Some Suggested Studies for Investigating Data/MC Differences which Could Affect fitQun Physics Results

Roman Tacik

Wednesday, August 19, 2015

University of Regina

For the last several years, I have been looking at SK Ni Calibration data. I have found a number of issues with the MC simulation of this data. In general, the Data/MC differences are quite small, and it is not clear whether they have any impact on HE physics results.

Today, I present some suggestions for fairly simple studies which could be done to study this question. They are not necessarily in order of importance. Those in red are more difficult.

For reference, I include slides from my presentation at the previous fiTQun Workshop in Dec 2014. These include more background details.

1 Dark Noise Rate

In the MC, all PMTs have the same dark noise rate,
specified with the variable DS-DARK in the input card file.

In reality, the rates for different PMTs vary significantly.
See the skofl/const/darkr/ directory.

It is easy to change this by adding a few lines to the
file dsnois_multi.F

Note that this issue has been studied informally before, and no impact on physics results found.

2 Water Quality (a) Scattering and (b) Absorption

Although the SK water quality has been quite stable during the SK4 period, it has varied somewhat. My understanding is that the HE group simulations ignore any variations.

Water quality parameters can be specified in card files, but it might be simplest to edit the `gtckov.F` file directly.

Scattering and absorption should be varied independently.

3 Top Bottom Asymmetry in Water Quality

In the simulation, the probability of absorption varies linearly with z in the SK tank.

The slope of the variation is specified with the variable DS-TBA in a card file.

I think all HE group simulations use a fixed value of 10. for this. In recent times, the best value from the calibration group has been varying slightly around a value of 16.

4 Top-Bottom Asymmetry Model

Some of my studies have led me to believe that the variation in water quality may not be simply linear.

In order to study this, the `gtckov.F` file would have to be rewritten.

5 Logical Flaw in Interaction Algorithm

This is the way the routine `gtckov.F` works:

For each Cerenkov photon being traced, an interaction length is obtained using a random number and water quality parameters. That length is compared to the distance of the next GEANT step.

If it is decided that an interaction will occur, the routine `sgabsc.F` is called. That routine decides whether the interaction should be scattering or absorption, based on the photon position at the interaction point, at the end of the current step.

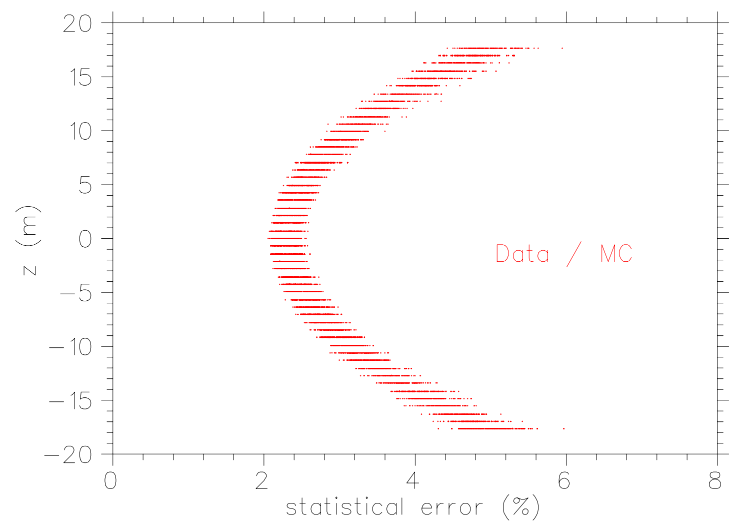
However, if the step involves a change in z, and the absorption probability varies with z, this leads to an incorrect weighting.

Fixing this properly involves a major rewrite of the two routines. But maybe a study could be made of the effect of choosing the initial rather than final position in `sgabsc.F`

6 PMT Quantum Efficiency

The PMT quantum efficiencies are specified in the file skofl/const/qetable3_0.dat

The values were determined from a direct comparison between data and MC during the SK3 period, at a time when the water quality was isotropic. Unfortunately, the MC used had poor statistics, and the statistical errors on the QE values are 2 to 4 %.



A simple study could be done by varying all the values in the QE table by 4%, say.
It would be slightly better to vary the values according to PMT z position.

9 Reflection Probability

I’ve never looked into this. I don’t think it is important for the Ni calibration data.

But I mention it for the sake of completeness, and because Zach has already started to look into it.

10 PMT Gain Change with Time

It is known that the PMT gains are changing with time.
This is not modeled in the simulations.

My investigations indicate that the rate of change varies from PMT to PMT, and that in general, the gain of old (accident survivor) tubes is increasing more slowly than that of new tubes.

I'm not sure what is the best way to study this.

One possibly interesting mock study would be to change the QE of only one type (old or new), while leaving the other alone.

Information on PMT type can be found in the files
`skofl/const/skpmt.dat` and `skofl/const/pmtinf.dat`

11 Acrylic Cover Geometry

In simulations, most photons reach PMT photocathodes by passing through the smooth parts of their acrylic covers.

In some cases, the photons pass through regions where the acrylic covers meet the FRP cases. In the tracking routine `gustep.F`, two routines are called to deal with these cases: `sgacredge.F` and `sggapedge.F`

My understanding is that these routines were implemented in the code without any sort of data validation.

I'm not suggesting they are wrong in any way, but I would be interested to know what the effect would be if photons were stopped in gustep at the point these routines are called.

12 Cylindrical vs Polygonal Geometry

In skdetsim, the model assumes that barrel PMTs all lie in a circle; whereas in real life, the PMTs were mounted on frames first, so that the actual geometry is polygonal.

It is probably impossible to change this in skdetsim.

But I understand that it should be possible to implement both geometries in the WCSIM version of SK.

It would be interesting to see how the choice of geometry affects physics reults.

For reference:

Next slides are from presentation given at the fiTQun Workshop in Boulder, on December 16, 2014

Data/MC Differences & MC Issues Which Could Affect fitQun

Roman Tacik

Tuesday, December 16, 2014

University of Regina

I have been mostly looking at Ni calibration data. This is a Cf neutron source inserted into ball made of NiO and polythene, where gammas of ~9 MeV are produced via (n,γ) reactions.

Data are typically taken every month at three vertical positions: $z = -12$, 0, and +12 m, with $(x,y) = (35.5, -70.7)$ cm; but there are also some special runs at other positions.

In the SK barrel, most PMTs were mounted on frames 4 wide X 3 high. There are (17 X 3) = 51 rows of PMTs, each with (37 X 4) + 2 = 150 tubes. So real geometry is polygonal, but

Skdetsim assumes a perfectly cylindrical geometry.

Ni calibration data taken starting in SK3 period, but with different sources and geometries, and with different trigger conditions.

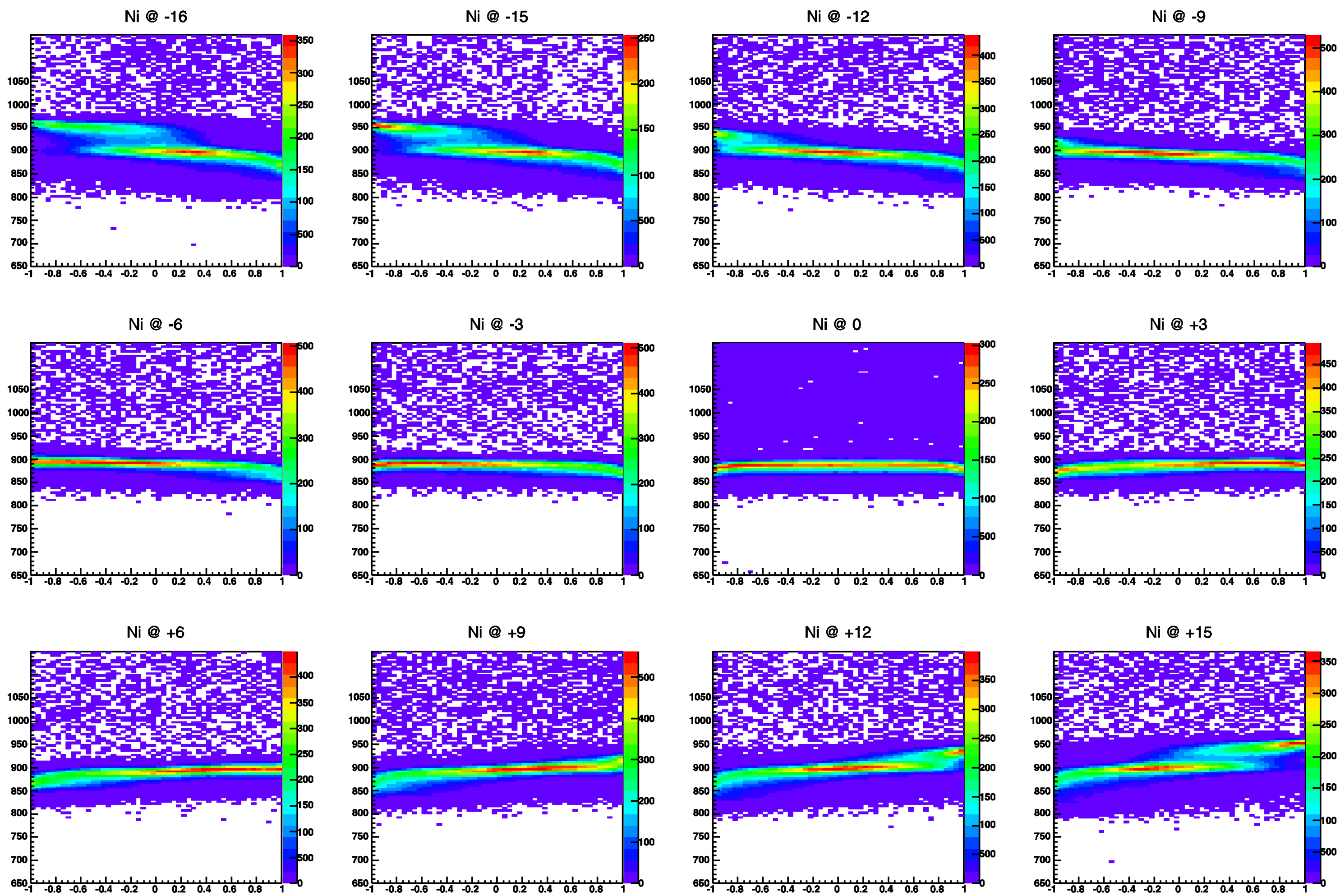
Latest Cf source and ball only in use since October 2011.

Fewer than 100 hit PMTs per event.

