Simulation studies for the calibration of the IWCD neutrino detector using an isotropic light source

Alie Craplet - October 2020

***Abstract*—This paper presents a calibration technique for the Intermediary Water Cherenkov neutrino detector using an isotropic light source located within the tank. Using the average response of all the PMTs over multiple events as a reference this analysis showed that PMTs can be calibrated with an accuracy of 0.096ns when all the PMTs' recorded times were offset by some random smearing values. In this same configuration any constant offset to the PMTs hit time can also be retrieved with an accuracy of 0.05ns and an offset of -0.12ns. It was shown that the accuracy of the calibration depends on the available statistics and that it was less accurate if only a fraction of the PMTs were smeared. This paper also presents the application of the same analysis to the Water Cherenkov Test Experiment and the issues that arise.**

I. Introduction

Water Cherenkov detectors are high-precision relatively low-cost experiments used notably in the field of Neutrino Physics. The Super-Kamiokande experiment is the one of most famous water Cherenkov detectors, these are composed of a very large tank of water surrounded by photomultipler tubes (PMTs). When a neutrino collides with a water molecule, a very energetic electron or muon is produced, depending on the neutrino's flavour. Those charged particles travel faster than light does in water which means that they emit Cherenkov radiations [1] as presented in Figure 1 below.

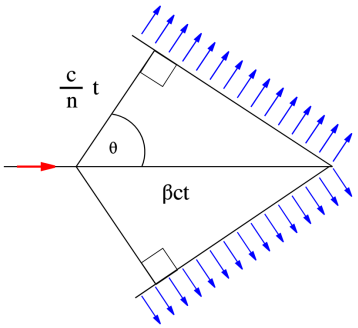

Fig. 1: Diagram of the Cherenkov radiation (in blue) emitted by a charged particle (red) travelling with a velocity $\beta c > \frac{c}{n}$ i.e. faster than light in this medium of refractive index n. Picture taken from [2].

The shape and size of the Cherenkov rings depend on the type and velocity of the charge particles and their detections by the PMTs allows a precise reconstruction of the position at which the neutrino collided with the water molecule.

Using this technique, the Super-K collaboration has shown with a 2 sigma confidence level that neutrinos violate CP-symmetry, which means that they could - at least partly - explain the matter-anti-matter asymmetry that is observed in the Universe [3].

In large experiments like the Water Cherenkov detectors, it is of upmost importance that each PMT is accurately located, oriented and connected to the rest of the experiment. To ensure this, the PMTs are regularly checked and calibrated. This paper presents how to use an isotropic light source, placed at different positions within the water tank, to calibrate the detector's PMTs. In this study we use datasets simulated for the Intermediate Water Cherenkov Detector (IWCD) to investigate the feasibility and accuracy of this calibration technique. As discussed in the last sections it is thought that the analysis presented here for IWCD can easily be adapted to other water Cherenkov geometries.

II. Method

A. Simulated datasets and initial correction.

The IWCD experiment was simulated using the Geant4 and WCSim softwares. The light source was simulated as a point emitting instantaneously an isotopic burst containing 10,000 photons. Obviously, this idealised light source is only a working approximation and further work will replace it by a more realistic approximation.

In order to meaningfully compare situations where the source is located at different positions in space, the Time Of Flight (TOF) of each photon was removed from the recorded hit time. The TOF of a photon is calculated as the time taken to travel from the source to the detecting PMT if the photon travels at a constant velocity of $2.234 \times 10^8 \mathrm{ms}^{-1}$. In a reflection-less, noiseless and scattering-free environment one should measure that all TOF-corrected recorded photon hit times are equal to zero. The results obtained in this study are presented in Figure 2.

