be a signal stemming from the radioactive decay in the glass, which gives a finite estimated efficiency. However, the relevant detection efficiency for light originating outside the detector is heavily oppressed.

ARCA.0009 is known to have sedimentation. The ratios data/MC of this string were compared to those of the other strings (ARCA.0010, ARCA.0011, ARCA.0013, ARCA.0014, and ARCA.0018). In addition, the data were separated by zenith angle: 0°- 45° and 45°- 90°. Since the sedimentation is mostly on top of the DOM, the efficiency of the PMTs may differ for different incidence angles of the photons.

The three potential biases mentioned above were analysed in section 3.2 for the ORCA detector. Because ORCA data are contaminated with bioluminescent signals, the single rates were not investigated. Only the time residuals have been used in this section. Other than that, the method described for ARCA was used for ORCA as well. For ORCA the DOMs possibly affected by the gel are DOMs 1-6 of ORCA.0001. Pictures of the strings after being employed in the water show that ORCA.0001, the string

that has been in the water the longest, has sedimentation accumulated on the DOMs. This string was compared to the other strings for 0°- 45° and 45°- 90° to investigate the effect of the sedimentation.

Results

The figures in section 3.1 are the analyses of the three potential biases in the detection efficiencies for ARCA. After that, the ORCA detector will be investigated in section 3.2 and compared to ARCA.

3.1 ARCA

3.1.1 Time residuals

In this section, the performance bias of the ARCA detector was evaluated by looking in general at the time residual distribution for simulated (indicated by 'MC') and measured muons (indicated by 'real data').

The peak of the distribution around $t_{res} = 0$ can be attributed to 'direct' light. The tails are mainly a result of scattered light, which does not follow a straight trajectory and thus takes a longer path. The size and shape of this tail are impacted by the properties of the medium, so they will be different for ice and water, for example.

Cut on triggered hits

All plots using time residuals were corrected for the difference in lifetime between real data and MC. This was achieved by scaling the MC by a weighting factor $\frac{lifetime_{data}}{lifetime_{MC}}$. For runs 9893-9894 the factor is 0.496.

The time residuals for the range $-100 \text{ ns} \leq t_{res} \leq 200 \text{ ns}$ for the real data and the MC are compared in figure 3.1, showing whether the simulation matches the data well.

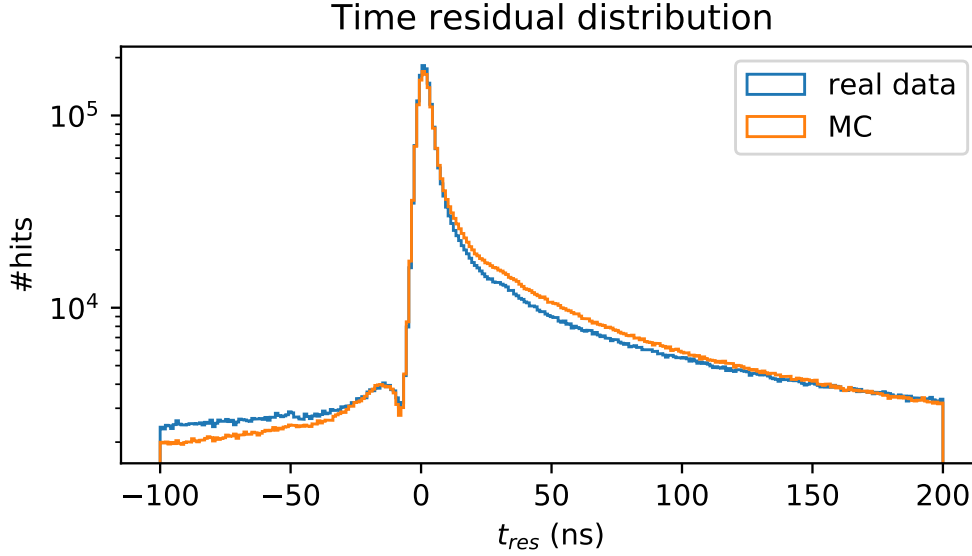


Figure 3.1: The time residual distribution for the real data (blue) and the Monte Carlo simulation (orange). The MC is weighted by $\frac{\text{lifetime}_{\text{data}}}{\text{lifetime}_{\text{MC}}}$ to correct for the difference in lifetimes of real data and MC.

The range of $-100 \text{ ns} \leq t_{\text{res}} \leq 200 \text{ ns}$ can be roughly divided in four parts: -100 to -20 ns , -20 to 10 ns , 10 ns to 150 ns and 150 to 200 ns . Data and MC agree best in the range -20 to 10 nanoseconds, corresponding to signals from direct light. The peak of the signal is slightly higher for the real data.

The bump at around -15 ns is an artefact of the reconstruction of the muon paths. It is simulated by the MC, so this is not a problem for the analysis.

For positive t_{res} , the MC has more hits than the data, but this difference gets smaller for even larger time residuals. At around 150 ns the data and MC have the same number of hits. Their distributions are close up to at least 200 ns .

From -100 to -20 ns the number of hits for real data is higher than for MC. These hits in the large $|t_{\text{res}}|$ region can be mainly attributed to noise.

To remove a part of the noisy events, a cut on the number of triggered hits can be made. Events that did not have a muon track (no time coincidence) can then be excluded from the time residual distribution.

The following figure contains the distribution of the number of triggered hits for 0-35 hits.

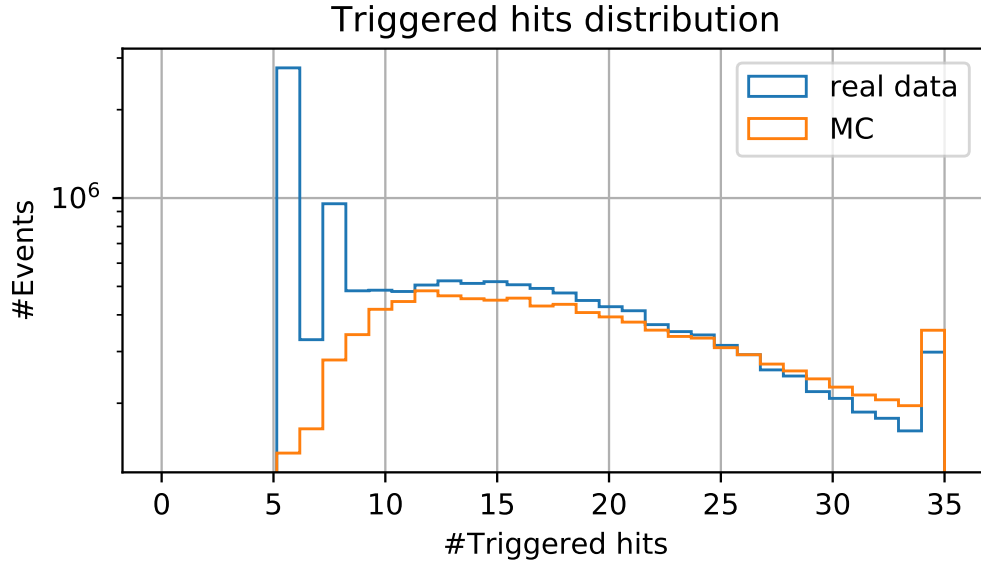


Figure 3.2: The distribution of the number of triggered hits for real data and MC. The data and MC differ for less than 15 triggered hits. Therefore, a cut of 15 triggered hits is applied to the time residual plots to exclude the events that have no muon track. Excluding these events gives a better agreement between real data and MC.

Fewer than 12 triggered hits gives a bad agreement between data and MC. From around 12 to 20 triggered hits a dip is visible for the MC compared to the data. After 20 triggered hits the MC gets slightly more events than the data. Since this figure does not show what the trend is after 35 triggered hits, a second plots is shown below, which shows 0-100 triggered hits.

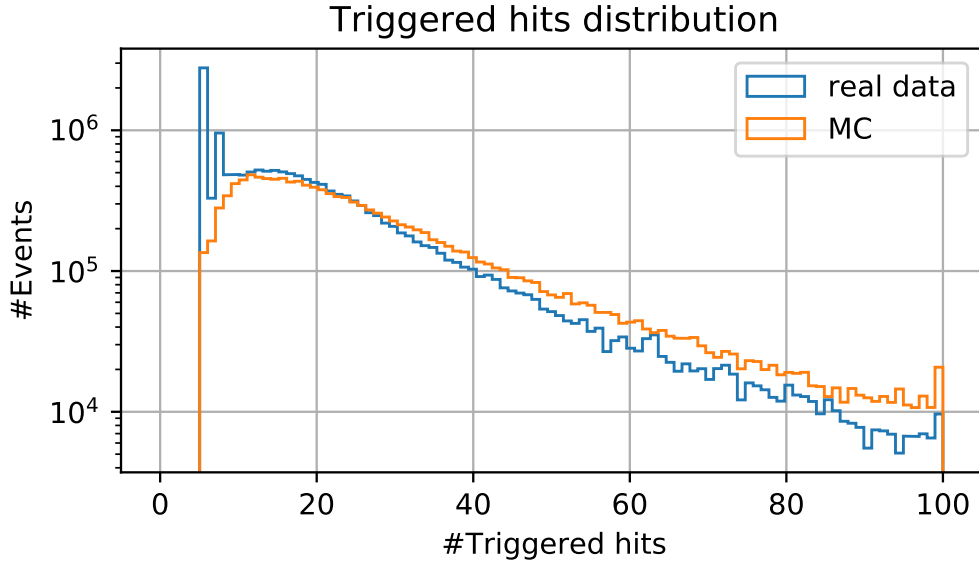


Figure 3.3: Extended version of 3.2. The distribution of the number of triggered hits for real data and MC. A cut of 15 triggered hits is applied to exclude the events that have no muon track. Excluding these events gives a better agreement between real data and MC.

It seems that the higher the number of triggered hits gets, the higher the discrepancy between real data and MC.

Although not all disagreement in data and MC can be removed by applying a cut on triggered hits, a cut on 15 triggered hits was applied to the rest of the ARCA plots. This cut excluded the largest discrepancies that are present for 0-10 triggered hits. Furthermore, from about 15 triggered hits the slopes of both data and MC do not change drastically, thus diminishing too large fluctuations in the next figures.

The cut was applied on the time residual plot of figure 3.1, resulting in figure 3.4.

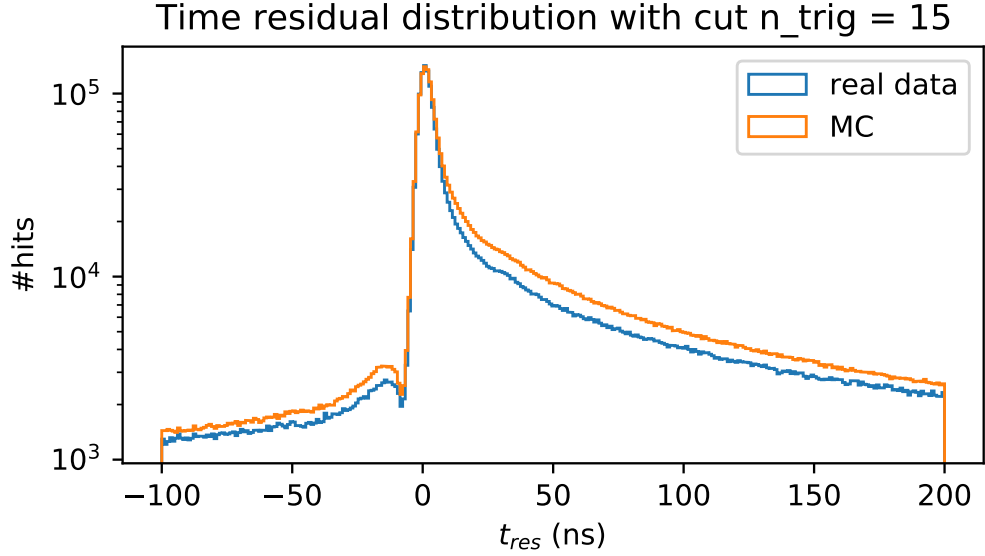


Figure 3.4: Time residual distribution after the cut on 15 triggered hits is applied. The shape of the real data and MC distribution are now more alike than before the cut.

The shape of the MC is more similar to the shape of the data than before the cut, especially in the range -100 ns to -10 ns. The MC is now higher than the data in this interval. For small $|t_{res}|$ there is little change due to the applied cut. The peak, however, is now the same height for the data and MC. Although the shape of the region 10 ns - 200 ns is now in better agreement, the difference in height has worsened by $\sim 10^3 - 10^4$ hits.

Applying a Kolmogorov-Smirnov test to two distributions gives a test statistic and a p-value. The likelihood that the two distributions are drawn from the same distribution can be obtained from these values with an α value chosen in advance. A KS-test (with $\alpha = 0.05$) of the distributions before and after applying the cut confirms that the cut does improve the similarity of data and MC slightly. The test statistic for the plots after the cut is lower (0.147, p-value 0.0030) than before (0.171, p-value 0.0003). It is thus more likely that the distributions from fig. 3.4 are drawn from the same distribution, but the p-value is still very low. The large discrepancies at the tails are probably the reason for this.

To verify that a cut on a greater number of triggered hits does not improve the similarity of data and MC, the KS-test is performed for a cut on 20 triggered hits. This gives a test statistic of 0.194 and a p-value of 0.00002. These outcomes confirm that a cut on 15 triggered hits is a good

choice.

The next step was to determine whether the described features are present at all DOM positions, specifically if there are systematic differences to the different rings (figure 2.1), or can be attributed to some rings only. The time residuals were plotted for the real data and MC just as above, but were now divided by rings A-F.

By ring

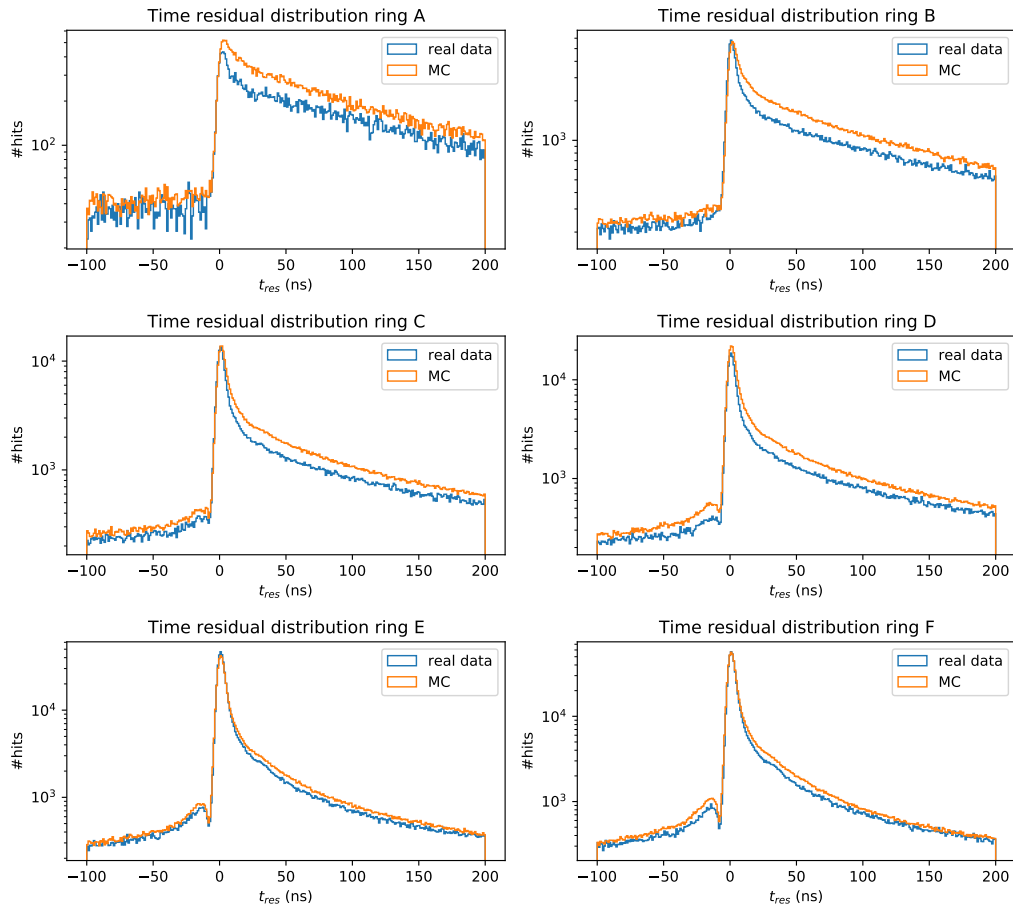


Figure 3.5: Time residual distributions sorted by ring.

- The first thing that stands out is the noise of ring A, both in the data and MC. A reason for this could be that there are less data for ring A than for rings B-F, since ring A consists of one PMT instead of six PMTs as is the case for rings B-F. The peak of the data is lower than

for MC and this discrepancy is higher than for other rings. Because ring A faces downwards it receives less light, which can be seen from the graph: the peak of ring A is in the order of $4 \cdot 10^2$ hits, but the other rings have a peak of at least 10^3 hits. For ring A the MC should have been lower in general. From -100 to 10 ns, the MC follows the data quite well. After that, however, the MC estimates too many hits. The shape of the MC does not follow the data as well as in the negative t_{res} region. It should have simulated less hits.

- As ring B does not face completely downwards, its peak has a higher number of hits than ring A, although it is not as high as for rings C-F. The distribution of the MC seems to be slightly shifted to the right compared to the data.
- What is the case for ring B is also the case for ring C, except the MC follows the data a bit better for ring C than for ring B after 10 ns.
- Of all rings, ring D follows the data the worst at -100 ns to -10 ns. The peak of the MC is also a bit too high.
- The MC of rings E and F follow the data the best.

The main deviation from the data for the MC is for rings A-D (bottom hemisphere of the DOM) from 10 ns and higher.

3.1.2 Possible biases due to change in gel transparency

The report 'Report on problems with data from ORCA-DU2' [16] mentions a lower efficiency for certain ORCA DOMs. The hypothesis is that these DOMs were made with gel of which its properties differ from the gel used for other DOMs. This then could have resulted in these DOMs having a lower efficiency, since the gel may not be as transparent for some wavelengths as for other wavelengths. In addition it can mean that the estimated efficiencies from the close-by K-40 light does not translate as expected to the efficiencies for muon light and for K-40 light travelling long distances.

For ARCA the suspicion is that DOMs 7-15 of ARCA.0010, all DOMs of ARCA.0012 and rings E and F of all ARCA.0013 DOMs show this difference in transparency. It is worth mentioning that the gel of DOMs 7-15 of ARCA.0010 comes from the same batch as for the investigated DOMs 1-6 of ORCA.0001.

In the next subsection, the ratios of number of hits for the full t_{res} range are plotted per ring for figure 3.5. Together with the estimated efficiencies, this gives the first results of the possibly affected PMTs compared to those which are not affected. In the subsection after that, the single rates will be used to examine the behaviour of the PMTs for light with long wavelengths.

Muon light

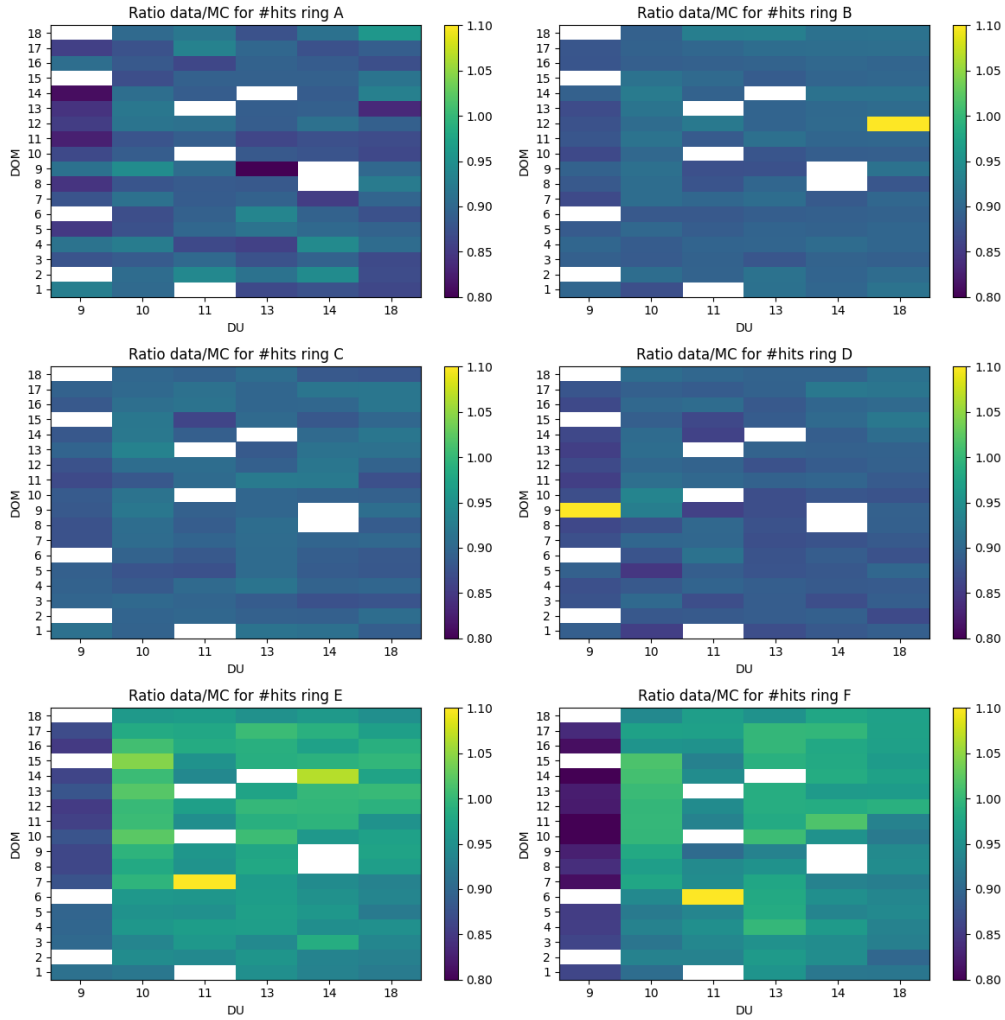


Figure 3.6: Ratios of data/MC plotted for the six rings. The x axis represents the string number and the y axis the DOM number. White rectangles show inoperative DOMs.

The ratio data/MC is plotted for the six rings, with the string on the horizontal axis and the DOM on the vertical axis. The colourbar next to the figures indicates the data/MC ratio. Some DOMs at certain positions are indicated by white rectangles: they are inoperative and thus left out of the analysis. The bright yellow rectangles are probably faulty DOMs.

Just as figure 3.5 already showed that the data and MC distributions for rings E and F are closer together than for rings A-D, this same feature is visible in the graph above. The ratio data/MC is the closest to one for rings E and F. For rings A-D the ratio is a little below 1, since fig. 3.5 indicates that the MC overestimates the number of hits for large values of t_{res} .

The feature that is most striking is the low value of ratio data/MC of ARCA.0009 for rings E and F, reflecting either an overestimation of the efficiency or a deficit in the received photons. This will be addressed in subsection 3.1.4. In this section the ratios for the possibly gel-affected PMTs of ARCA.0010 and ARCA.0013 are examined.

The ratios for ARCA.0010 DOMs 7-15 are a bit higher than other strings and DOMs in the same string in all rings, although this is best visible in rings E and F. The efficiencies for these DOMs are estimated to be lower than for other DOMs in the same string (figure 3.7).

Similarly, the efficiencies of all ARCA.0013 DOMs are low as well, with the exception of DOMs 8, 9, and 18. This results in a ratio data/MC being close to 1. Compared to other strings this ratio is slightly higher.

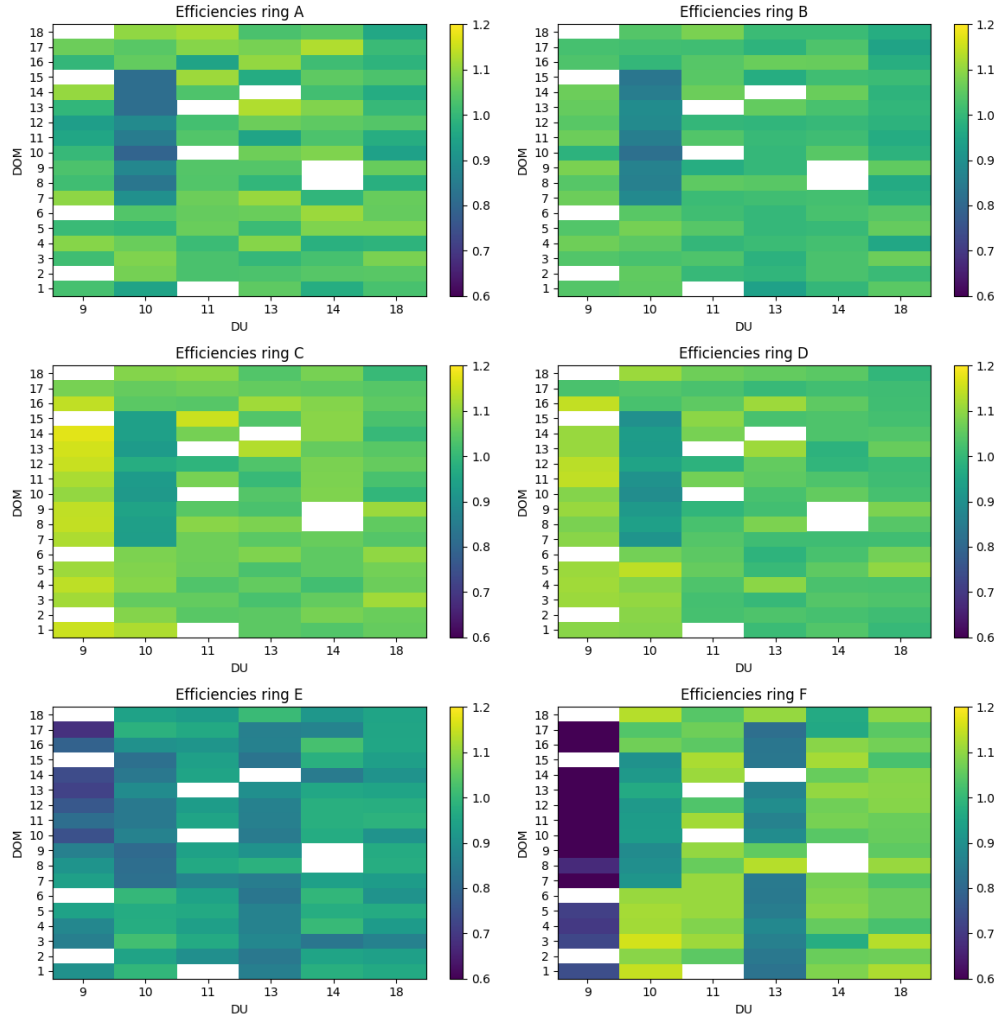


Figure 3.7: Efficiencies as a function of string (horizontal axis) and DOM (vertical axis) for rings A-F. There are a few PMTs with zero efficiency. If all six PMTs of a DOM have zero efficiency, the bin is white. If only one or two have zero efficiency, the efficiencies of these PMTs of zero are neglected and the average of the other four or five are plotted. For rings B-F the average efficiency of the four / five or six PMTs is used. (Ring A is a single PMT, so this is already the average.) Inoperative DOMs are shown as white rectangles.