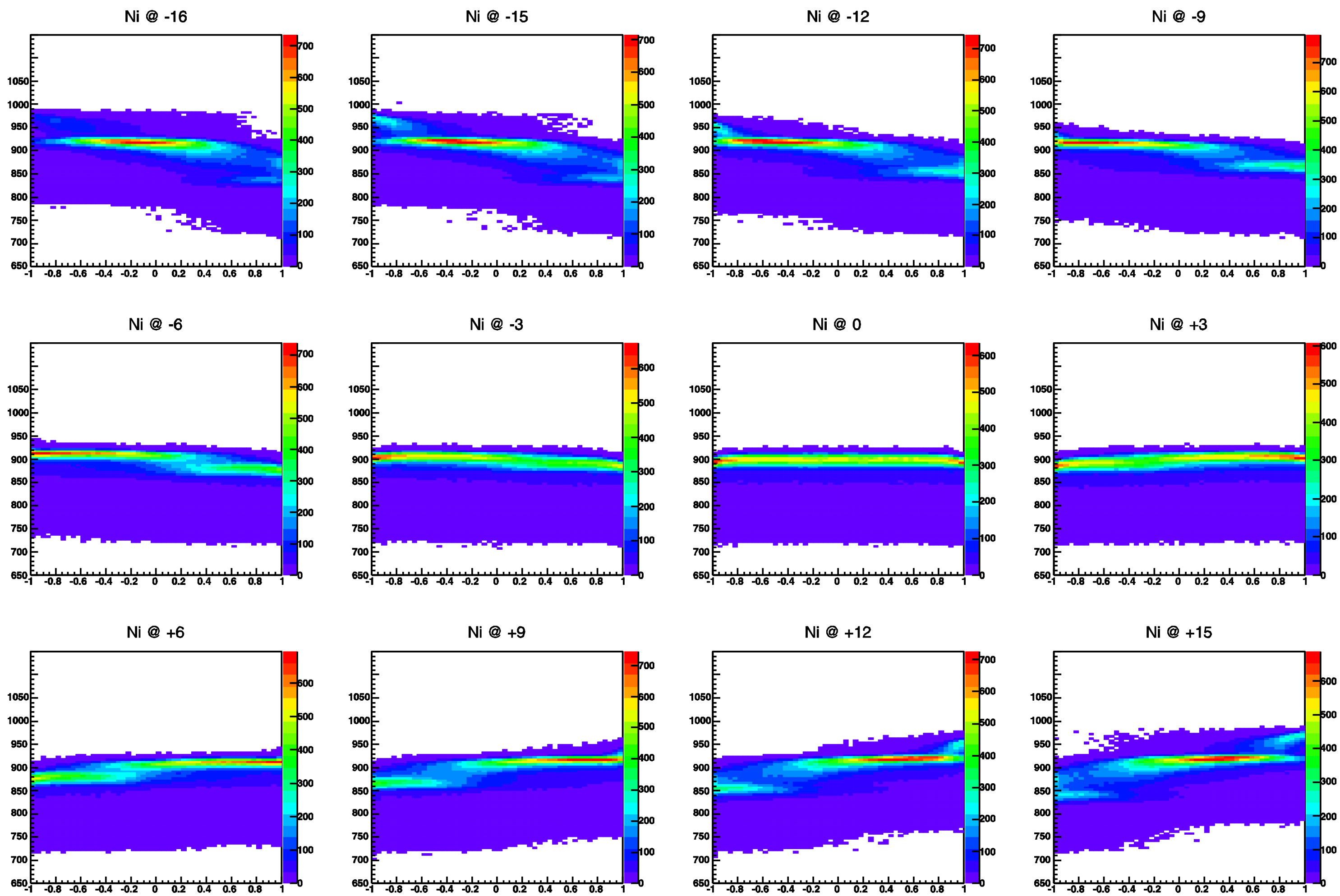
Processed with low-energy time likelihood fitter called Bonsai, which returns (x,y,z,t0).

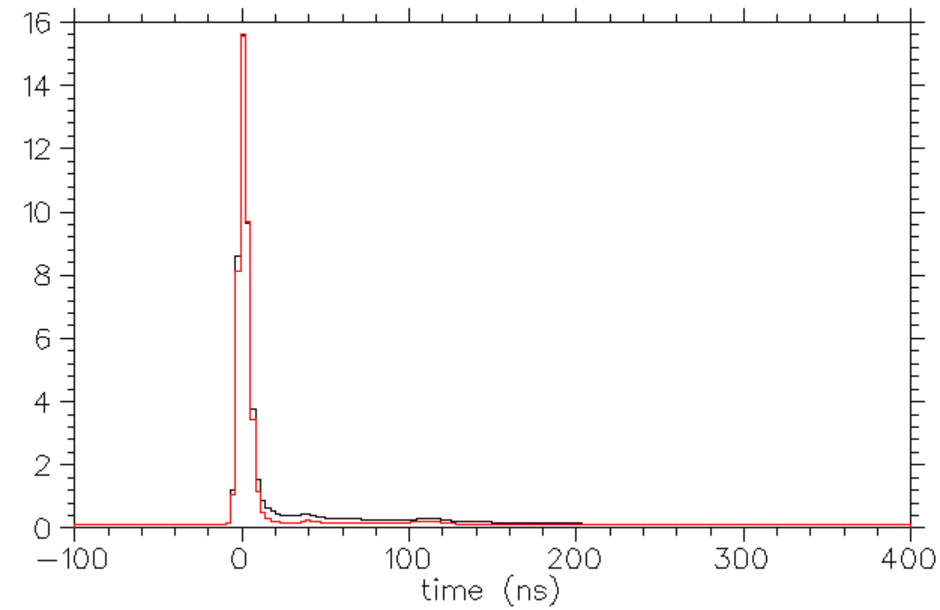
Also get E and (θ,ϕ) direction.

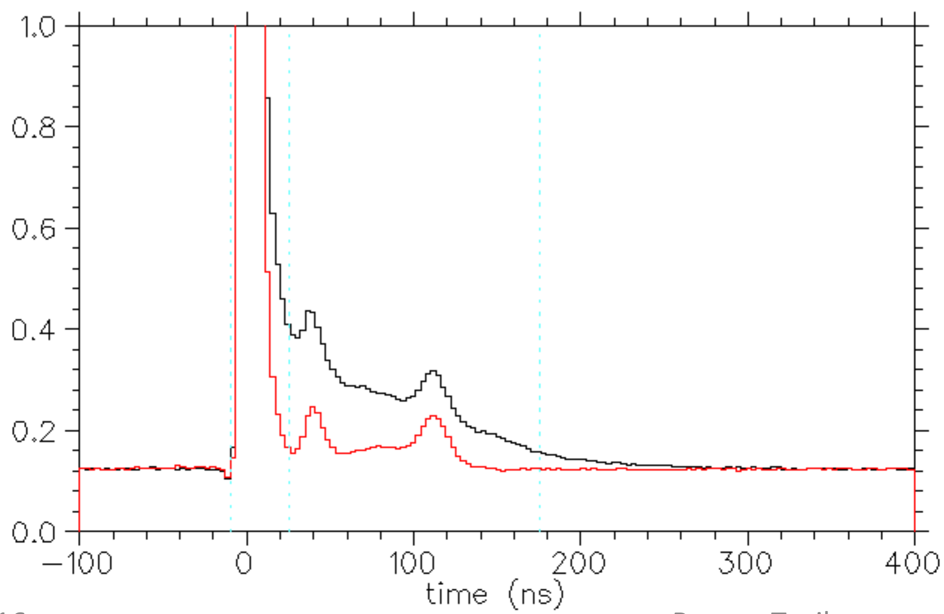
t0 vs cos(θ) for various Ni Ball positions **DATA**



t0 vs cos(θ) for various Ni Ball positions **MC**







MC
Ni (black)
ScatToAbs (red)

Histo of time of all hits, accounting for t0 from fit, corrected for tof

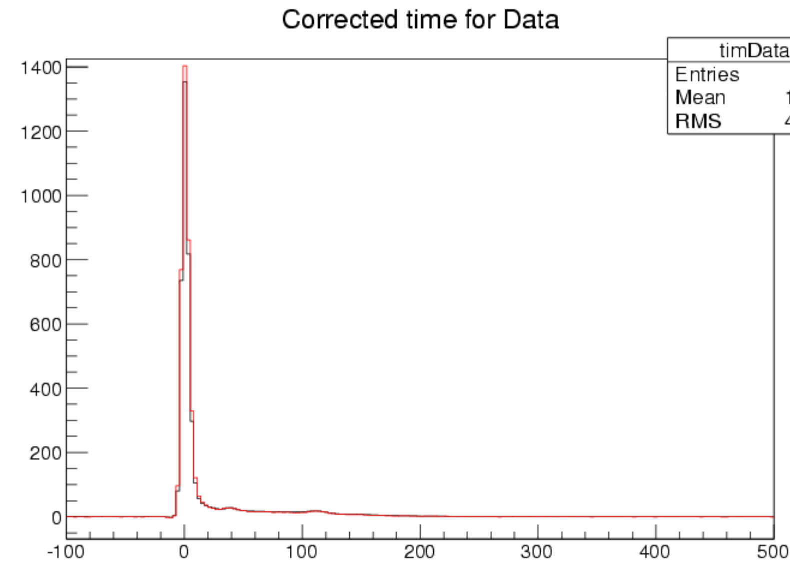
Define time regions of interest:
-10 < t < 25 : "prompt"
25 < t < 175 : "late"

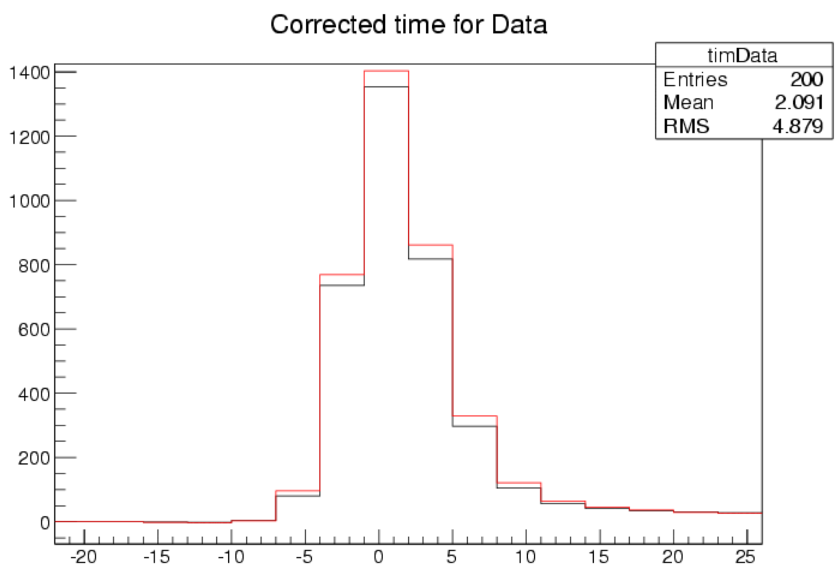
Comparison of Data (black) and MC (red) for all PMTs