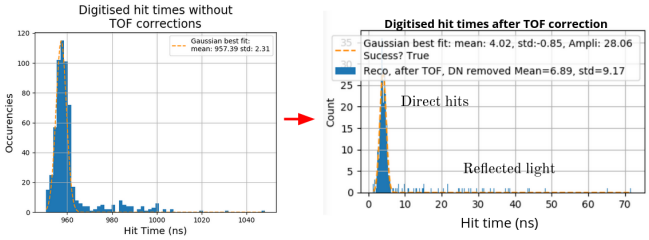

Fig. 2: Digitised hit times before (left) and after TOF correction and offset removal (right). The Gaussian fits indicate that TOF corrections reduces the spread in hit times from 2.31ns to 0.85ns. The figure also highlights the scattered and reflected photons that are not included in the fit.

B. A Gaussian smearing

The random noise that the PMTs experience was then added to the recorded hit times. These random PMT-dependant offsets were drawn from a Gaussian distribution of known mean and standard deviation and then added to each PMT's recorded hit times over every one of the 1,000 iterations of the simulation. These random offsets simulate a fault in the wiring, a misplacement of that PMT or any other kind of experimental noise. The user could choose which fraction of PMTs was smeared. If a fraction below 1 was chosen then the choice of which PMT should be smeared was made randomly. An example of the time distribution after smearing of the hit times is shown in Figure 3.

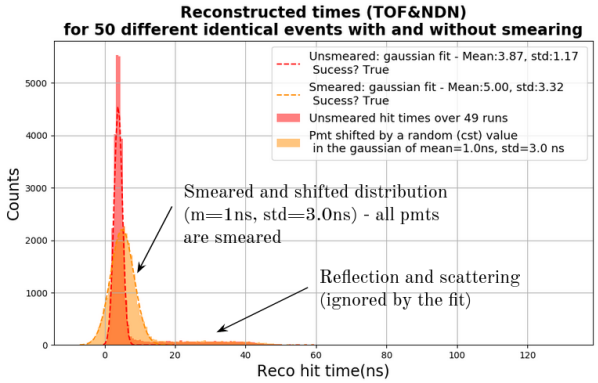

Fig. 3: Digitised hit time distribution recorded by all PMTs over 50 iterations of the same experiment. In red, before any smearing and in yellow after all the PMTs have been smeared by random offsets taken from a Gaussian of mean 1.0ns and standard deviation 3.0ns. The dotted red and yellow line indicate the Gaussian fit to the time distributions before and after smearing respectively.

C. The Calibration Process

A Gaussian distribution is fitted to the smeared hit times recorded by all the PMTs over all of the simulated events, as shown in yellow in Figure 3. The mean of this Gaussian fit is called TotalMean. It is chosen a reference for the calibration. The idea is then to highlight the PMTs whose average hit times are too far from the global mean as having probably been smeared. Only the PMTs which have recorded at least 100 hits out of the 1,000 events simulated can be accurately calibrated. The Gaussian fit to these PMTs hit times are shown in Figure 4 below.

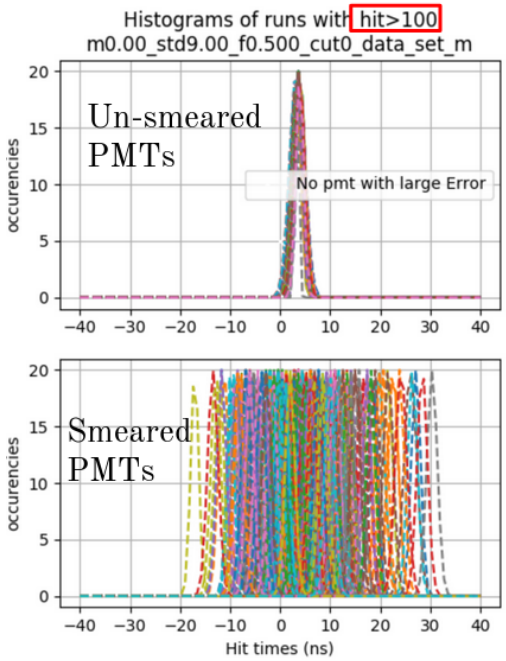

Fig. 4: Top: Gaussian fits to the un-smeared PMTs' hit times distribution. The spread of each PMTs recorded hit times is approximately constant and in this case where the PMTs have not been smeared, their mean hit times are all very close to the TotalMean value. Bottom: Gaussian fits to the smeared PMTs' recorded hit times. The smearing values are taken from a Gaussian of mean 0ns and standard deviation 9ns.