Single rates

To test the hypothesis that the gel affects photons from far away differently than light coming from nearby sources, the single rates and detection efficiencies of eight strings are plotted and compared. Single rates are hits from photons from far away. The detection efficiencies were evaluated using light from nearby K40 decay. If the transmission properties of the gel are wavelength dependent, the relationship between the single rates and the efficiency for the potentially affected PMTs is altered.

The data with an efficiency of zero were excluded. The range of the single rates used is 2-11 kHz. This got rid of the outliers in the data. The error used for both the efficiency and the rate was 5%, which is on the high side and thus safe to assume.

In figure 3.8 the black data points represent data from PMTs which are not affected by the gel transparency bias. Blue points indicate data from PMTs which may be affected. The orange lines are fits through the black data points and the blue lines are fits of the blue data.

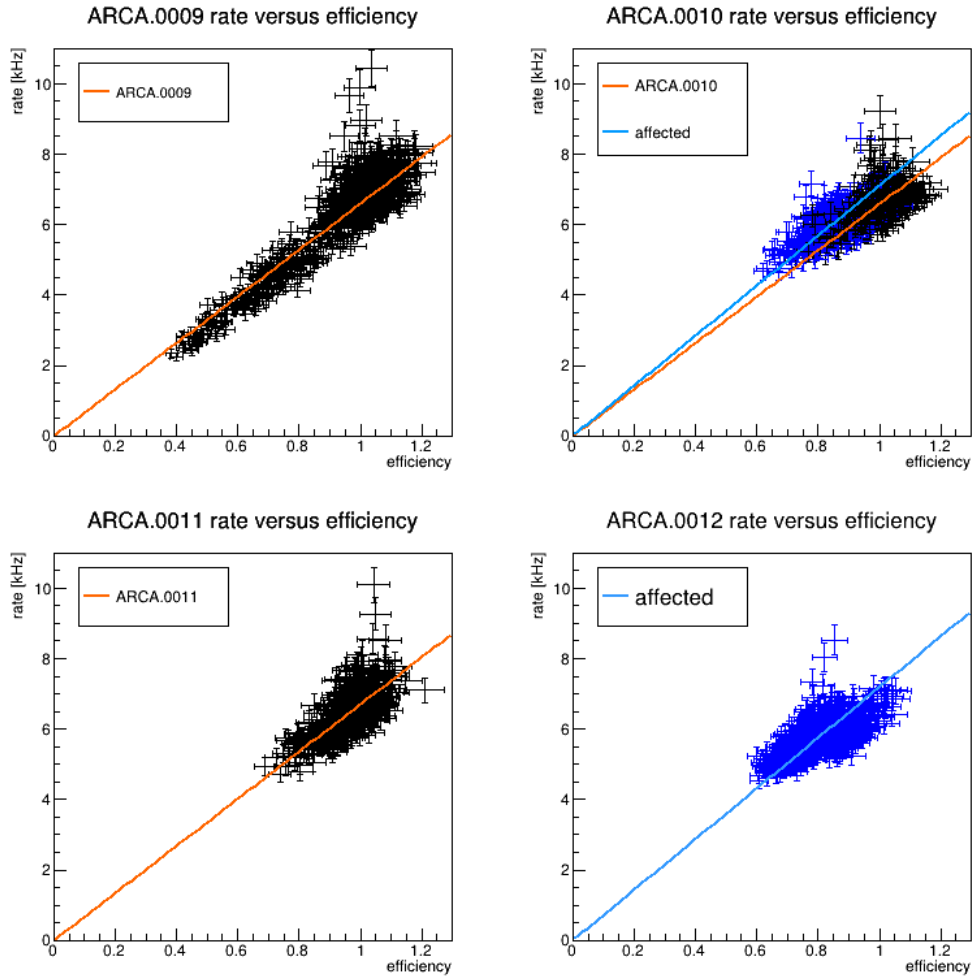


Figure 3.8: All data plotted, with on the x axis the efficiency and the y axis the single rate. The PMTs possibly affected by the gel are blue, the other data are shown in black. The orange fit indicated by 'ARCA.00x' shows the fit through all data, excluding the blue data points. For ARCA.0010, ARCA.0012, and ARCA.0013 the blue line corresponds to the fit through only the blue data points ('affected'). Continued on next page.

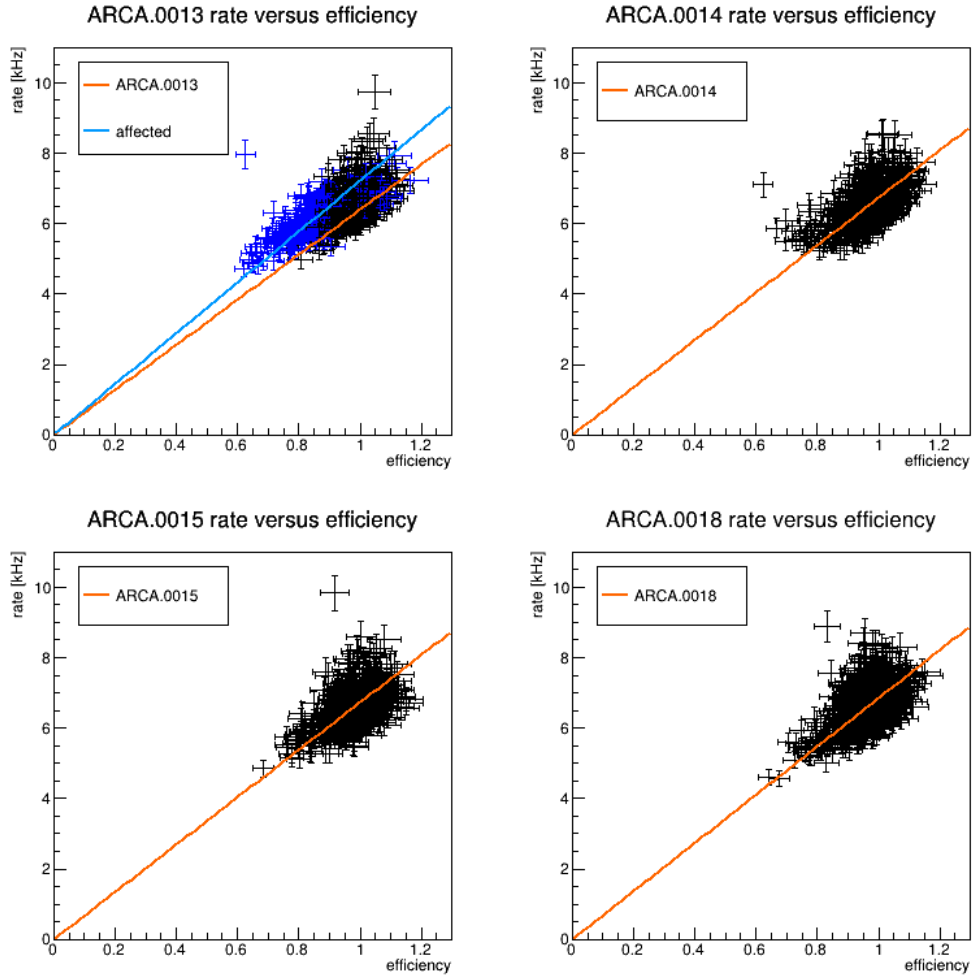


Figure 3.8: All data plotted, with on the x axis the efficiency and the y axis the single rate. The PMTs possibly affected by the gel are blue, the other data are shown in black. The orange fit indicated by 'ARCA.00x' shows the fit through all data, excluding the blue data points. For ARCA.0010, ARCA.0012, and ARCA.0013 the blue line corresponds to the fit through only the blue data points ('affected'). Continued from previous page.

The fits through the data points are fits of the form $y = ax$, meaning that the offset was set to zero. The reason for this is that the offset in equation 2.1 would be negative otherwise, which does not make sense physically: the offset should be the baseline rate. This constant baseline rate is caused by thermal(electric) noise of the equipment itself, so there will always be a signal of single rates measured. Therefore, the offset in

the calculation for the corrected rate was set to zero.

Compared to other strings, ARCA.0009 has a wide range of rates and efficiencies. This is due to ARCA.0009 being in the water the longest. The longer a string is in operation, the more sedimentation has accumulated on the DOMs. This results in less light being captured by the PMTs and the efficiency decreases.

ARCA.0010 and ARCA.0013 confirm the hypothesis: the possibly affected PMTs have on average a steeper slope than the PMTs that are not affected. Therefore, the relation between single rates and efficiency has altered for the potentially affected DOMs. This means that the transmission properties of the used gel may depend on the wavelength of the light.

Corrected rate per string

In this section, the corrected rate per string is calculated using eq. 2.1. A bias in the efficiency estimates would translate to differences in corrected rates between the PMTs. The results of the corrected rates are plotted by string in figure 3.9. Each rectangle represents a certain PMT on a DOM in ascending order from facing downwards to upwards. Some DOMs are indicated by white rectangles: they are inoperative. The same corrected rates are plotted as 1D distributions (see fig. 3.10). For DOMs 7-15 of ARCA.0010, all DOMs of ARCA.0012, and rings E and F of ARCA.0012 the distributions are drawn with a red line instead of a blue line, indicating that these are the possibly affected DOMs. The inoperative DOMs are excluded from the plot.

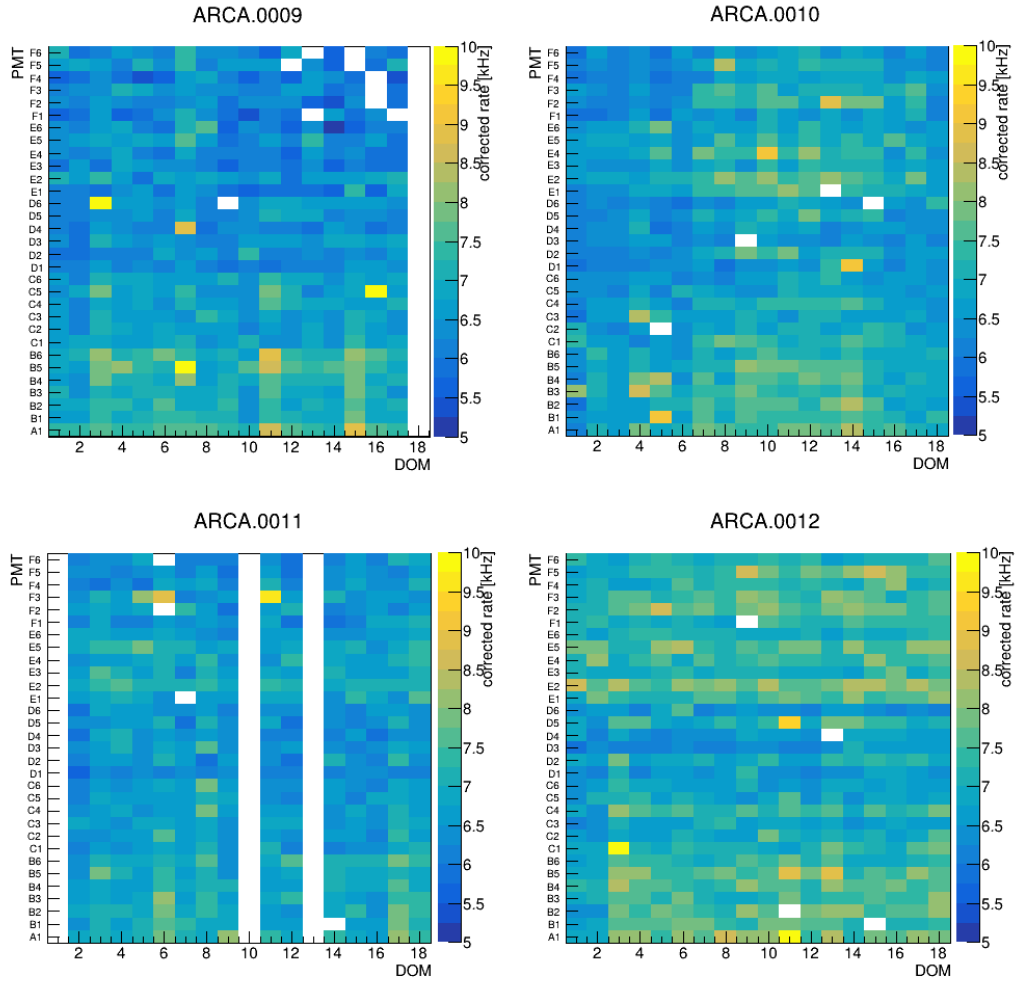


Figure 3.9: This histogram shows the corrected rate per string (see eq. 2.1). The x axis shows the DOM and the y axis the position of the PMT in the DOM. Inoperative PMTs are shown as white rectangles. Continued on next page.

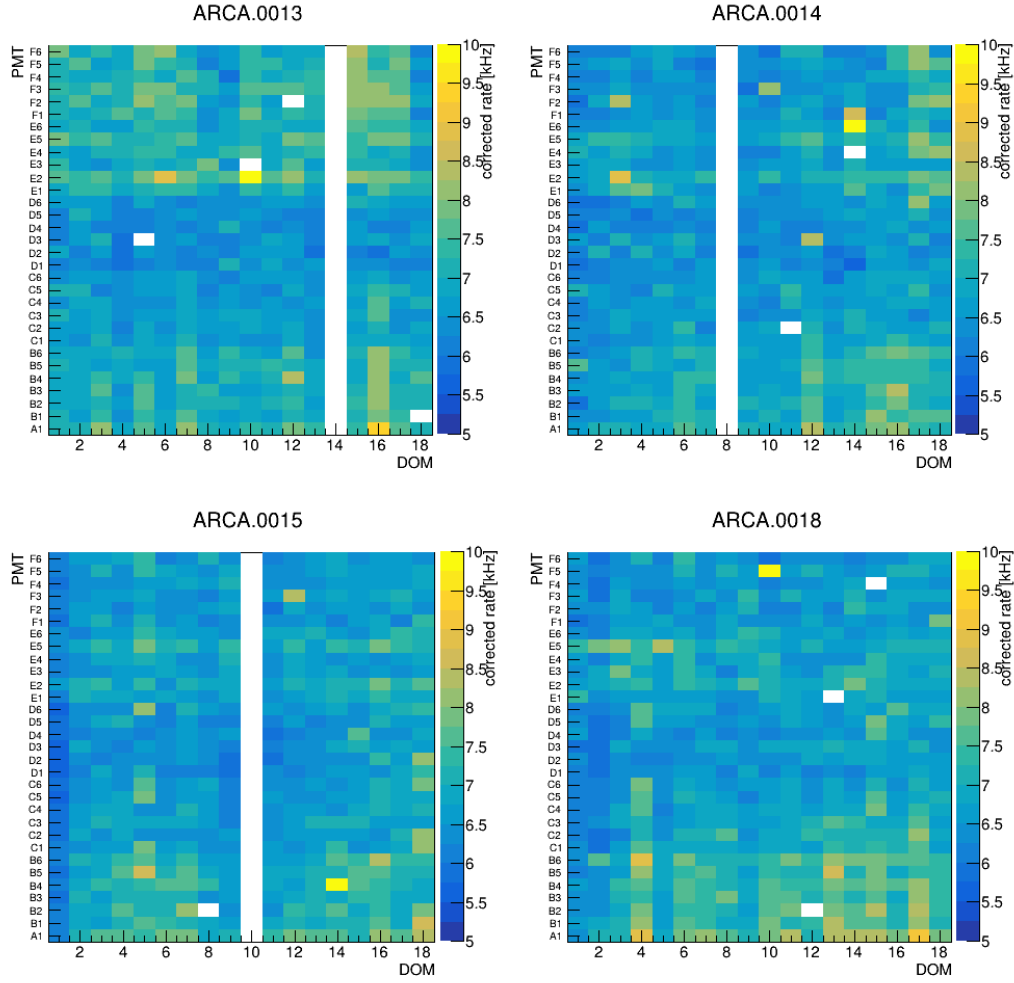


Figure 3.9: This histogram shows the corrected rate per string (see eq. 2.1). The x axis shows the DOM and the y axis the position of the PMT in the DOM. Inoperative PMTs are shown as white rectangles. Continued from previous page.

The corrected rates of DOMs 7-15 of ARCA.0010 are higher than the other DOMs. For ARCA.0012, the middle PMTs have lower corrected rates than the other PMTs. This is visible in ARCA.0010 to a certain extent, as well. For ARCA.0013 the hypothesis is that the upper PMTs (ring E and ring F) are affected by the gel transparency. The more yellowish colour of the figure supports this. This string also has lower rates for the middle DOMs. ARCA.0010 and ARCA.0012 seem to have lower rates for the middle DOMs as well.

To better distinguish the corrected rates of the affected PMTs and those of the non-affected ones, the rates of figure 3.9 are plotted in a one-dimensional distribution. The possibly affected PMTs are indicated by a red line; the PMTs that are not affected by a blue line.

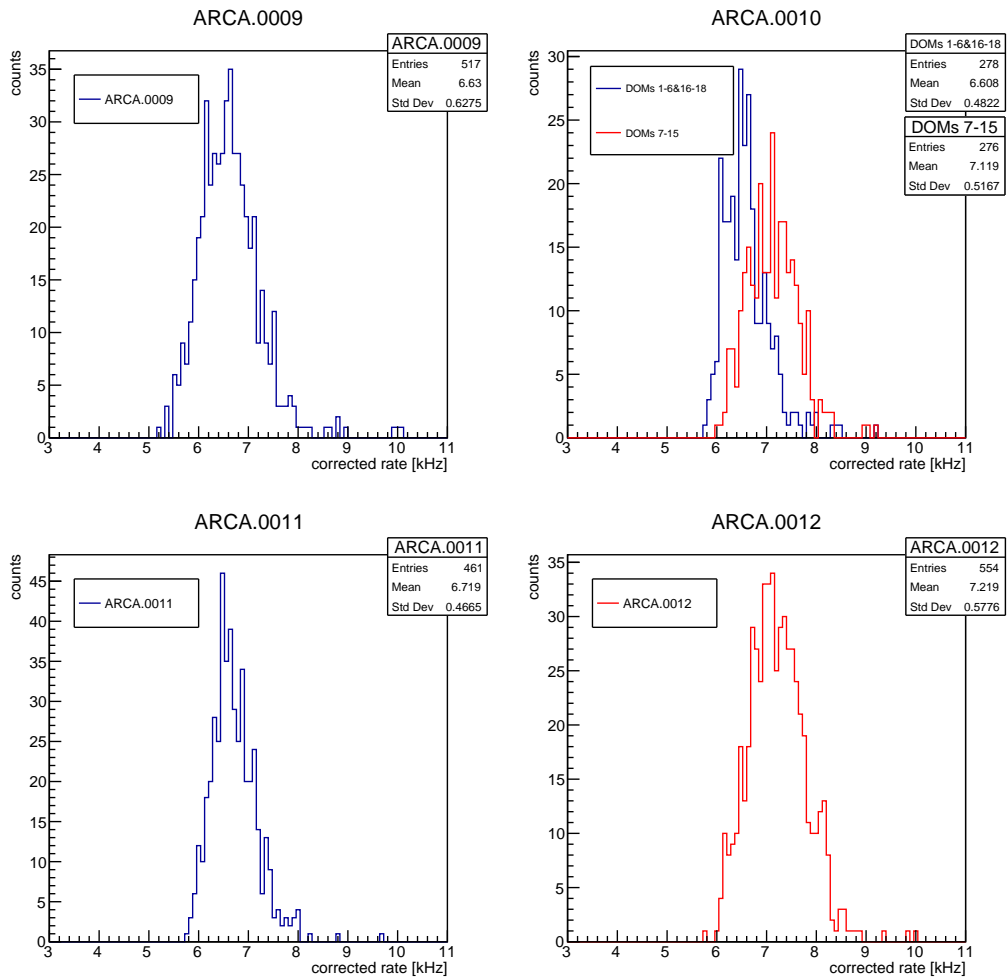


Figure 3.10: The distribution of the corrected rates per string. The blue coloured distributions are for the non-affected efficiencies, thus excluding the corrected rates of the DOMs and PMTs possibly affected. The possibly affected are depicted by the red coloured distributions. The corrected rate (in kHz) is plotted against the number of times the rates occur. Continued on next page.

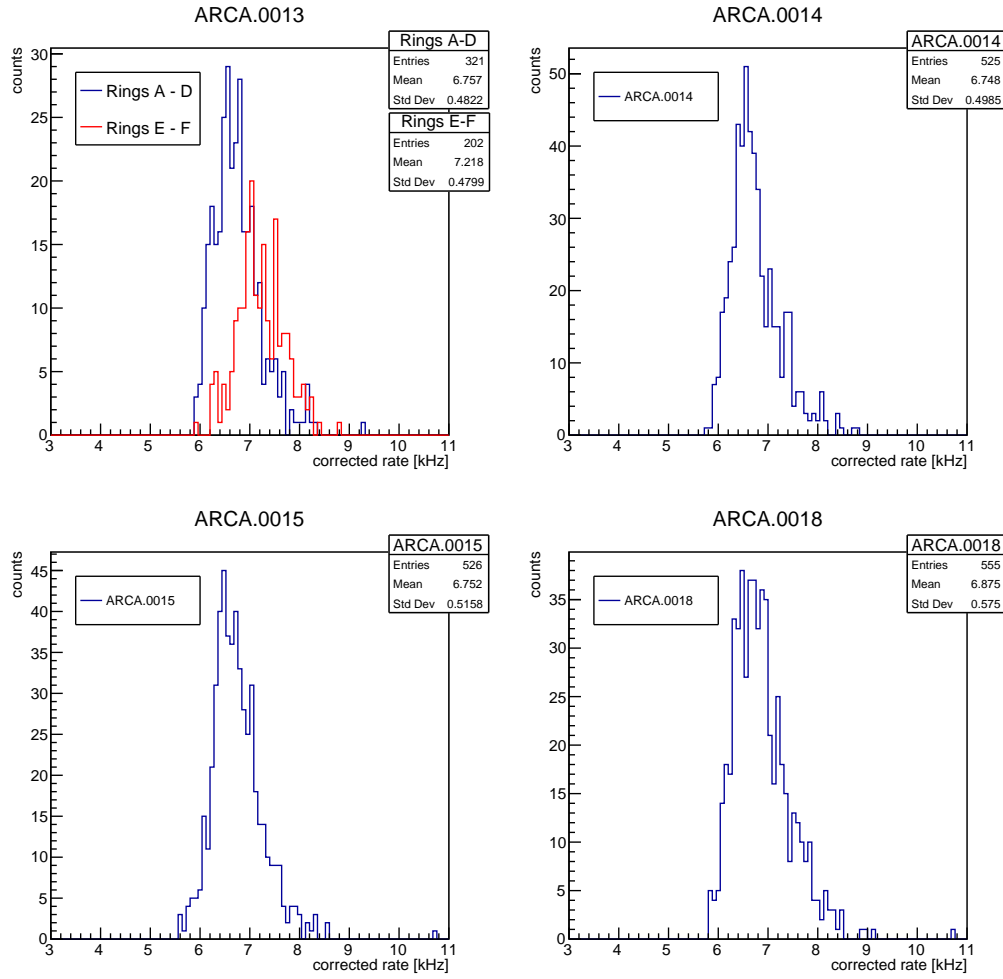


Figure 3.10: The distribution of the corrected rates per string. The blue coloured distributions are for the non-affected efficiencies, thus excluding the corrected rates of the DOMs and PMTs possibly affected. The possibly affected are depicted by the red coloured distributions. The corrected rate (in kHz) is plotted against the number of times the rates occur. Continued from previous page.

The distribution shows that ARCA.0010 affected PMTs (box 'DOMs 7-15') have a higher average corrected rate (7.119) than those not affected (6.608, box 'DOMs 1-6&16-18'). The same is true for ARCA.0013: the affected PMTs (box 'Rings E-F') have an average of 7.218, whereas the 'normal' PMTs (box 'Rings A-D') have an average of 6.757. To verify that the red and blue distributions are not the same, a Kilmogorov-Smirnov test is performed. For ARCA.0010 the test statistic is 0.4908 and the p-value is

3.489e-31. The test results in a KS test statistic of 0.5049 and p-value of 0.0 for ARCA.0013. Both ARCA.0010 and ARCA.0013 distributions are therefore probably different. The PMTs which are suspected to be affected by the gel are thus probably affected. The estimated efficiency seems to underestimate the real efficiency (for muon light) by 7.5%, using the ratio of the mean corrected rates.