(normalize to same number of events)

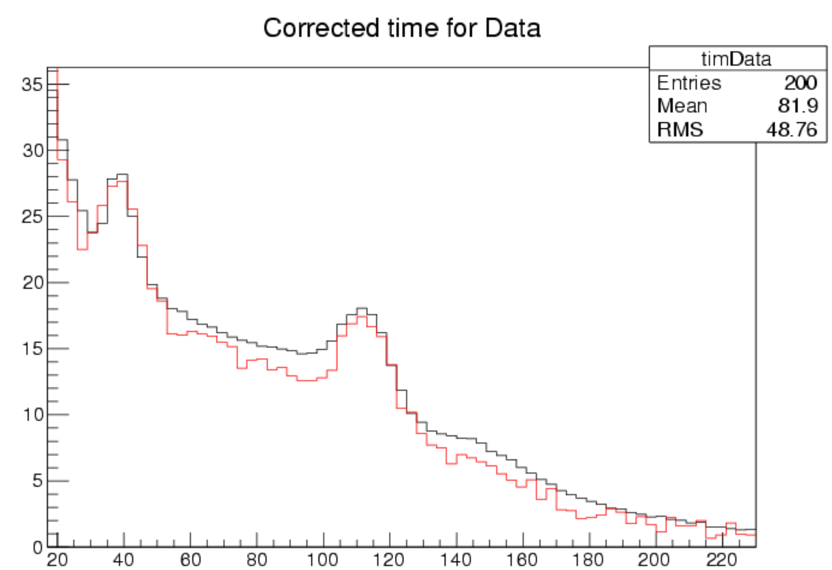
Average number of hits per event is similar (22.5 vs 21.9) but more prompt hits in MC more late hits in Data

all





prompt



late

Consider count rate distribution as a function of z:

There are 150 PMT's at each of 51 z positions.

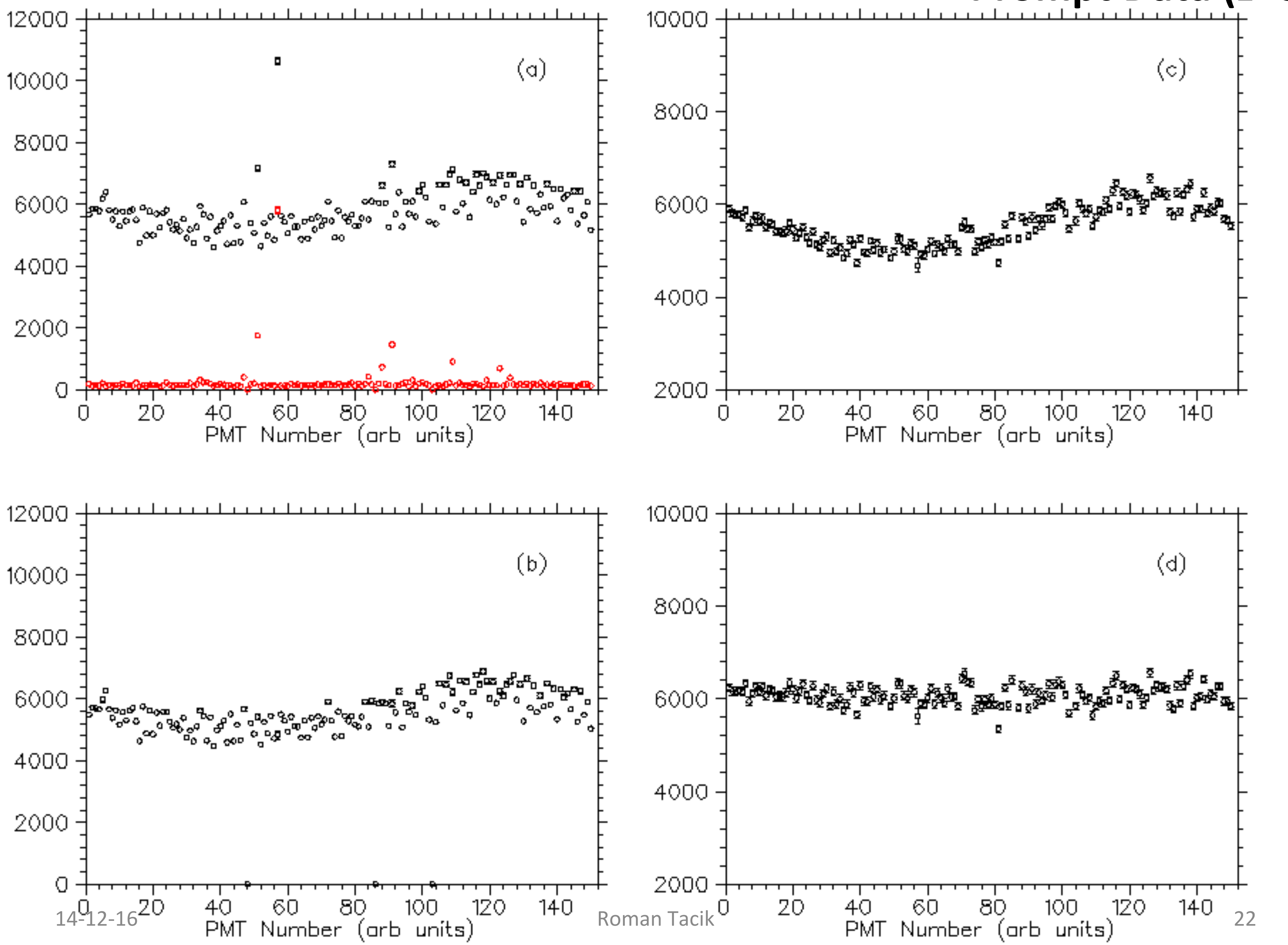
(a) Histogram counts in each PMT, separately for prompt and late times, and separately for dark noise background (using time windows with $t > 400$ ns).

(b) Subtract background from foreground.

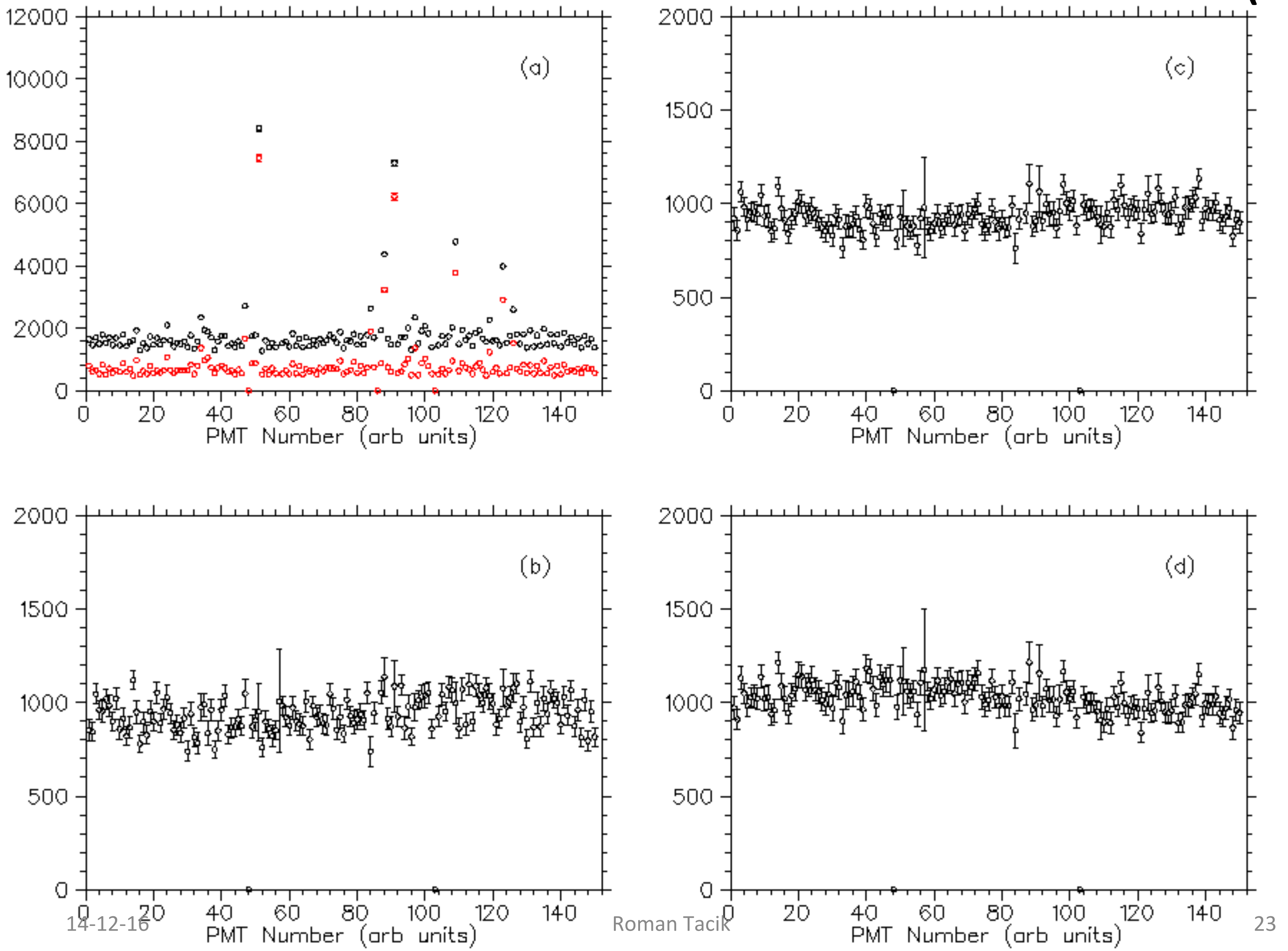
(c) Correct for QE.

(d) Correct for solid angle by multiplying by $(r/r_{min})^2$, where r is distance from average vertex and $r_{min} = 1611$ cm.