The non-smeared PMTs (top plot in Figure 4) all have means similar to TotalMean. On the bottom plot of this figure the PMTs' mean hit times are more spread, reflecting the smearing offset that has been added to each one of them. The time separation between the TotalMean reference and each PMT's mean hit time is called DeltaMean. For each PMT, DeltaMean was then compared the the known smearing value as presented in Figures 5 and 6.

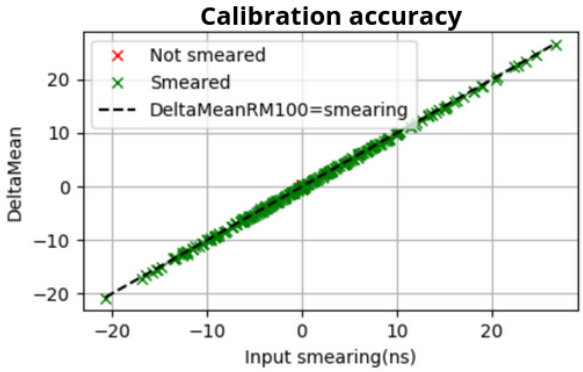

Fig. 5: Comparison, for each PMT, of DeltaMean against the smearing value showing that DeltaMean roughly equals the PMT's smearing value. The y=x line is shown for reference.

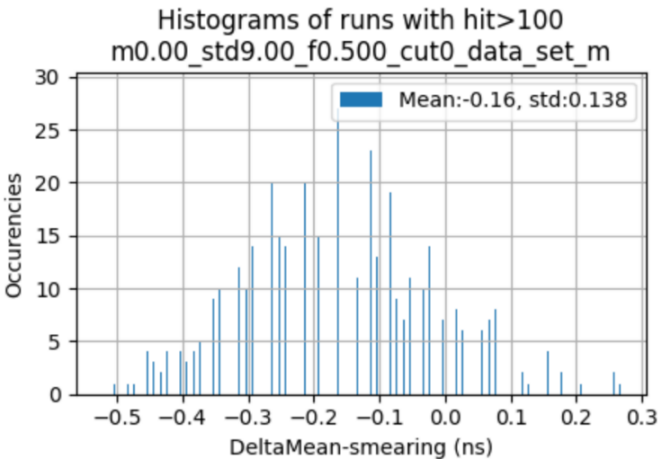

Fig. 6: Histogram of the difference between the PMT's smearing offset and their DetlaMean value showing that the PMTs can be calibrated with an accuracy of 0.138ns with an offset of -0.16ns.

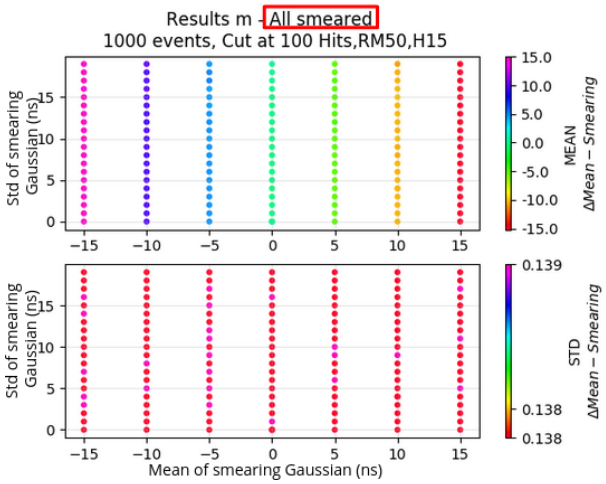

Fig. 7: Mean (top) and standard deviation (bottom) of the DM-SV distribution over all of the PMTs which recorded at least 100 hits out of the 1,000 simulated events. The mean and standard deviation of the smearing Gaussian are indicated in x and y respectively. In this case all the PMTs are smeared by a random offset drawn from the smearing Gaussian.