3.1.3 Shadowing and enhanced dark rate

As can be seen in figure 2.1 some PMTs are shadowed by the equator tape and/or by the titanium collar.

The efficiency of each PMT is evaluated as an average of the efficiency in all directions (where there are PMTs) and is thus mostly from an isotropic distribution. Single rates come from ^{40}K decays up to far away distances and is thus also an isotropic source. For that reason no bias is expected for shadowed PMTs with regard to the single rates and efficiency relation. When looking at certain incoming angles of the light only, there might be differences for shadowed PMTs.

Single rates

In this section, PMTs which may be affected by the gel are left out of the graphs and the fits. This prevents the bias of the gel impacting the possible biases due to shadowing.

The figure below shows the single rate-efficiency fits for the PMTs affected by the titanium belt, which are C2, C5, E2, and E5. This fit is compared to the average fit of the non-affected PMTs from figure 3.8. The orange line indicated by 'titanium belt' is the fit through all data of the PMTs shadowed by the titanium belt of that string. The blue line corresponds to the fit of all PMTs of that string. A bias caused by the titanium belt would lead to a different fit for the shadowed PMTs.

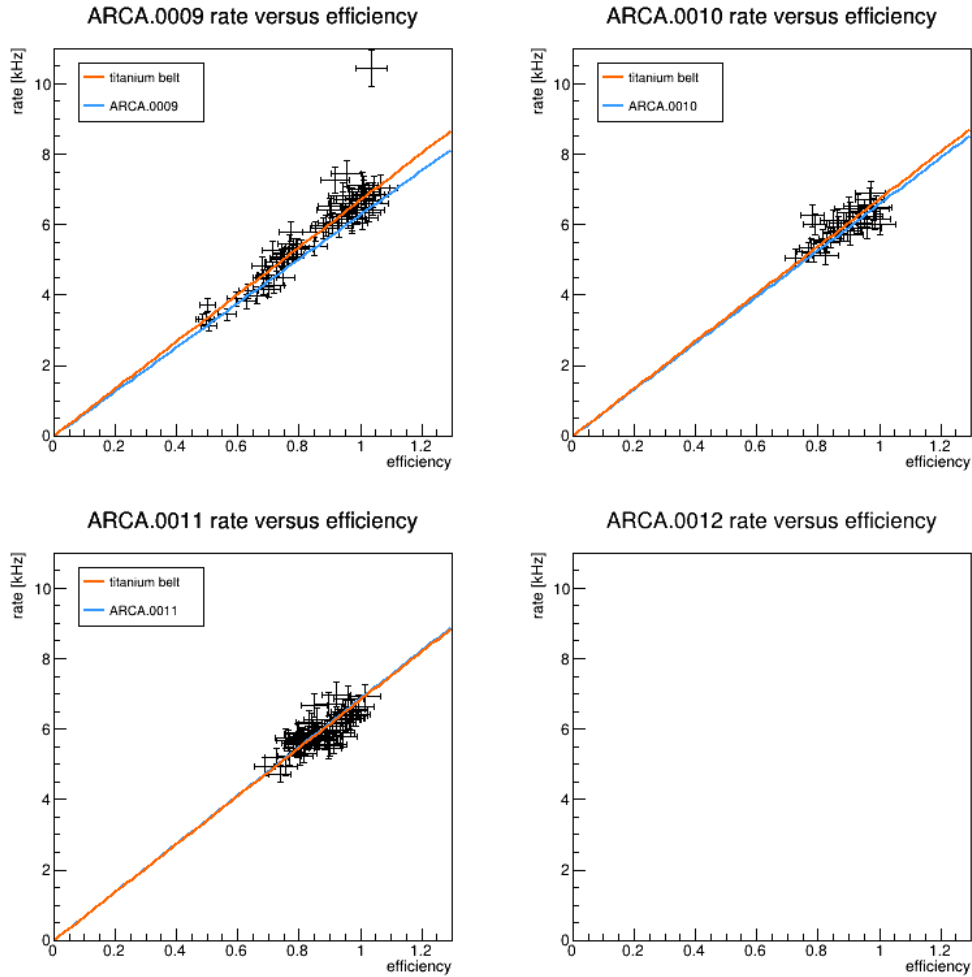


Figure 3.11: Single rates and efficiencies for PMTs shadowed by the titanium collar (C2, C5, E2, E5). The data from the PMTs affected by the gel are left out. The orange fit shows the fit through the data points of PMTs C2, C5, E2, and E5 of that specific string. The blue string is the fit through data points of that string for all data of that string (except possibly gel-affected PMTs) from figure 3.8. Continued on next page.

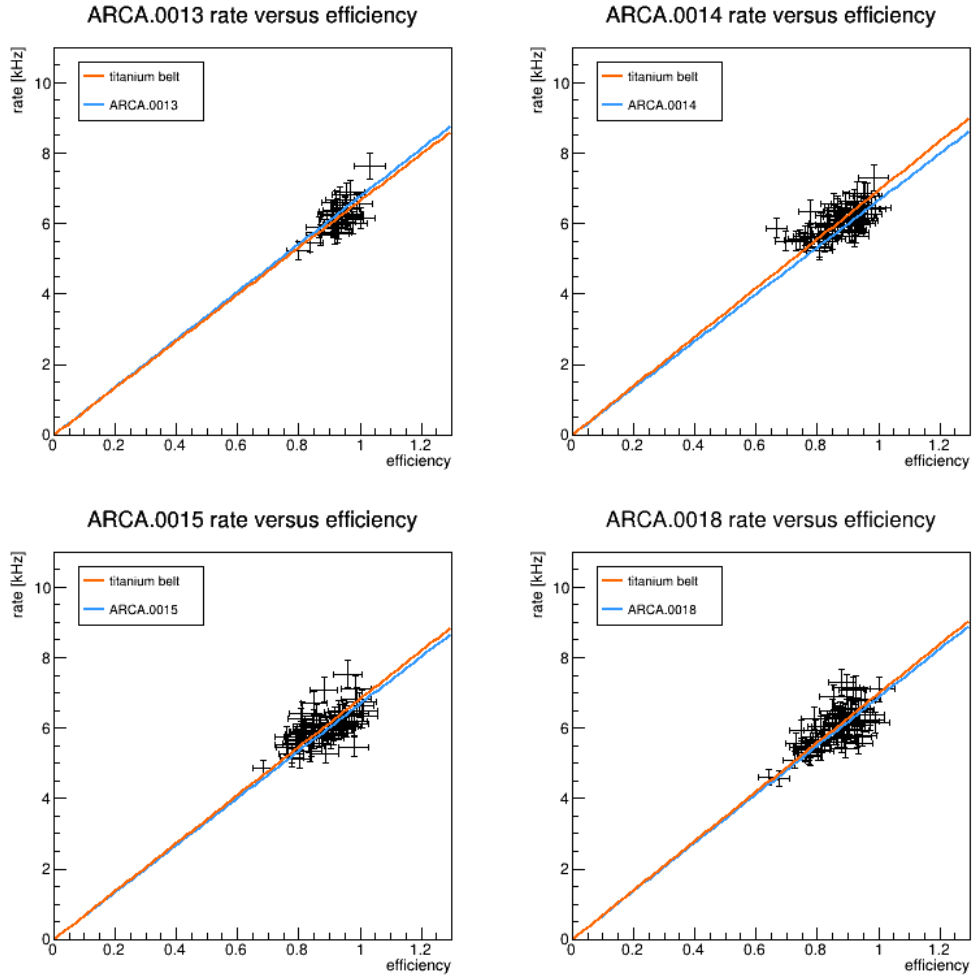


Figure 3.11: Single rates and efficiencies for PMTs shadowed by the titanium collar (C2, C5, E2, E5). The data from the PMTs affected by the gel are left out. The orange fit shows the fit through the data points of PMTs C2, C5, E2, and E5 of that specific string. The blue string is the fit through data points of that string for all data of that string (except possibly gel-affected PMTs) from figure 3.8. Continued from previous page.

Figure 3.11 shows the fit per string of the PMTs affected by the titanium belt.

The fits are the most similar for ARCA.0011, the least for ARCA.0009. The slopes of the titanium belt data is for all strings a little steeper, except for ARCA.0013, where the slope is less steep.

However, these differences are insubstantial. Shadowing caused by the

titanium collar does not result in a significant change in efficiency for the single rates.

The equator tape blocks light for rings D (mostly from above) and ring E (mostly from below). Figure 3.12 shows the effect of this structure on the rate-efficiency relationship.

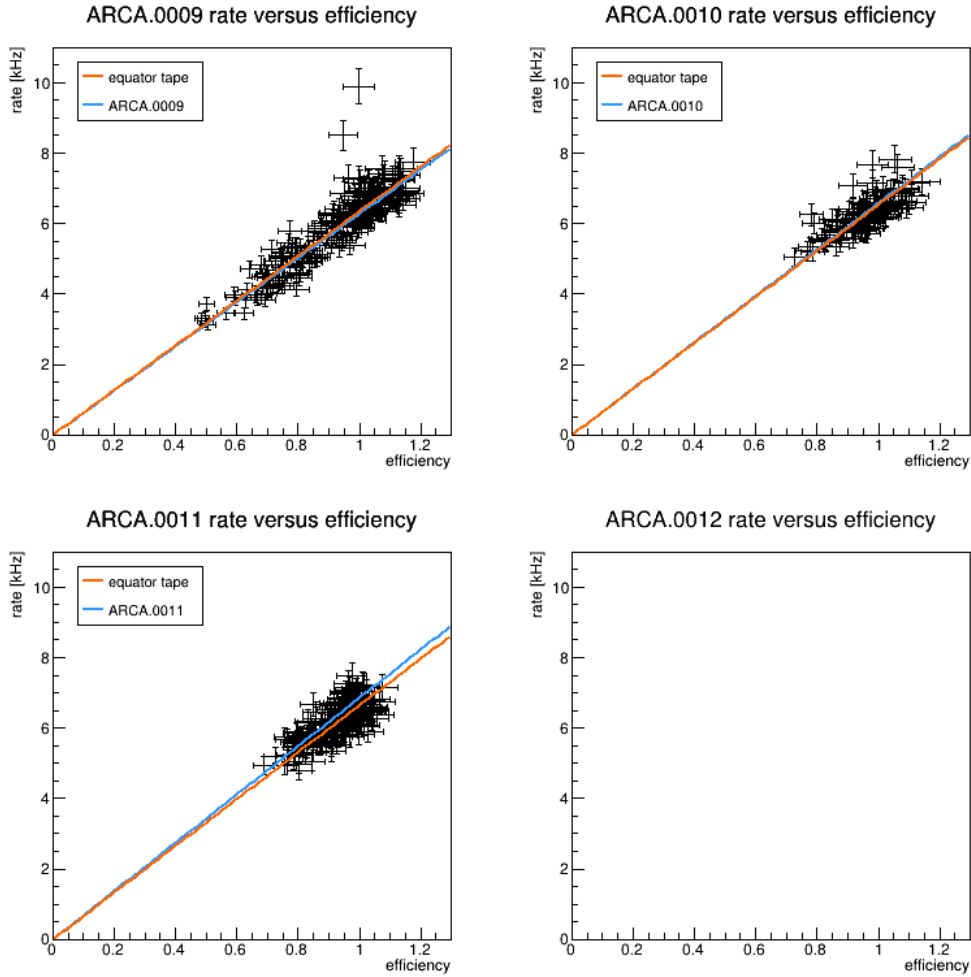


Figure 3.12: Single rates and efficiencies for PMTs shadowed by the equator tape (rings D and E). The data from the PMTs affected by the gel are left out. The orange fit shows the fit through the data points of the PMTs affected by the equator tape of that specific string. The blue string is the fit through data points of that string for all data of that string (except possibly gel-affected PMTs). Continued on next page.

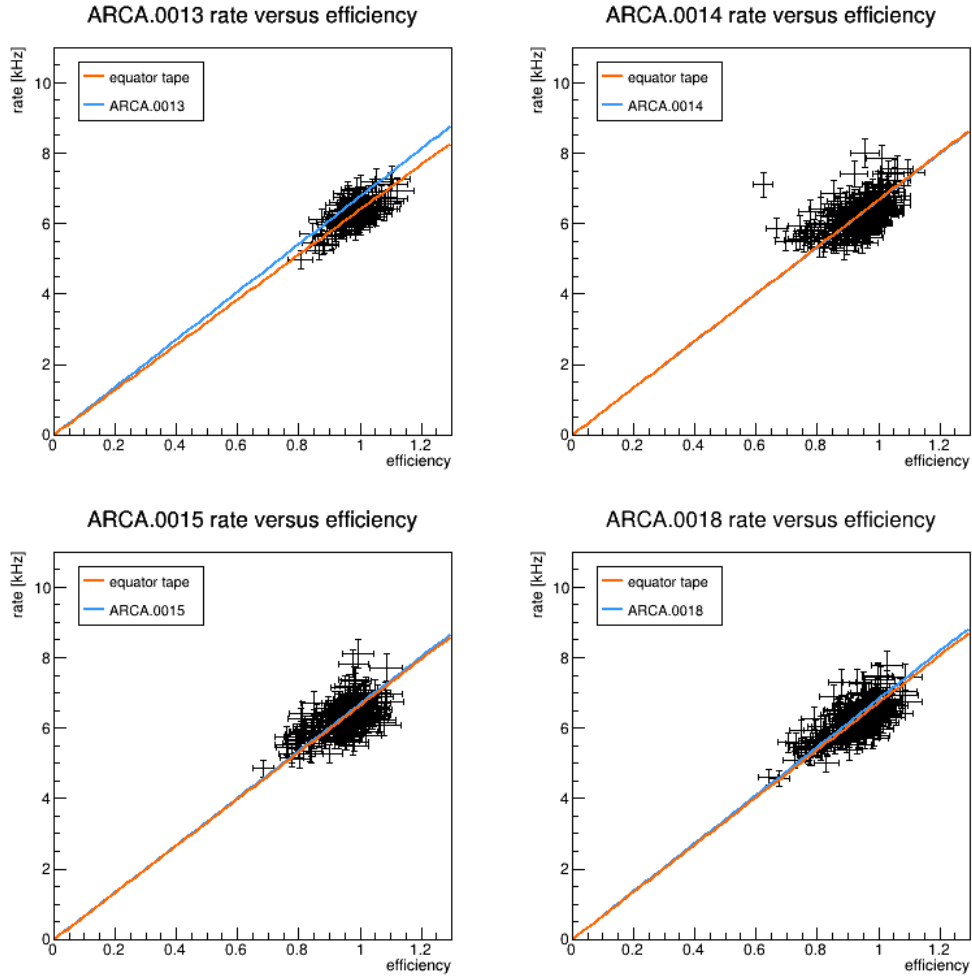


Figure 3.12: Single rates and efficiencies for PMTs shadowed by the equator tape (rings D and E). The data from the PMTs affected by the gel are left out. The orange fit shows the fit through the data points of the PMTs affected by the equator tape of that specific string. The blue string is the fit through data points of that string for all data of that string (except possibly gel-affected PMTs). Continued from previous page.

For ARCA.0009, the blue 'average' line of all ARCA.0009 data lies just below the orange line for the shadowed rings. The two lines overlap for ARCA.0014. Lastly, for strings 10, 11, 13, 15, and 18, the non-shadowed line (blue) is steeper than the shadowed one (orange). This difference is the largest for ARCA.0011 and ARCA.0013.

Again, the small differences in the fits are not significant and thus no

bias is present in the single rates and efficiencies due to shadowing of the equator tape.

In DOM acceptance tests on shore in a dark room it was recently observed that the PMTs A1, B5, B6 had higher dark rates than usual. These specific PMTs were plotted and fitted. Their fits were then compared to the fits through all data to see if the PMTs show an enhanced dark rate.

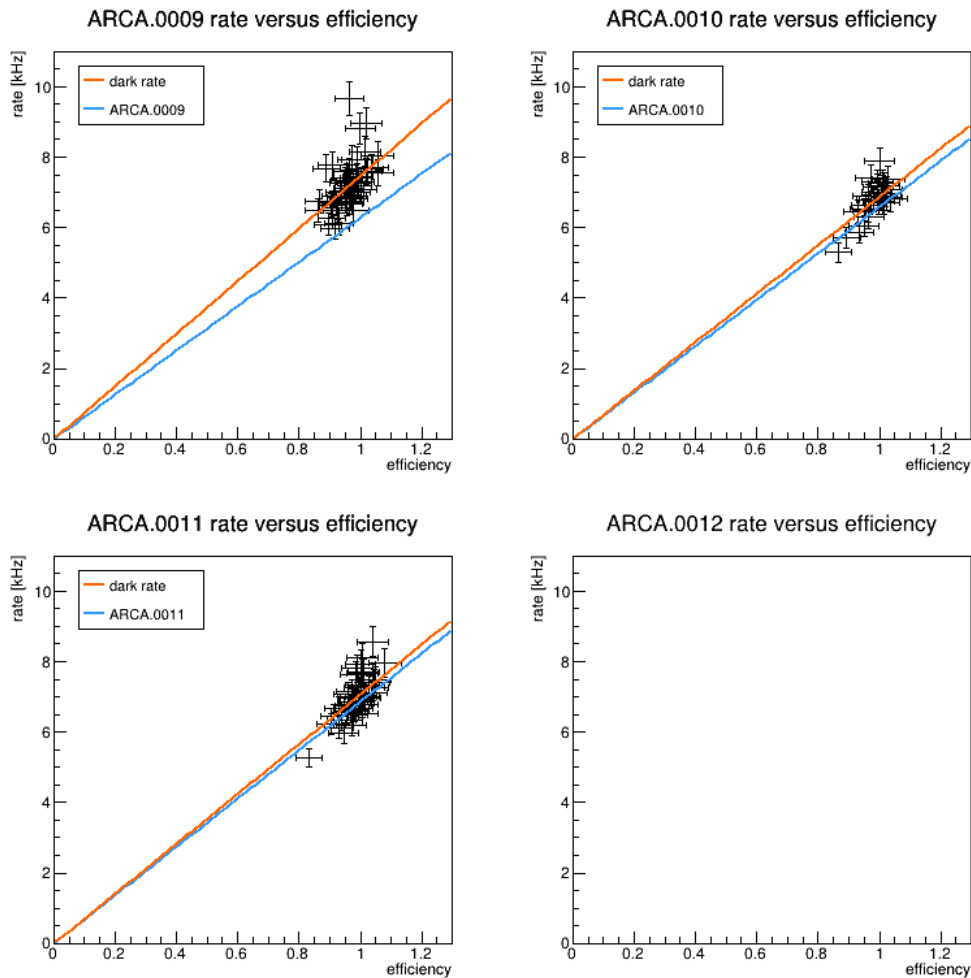


Figure 3.13: Single rates and efficiencies for PMTs with a possibly enhanced dark rate (A1, B5, B6). The orange fit shows the fit through the data points of the PMTs with a suspected higher dark rate for that specific string. The blue line is the fit through all data of that string. Continued on next page.

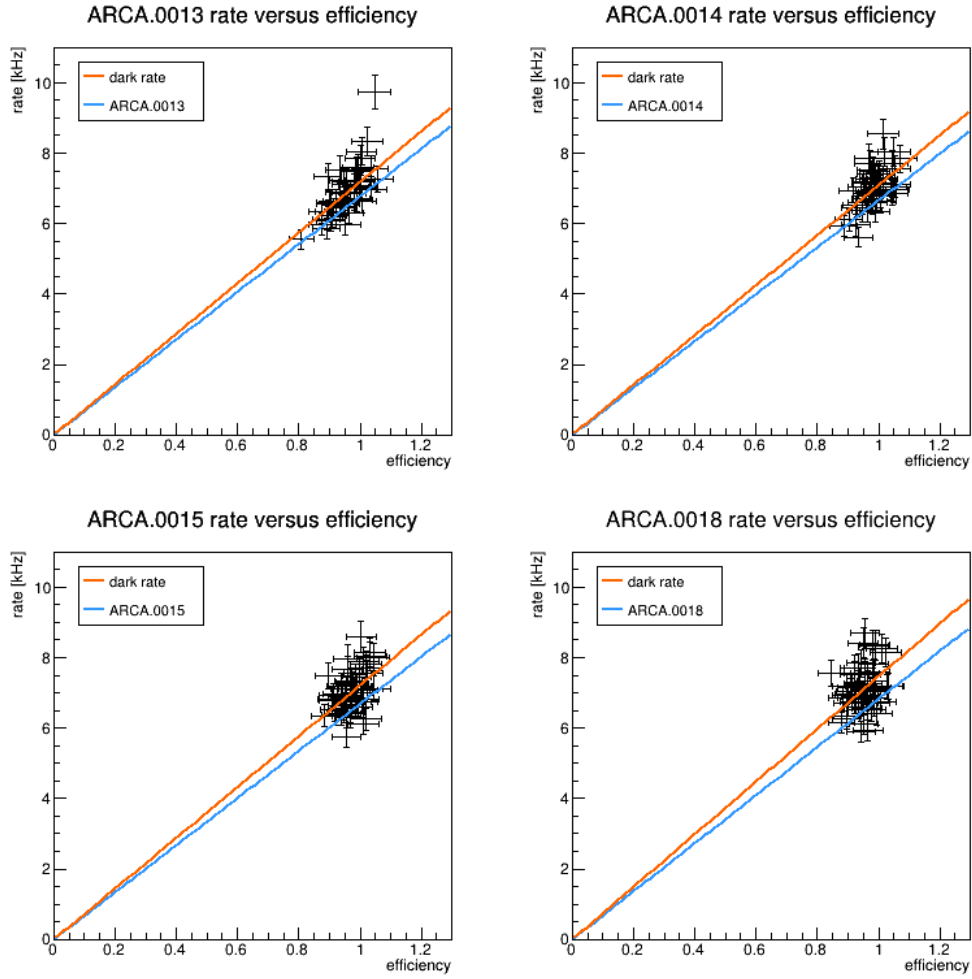


Figure 3.13: Single rates and efficiencies for PMTs with a possibly enhanced dark rate (A1, B5, B6). The orange fit shows the fit through the data points of the PMTs with a suspected higher dark rate for that specific string. The blue line is the fit through all data of that string. Continued from previous page.

All strings show that the fits for these PMTs ('dark rate') are steeper than the fits of all PMTs of the same string. This means that the single rates are higher than usual. Therefore, these PMTs show indeed an enhanced dark rate. The effect is the largest for ARCA.0009. A possible cause is the sedimentation and its influence on the detection efficiency. As the light from outside the DOM is oppressed, the signals from the dark rate might become an important component of all measured signals. The same effect involving radioactive decay from the glass is investigated in section 3.1.4.

Muon light

Whereas it was not expected that the equator tape and titanium belt affects the single rates as that is isotropic light, it is expected that the structures influence photons at certain angles. The effect of the titanium belt can be examined by comparing the shadowed and non-shadowed PMTs of ring C and E.

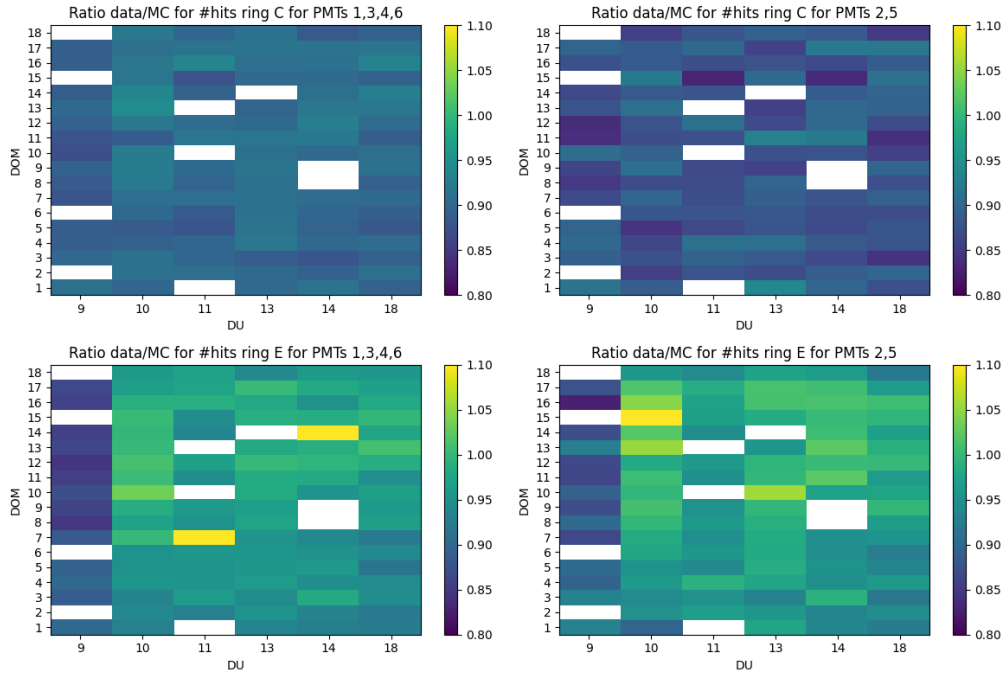


Figure 3.14: The ratio data/MC compared for the shadowed (right) and non-shadowed (left) PMTs of rings C and E. In addition to the shadowing of the titanium belt, all PMTs of ring E are shadowed by the equator tape. Inoperative DOMs are shown as white rectangles.

The figure shows the difference in ratio data/MC for PMTs which are affected by the titanium collar (C2, C5 and E2, E5) and those which are not affected (C1, C3, C4, C6 and E1, E3, E4, E6).

- A clear difference is visible for ring C: the affected PMTs have a ratio of around 0.9, whereas the ratios for C2 and C5 are closer to 0.85-0.9. The MC overestimates the amount of light hitting the shadowed PMTs more than those which are not shadowed.
- For ring E the contrast between the affected and unaffected PMTs is smaller. The main difference is visible in ARCA.0009: the ratio

of ARCA.0009 is considerably higher for E2 and E5 than for E1, E3, E4, E6. The other strings do not show a significant difference. It seems the hypothesis that the overestimation of shadowed PMTs differs from non-shadowed PMTs should be rejected for ring E.