Prompt Data (z=0)



Late Data (z=0)

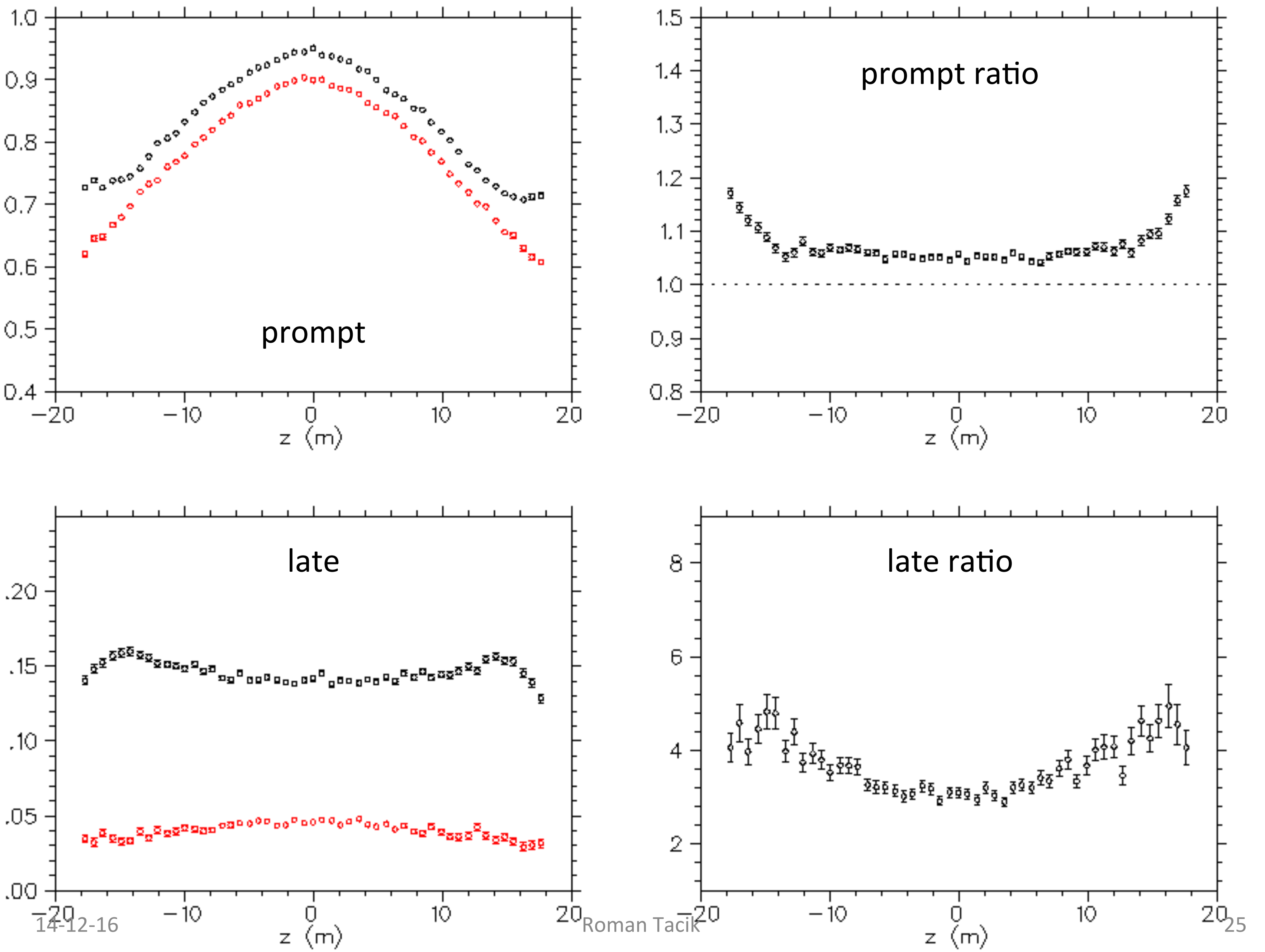

(a)
(b)
(c)
(d)
PMT Number (arb units)
PMT Number (arb units)
PMT Number (arb units)
PMT Number (arb units)

Next, sum up all the counts in the corrected distributions (i.e. the total in the (d) subplots), correcting for the minor effect of dead PMTs, separately for each of the 51 z rows.

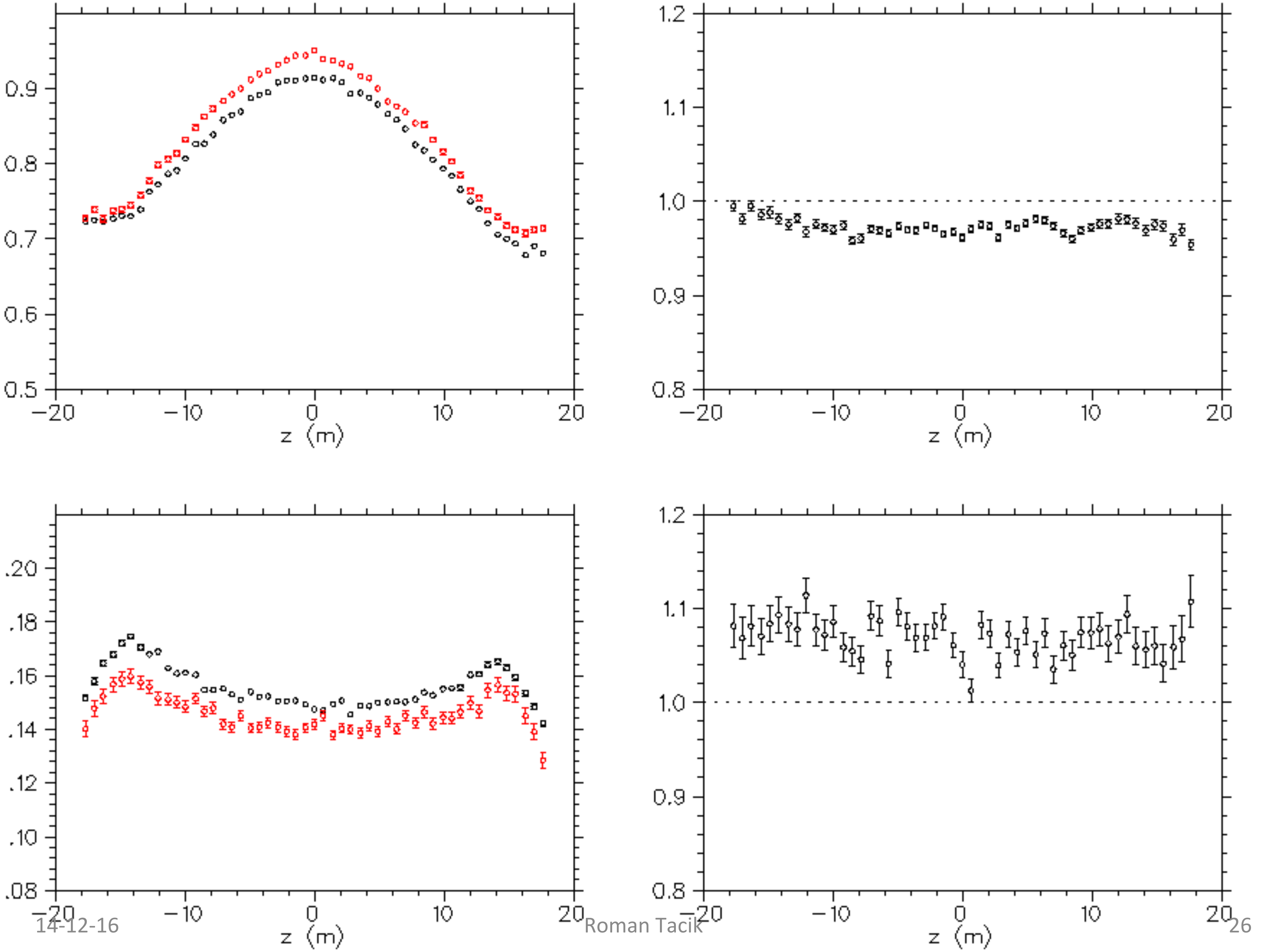
Plot z distributions; compare Data and MC.

First though, compare standard Ni MC with ScatToAbs = 1 case.

MC: Ni (black) ScatToAbs (red)



Ni @ z=0 m Data (black) MC (red)

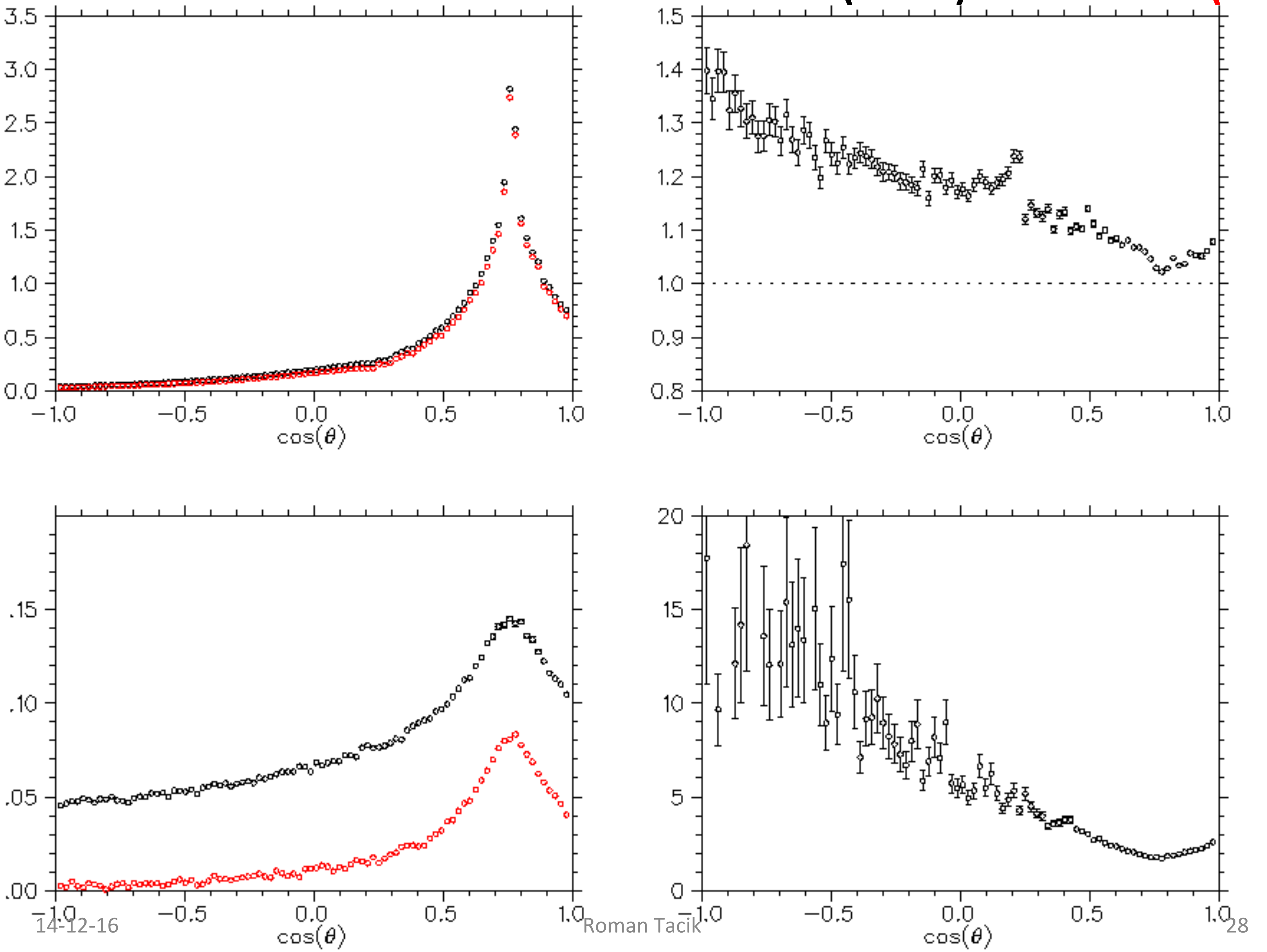


Finally, consider $\cos(\theta)=p\cdot h$, where p is Bonsai fit direction, and h is direction from vertex to hit PMT