In this section we have shown that when working with 1,000 simulations of the same event, one could calibrate all the PMTs which recorded at least 100 events. If the smearing parameters for the PMTs are drawn from a Gaussian of mean 0ns and standard deviation 9ns then the PMTs can be calibrated back to their true value by removing DeltaMean to all of their hit times. We will see in the next section how the calibration quality varies when different smearing parameters are chosen, when more statistics are used and finally what happens when only a fraction of the PMTs is smeared.

III. Results and Discussion

A. All PMTs smeared by Gaussians of different parameters

The quality of the calibration for each smearing configuration was drawn from the distribution of the difference between DeltaMean and the corresponding PMT's smearing value, as was presented in Figure 6. A Gaussian distribution was then fitted to the DeltaMean-Smearing Value (DM-SV) distribution. The mean and standard deviations of the Gaussian fit to the DM-SV distributions are presented as a function of the smearing parameters in Figure 7. This figure highlights two very interesting results of the calibration process. Firstly one can see that the standard deviation of the DM-SV distribution is independent of the input smearing parameters with a value of 0.138 ± 0.001ns. Secondly, one can see that the mean of the DM-SV distribution seems to be equal to the negative of the mean of the input smearing value. Figure 8 presents the distribution of the difference between those two values (i.e. the difference between the z and x coordinates in Figure 7 (top)).

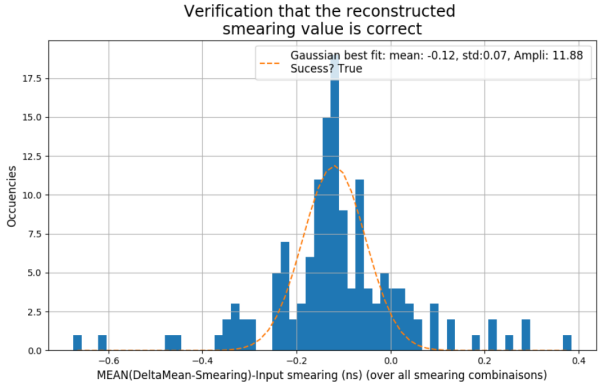

Fig. 8: Histogram of the difference between the mean of the smearing Gaussian and the mean of the DM-SV distribution over all of the smearing configuration considered. The dotted line presents the Gaussian fit to this data which has a mean of -0.12ns and standard deviation of 0.07ns.

The 1:-1 almost perfect correspondence between the mean of the smearing Gaussian and the mean of the DM-SV distribution is due to the fact that if all the pmts are shifted by a Gaussian of mean non-zero then the total time distribution, TotalMean, is also shifted. Since the t=0s point is arbitrary in this study this artefact of the calibration is too relevant. Indeed all the hit times can simply be shifted by the same value back in time.

B. An issue when only a fraction of the PMTs are smeared

Special care has to be taken when only a fraction of PMTs are smeared by a Gaussian with a non-zero mean or if different portions of the PMT population are smeared by Gaussians of different means. In that case two or more population arise, with significantly different hit times, as is shown in Figure 9.

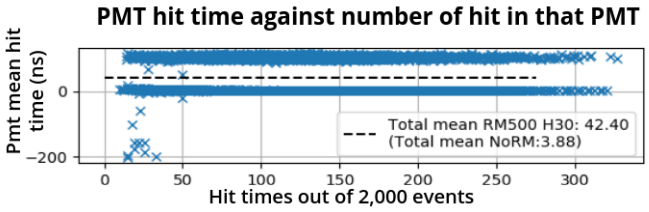

Fig. 9: Mean of the Gaussian fit to the hit times of each PMTs as a function of how many hit it records out of 2,000 simulated events. Half of the PMTs are smeared by a Gaussian of mean 100ns and standard deviation 4ns, the other half is not smeared. The TotalMean parameter is indicated in dotted black. It is interesting to note that the Gaussian fit fails for PMTs which recorded very few hits which explains the cuts applied previously.