- With regard to the equator tape, the ratio data/MC plot of figure 3.6 shows bins that are just a little bit darker for ring D. This may be due to the shadowing of the equator tape. The efficiencies for ring D are similar to those of rings A-B, indicating that the effect of the equator tape is either well-estimated by the MC or the effect is not sizable to begin with. This is supported by the fact that the ratios for PMTs E1, E3, E4, E6 are close to one. These PMTs are affected by the equator tape, but are either well estimated by the MC or not affected much.

The distribution of the ratios in figure 3.15 confirms the statements above. The overestimation of the shadowed PMTs of ring C is approximately 14%, whereas the overestimation for the non-shadowed PMTs is about 11%. The MC overestimates all PMTs of ring E by $\sim 4\%$.

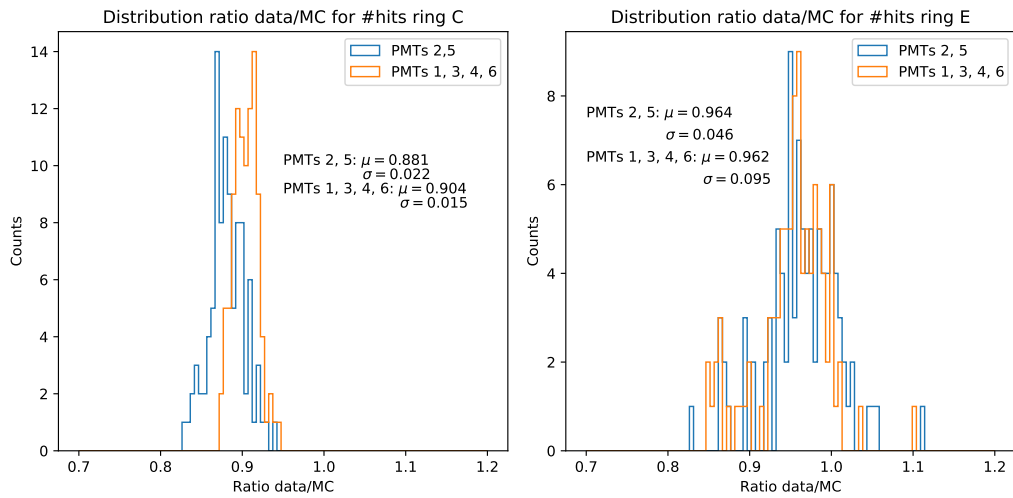


Figure 3.15: The distribution of ratio data/MC compared for the shadowed (blue) and non-shadowed (orange) PMTs of rings C and E. The average values and standard deviations are provided in the plots.

Reconstructed muon angle

A logical next step would be to study the incidence angle of the photons. This information is not easily available, so for now the focus lies on prob-

ing the effects on data parts, for which on average different angular incidences are expected.

The reconstructed muon angles are divided in zenith angles of 0° - 45° and 45° - 90° . This gives a rough distinction between the muons coming from above and those coming from the sides of the DOMs. It should be mentioned beforehand that there is not yet a good representation of the actual angular acceptance, energy, and muon multiplicity spectrum of the atmospheric muon flux. This is something that is being improved at KM3NeT at the moment. The consequence is that it may create a more prominent mismatch in certain angular ranges. Relative comparisons among PMT positions provide therefore the most meaningful conclusions.

As ring E is shadowed by the equator tape from below, difference incidence angles might give different ratios data/MC. As the titanium belt covers some PMTs from the sides, some from above or below, it is expected that a difference will be visible between the shadowed and non-shadowed PMTs for all angles, but less clear than for the equator tape.

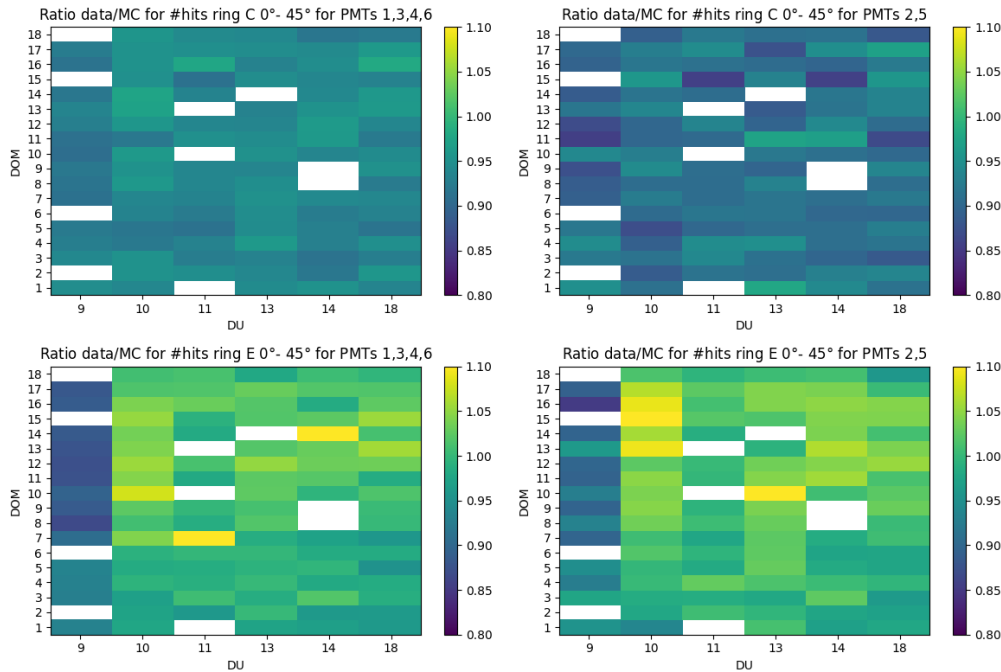


Figure 3.16: The ratio data/MC for rings C and E divided by the PMTs that are shadowed and those which are not shadowed for muons with a reconstructed angle of 0° - 45° . Inoperative DOMs are shown as white rectangles.

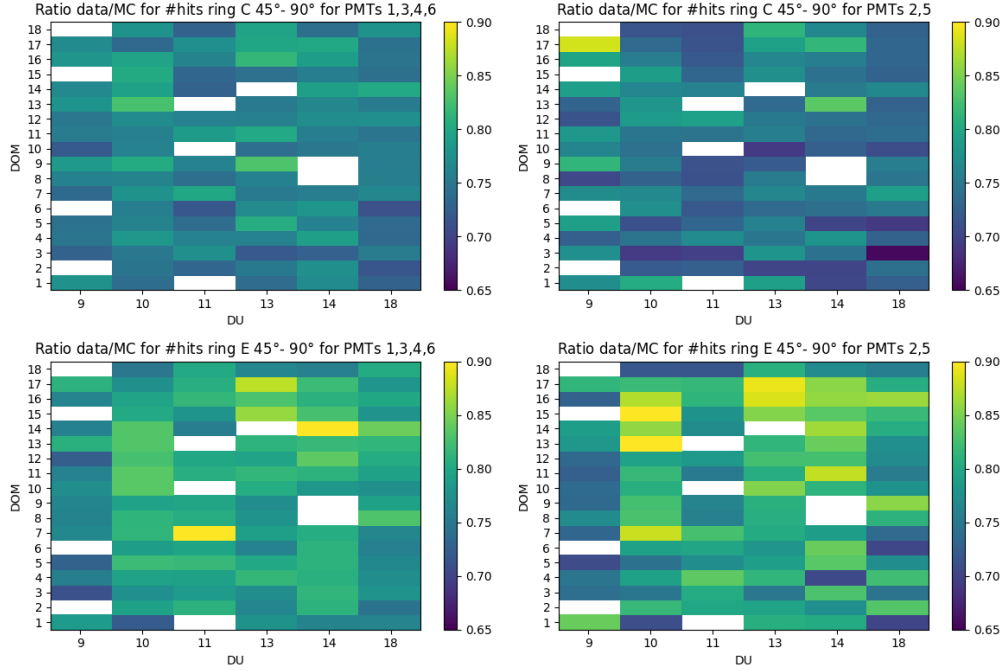


Figure 3.17: The ratio data/MC for rings C and E divided by the PMTs that are shadowed and those which are not shadowed for muons with a reconstructed angle of 45° - 90° . Inoperative DOMs are shown as white rectangles.

For 0° - 45° , the behaviour of the general ratio data/MC is visible. The lower ratio for ARCA.0009 is visible for both ring C and E. The shadowed PMTs of ring C show more overestimation of light than the ones which are not shadowed. This distinction is not as much visible for ring E.

The shadowed PMTs of ring C are again more overestimated than PMTs 1, 3, 4, 6 for 45° - 90° . For ring E there is more fluctuation between DOMs in this range than for smaller angles. All in all the difference between the shadowed and non-shadowed PMTs of ring E has increased slightly for 45° - 90° with respect to 0° - 45° . For ring C this difference is even smaller. What can be concluded is that the effect of the sedimentation on ARCA.0009 has decreased for larger angles for ring E. Section 3.1.4 will further investigate this.

Because the most meaningful conclusions are drawn from comparisons between rings, the non-shadowed PMTs of ring E can be compared to the non-shadowed PMTs of ring C to estimate the size of the bias of the equator tape. The difference between those PMTs for rings C and E decreases to some extent for larger angles. As this is not a substantial difference, it is

not clear if the equator tape shadows ring E less for larger angles. Furthermore, it would not be a logical conclusion: the equator tape shadows ring E from below, not from above. This result might be explained by the sedimentation on top of rings E and F, which is discussed in the next section.

3.1.4 Radioactivity and sedimentation

Muon light

The third and last possible bias to be investigated is the radioactivity. When sedimentation has accumulated on the upper hemisphere of the DOM, light from outside is suppressed. As a result, the light from radioactive decay in the glass increases with respect to signals coming from outside. Because the signals from reactions in the glass are not taken into account in the efficiency estimations, these estimations might be inaccurate. This effect is probably most visible in ARCA.0009, which has been in the water the longest. Sedimentation on a string affects the top half of the DOM. Thus, in this analysis rings E and F will be compared to rings A-D. ARCA.0009 will be compared to the other strings as well.

Figure 3.6 indicates that the ratio data/MC is considerably lower for ARCA.0009 for rings E and F. For the other rings the difference is smaller, but still notable. Especially for ring F the estimated efficiencies (see fig. 3.7) of ARCA.0009 are considerably lower than for other strings.

ARCA.0009 is compared to the other strings in a one-dimensional distribution depicting the data/MC ratios of figure 3.6 (fig. 3.18). This is done per ring, as to get the distinction between the rings simultaneously. The ratio data/MC for each ring is separated in 'ARCA.0009' and 'other'. 'Other' means all strings excluding ARCA.0009. The shapes of the distributions, their average values and standard deviations provide information about the difference between ARCA.0009 and the other strings.

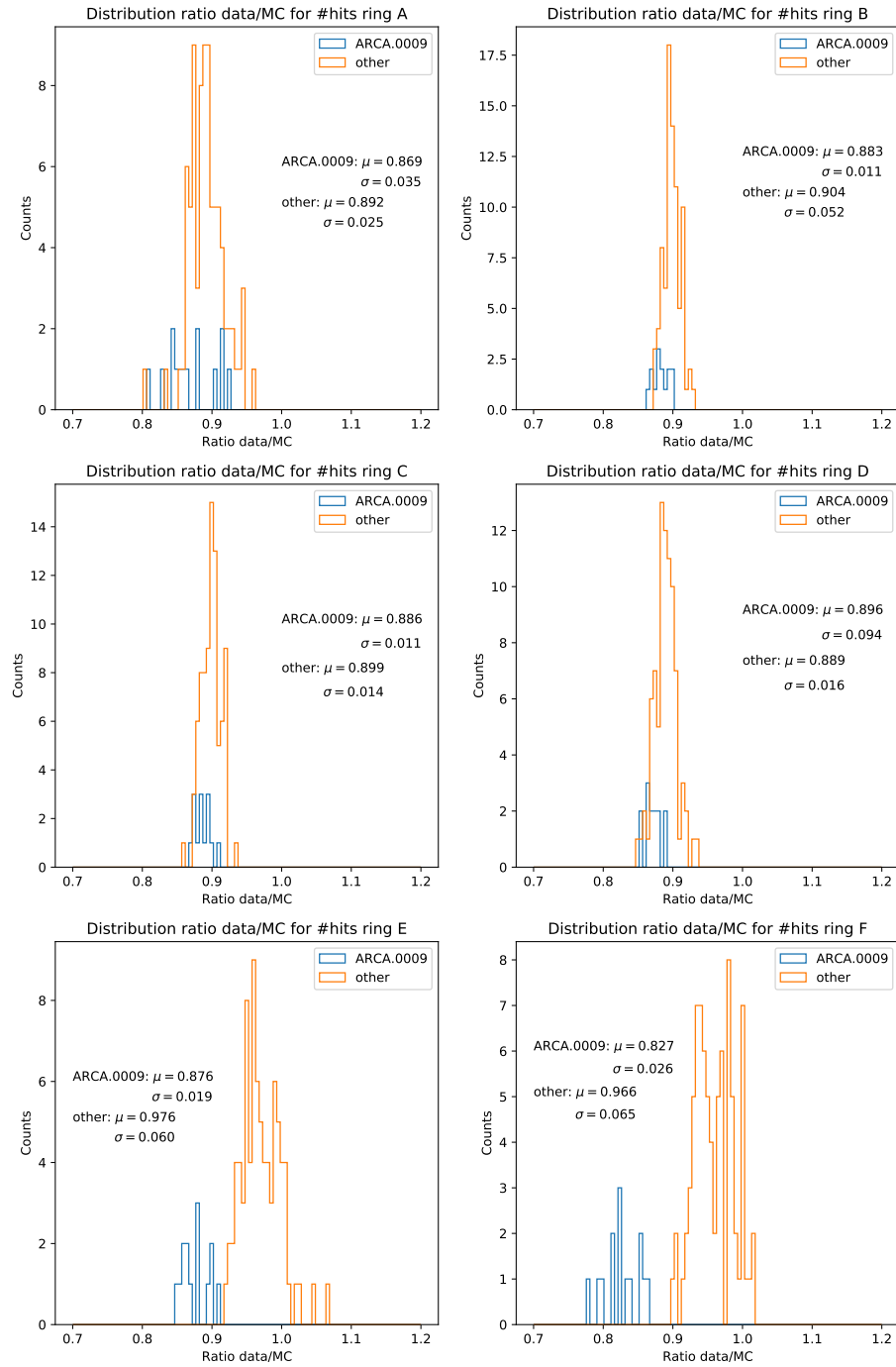


Figure 3.18: Distribution for ratio data/MC. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

It is evident that the distribution of ARCA.0009 is not the same as that of the other rings; this is the case for both rings E and F. The average value of ARCA.0009 is lower than that of the other strings. A KS-test supports this statement. For both ring E and F: the KS test statistic is 1.0, and the p-value 0.0. For the other rings it is not that clear. Their average values are much closer and by eye the distributions cover the same range of ratios data/MC. The KS-test shows that, for an α of 0.05, also the distributions for rings A-D are different. However, their test statistics are considerably lower and their p-values higher.

To compare the rings better, the overestimations of MC are calculated for the rings, both for ARCA.0009 and for the other strings. The results are presented in table 3.1.

Table 3.1: Overestimation of MC in data/MC plots of figure 3.6. Except for ring D, the overestimation of MC is higher for ARCA.0009 than for other strings.

	ARCA.0009 (%)	other strings (%)
Ring A	15.1	12.1
Ring B	13.3	10.6
Ring C	12.9	12.5
Ring D	11.6	12.5
Ring E	14.2	2.5
Ring F	20.9	3.5

The most substantial difference in overestimation is for rings E and F: ARCA.0009 is overestimated by 14.2% and 20.9%, respectively. The other strings are overestimated by 2.5 % (ring E) and 3.5% (ring F).

Compared to the the other strings, the difference in overestimation of ARCA.0009 is on the order of only a few percentage points for rings A-D. The difference is smallest for ring C. The only ring where the overestimation for ARCA.0009 is smaller than for the other strings, is for ring D: 11.6% for ARCA.0009 and 12.5% for other strings. One possible explanation is that there is one DOM for ARCA.0009 (see fig. 3.6) with a ratio data/MC far higher than all other DOMs in this ring. This may affect the average ratio data/MC relatively much, since the average of ARCA.0009 is only calculated using 18 DOMs while the five other strings total 90 DOMs. This makes ARCA.0009 more susceptible to outliers.

There may be a DOM-dependency for this as well. It could be that DOMs closer to the surface behave differently to DOMs closer to the ground. Figure 3.6 shows that for rings A, E, and F the top DOMs have lower data/MC ratios than the bottom DOMs.

Reconstructed muon angle

Dividing the data by zenith angle allows for a rough comparison between light that is expected to come from above and from the side. For DOMs that have no or little sedimentation a minimal difference between zenith angles is expected. When a significant amount of sedimentation covers the DOMs it prevents mostly photons from above from hitting the PMTs. In that case it is expected that the ratio data/MC decreases mostly for the smallest angles.

The ratio data/MC is plotted for all six rings for 0° - 45° (figure 3.19) and 45° - 90° (figure 3.20).

Similarly to the reconstructed muon angle plots for rings C and E, the behaviour of 0° - 45° does not deviate much from fig. 3.6. While the over- and underestimations of MC of the 0° - 45° plot closely resemble those of fig. 3.6, the MC overestimates the amount of light for all rings more for 45° - 90° because of the incorrect simulation of the angular distribution of the muon flux.

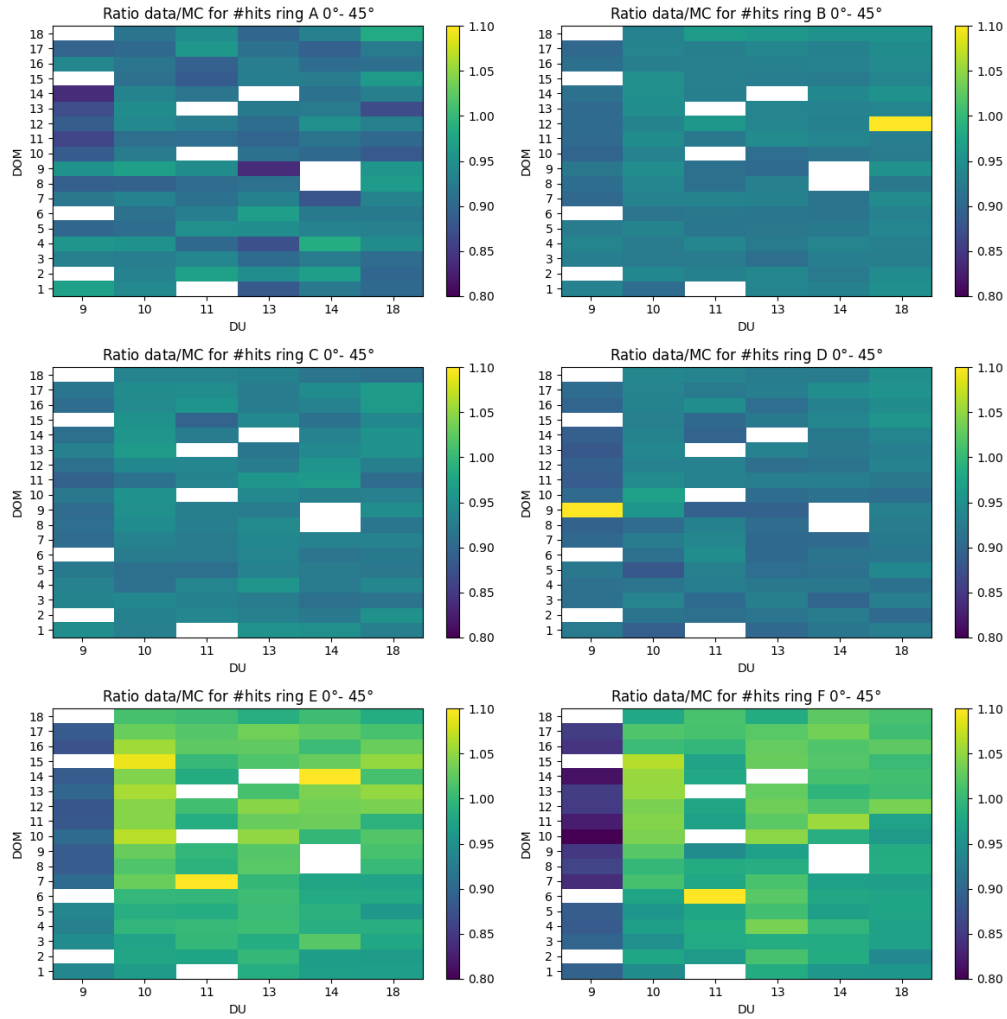


Figure 3.19: Ratio data/MC for 0°- 45° per ring. The x-axis shows the strings and the DOMs are on the y-axis. Inoperative DOMs are shown as white rectangles.

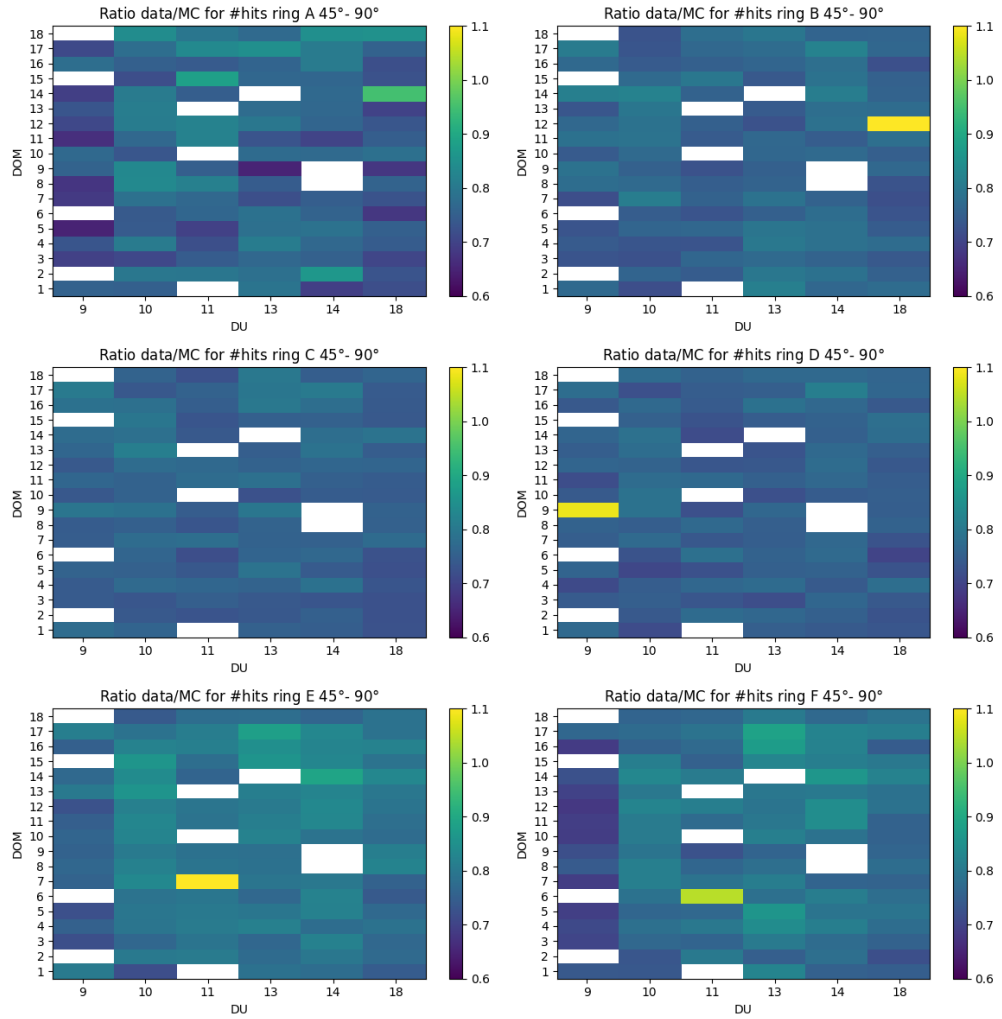


Figure 3.20: Ratio data/MC for 45°- 90° per ring. The x-axis shows the strings and the DOMs are on the y-axis. Inoperative DOMs are shown as white rectangles.

The difference in the MC overestimation of ARCA.0009 and the other strings decreases for 45-90 ° compared to 0-45 °.