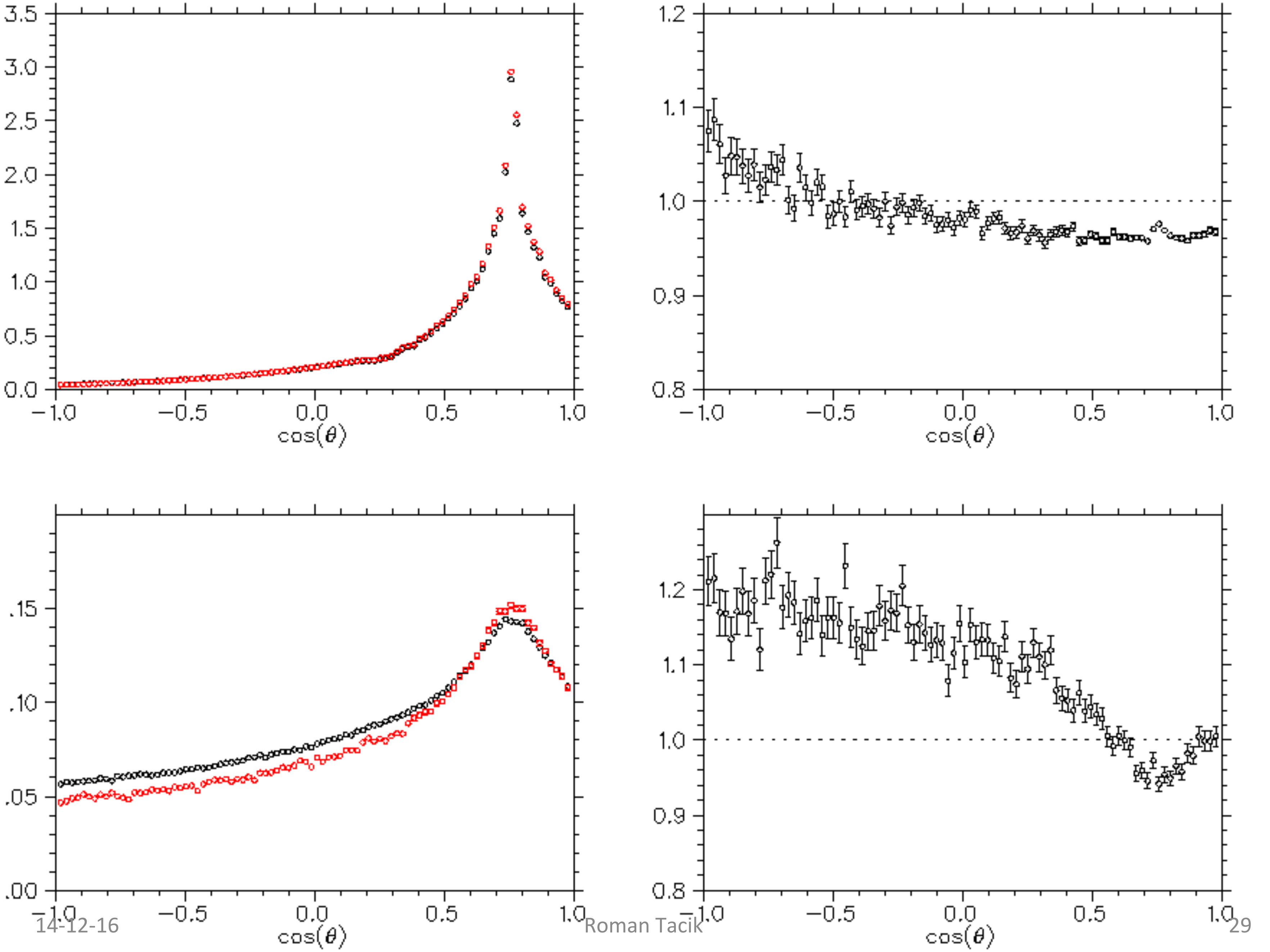
Expect prompt light to be mostly directed in the forward cone, while scattered light will be distributed more isotropically .

Check this by first comparing standard MC with ScatToAbs = 1

Ni MC (black) ScatToAbs (red)



Ni @ z=0 m Data (black) MC (red)



The details of the Data/MC disagreement shown on last slides will depend on water quality.

In skdetsim, Cerenkov photon scattering is dealt with in three separate routines:

WSYMSCSG -> $(1+\cos^2(\theta))$ Rayleigh & Symmetric Mie

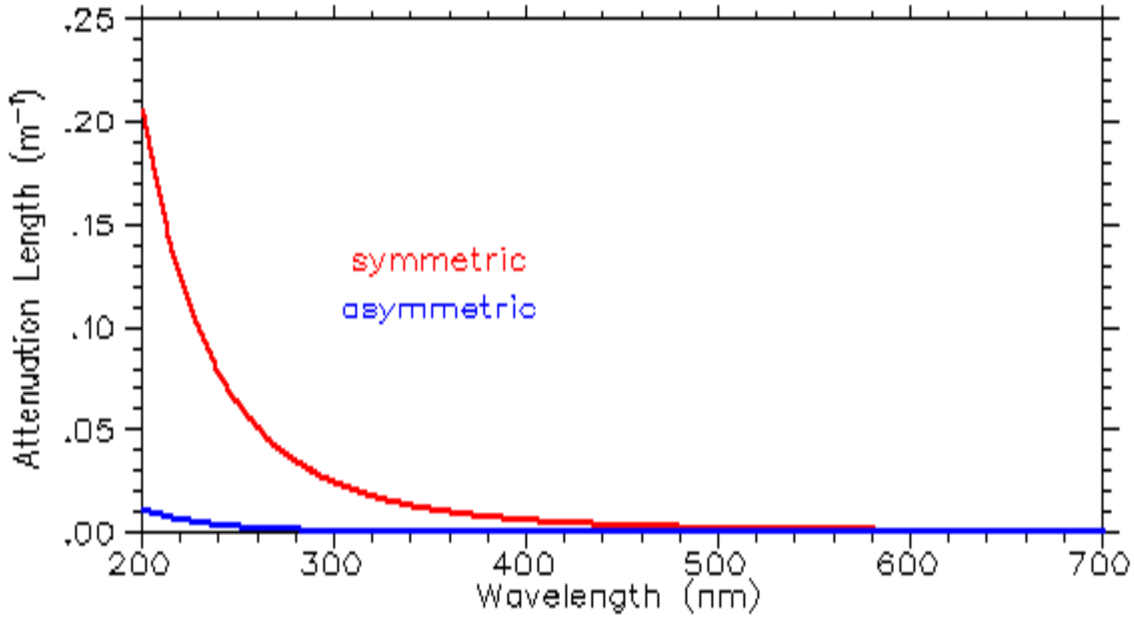
WASYSCSG -> (forward only) Asymmetric Mie

WABSSG -> Absorption

All three are functions of photon wavelength.

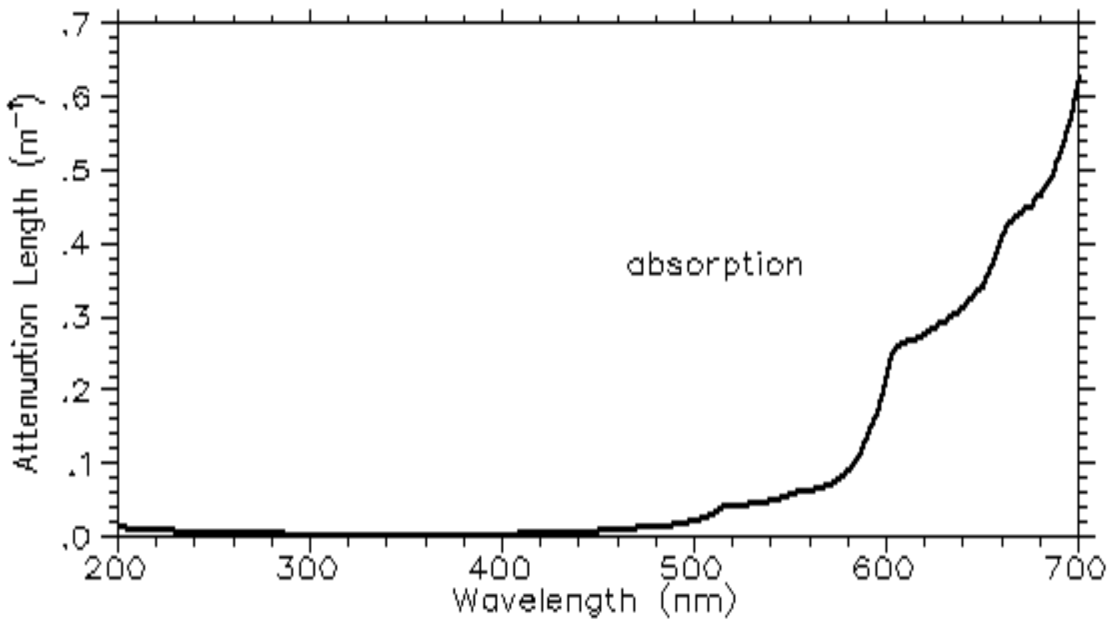
Parameters (read in from tables) vary with time in low energy analyses, but fixed in atmpd analyses.

Elastic scattering is important at shorter wavelengths.



Absorption is important at longer wavelengths.

Varies with z.