Clearly the two PMTs populations presented in Figure 9 (one shifted in time compared the other) have different mean hit times and the TotalMean value is not a meaningfull calibration reference anymore. Indeed, using this TotalMean value as a reference means calibrating all the PMTs to this intermediary mean value rather than to to the proper t_0. In this case, the mean of the DM-SV distribution does not reflect the input Mean smearing anymore, as can be seen on Figure 10 below.

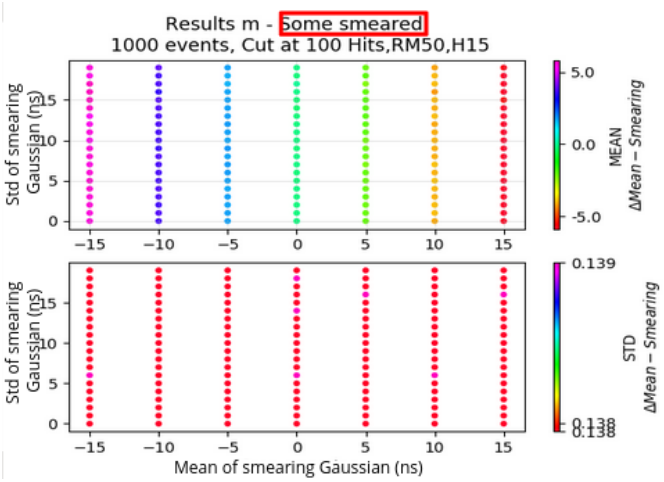

Fig. 10: Results of the calibration when only half of the PMTs are smeared. It shows that the mean of the DM-SV distribution (top) is not the opposite of the mean of the the smearing Gaussian, unlike when all of the PMTs are smeared.

It is thought that this issue with TotalMean could easily be fixed by scanning through the mean recorded hit time of all PMTs and splitting them into subgroups of similar recorded hit times. The idea of coding a Neural Network to make this choice has been put forward. This issue could be encountered in real-world situation, take the example a damaged local clock. All of the PMTs dependent on this clock will record a significantly later time but not the others. The calibration would be more accurate if all of these PMTs were first flagged as having an offset, then if this offset was removed, shifting these PMTs recorded times back to t_0 before finally calibrating them with the other PMTs.

C. Investigating how the amount of statistics available impacts the quality of the calibration

We next investigated how the calibration's quality would improve when working with a larger dataset. Figures 4 to 8 presented the calibration results for 1,000 repetitions of the same event with the source located at the centre of the detector. In this case only the PMTs recording at least 100 hits were kept for calibration. Figure 11 presents the calibration quality when all PMTs are smeared, 2,000 events are simulated and only the PMTs recording at least 200 events are kept for analysis.

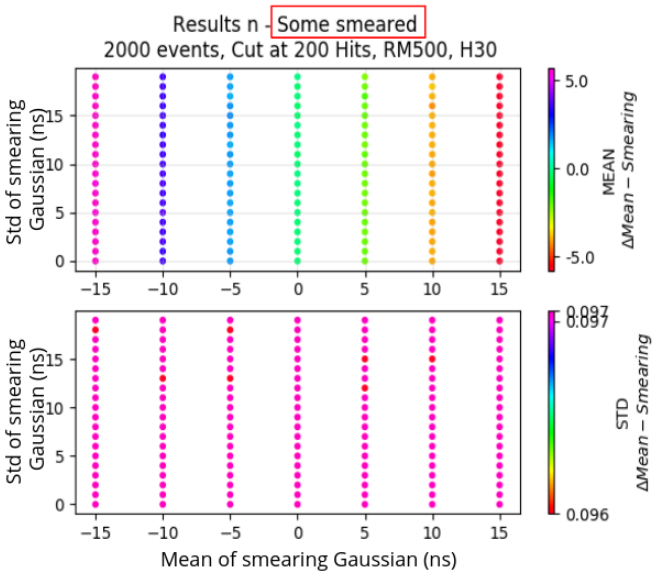

Fig. 11: Mean (top) and standard deviation (bottom) of the DM-SV distribution over 2,000 iterations of the same event, keeping only the PMTs having recorded at least 200 hits. In this case all PMTs are smeared by Gaussians of different parameters.