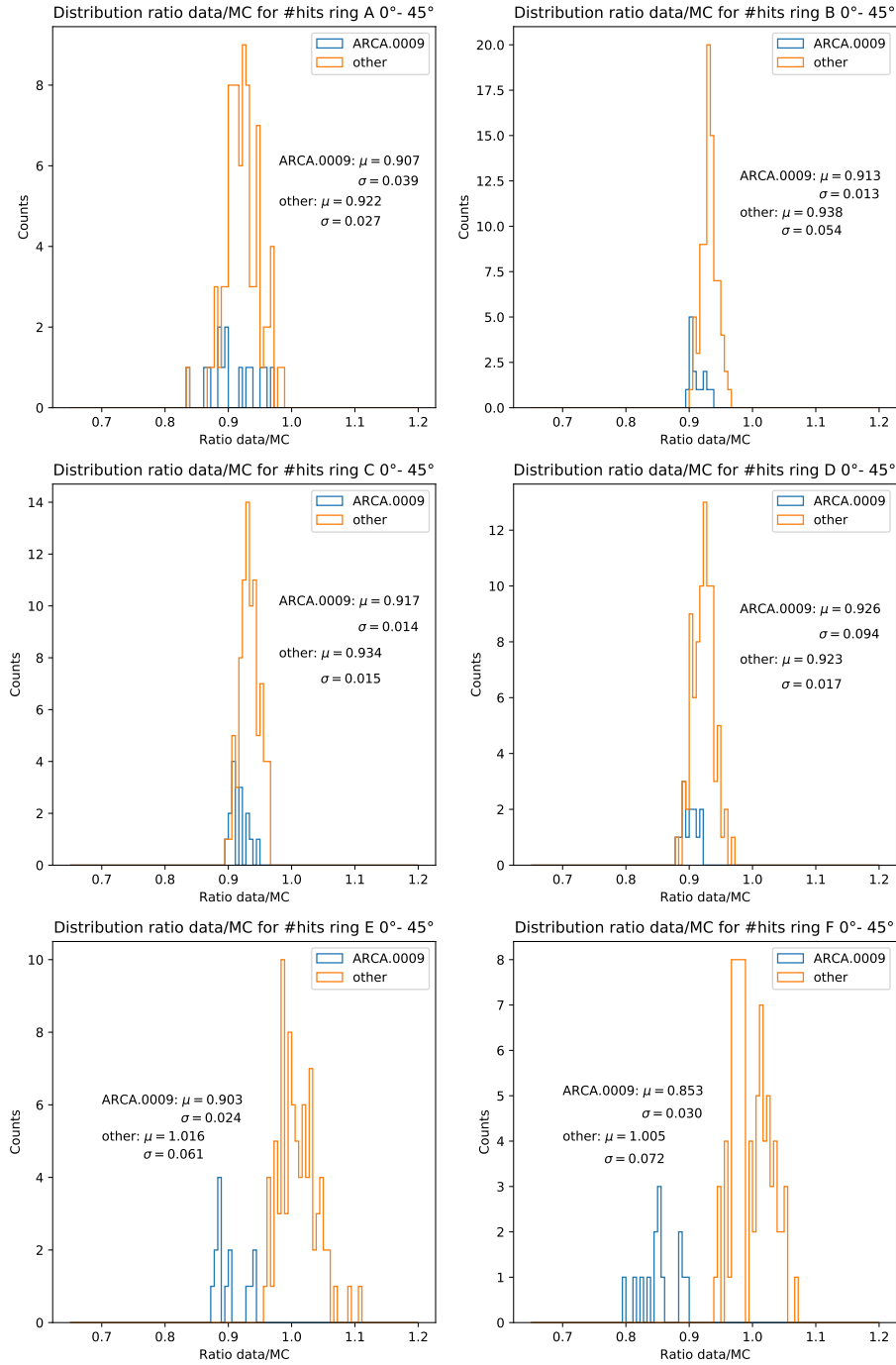


Figure 3.21: Distribution for ratio data/MC for 0°- 45°. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

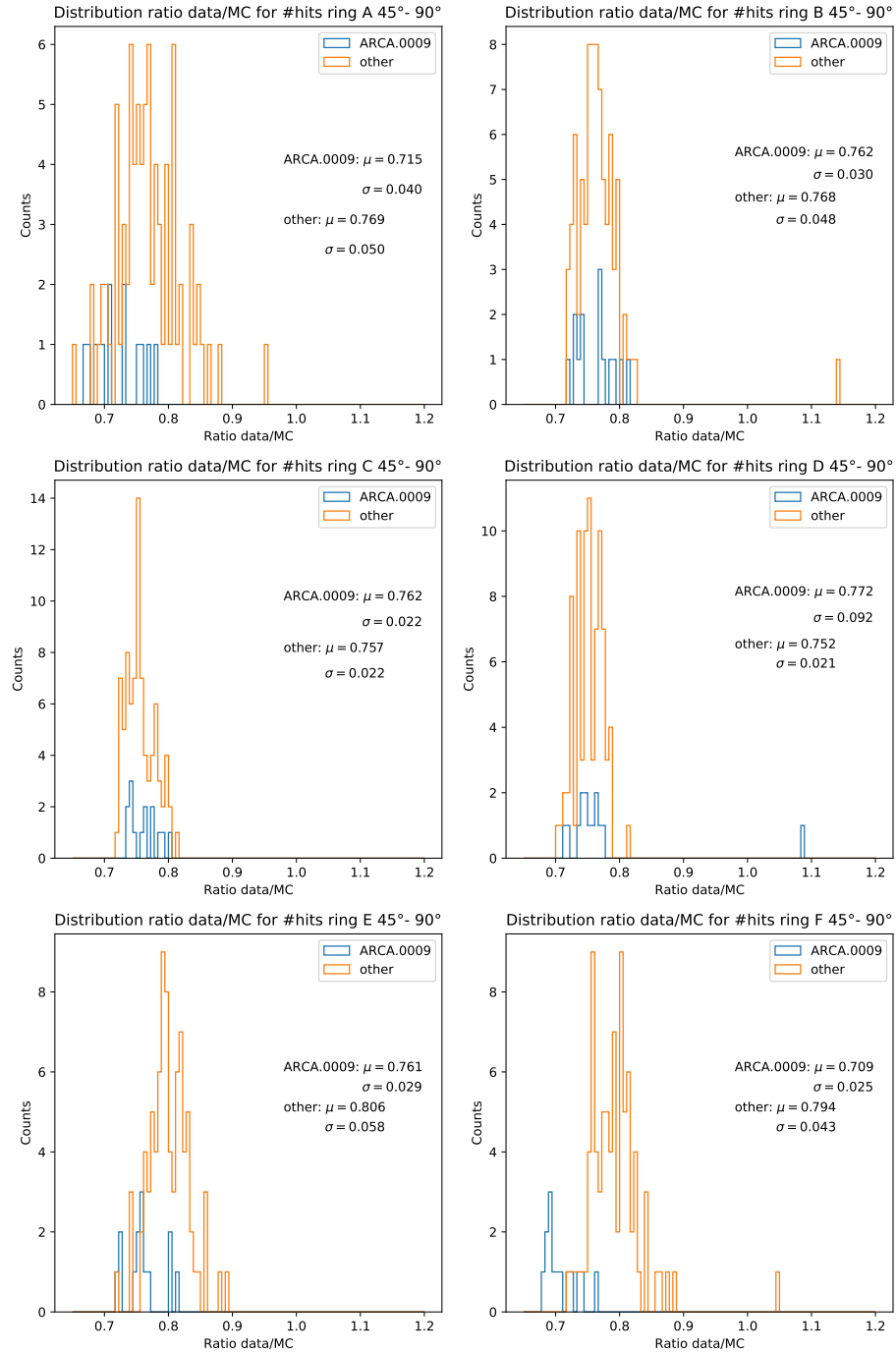


Figure 3.22: Distribution for ratio data/MC for 45°- 90°. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

The shift of ratio data/MC to the left is present for all rings; this is expected because the MC with data agreement is worse for more muons coming from the sides, as mentioned before. However, the fact that rings E and F behave differently compared to other rings has not changed.

Table 3.2: Ratio mean ARCA.0009/mean other strings from fig. 3.21 (0° - 45°) and fig. 3.22 (45° - 90°) respectively.

	0° - 45°	45° - 90°
Ring A	0.984	0.930
Ring B	0.984	0.992
Ring C	0.982	1.007
Ring D	1.003	1.027
Ring E	0.889	0.944
Ring F	0.849	0.893

For all rings except ring A the effect of ARCA.0009 diminishes for larger angles (table 3.2). In figure A.11 the ratio of the means of ARCA.0009/other is 0.987 for 0° - 45° and figure A.12 shows a ratio of 0.986 for 45° - 90° . As this is not a significant difference, and ring A faces downwards, it can be concluded that the outcomes of ring A do not necessarily reflect the behaviours of the other rings. Calculating the ratios mean ARCA.0009/other for figures A.11 and A.12 does give results consistent with 3.2 for rings A-F.

This means that, although the effect decreases for rings E and F as well, the effect is not eradicated completely. This can be seen by examining the distributions themselves, especially for ring F. These results can be an indicator that the sedimentation covers the whole PMTs, but mostly the top of the PMTs.

3.2 ORCA

This section contains mostly the same analysis as the previous one, but for ORCA6. However, since the bioluminescence in ORCA cannot be neglected, it takes more effort to make the rate versus efficiency plots. For ORCA6 the choice was made to focus on the comparison between measured and simulated muons.

Just as for ARCA, first the time residual plots are presented and the best cut on the number of triggered hits is chosen. This cut is then used throughout the rest of the section.

3.2.1 Time residuals

The time residual distribution of figure 3.23 compares the time residuals for data and MC. The MC are weighted by a weighting factor of $\frac{lifetime_{data}}{lifetime_{MC}} = 3.087$. This corrects for the different data and MC lifetimes.

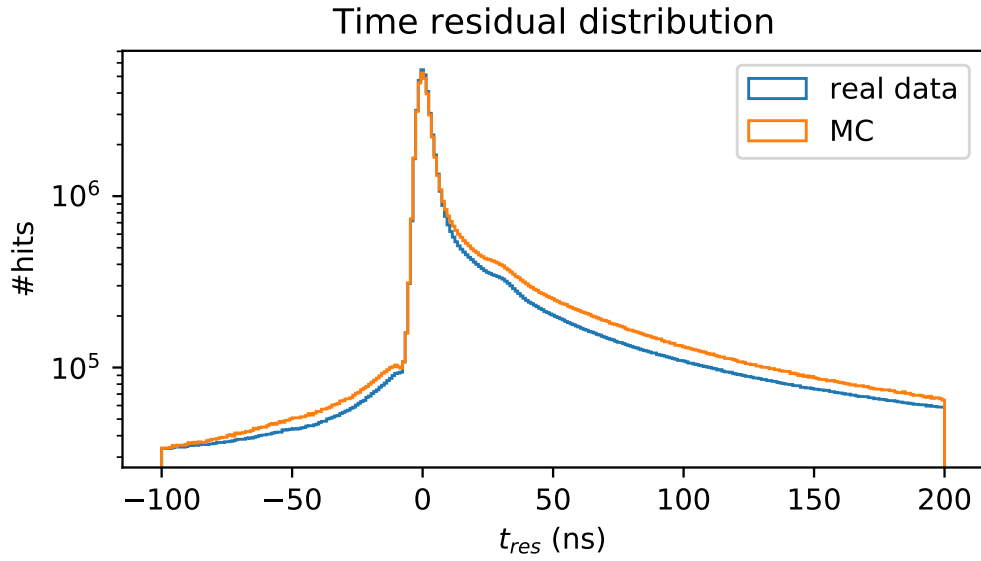


Figure 3.23: The time residual distribution for the real data (blue) and the Monte Carlo simulation (orange). The MC is weighted by $\frac{lifetime_{data}}{lifetime_{MC}}$ to correct for the difference in lifetimes of real data and MC.

From -10 ns to 10 ns, the MC follows the data well. The peaks of data and MC lie at nearly the same number of hits. From 10 ns to 200 ns the MC distribution lies higher than that of the data, but it matches the shape of the data-distribution. There is a small bump at around -10 ns and one at 25 ns, which are tails from the Transit Time transcending the scattered and background light. For negative t_{res} the MC lies above the data as well, but this difference diminishes at -100 ns. All in all the MC is closer to the data than for ARCA.

To see if a cut on the number of triggered hits may still remove some of the noise, the number of triggered hits is plotted. The extended version of this plot can be found in the appendix (fig. A.13).

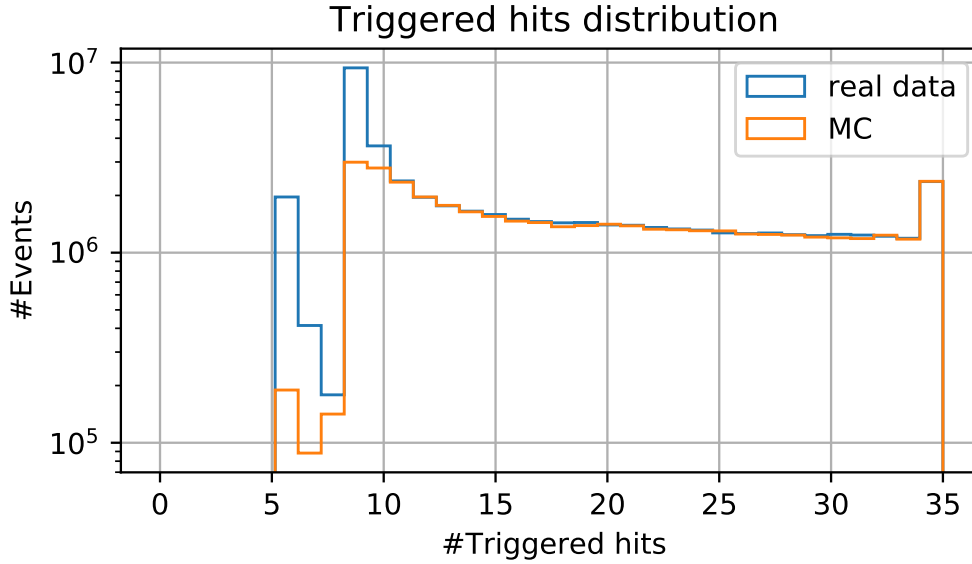


Figure 3.24: The distribution of the number of triggered hits for real data and MC. A cut of 11 triggered hits is applied to exclude the events that have no muon track. Excluding these events gives a better agreement between real data and MC.

It is immediately clear from the graph that the MC of ORCA follows the ORCA data much better than the ARCA MC did for ARCA. However, for a low number of triggered hits there are some peaks in the real data which are not simulated by the MC. Therefore a cut on the triggered hits can still improve the likeness. A cut of 11 triggered hits has been chosen, because after that the MC and data agree well. This excludes most of the events without a muon track, as they do not provide useful information in the analysis. For less than 11 hits there are still more events for the data than for the MC.

The time residual distribution is plotted again with the cut on 11 triggered hits applied.

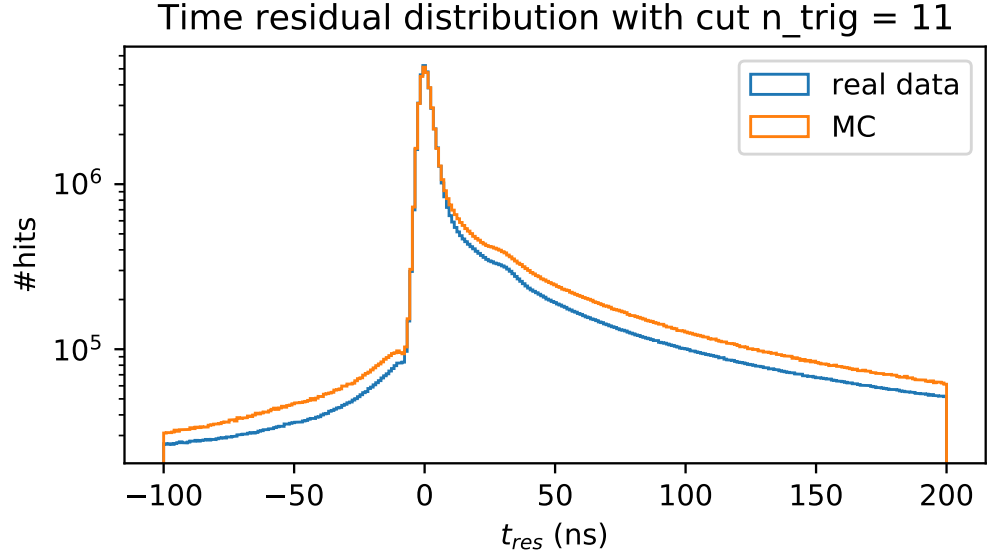


Figure 3.25: Time residual distribution after the cut on 11 triggered hits is applied. The shape of the real data and MC distribution are now more alike than before the cut.

From -10 ns to 10 ns no difference is visible with the applied cut. In the region 10 ns - 200 ns the cut has not changed anything either. However, from a t_{res} of -10 ns to -200 ns, the MC distribution is higher compared to the data than before. This results in the shapes of the data and MC matching better, but the number of triggered hits differs more now. Nevertheless, the cut on 11 triggered hits will be applied throughout the rest of the plots, since this does remove events that do not have a reconstructed track and thus contain no relevant information. In turn, this should give more accurate results.

In order to analyse the similarities of the data and MC distribution better, the time residuals are plotted per ring.

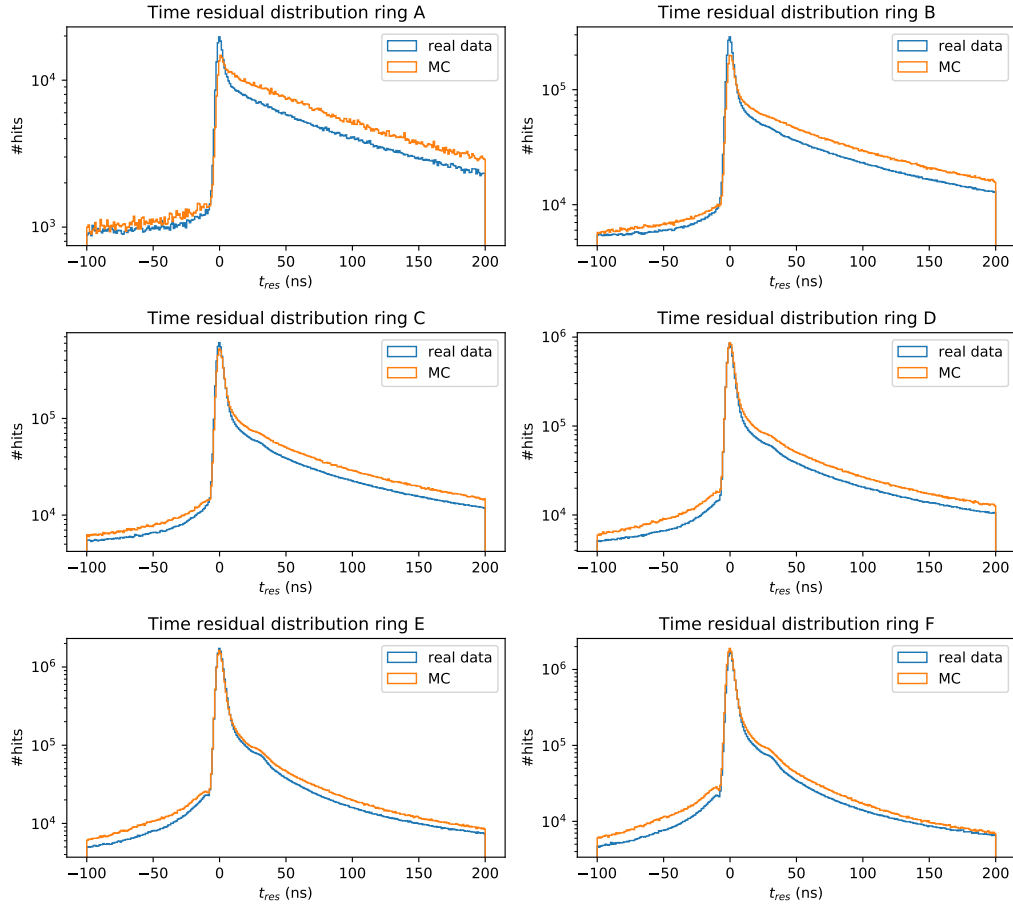


Figure 3.26: Time residual distributions sorted by ring.

Just as is the case for ARCA, ring A has the worst agreement between data and MC. Moreover, it is also the ring with the most noise. Where the MC follows the data quite well in the negative t_{res} region, the peak of the MC is considerably lower than that of the data. Although the shape of the MC is comparable to the data from 10 ns to 200 ns, the MC is here higher by a sizable amount.

3.2.2 Possible biases due to change in gel transparency

Previous investigations indicated that DOMs 1-6 of ORCA.0001 may have a different gel transparency. The gel used for these DOMs comes from the same batch as the gel of DOMs 7-15 of ARCA.0010.

The following plots use the data/MC ratio with the full range of t_{res} hits.

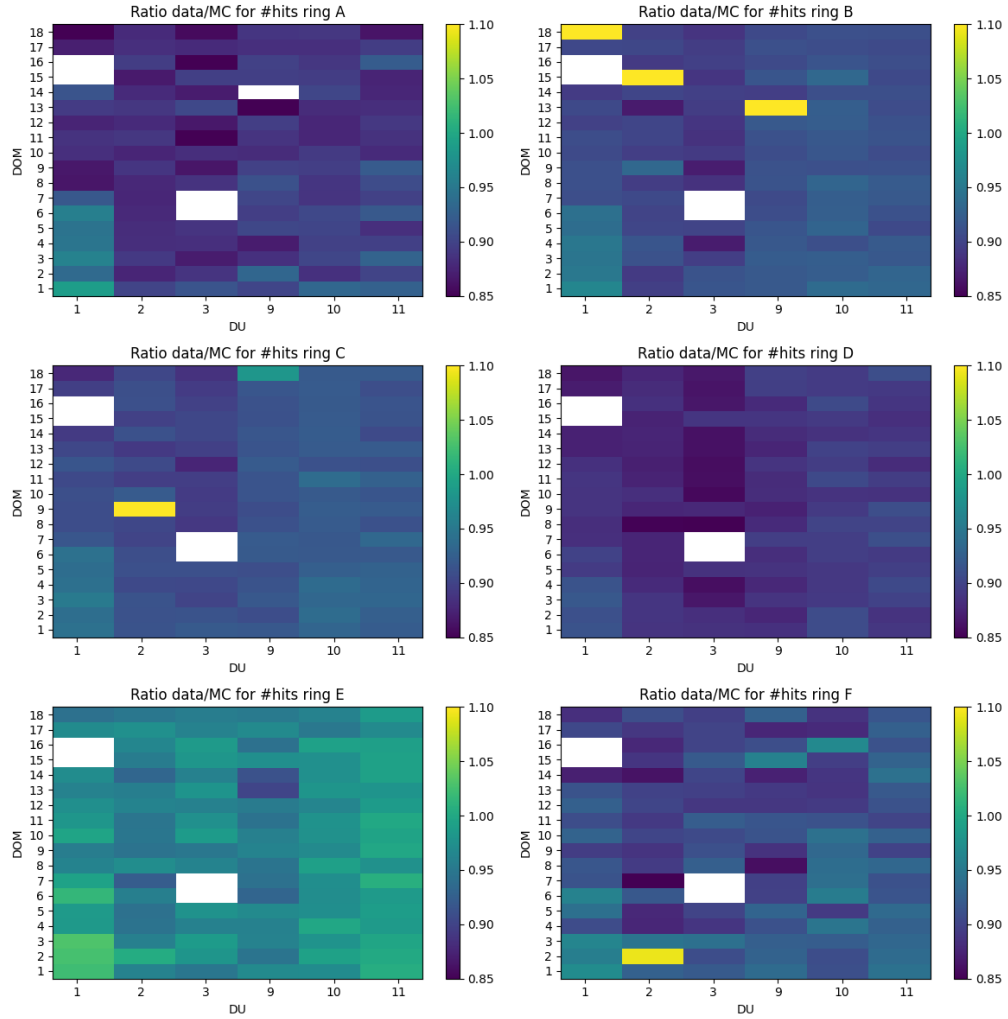


Figure 3.27: Ratio data/MC sorted ringwise. The x-axis shows the strings, the y-axis the DOMs. Inoperative DOMs are shown as white rectangles.

As expected the ratio data/MC is lower for ORCA.0001 DOMs 1-6 (on average a 5% overestimation of MC) than DOMs 7-18 ($\sim 10\%$ overestimation) of the same string. This feature is present in all six rings.

Using the same division by ring, the efficiencies are shown below. They may reflect the relatively high overestimation of MC as lower efficiencies for DOMs 1-6.

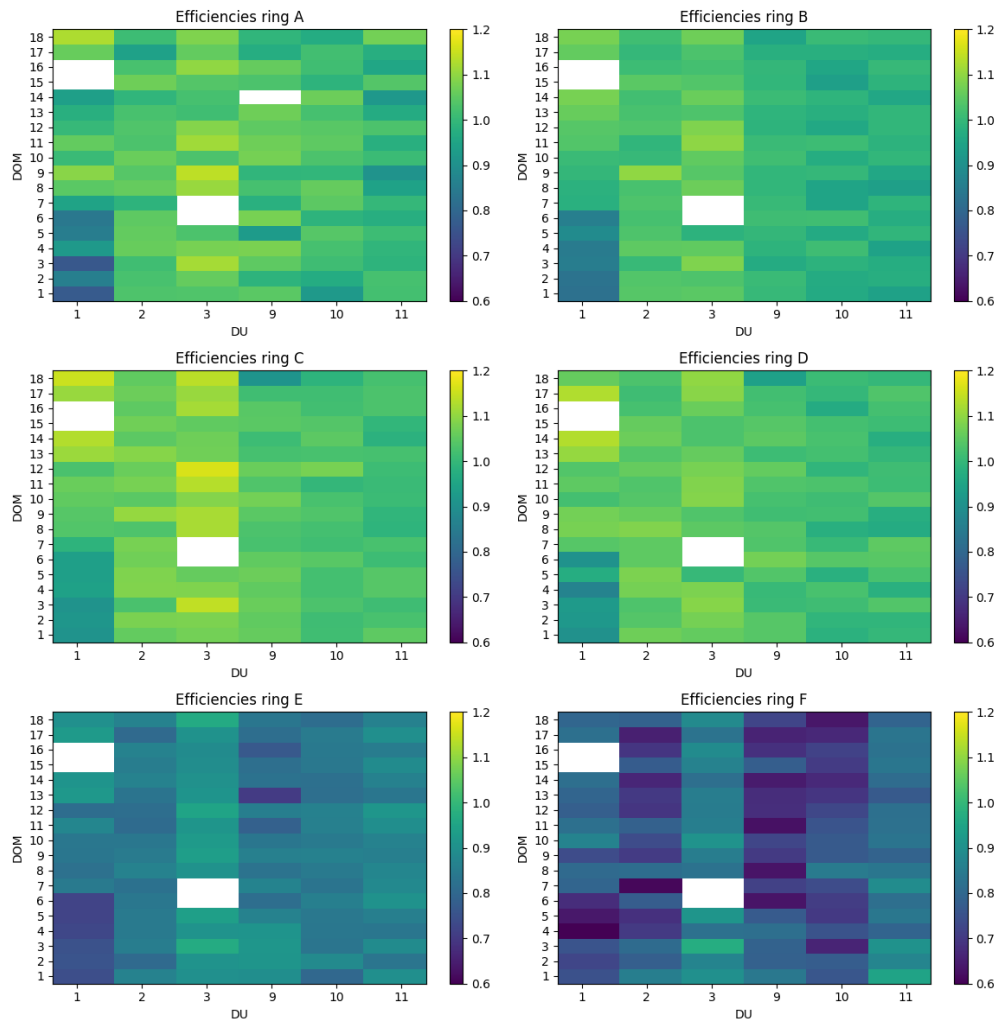


Figure 3.28: Efficiencies sorted ringwise. The x-axis shows the strings, the y-axis the DOMs.