The probability P of interaction in a distance dr is:

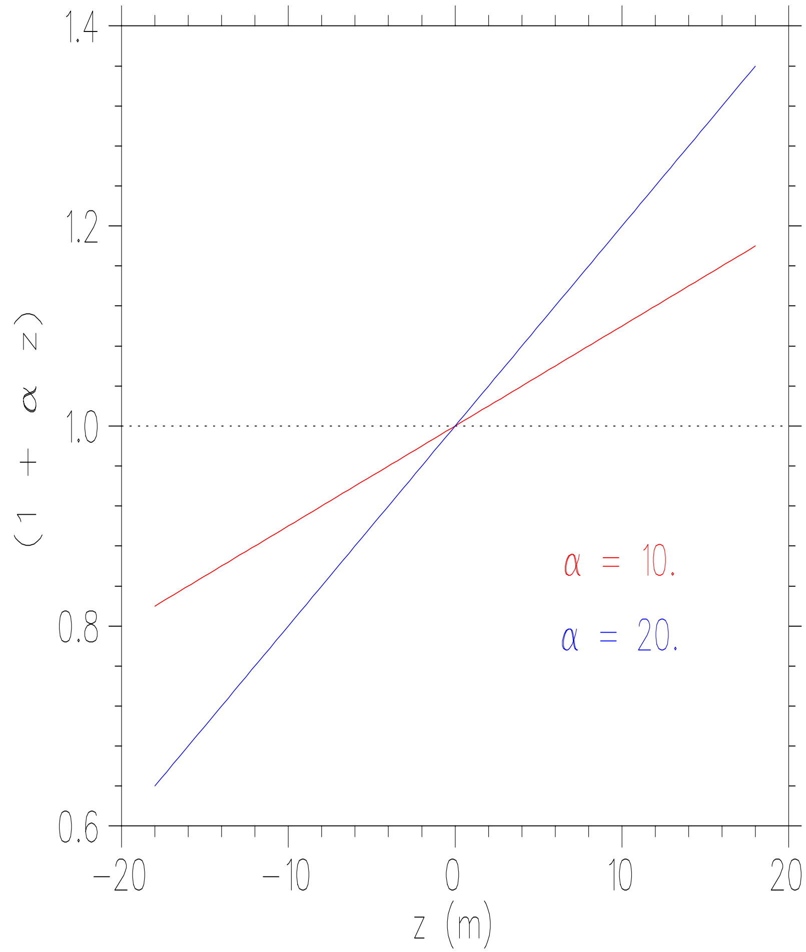
$\ln(P) = [\, L_{sym} + L_{asym} + L_{abs}(1+\alpha z)\,]\,dr$

L's are inverse lengths, so larger values of α mean clearer water at the bottom of the tank, and less clear at the top.

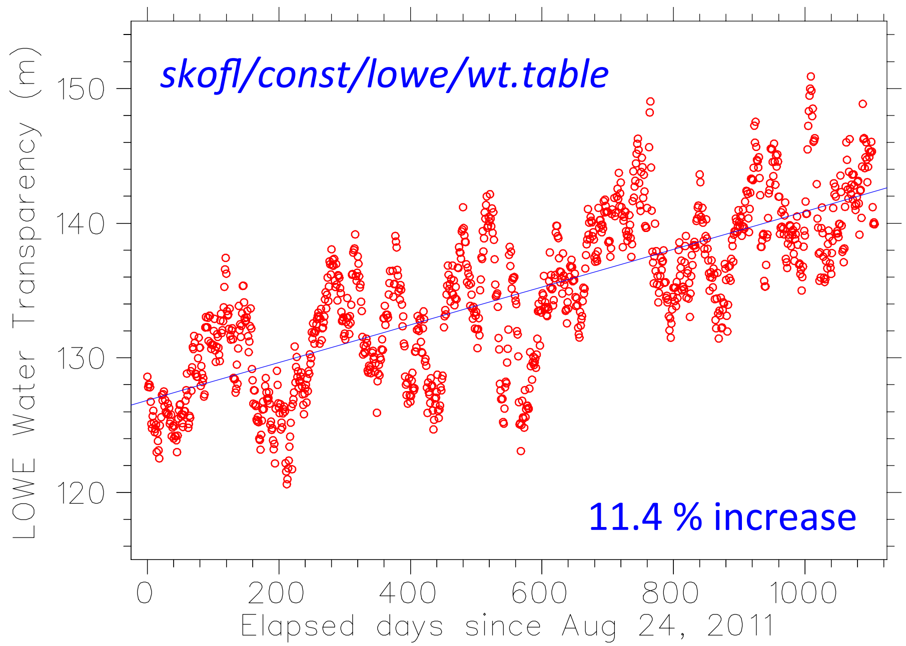
α is related to the top-bottom asymmetry, *tba* (daily values in file `tba_ax.dat`):

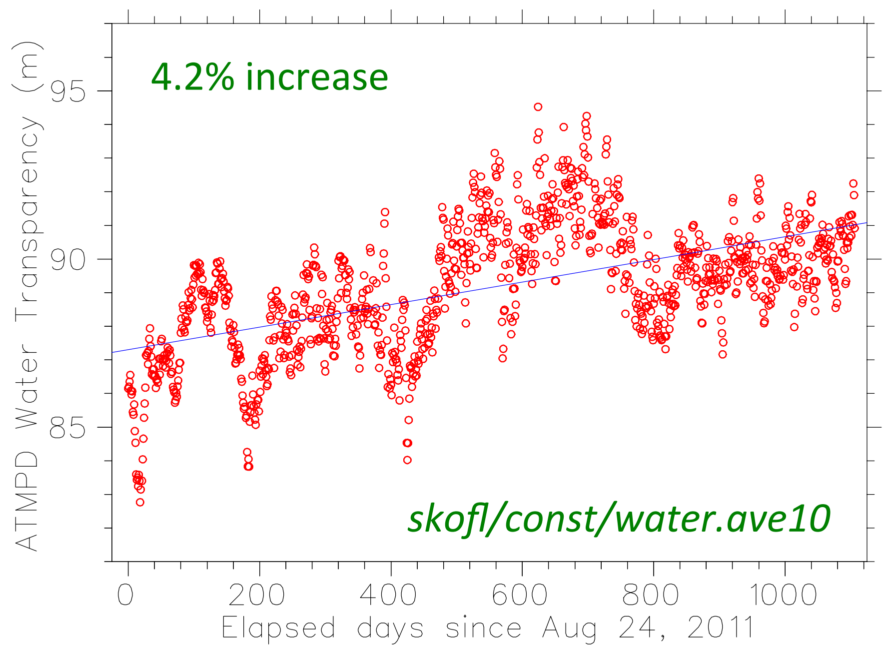
$\alpha = [-0.163(100\,tba)^2 - 3.676(100\,tba)] \times 0.00001$

z dependence is constant for z < -11 m



My analysis is consistent with a z variation in the water quality, but I haven't been able to confirm a strictly linear variation, or the official tba values.



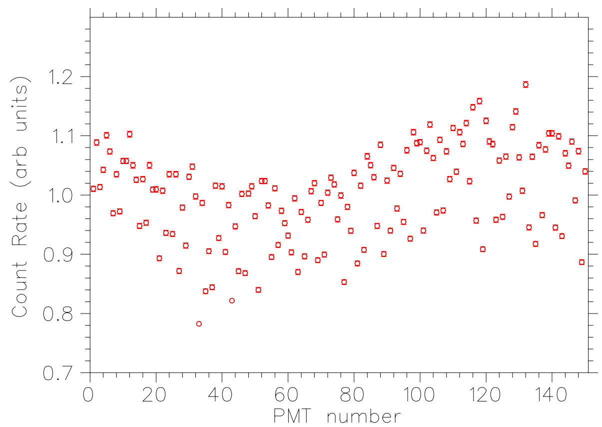


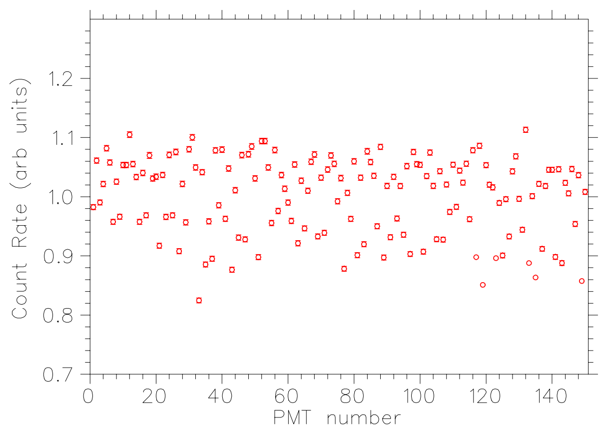
Both LOWE and ATMPD groups report an increase in water transparency over the last three years, which they successfully use to correct their data.

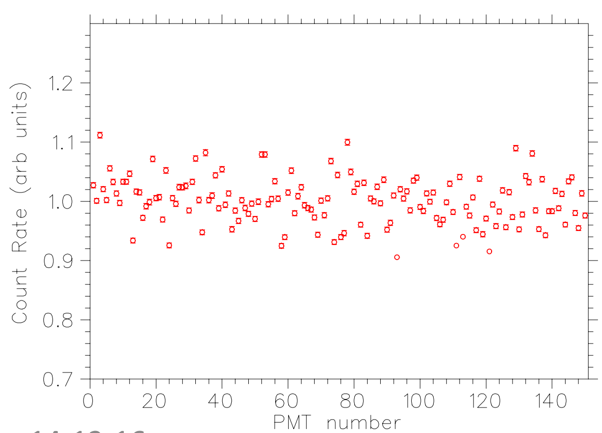
But the absolute value of the transparency, and the amount of change, is different.

Have a closer look at the type of analysis shown on slides 9 and 10:

DATA







Consider Ni Calibration run with Ni at z=0.
Find count rate for all PMTs in same row.
Plots for typical case of row at z=-10.6 m.

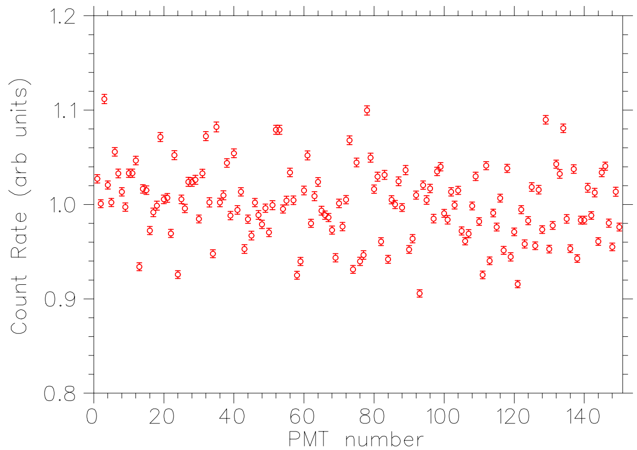
Consider on-time hits and subtract dark noise using off-time hits.

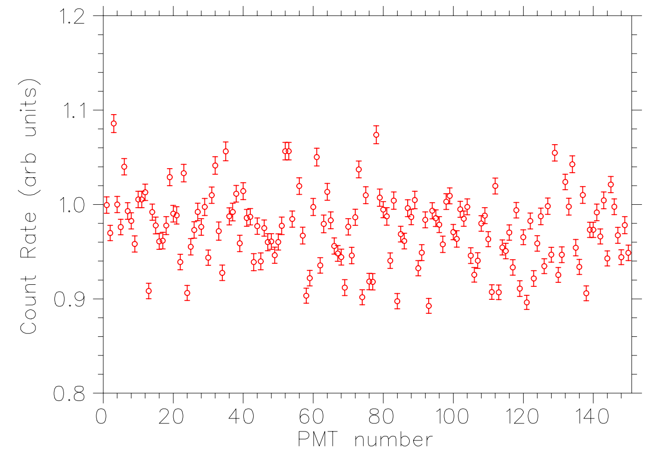
Account for small variation in solid angle by correcting by r^2, where r is distance from center of Ni ball to PMT.