The result is very similar to the one obtained with 1,000 iterations of the simulation. The standard deviation of the DM-SV distribution is again constant but this time equal to 0.097ns which is 0.703 times the value it was when working with half the sample size. This seems to indicate that the calibration's quality is correlated with the size of the statistical population by a $1/\sqrt{n}$ dependency ($1/\sqrt{2} = 0.707$). More research on larger statistical populations are still required to validate this hypothesis. In this case again, the mean of the DM-SV distribution reflects the mean value of the smearing Gaussian, the difference between these two variables presented in Figure 12.

Here again the 2,000 events simulation presents a better accuracy than the 1,000 events one by a factor of $0.71 \approx 1/\sqrt{2}$. It is also thought that a lot of calibration precision could be gained from using data collected with the source located at different positions in the tank but this work still remains to be done.

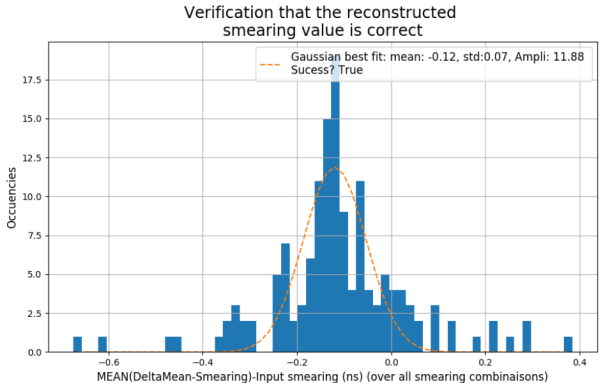

Fig. 12: Histogram of the difference between the mean of the DM-SV distribution and the mean of the smearing Gaussian when working with 2,000 simulated events.

D. Applying the same calibration procedure to the WCTE geometry

The analysis method presented in this paper was then applied onto the simulated Water Cherenkov Test Experiment (WCTE) geometry. Even if the position and number of PMTs are different between the IWCD and WCTE experiments the calibration procedure should function in a similar way. Some errors were however encountered, notably the TOF corrected datasets presented some unexpected very early hits as presented in Figure 13.

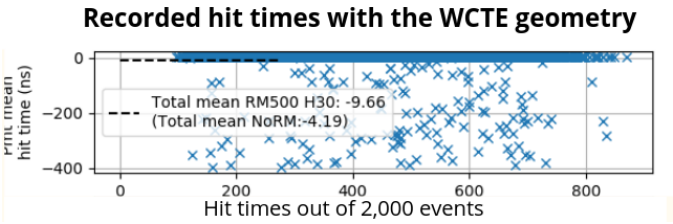

Fig. 13: Simulated mean hit times of the PMTs in the WCTE geometry showing some unexplicable early hits.

Table I summarises potential sources to the issue observed when we are applying the calibration sequence onto the WCTE geometry.

Potential issues	Checks to be performed
Different offset value	Get back to detector specifications
Different time delay needed to be considered a scattered photon	Calculate with respect to the detector size
Different statistical threshold to be considered a successful fit	Increase the threshold since we detect more hits per PMT in WCTE geometry
Multiple hits recorded by the same PMT in a single event	Lower the amount of photons that get sent per simulated light burst

TABLE I: Potential reasons explaining the issue observed on the WCTE geometry and means to fix them.

As well as being performed on the original WCTE geometry this calibration process has also been applied onto a WCTE geometry modified by Dr Anthony to include a support arm for the calibration source. The calibration quality with and without the arm is shown in Figure 14 and 15 respectively. It is interesting to note that the shadowing caused by the arm does not significantly impact the quality of the calibration.