In all rings the efficiencies of DOMs 1-6 of ORCA.0001 are lower than DOMs 7-18. These findings are consistent with the efficiencies of DOMs 7-15 of ARCA.0010 and rings E and F of ARCA.0013, which had relatively low efficiencies as well.

Noteworthy are the low efficiencies for rings E and F, especially ring F strings 2, 9, and 10. This will be discussed in section 3.2.4

To investigate the difference in ORCA.0001 DOMs, the one-dimensional distribution of the ratios data/MC are plotted per ring. DOMs 1-6 are compared to DOMs 7-18.

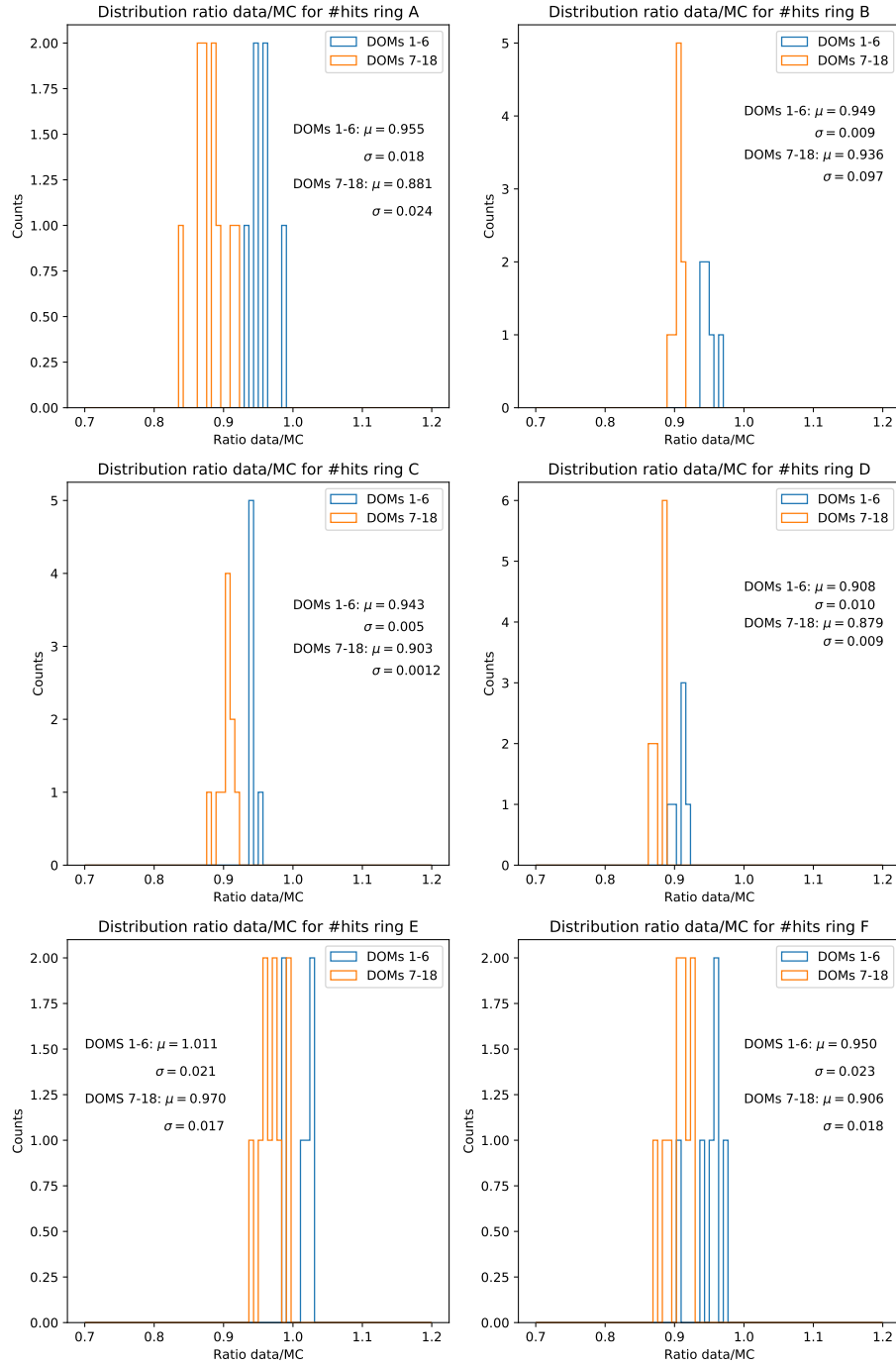


Figure 3.29: Distribution ratio data/MC for ORCA per ring. The distributions for ARCA.0001 DOMs 1-6 (blue) and the DOMs 7-18 of ORCA.0001 (orange) are shown per string. The average values and standard deviations are presented in the graphs.

On average, hits of DOMs 1-6 of ORCA.0001 are overestimated less than the other DOMs of that string. As this can be seen in all rings, it is a systematic bias. Ring A has the most substantial difference. Although this ring consists of only one PMT, the difference is probably not caused by fluctuations alone. The validity of comparing DOMs 1-6 to DOMs 7-18 is lower because only six respectively twelve ratios are used in the distribution and in calculating the mean values. Furthermore, ring A is a single PMT and has thus already less data than rings B-F. However, the difference in the sets of DOMs is found for all six rings, so the results are significant enough to call it a bias.

Table 3.3: Overestimation of MC in data/MC plots of figure 3.27. For all rings ORCA.0001 DOMs 1-6 are more overestimated than the other DOMs of ORCA.0001 and other strings. The minus sign in the second column of ring E means that these DOMs are underestimated.

	ORCA.0001 DOMs 1-6 (%)	ORCA.0001 DOMs 7-18 (%)
Ring A	4.7	13.5
Ring B	5.4	6.8
Ring C	6.0	10.7
Ring D	10.1	13.8
Ring E	-1.1	3.1
Ring F	5.2	10.4

Table 3.3 shows not only difference between the affected and non-affected DOMs, but also between rings.

3.2.3 Shadowing

As mentioned before, the equator tape shadows rings D and E. PMTs C2, C5, E2, and E5 are shadowed by the titanium collar. The ratios data/MC for the shadowed and non-shadowed PMTs of rings C and E are separated to investigate the effect of these structures.

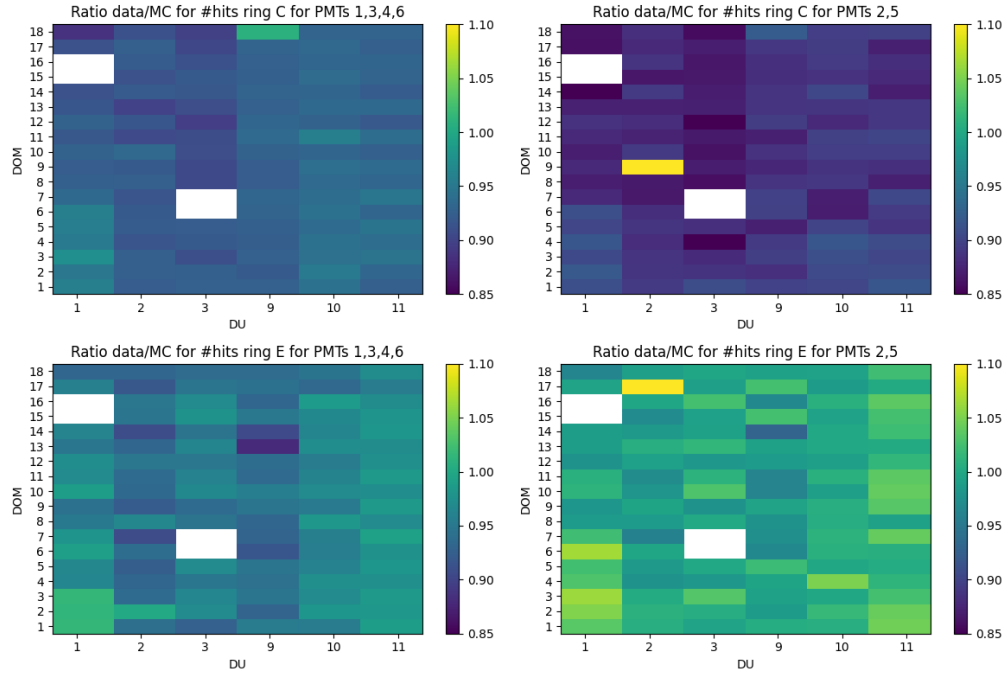


Figure 3.30: The ratio data/MC compared for the shadowed (right) and non-shadowed (left) PMTs of rings C and E. In addition to the shadowing of the titanium belt, all PMTs of ring E are shadowed by the equator tape. Inoperative DOMs are shown as white rectangles.

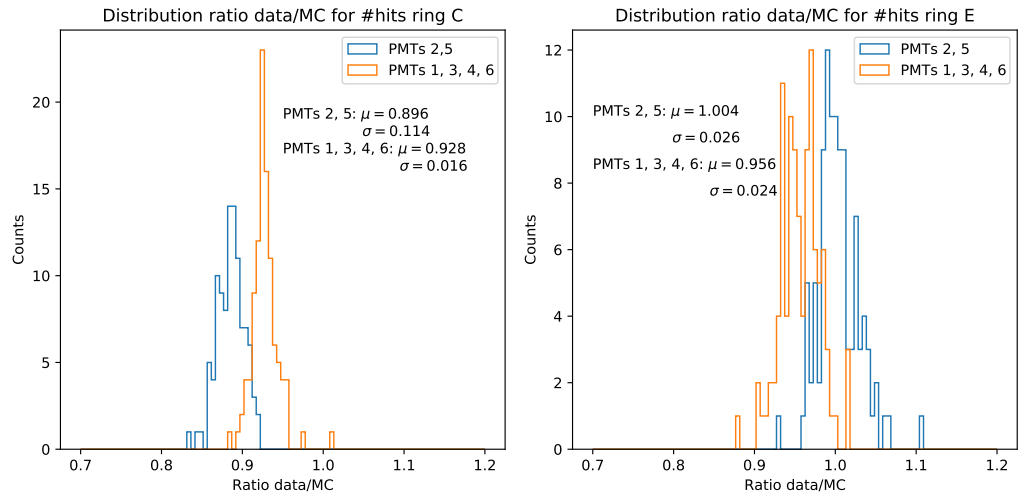


Figure 3.31: Distribution of ratio data/MC for shadowed and non-shadowed PMTs. The titanium belt shadows PMTs C2, C5, E2, E5. The titanium collar shadows all of ring E.

For both ring C and ring E, ORCA.0001 has a lower ratio data/MC for its six lowest DOMs. This is present in PMTs 1, 3, 4, 6 and PMTs 2, 5.

- While the non-shadowed PMTs of ring C show an overestimation of about 4% to 10%, the average of PMTs C2 and C5 is 0.896: an overestimation of 10% of MC. This means that these PMTs receive less light than expected, so the efficiencies of C2 and C5 should be lowered. This seems different from ARCA, which is to first order not expected, as the shadowing should affect the detectors similarly.
- PMTs E2 and E5 are not only shadowed by the titanium belt, but also by the equator tape. That is why they have on average higher ratios data/MC than the non-shadowed PMTs. Comparing the two distributions of fig. 3.31 allows one to quantify the effect of the shadowing of the equator tape and titanium belt. PMTs E1, E3, E4, and E6 are only shadowed by the equator tape, which gives an overestimation of 5%. Due to the titanium belt this overestimation is balanced to a ratio of 1.004 for PMTs E2 and E5. The underestimation of the titanium belt is therefore around 5%. In conclusion, the efficiencies of PMTs E2 and E5 are well estimated, but the efficiencies of the other PMTs in ring E are actually lower than assumed now.
- In ring D the equator tape results in an overestimation of approximately 10%. For the DOMs affected by the change in transparency of the gel it is even higher (13.1%). Figure 3.28 shows that the efficiencies for ring D are of the same order as those of ring A - C. The extra overestimation of ring D shows that the equator tape causes a bias of around 4 percentage points compared to rings A - C and F.

3.2.4 Radioactivity and sedimentation

Sedimentation affects predominantly the top of the DOMs, i.e. rings E and F. Less photons from outside the DOMs would be able to reach the covered PMTs. As a consequence, the light from radioactive decay in the glass of the DOM increases relatively to the light from outside. The efficiency estimations do not consider this radioactive signal and might thus be inaccurate.

This would translate to a difference in ratio data/MC between rings A-D and E-F. Figure 3.27 showed the general data/MC ratio. They have been divided by zenith angle (0° - 45° and 45° - 90°) in figures 3.32 and 3.33 to probe the angular dependency of the sedimentation. It was expected that the ratio data/MC for 0° - 45° would be close to that of 3.27,

as muons mostly come from above. The most distinct difference between rings would be in this range.

Figure 3.27 (see section 3.2.2) shows that the least amount of overestimation of MC is found in ring E. Rings B, C, and F have approximately the same ratio data/MC. The time residuals for ring F (fig. 3.26) show that the MC is not as close to the real data as for ring E, for example. However, this difference is not substantial enough that it explains the lower ratio data/MC.

Section 3.2.2 already mentioned the low efficiencies for rings E and F (see fig. 3.28). It is plausible that this is because sedimentation covers mostly the upper hemisphere of the DOM and this is already considered in the estimation of the efficiencies. This would also explain that the efficiencies for ring F are lower than for ring E, as ring F is the uppermost ring. The lower efficiencies for ring E may also be partially due to the shadowing discussed in the previous section.

Efficiencies for ORCA rings A-D are comparable to those of ARCA rings A-D (fig. 3.7), while the efficiencies of rings E and F are considerably lower for ORCA than for ARCA. Comparing the ratios data/MC for ARCA and ORCA, one finds that only the ratios for ring F are significantly dissimilar: ARCA ring F is overestimated by 3.5% but ORCA ring F by 9.5%. MC overestimates the number of hits in ring E by 2.5% and 3.1% for ARCA and ORCA, respectively. For the other rings the overestimations of ARCA and ORCA are all within 1.3 percentage points of each other.

When for ORCA the data/MC ratios are divided in 0° - 45° and 45° - 90° a clear dependence on zenith angle is visible. In addition to investigating the angle dependency, the data from ORCA.0001 DOMs 7-18 were compared to those of the other strings as ORCA.0001 has been deployed the longest, so the most sedimentation is probably on this string. DOMs 1-6 of ORCA.0001 were excluded from the one-dimensional distributions because hits of these DOMs overestimated more by MC due to the change in gel transparency.

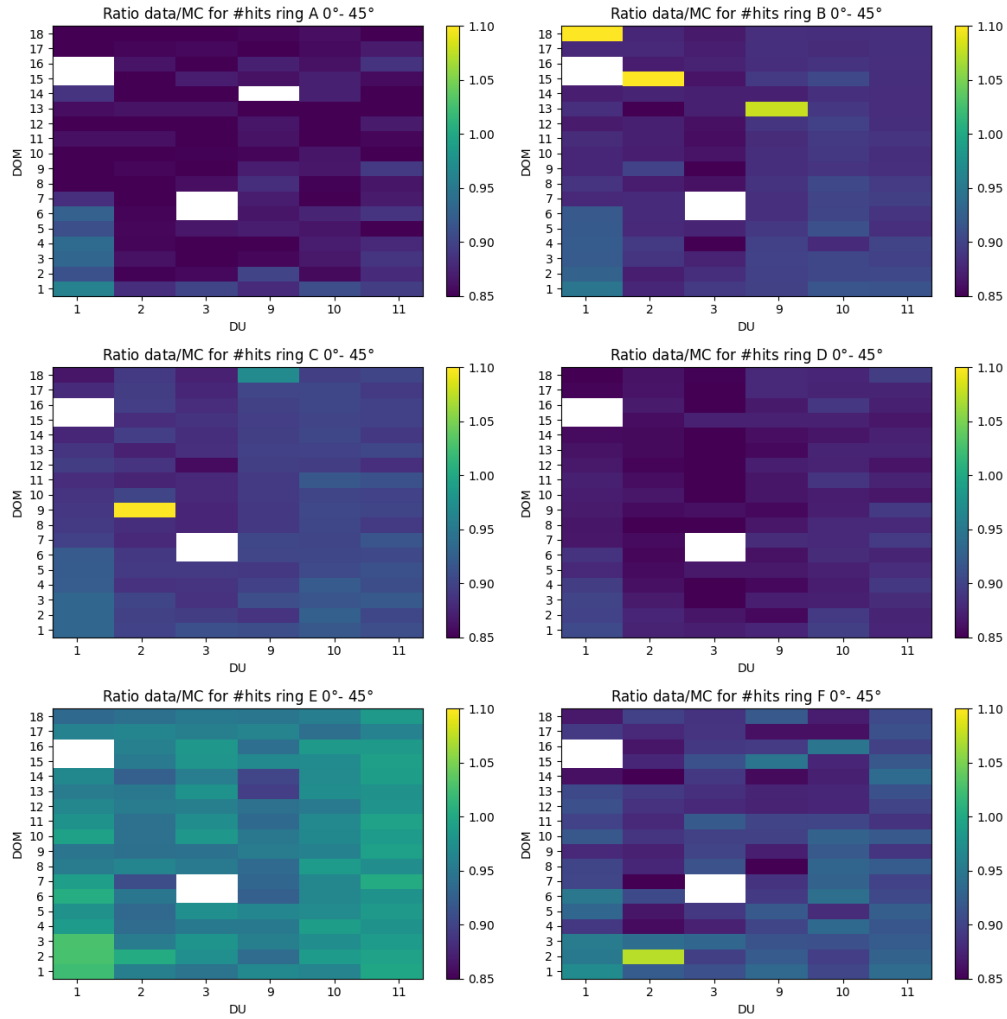


Figure 3.32: Ratio data/MC for 0°- 45° per ring. The horizontal axis shows the strings and the DOMs are on the vertical axis. Inoperative DOMs are shown as white rectangles.

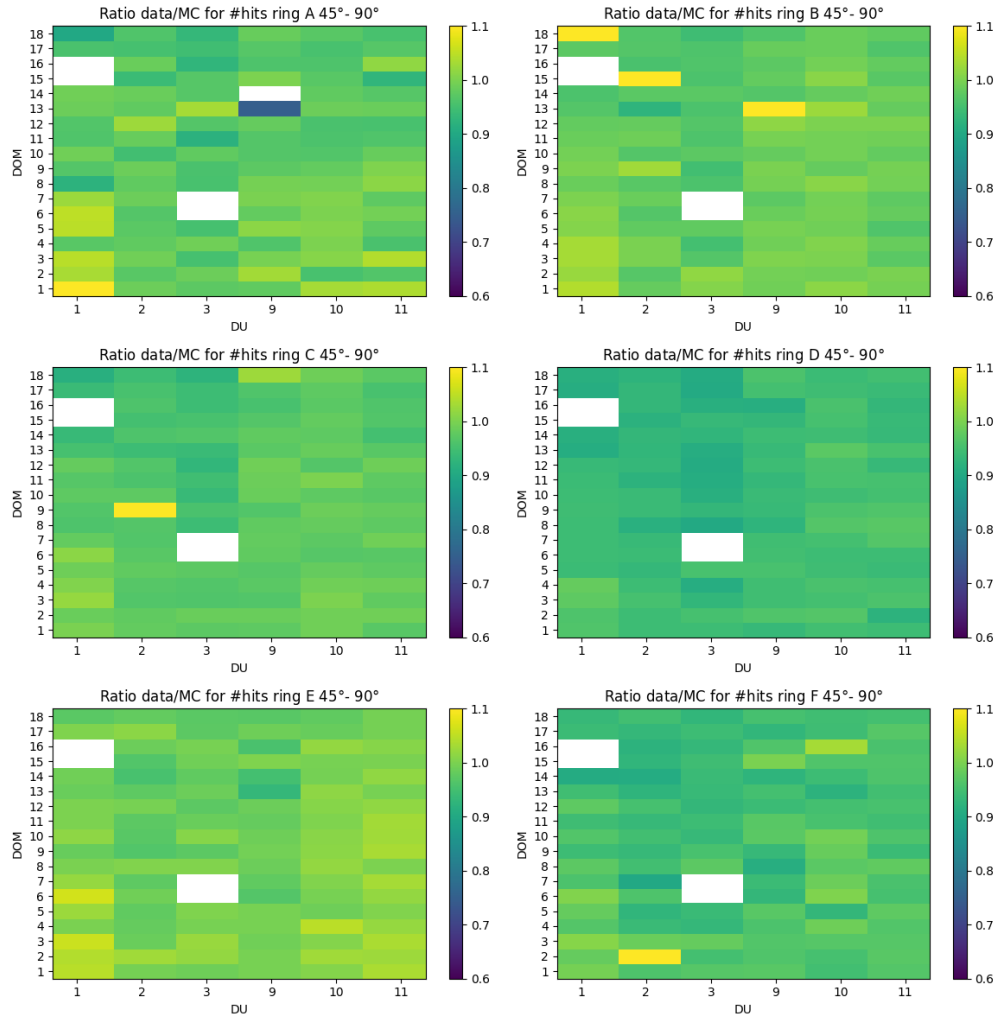


Figure 3.33: Ratio data/MC for 45°- 90° per ring. The horizontal axis shows the strings and the DOMs are on the vertical axis. Inoperative DOMs are shown as white rectangles.

It was correctly hypothesised that the ratios data/MC of 0°- 45° would resemble those of all ratios (fig. 3.27) more than 45°- 90° does, although this might be a consequence of the simulations that do not accurately describe the angular distribution.

For 0°- 45° the distinction between ring E and the other rings is still noticeable. The similarity between ring F and rings B and C has not changed either. Something interesting is found when ring E is compared to the other rings for 45°- 90°. The difference in ratios between them decreases in contrast to 0°- 45°.

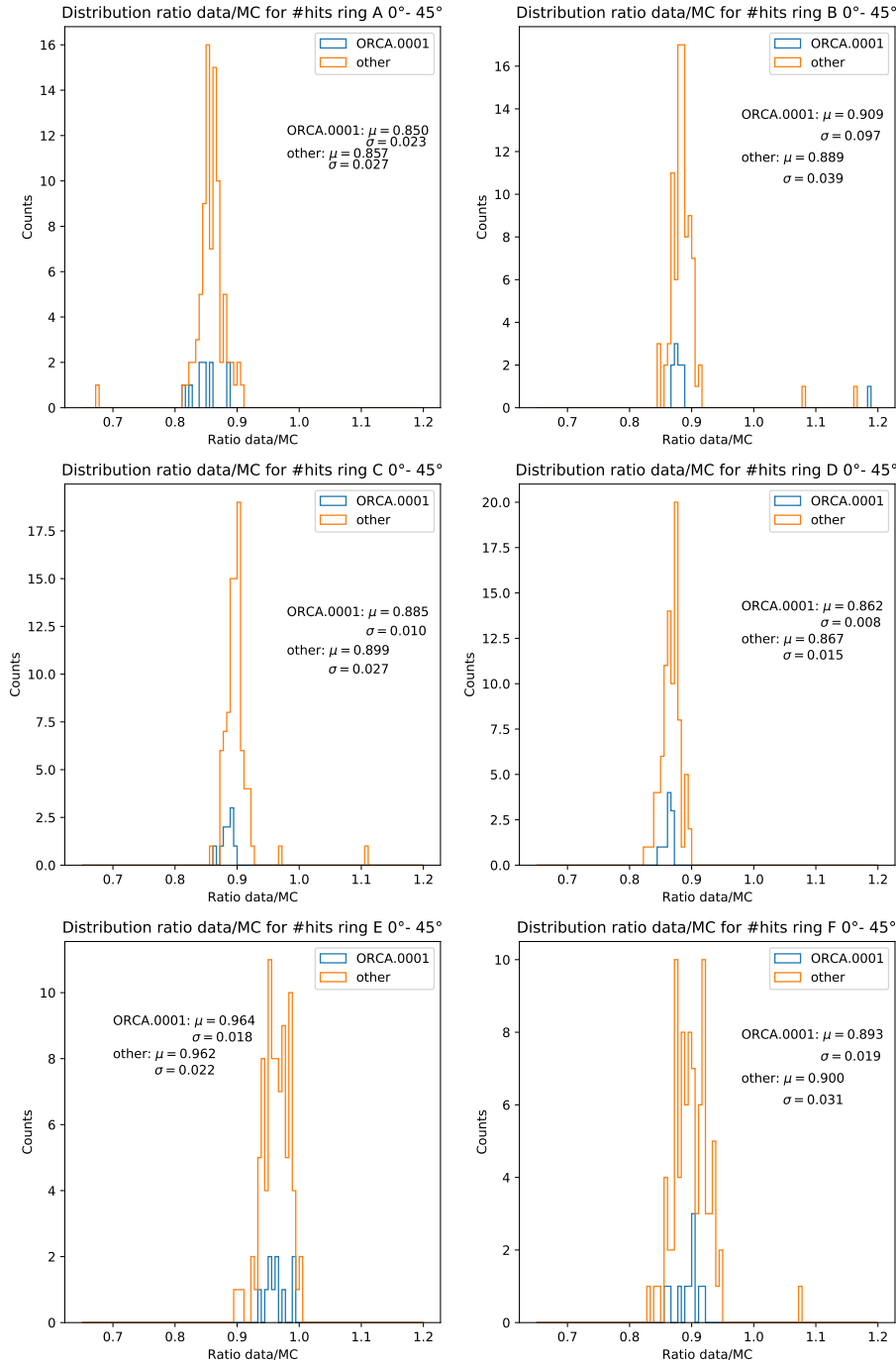


Figure 3.34: Distribution of ratio data/MC for 0°- 45°. The distributions for ORCA.0001 DOMs 7-18 (blue) and the other strings (orange) is shown per ring. The average values and standard deviations are presented in the graphs.