Account for variations in QE using values from *skofl/const/qetable3_0.dat*

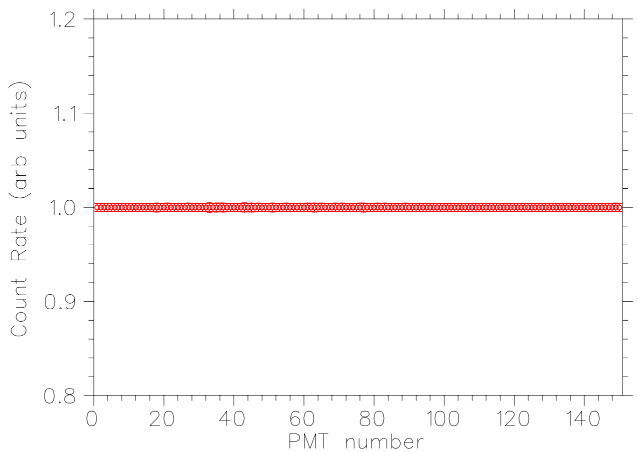
Expect all PMT to have the same corrected rate, but in fact find much larger variation than expected from statistical errors.

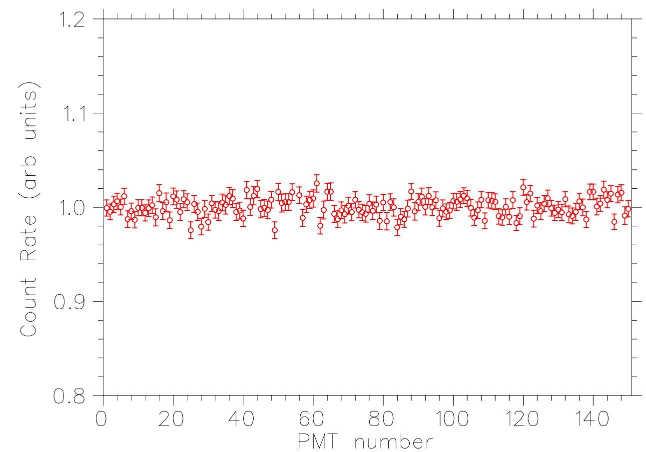
Make additional correction: construct new QE Table (qetable') whose values ensure that all PMTs in same row do have exactly the same count rate for some reference Run (choose 68586 because of large number of subruns). Look at effect on all later Runs with Ni at z=0: find that new correction gives improved results.





qetable3_0





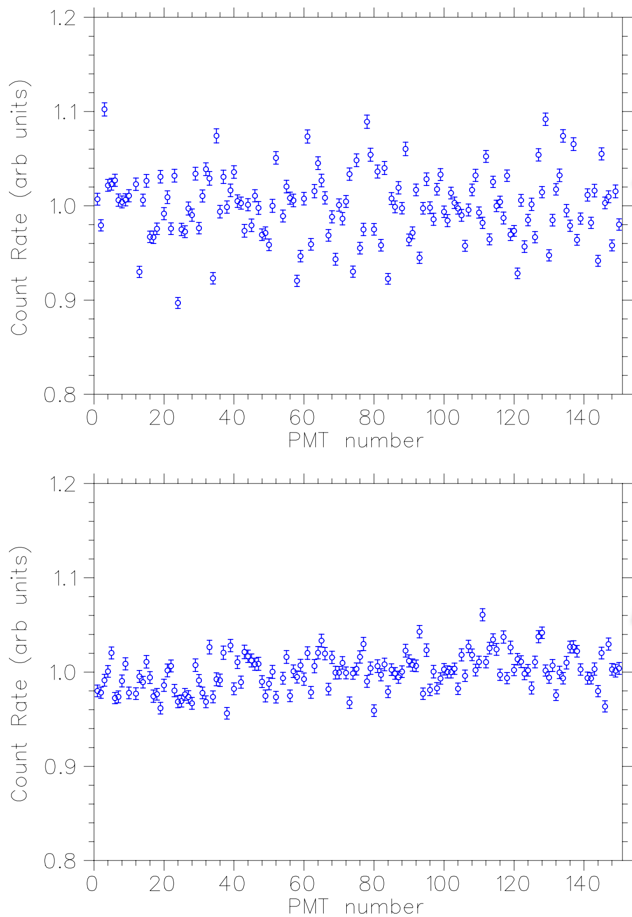
qetable'

Reference Run (68586)

DATA

Typical later Run (70821)

The values in qetable3_0 are based on a Data/MC comparison using SK3 Ni Calibration Runs taken at a time when the water quality was uniform throughout the SK tank. Re-analyze these (Runs 31436 and 31443, taken on October 10, 2006) using same technique: look at count rate of all PMTs in the same row:



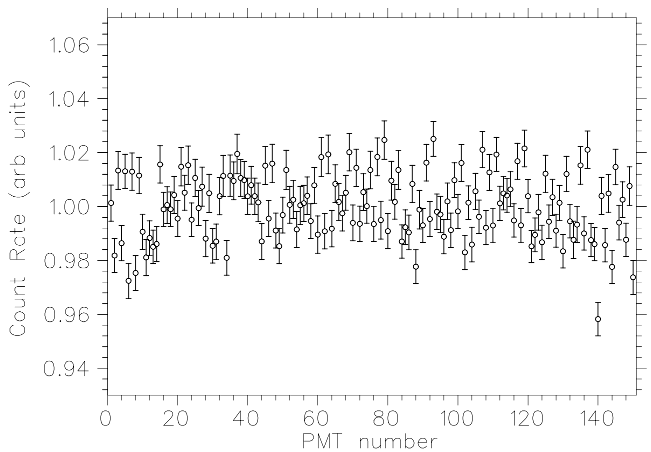
Contrary to expectations, find large variation in rate for PMTs in same row when using qetable3_0 corrections.

The variation is reduced, but still exceeds expectations from statistics, when qetable' is used for correction.

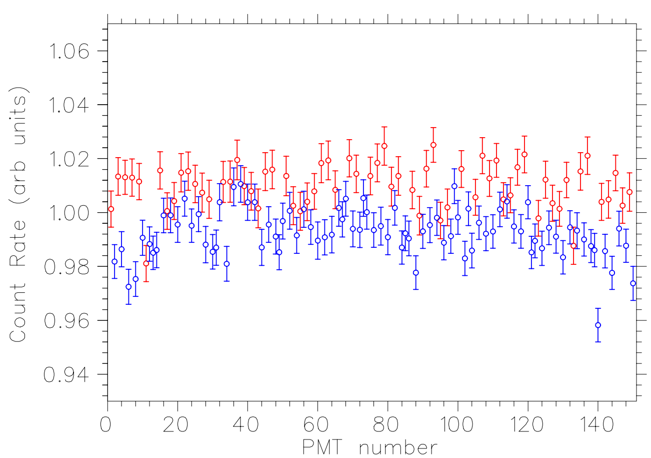
There is some systematic difference between SK3 and SK4 Ni calibration data.

SK3 DATA

Test the assumption that the count rate for all PMTs in same row should be the same by looking at skdetsim simulation:



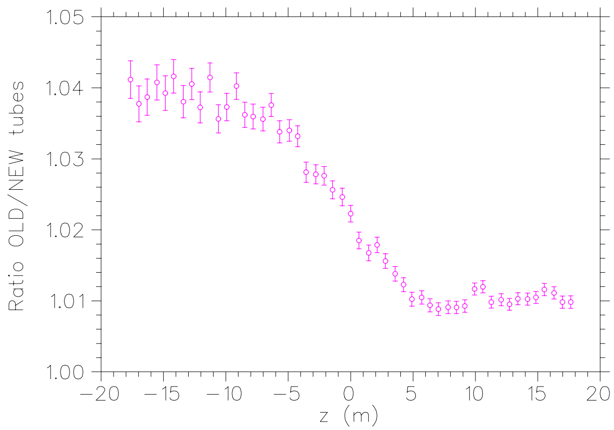
Surprised to find a small, but statistically significant, variation. Plot at left is for same row (z=-10.6) as data plots in previous slides.



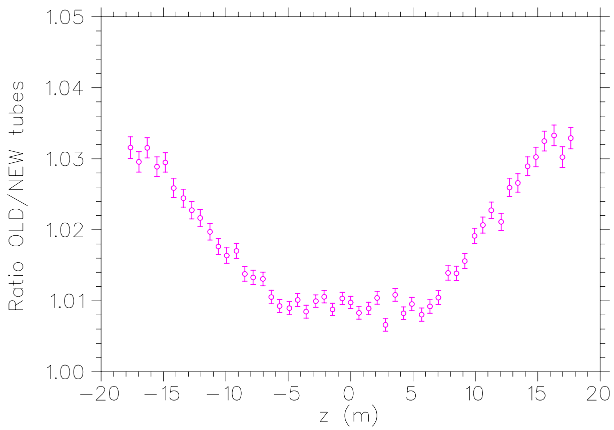
Upon closer inspection, find that variation is actually due to different rates for old (red points) and new (blue points) PMTs.