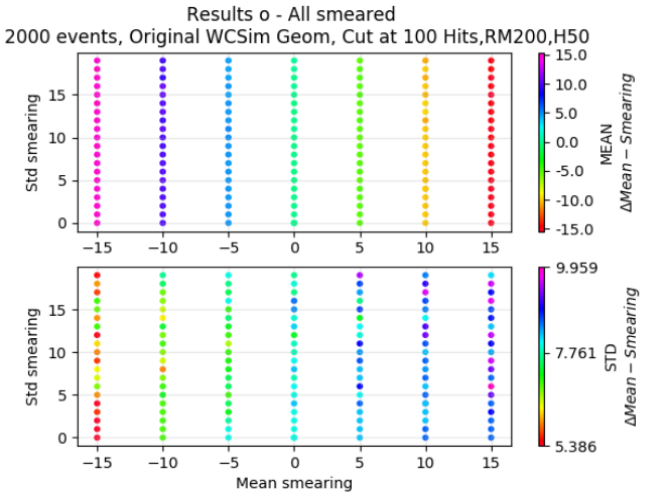

Fig. 14: Calibration quality for the original WCTE geometry showing that the standard deviation of the DM-SV distribution is much larger and less constant than in the IWCD case.

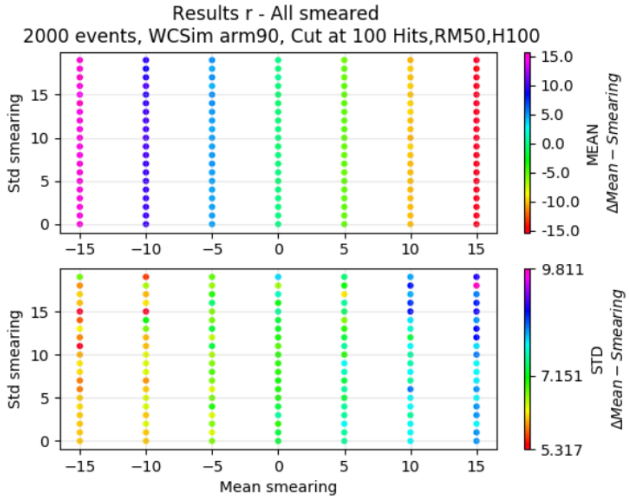

Fig. 15: Calibration quality for the WCTE geometry modified by Dr Anthony to include the support arm. It shows that the quality of the calibration is not affected by the presence of the arm very much as the standard deviation of the DM-SV distribution is similar to what it is in the original WCTE geometry.

It is thought that this calibration program, if modified to correspond better to the WCTE geometry requirements could produce results at least as good if not better than what has been obtained for the IWCD geometry. If it is the case then this is a strong indication that this calibration process could be applied to any Water Cherenkov experiment including the Super- and Hyper-Kamiokande detectors.

IV. Conclusion

This paper presents a method for calibrating the photomultiplier tubes of Water Cherenkov neutrino detectors.

It is based on the use of a uniform light source located in the water tank. This study has been performed on datasets simulated with the Geant4 and WCSim softwares. The calibration technique presented showed that IWCD's PMTs can be calibrated with an accuracy of 0.096ns when working with 2,000 iterations of the same experiment, regardless of the mean and standard deviation of the smearing Gaussian. The results point out that increasing the statistical sample size would increase the accuracy of the calibration, probably by a factor of $\sqrt{n}$.

Some issues have been highlighted when only a portion of the PMTs is smeared and means to fix them are presented in this paper. This calibrating method, originally written for the IWDC geometry was also applied to the WCTE geometry, with and without the additional light source's support arm. The issues that arose there are discussed and some potential solutions are presented.

REFERENCES

[1] B. M. Bolotovskii, "Vavilov cherenkov radiation: its discovery and application," *Physics-Uspekhi*, vol. 52, no. 11, pp. 1099–1110, 2009.
[2] A. Horvath, "Wikimedia commons." [Online]. Available: https://commons.wikimedia.org/w/index.php?curid=637092
[3] T2K Collaboration, "Constraint on the matter-antimatter symmetry-violating phase in neutrino oscillations," 2020.