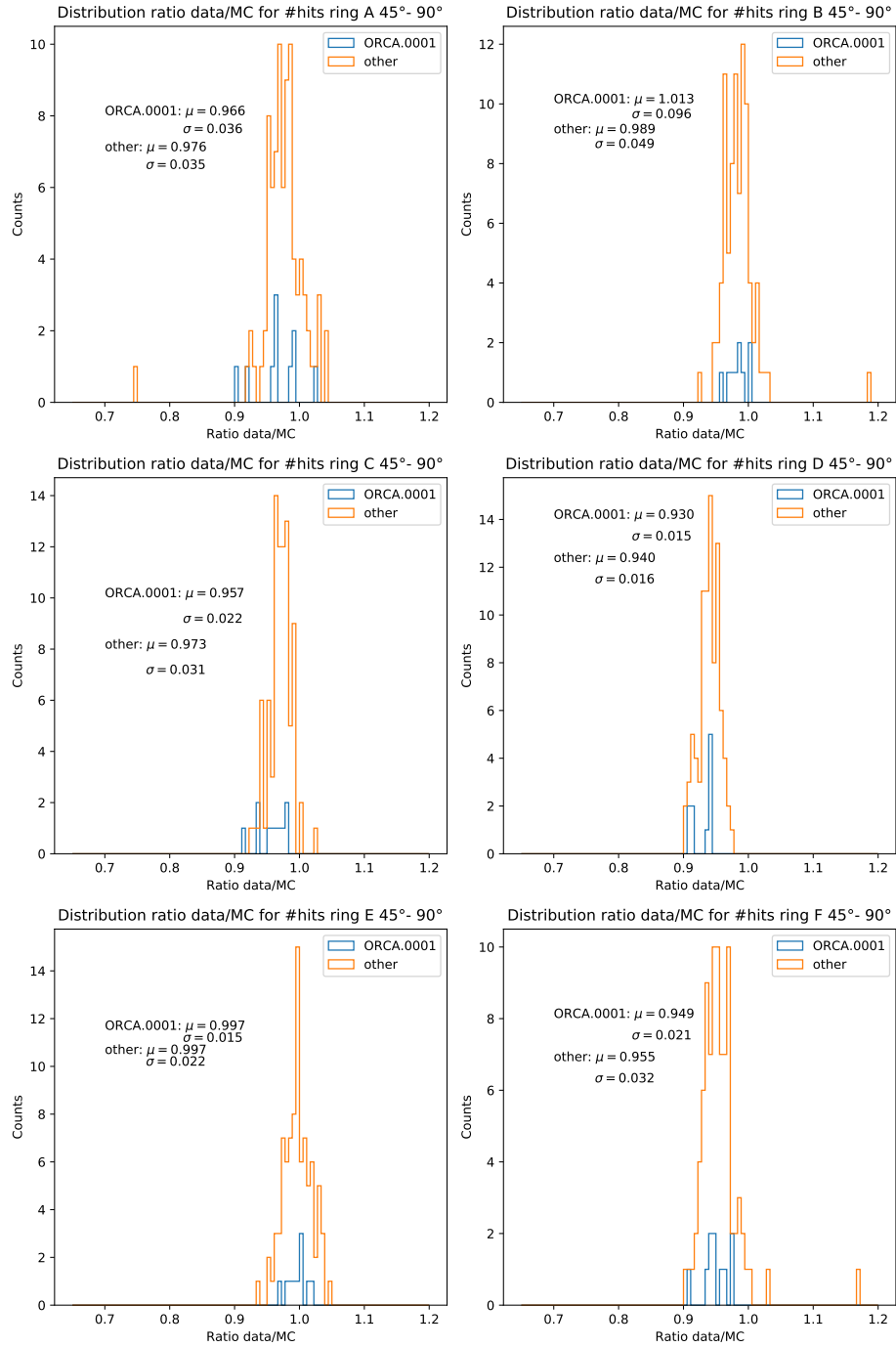


Figure 3.35: Distribution of ratio data/MC for 45°- 90°. The distributions for ORCA.0001 DOMs 7-18 (blue) and the other strings (orange) is shown per ring. The average values and standard deviations are presented in the graphs.

The strings excluding ORCA.0001 show an overestimation between 4.0% - 16.7% of light for 0° - 45° . The DOMs of ORCA.0001 which are not affected by the gel are overestimated by 3.7%-17.6%. Ring E shows the least overestimation for both ORCA.0001 and the other strings.

In the range 45° - 90° an overestimation of 0.3%-6.4% is found on average for the combination of ORCA.0002, ORCA.0003, ORCA.0009, ORCA.0010, and ORCA.0011. For ORCA.0001 DOMs 7-18 this varies from 0.3%-7.5% for rings A and C-F, while ring B is underestimated by 1.3%. The latter result is likely due to DOM 18, which shows a ratio data/MC of at least 1.1. This is probably a faulty DOM and is also visible in the 0° - 45° range. Since the average of ORCA.0001 is calculated for twelve DOMs and the other strings total 90 DOMs, the average of ORCA.0001 is more influenced by outliers than the other strings. The outliers are visible in the one-dimensional distributions below (fig. 3.34 and fig. 3.35).

Two things can be concluded from the above data. Firstly, the discrepancy in overestimation for ORCA.0001 DOMs 7-18 and the other strings is too small to call this a bias. Both inspected ranges of angles support this.

Secondly, the sedimentation on top of ORCA DOMs does not result in the same loss of efficiencies as for ARCA. Where for ARCA the ratios data/MC are lower for 45° - 90° than for 0° - 45° , the opposite happens for ORCA, including all DOMs of ORCA.0001.

Signals from muons coming sideways are less underestimated than signals from muons closer to zenith for ORCA, while for ARCA it is reversed. A possible explanation is that there is a difference in the efficiency estimations or propagation simulations for ARCA and ORCA. Another possibility is that this difference can be attributed to the sedimentation itself; that this is a different kind of sedimentation thus affecting the PMT efficiencies differently. The time a string has been in the water correlates directly with the amount of sedimentation. Although ARCA.0009 has been in the water longer than ORCA.0001 (since 2015 respectively 2017), this difference does not translate to the differences in ratios for the strings. Movements of the strings (for example by strong currents) decrease the amount of sedimentation. Still, this would probably not result in such differences for ARCA and ORCA. Perhaps the simulation or the different composition of the sedimentation accounts for the dissimilarity in biases.

Conclusion

This research aimed to identify and evaluate potential biases in the evaluated detection efficiencies of the KM3NeT ARCA and ORCA detectors. For that purpose two kinds of analyses were performed. One is the relationship between evaluated efficiency (from nearby photons from K40 decay) and the detection of isotropic light called single rates (from far away). The second is the comparison of simulated muon light (MC) and measured muon light (muon light identified via a selection in the time window of the time residuals). The muon light is not isotropic but comes from above.

Three potential biases were investigated: 1) The possible bias due to a change in the gel transparency of some DOMs (ARCA.0010 DOMs 7-15, all DOMs of ARCA.0012 and rings E and F of ARCA.0013). 2) Shadowing of rings D and E by the equator tape shadowing of PMTs C2, C5, E2, E5 by the titanium collar. A potentially enhanced dark rate was investigated as well. 3) Sedimentation on strings prevents light from hitting the covered PMTs. The radioactivity in the glass of the DOM is not considered in the efficiency evaluation, but might be of importance when there is less light from outside the DOM. Strings ARCA.0009 and ORCA.0001 have been studied for this effect.

For ARCA, a bias in the efficiency estimates has been found for ARCA.0010 DOMs 7-15, all DOMs of ARCA.0012 and rings E and F of ARCA.0013, which are the DOMs for which some difference in the transparency of the gel is suspected. The relation between single rates and evaluated efficiency confirmed this as the corrected rates of these PMTs were higher than for other PMTs. The comparison of simulated and measured muon light showed for ARCA.0010 and ARCA.0013 that the affected DOMs had a higher ratio data/MC than other DOMs.

The figures showing the ratio data/MC and their distributions resulted in an overestimation of the MC of 14% in ring C for the titanium belt, on average. The non-shadowed PMTs were overestimated by $\sim 11\%$. This effect was not found in ring E. For 0° - 45° the results for shadowed and non-shadowed PMTs followed those for the whole zenith angle. However, the ratio data/MC decreased for 45° - 90° . Together with the fact that the relationship between single rates and efficiencies did not change, this emphasises the angular dependence of the bias.

Thirdly, the sedimentation on top of ARCA.0009 affects the detection efficiencies, especially in rings E and F. The DOMs were overestimated by 14% (ring E) and 20% (ring F) for ARCA.0009, compared to 2.5% (ring E) and 3.5% (ring F) for the other strings. When the ratios were sorted by angle, those in the range 0° - 45° were almost exactly the same as seen for the data of the whole angular range. The difference in overestimation for ARCA.0009 and other strings decreased for 45° - 90° .

ORCA.0001 showed that DOMs 1-6 are less overestimated by MC than DOMs 7-18 of the same string for all six rings. Analysing the single rates of these DOMs could provide more information on the possibility that the transmission properties of the gel are wavelength dependent.

The shadowing of the titanium belt caused an overestimation of 12 percent for shadowed PMTs of ring C, compared to 8 percent for non-shadowed PMTs. In ring E the opposite happened: PMTs 2 and 5 were underestimated by 0-1% and PMTs 1, 3, 4, 6 were underestimated by $\sim 5\%$. This means that the effect of the equator tape and titanium belt together are estimated well, but not the equator tape by itself.

Ring D (shadowed by the equator tape) was overestimated by 11%.

The MC does not estimate the number of hits for ORCA.0001 differently than strings that have likely less sedimentation. Although it is not clear if and how the sedimentation on top of ORCA DOMs affects the efficiencies, one can conclude that the sedimentation on ORCA has a different effect than on ARCA.

Future studies can be performed on the subject of the sedimentation. It is evident that the performance of the ARCA detector is impacted by the sedimentation in some way. What is not yet clear is the reason why the sedimentation on top of ARCA DOMs has a different effect than the sedimentation on the ORCA DOMs. One could argue that the composition of the sedimentation may provide an answer to this question. The amount of sedimentation could be monitored during a certain time period to investigate after what time duration in the water the effect of the

sedimentation is too substantial and the radioactivity cannot be neglected anymore. Whether the sedimentation reaches a certain level that will not be exceeded can be researched, as well.

The effect of the sedimentation, but also that of the change in gel-transparency and the shadowing and enhanced dark rate should be taken into account for the analysis of the data that ORCA and ARCA will eventually take. This will improve the efficiency of the detectors, and, in time, contribute to answering the most burning questions about neutrinos.

Acknowledgements

I would like to thank Dorothea Samtleben for her support and insights for this project. Special thanks to Daniel Guderian and the KM3NeT group at Nikhef.

Appendix

A.1 ARCA runs 9791-9796

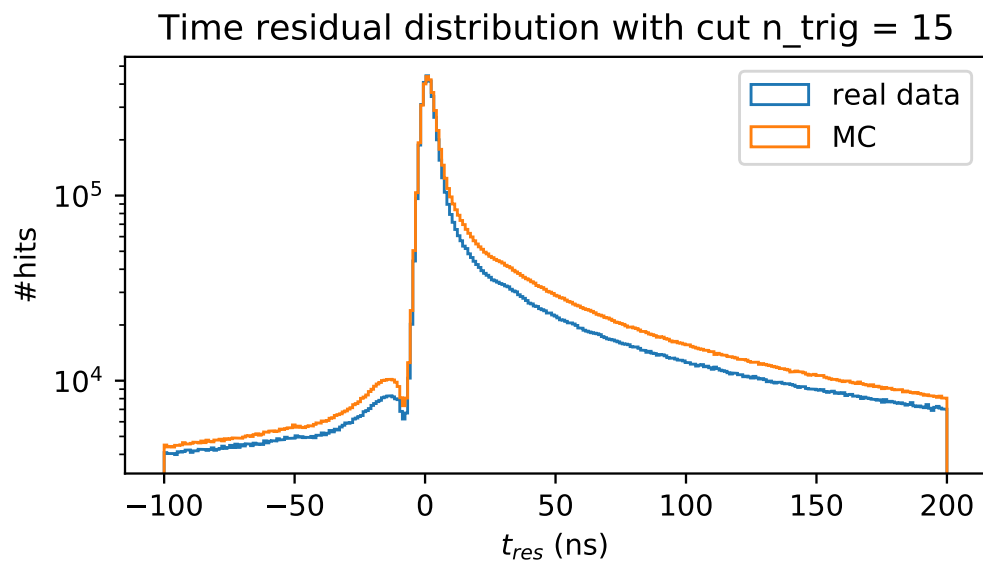


Figure A.1: The time residual distribution for the real data (blue) and the Monte Carlo simulation (orange). The MC is weighted by $\frac{lifetime_{data}}{lifetime_{MC}} = 0.496$ to correct for the difference in lifetimes of real data and MC.

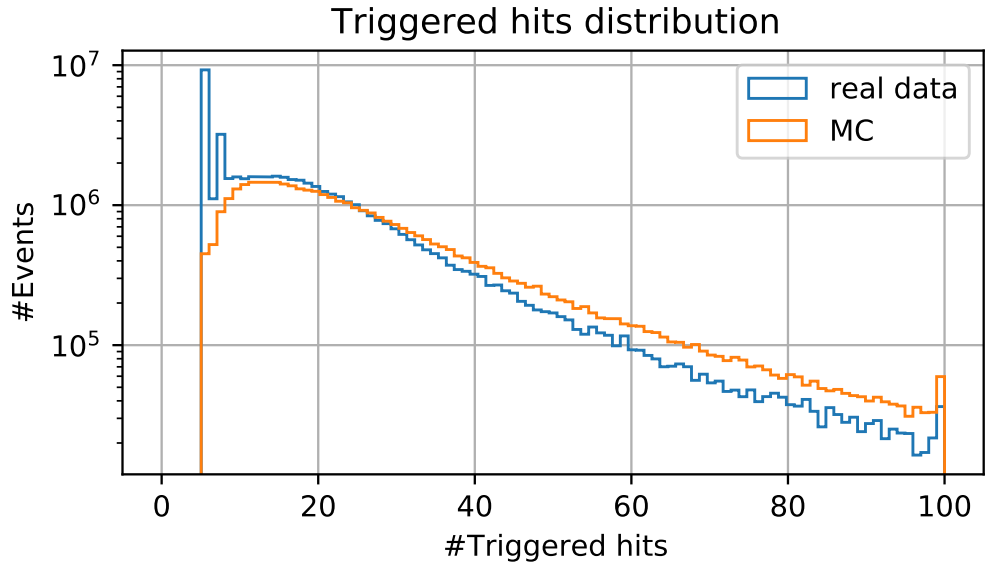


Figure A.2: The distribution of the number of triggered hits for real data and MC. The data and MC differ in the region where the number of triggered hits is less than 15. Therefore, a cut of 15 triggered hits is applied to the time residual plots to exclude the events that have no muon track. Excluding these events gives a better agreement between real data and MC.

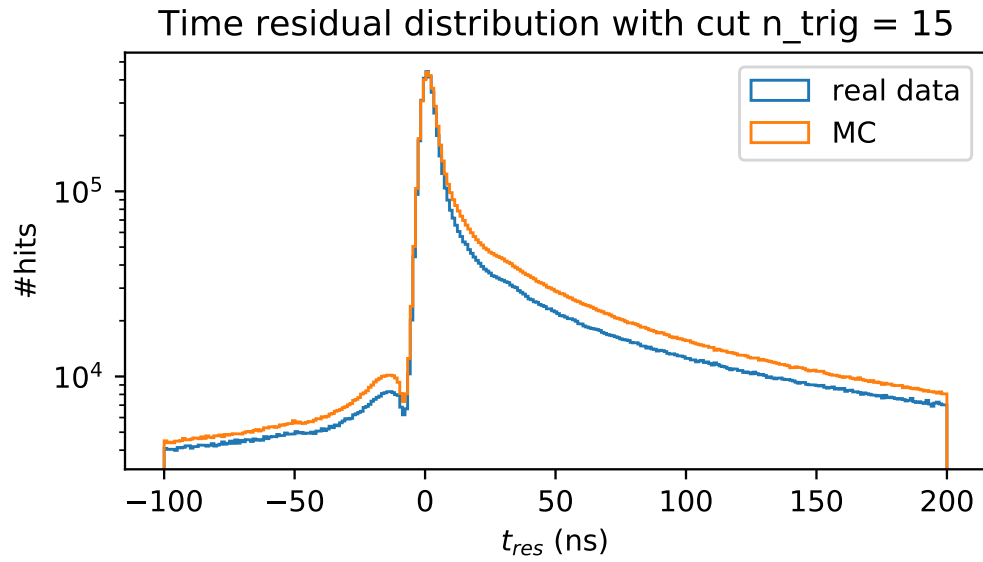


Figure A.3: Time residual distribution after the cut on 15 triggered hits is applied. The shape of the real data and MC distribution are now more alike than before the cut.

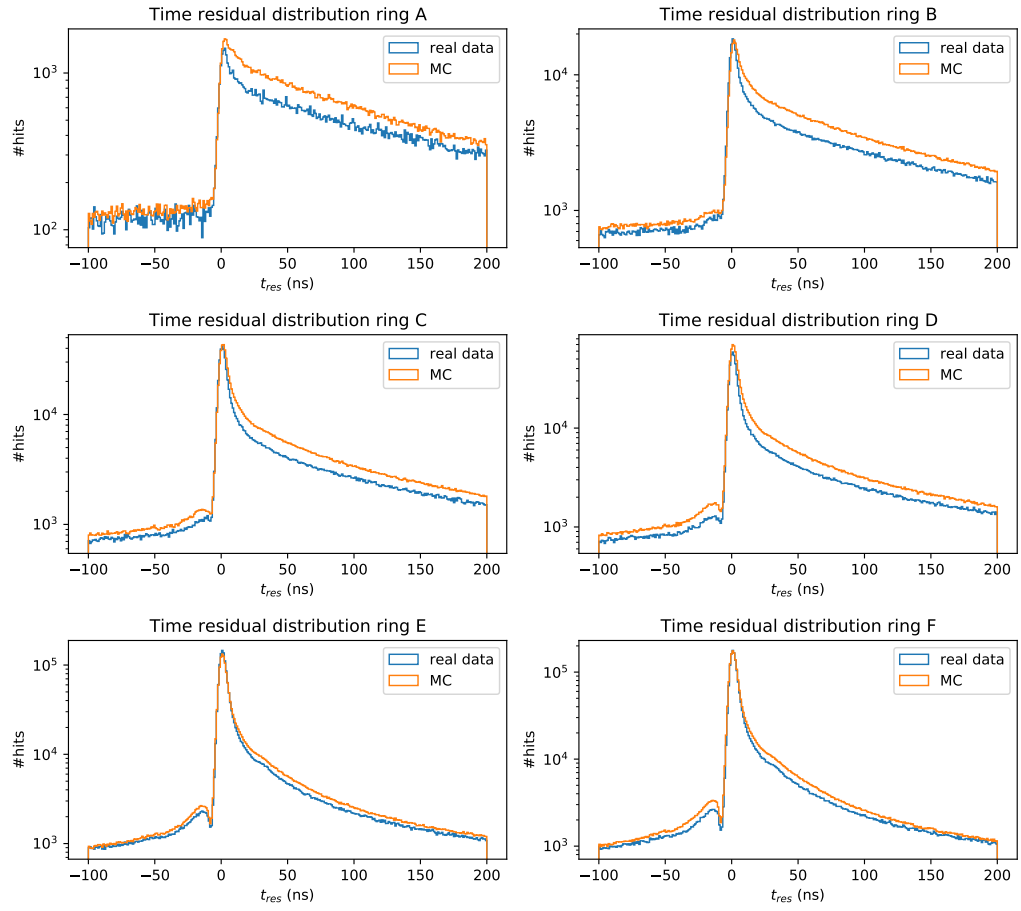


Figure A.4: Time residual distributions sorted by ring.

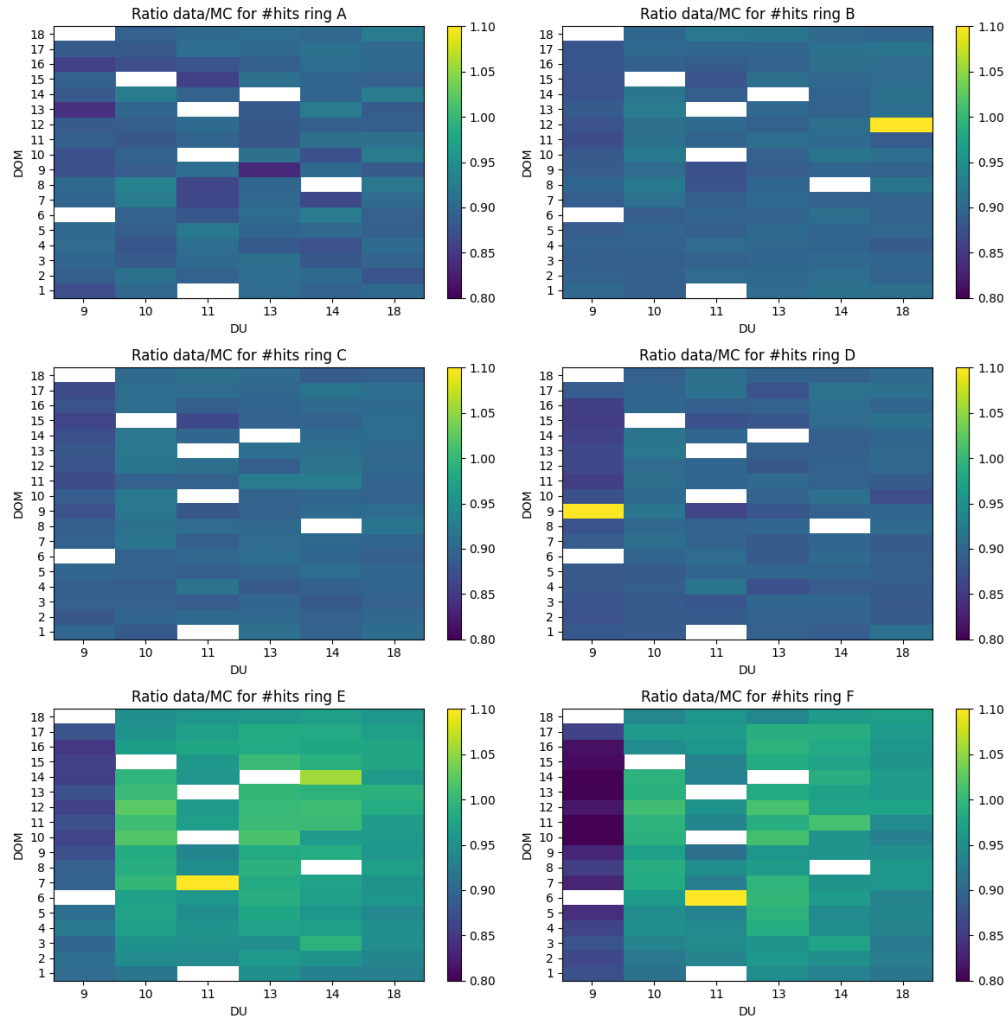


Figure A.5: The ratio data/MC sorted per ring. Inoperative DOMs are shown as white rectangles.

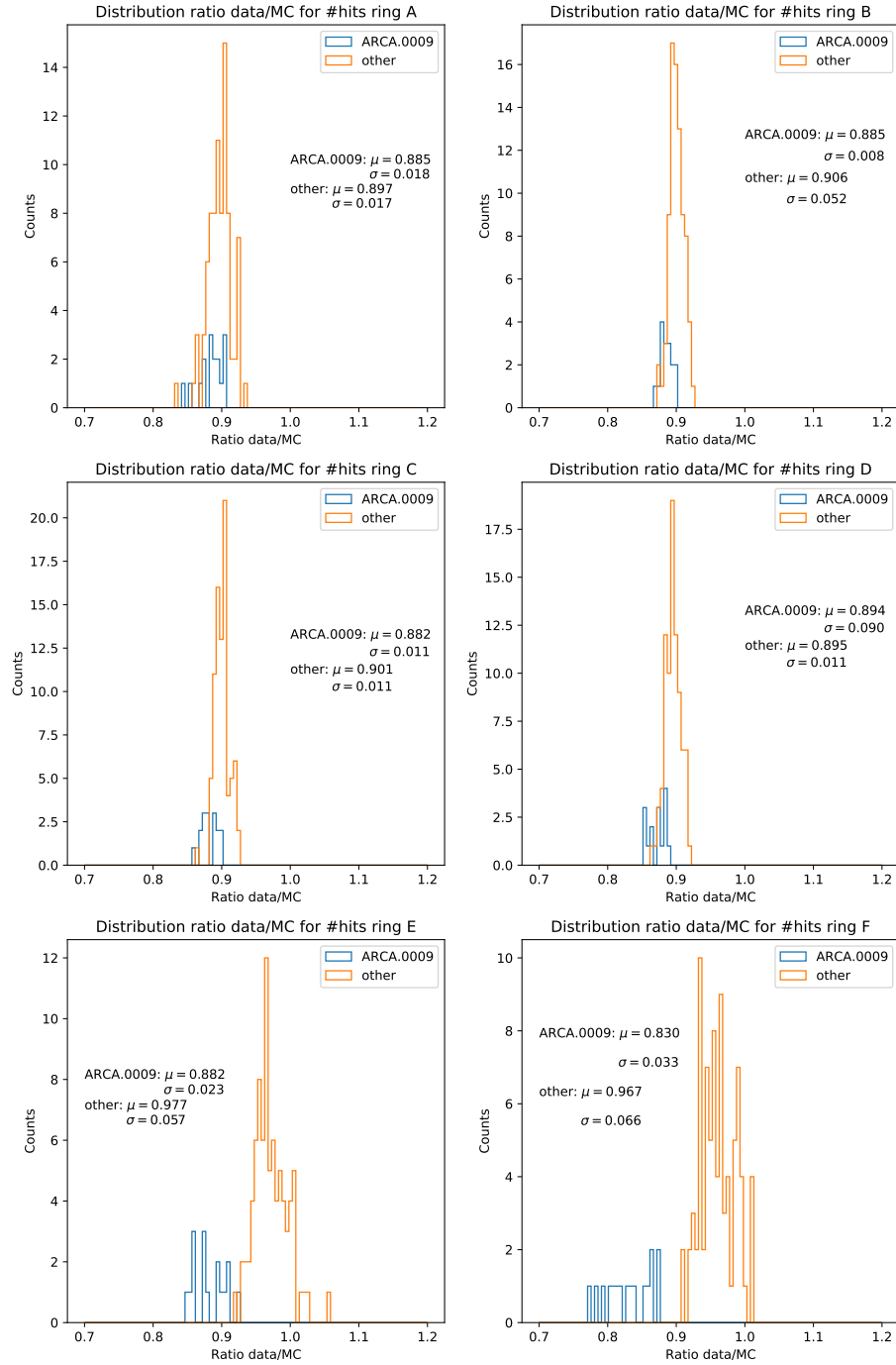


Figure A.6: Distribution for ratio data/MC. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

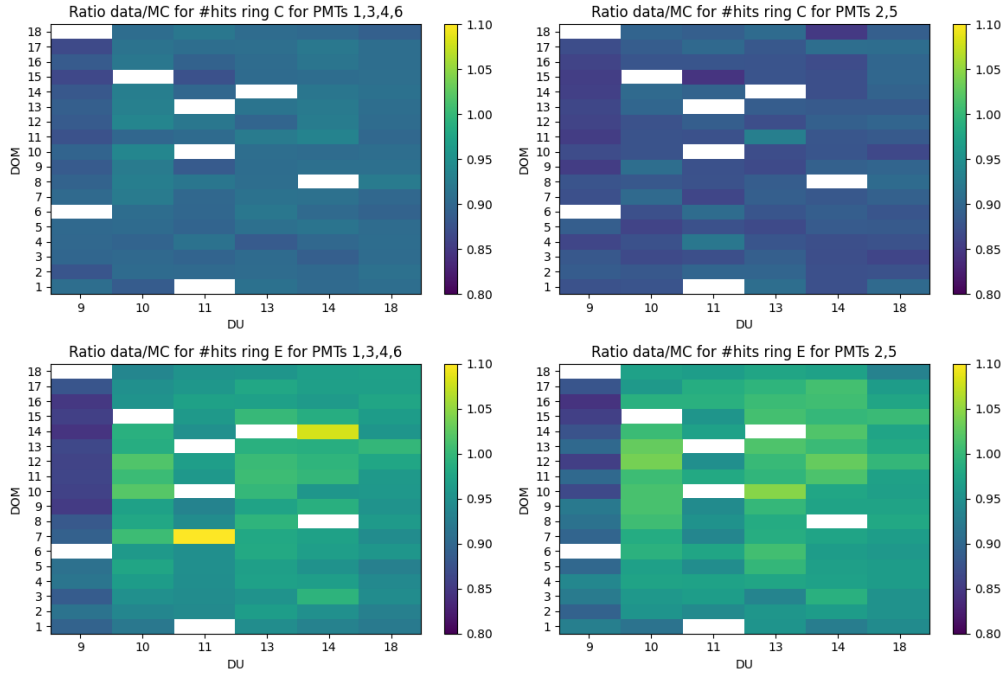


Figure A.7: The ratio data/MC compared for the shadowed (right) and non-shadowed (left) PMTs of rings C and E. In addition to the shadowing of the titanium belt, all PMTs of ring E are shadowed by the equator tape. Inoperative DOMs are shown as white rectangles.

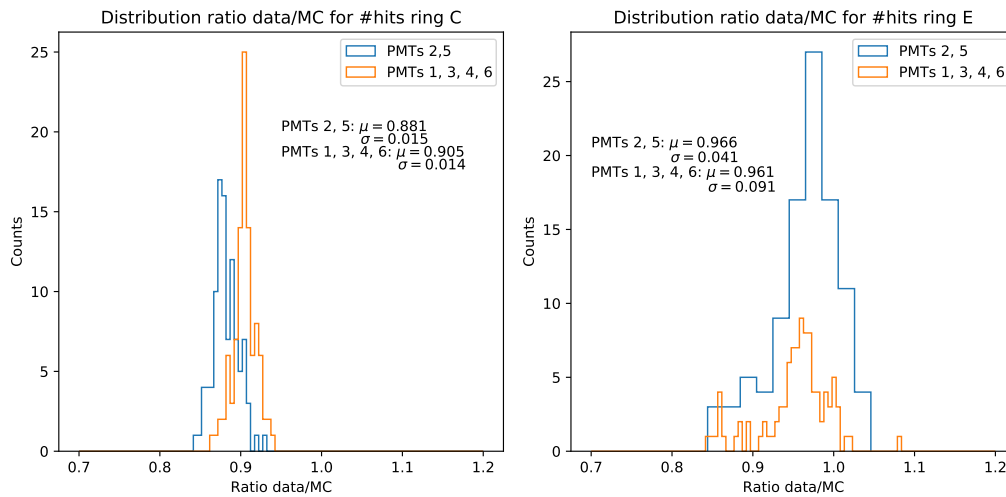


Figure A.8: The distribution of ratio data/MC compared for the shadowed (blue) and non-shadowed (orange) PMTs of rings C and E. The average values and standard deviations are provided in the plots.

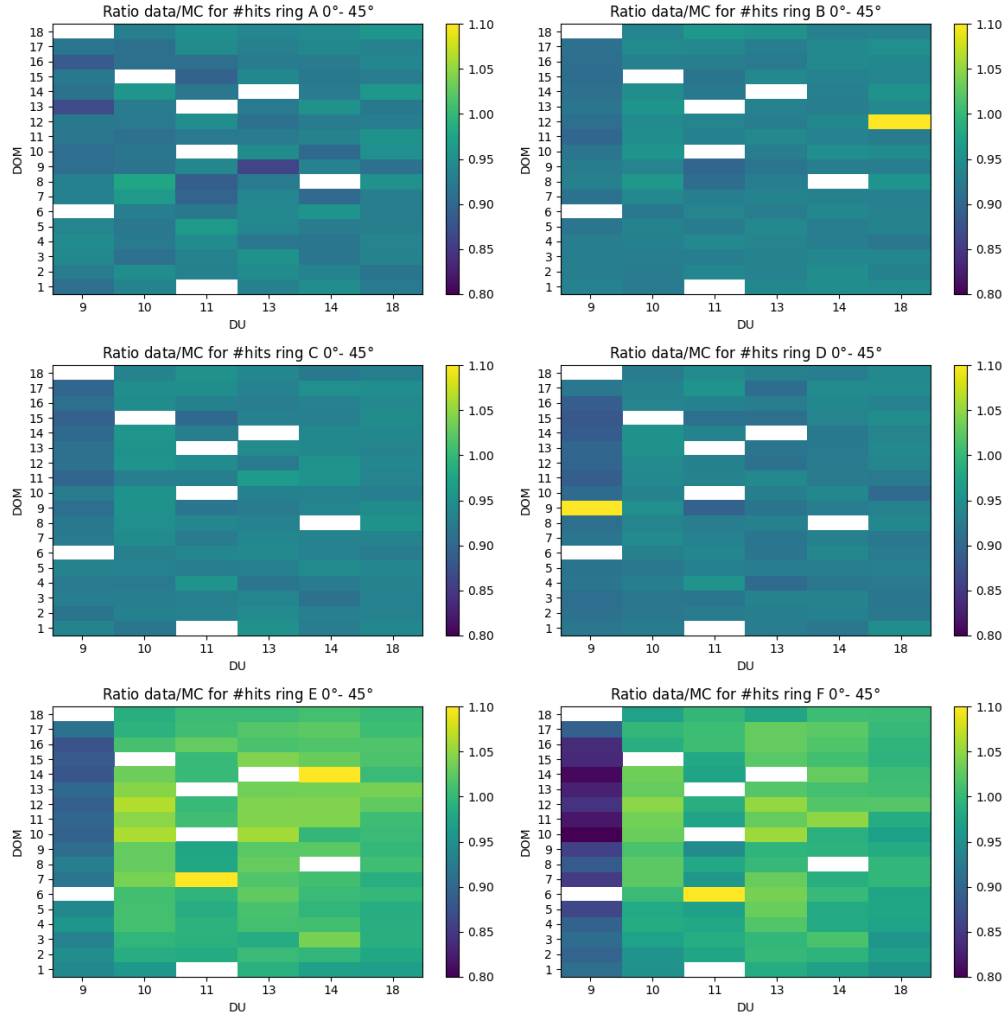


Figure A.9: Ratio data/MC for $0^\circ - 45^\circ$ per ring. The x-axis shows the strings and the DOMs are on the y-axis. Inoperative DOMs are shown as white rectangles.

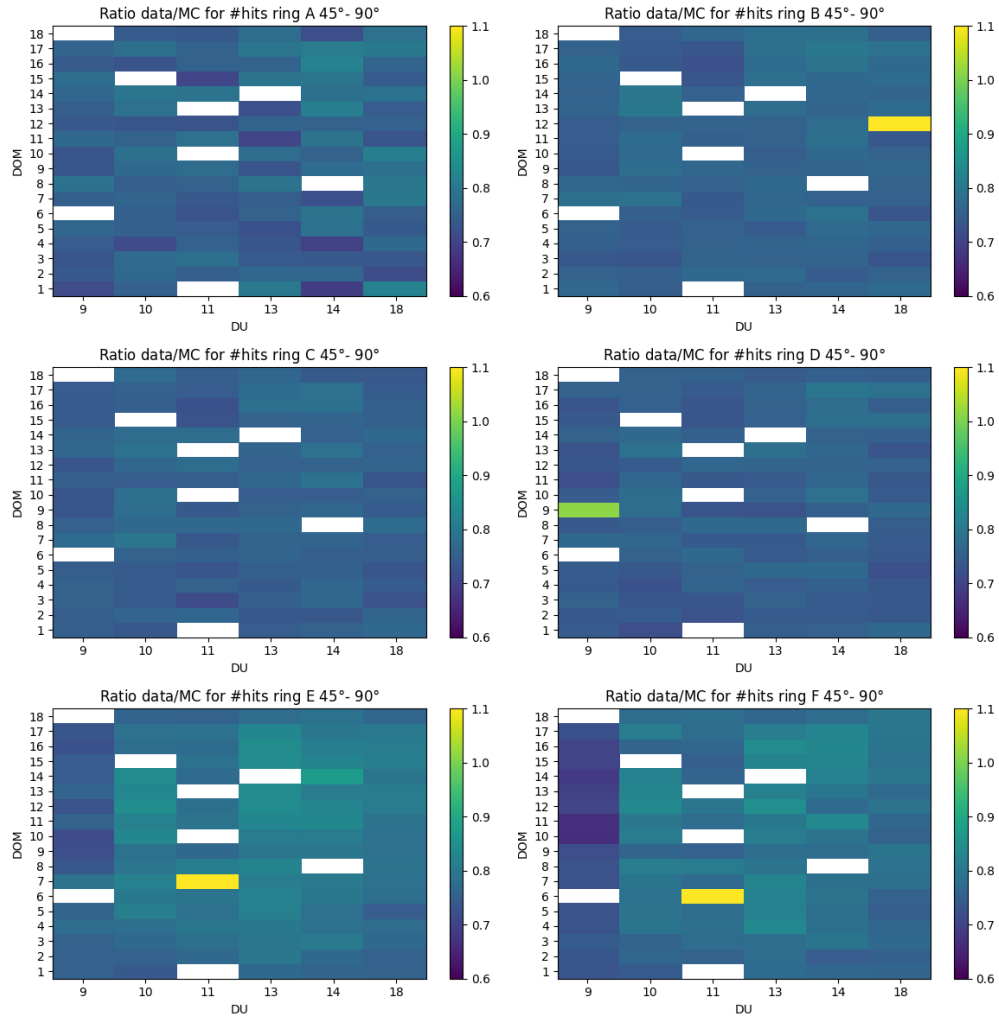


Figure A.10: Ratio data/MC for 45° - 90° per ring. The x-axis shows the strings and the DOMs are on the y-axis. Inoperative DOMs are shown as white rectangles.

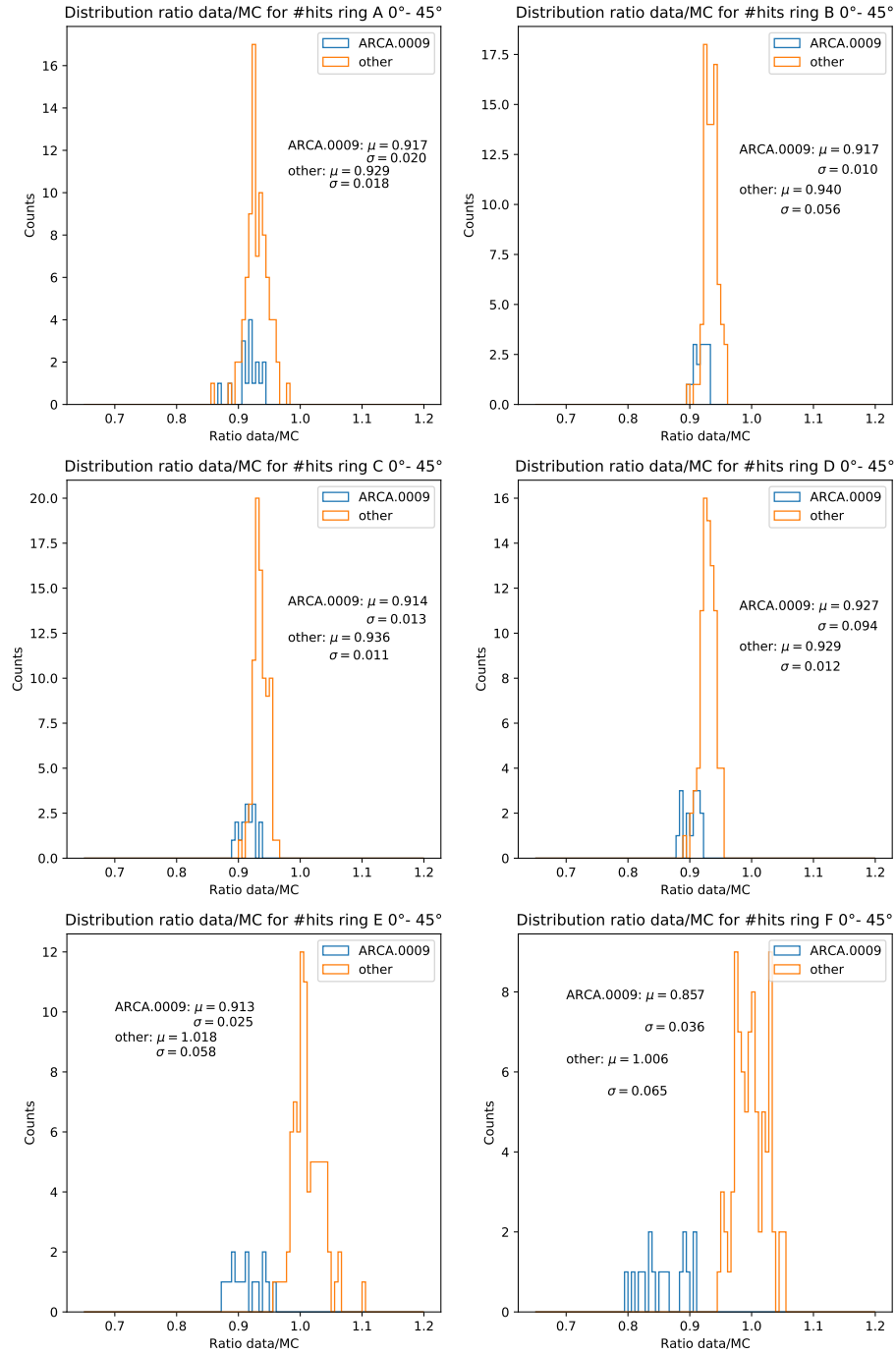


Figure A.11: Distribution for ratio data/MC for 0°- 45°. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

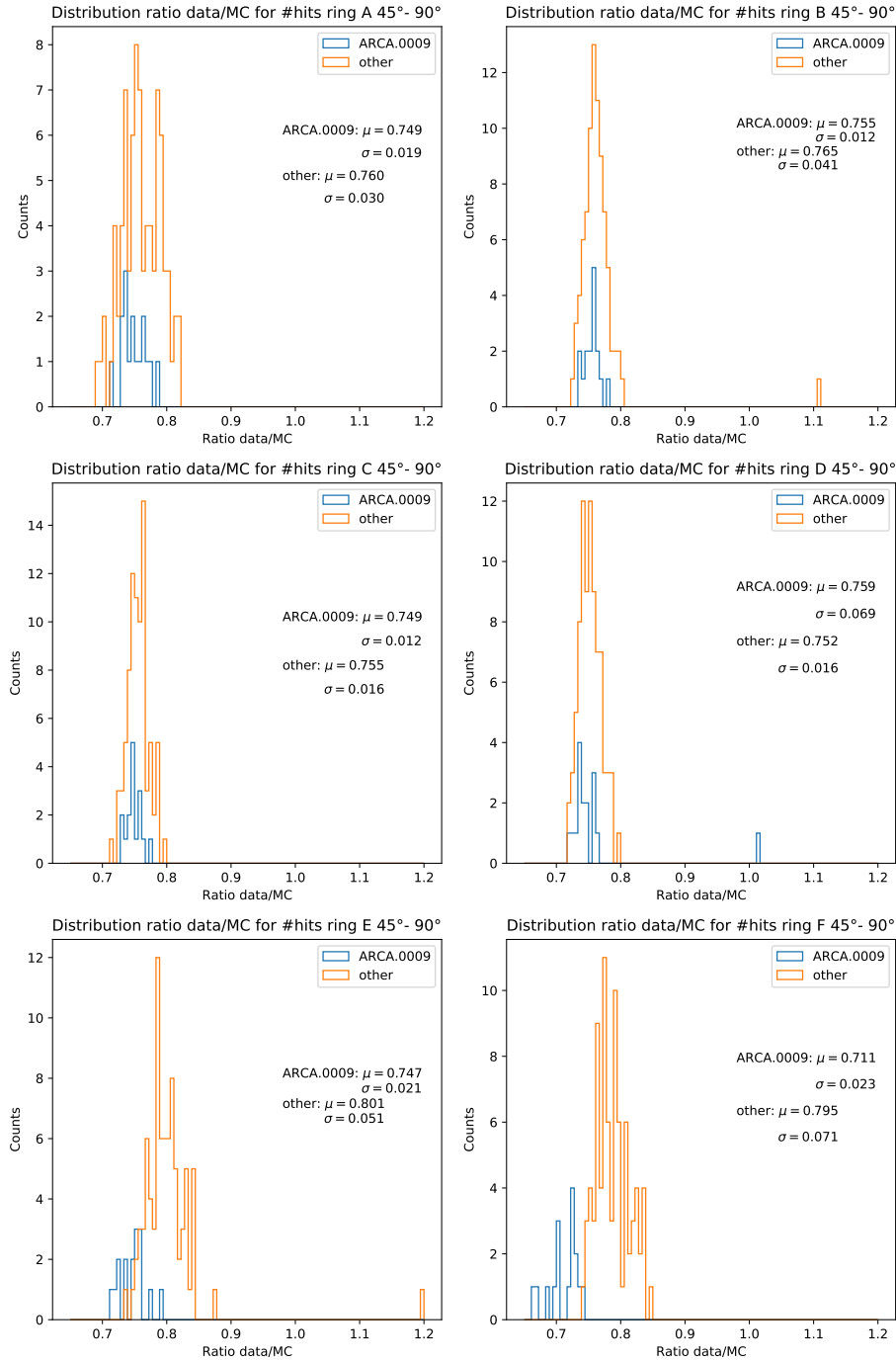


Figure A.12: Distribution for ratio data/MC for 45°- 90°. The distributions for ARCA.0009 (blue) and the other strings (orange) are shown per string. The average values and standard deviations are presented in the graphs.

A.2 ORCA

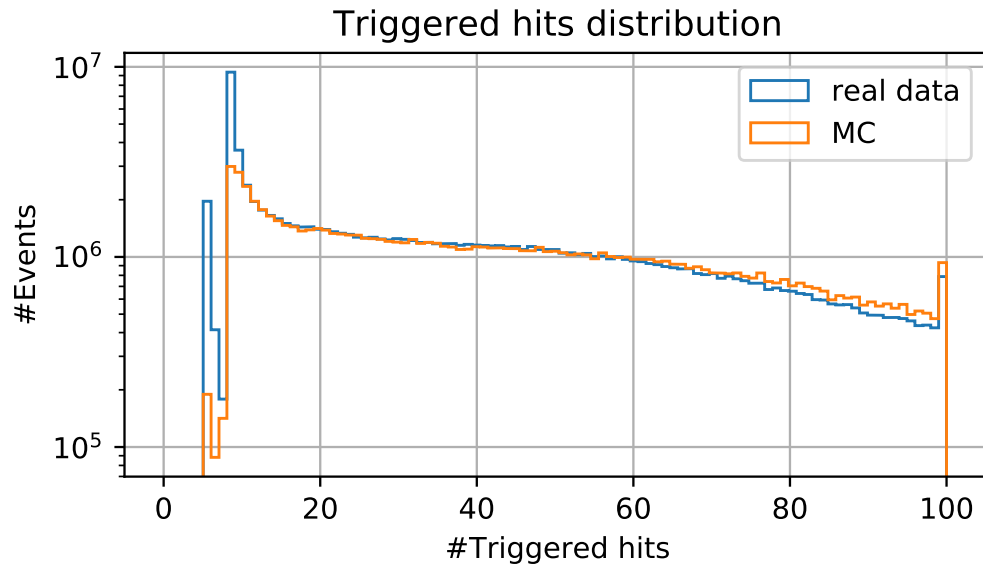


Figure A.13: The extended version of the distribution of the number of triggered hits for real data and MC. A cut of 11 triggered hits is applied to exclude the events that have no muon track.

Bibliography

- [1] A. Mann, *What are neutrinos?*, <https://www.livescience.com/64827-neutrinos.html>, 2019, Accessed 8 June 2022.
- [2] L. Bergström and A. Goobar, *Cosmology and Particle Astrophysics*, Springer, second edition, 2006.
- [3] The KATRIN Collaboration, *Direct neutrino-mass measurement with sub-electronvolt sensitivity.*, Nat. Phys. **18**, 160 (2022).
- [4] Borexino Collaboration, *Observation of Geo-Neutrinos*, Phys.Lett.B **687** (2010).
- [5] KM3NeT, *Particle physics with ORCA*, <https://www.km3net.org/research/research-infrastructure/km3net-fr-site/>, Accessed 9 June 2022.
- [6] KM3NeT, *Astroparticle physics with ARCA*, <https://www.km3net.org/research/physics/astronomy-with-arca/>, Accessed 9 June 2022.
- [7] R. Bruijn and D. van Eijk, *The KM3NeT Multi-PMT Digital Optical Module*, page 1157, ICRC2015, 2015.
- [8] KM3NeT, <https://www.km3net.org/research/detector/sensors/>, Accessed 8 June 2022.
- [9] P. J. Kiger, <https://science.howstuffworks.com/muon.html>, 2021, Accessed 7 June 2022.
- [10] K. W. Melis, *Studying the Universe from -3000m N.A.P.*, PhD thesis, Universiteit van Amsterdam, 2021.

- [11] KM3NeT, *Artists impression of a Cherenkov neutrino telescope*, <https://www.km3net.org/wp-content/uploads/2021/12/Cherenkov-detector-edw.png>, 2021, Accessed 8 June 2022.
- [12] J. Carr, G. Hallewell, A. Holford, U. Katz, P. Kooijman, L. Moscoso, P. Piattelli, and P. Rapidis, *Conceptual Design for a Deep-Sea Research Infrastructure Incorporating a Very Large Volume Neutrino Telescope in the Mediterranean Sea*, (2015).
- [13] S. Aiello et al., *Characterisation of the Hamamatsu photomultipliers for the KM3NeT Neutrino Telescope*, JINST **13** (2018).
- [14] KM3NeT Collaboration et al., *Dependence of atmospheric muon flux on seawater depth measured with the first KM3NeT detection units The KM3NeT Collaboration*, European Physical Journal C **80** (2020).
- [15] A. de Benedittis, *Overview of efficiency/gain D0ARCA009 & D0ORCA010*, 2022.
- [16] J. Höfestad, D. Samtleben, *Report on problems with data from ORCA-DU2*, 2019.
- [17] Calibration group, *PMT obscuration observed in the sea*, Technical report, KM3NeT, 2019.
- [18] Sippens Groenewegen, Lieke, *Characterisation of the KM3NeT Detector Efficiencies Using Atmospheric Muon Measurements*, 2021, Bachelor thesis.
- [19] S. Aiello et al., *Architecture and performance of the KM3NeT front-end firmware*, J. Astron. Telesc. Instrum. Syst. **7** (2021).
- [20] K. Melis, A. Heijboer, and M. de Jong, *KM3NeT/ARCA Event Reconstruction Algorithms*, ICRC2017, 950 (2017).