MC

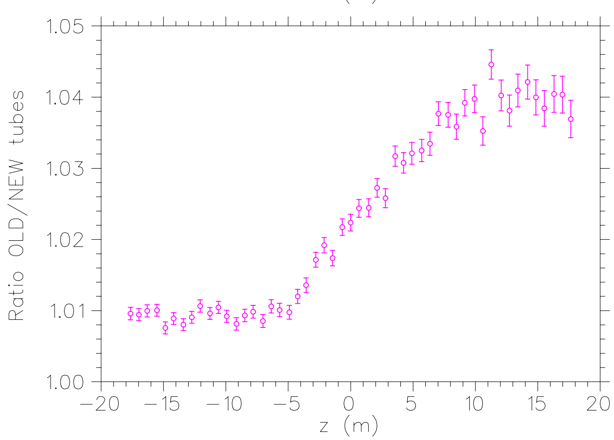
Ni @
z=+12



Ni @
z=0



Ni @
z=−12



MC

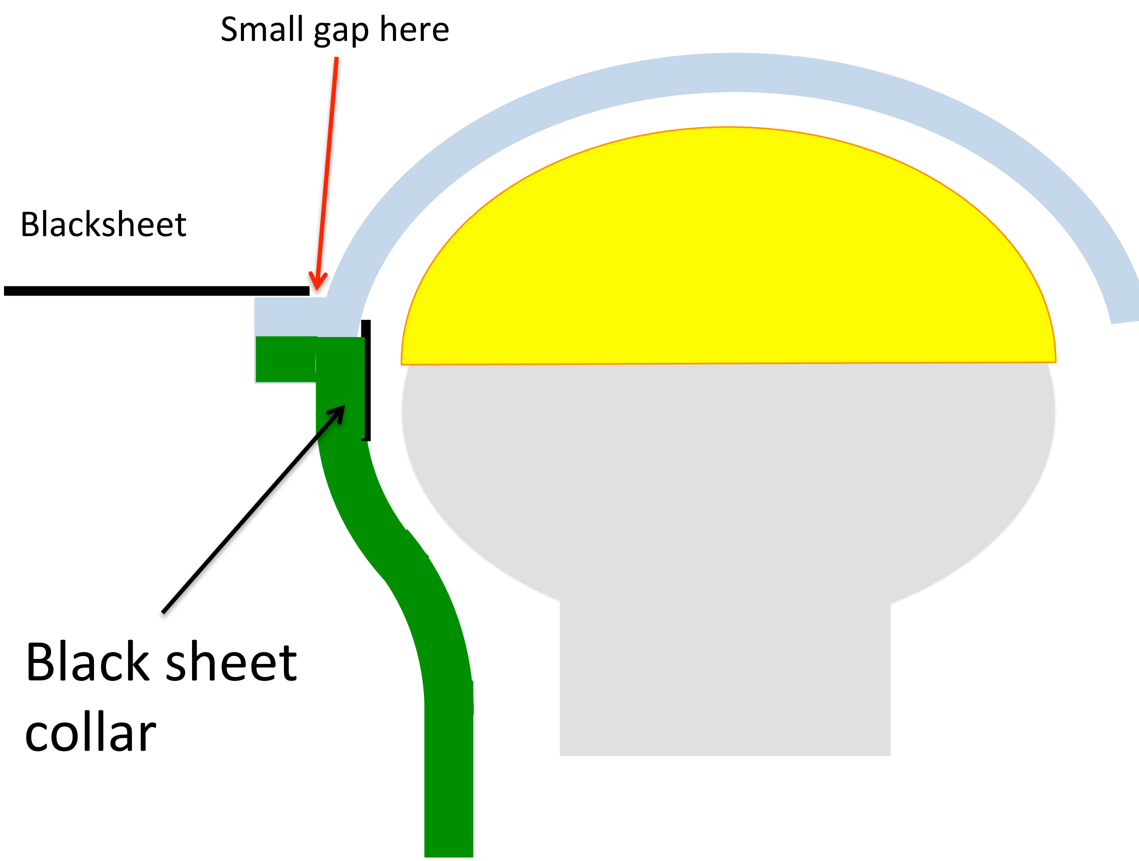
Take average rate of all old PMTs and all new PMTs separately for each row, find ratio, and plot this as a function of z, for all 51 rows.

Do this for Ni runs at the three usual monthly positions: +12, 0, and −12 m.

Find that the MC old/new ratio is constant near the position of the Ni source, but then increases.

Trace this behavior to skdetsim routine *sgacredge*, which treats old and new PMTs differently.

Ask Kameda-san about this. He explains that there is a physical difference between old and new PMTs, having to do with the presence of a black sheet collar, which was installed to prevent light from FRP cases reaching the bottom of photocathode.



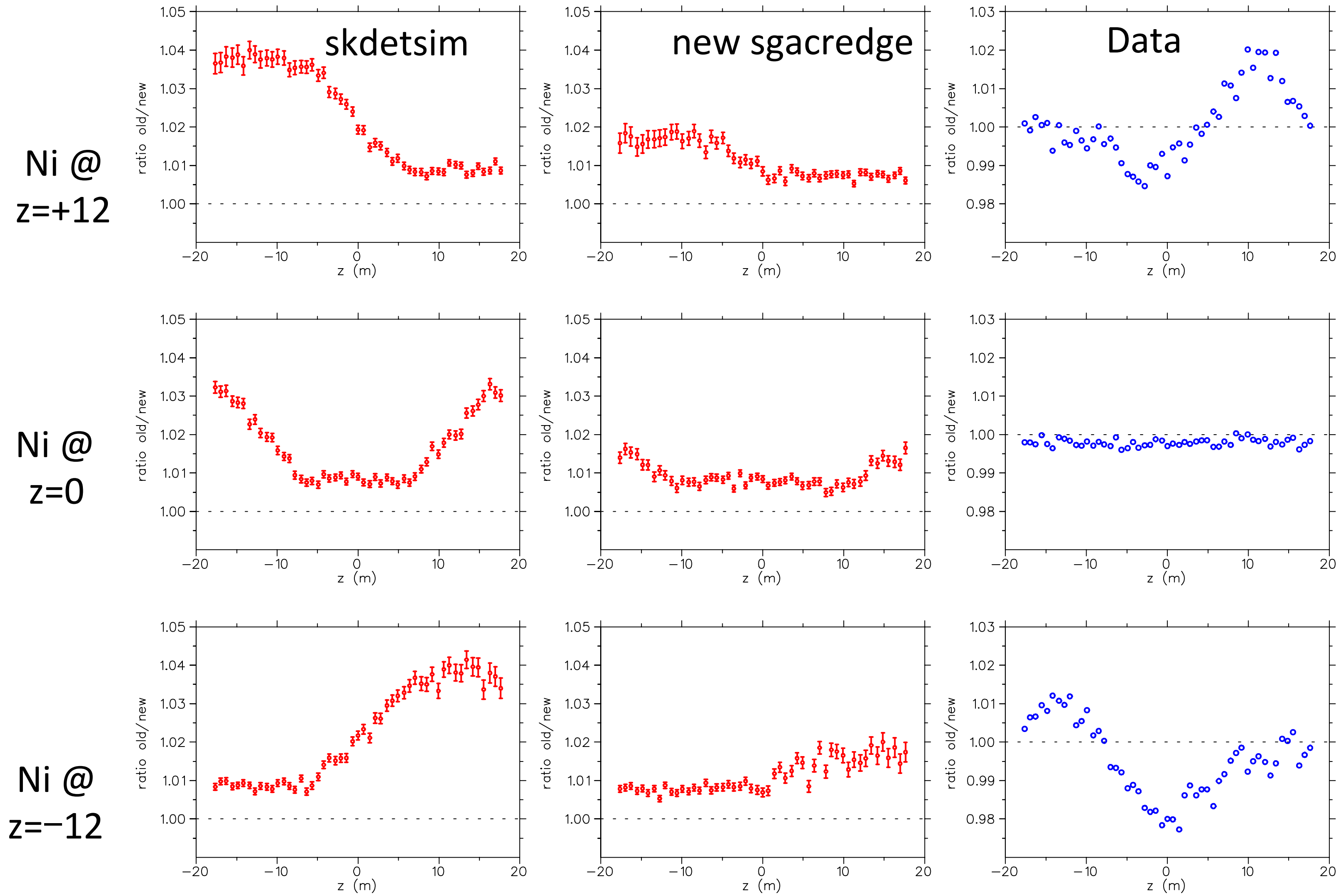
But when he investigates further, he discovers a bug in the *sgacredge* routine: the same variable name was incorrectly used twice.

Kameda-san fixes this bug and produces a new version of *sgacreadge* for testing.

Only SK-III PMTs have “black sheet collar”.
This is to reduce Chemiluminescence effect from FRP case.

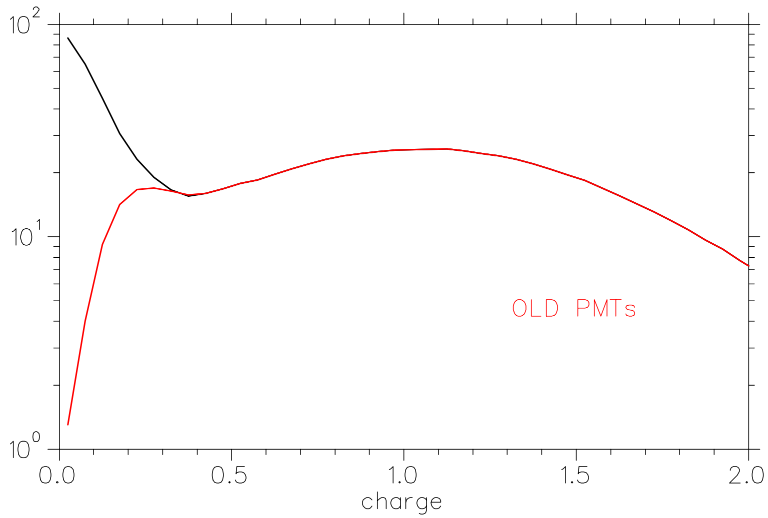
See Kameda-san’s presentation to Calibration Group on June 26, 2014 for more details.

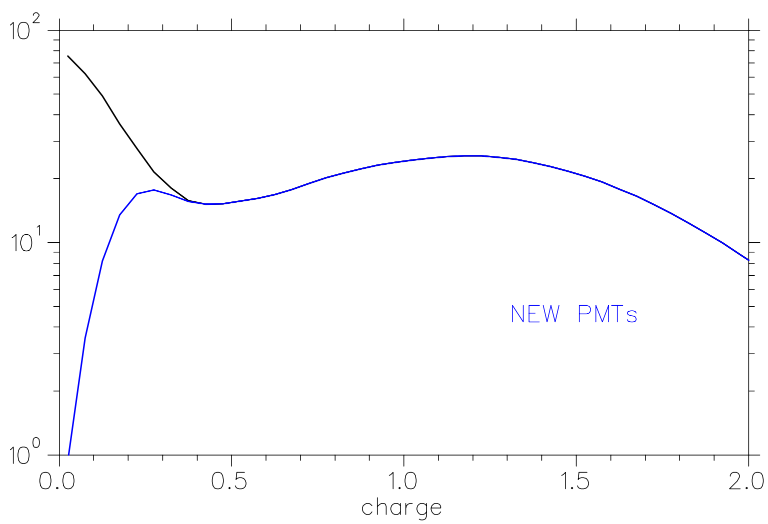
With this new bug-fixed version of sgacredge, the effect on the old/new ratio is reduced. But **the MC distributions to not match the data.**



Notice that even in regions of z where the MC old/new ratio is flat, it is not exactly = 1.0. There are three skdetsim routines which take care of simulating p.e. pulse heights and applying thresholds:

skrn1pe_sk3.F, *dsthr_sk4.f*, and *dsreso_qb.f*





Old and new PMTs are treated differently in these routines.

The black lines in the plot show the shapes of the generated charge distributions, while the colored lines show shapes after threshold and resolution effects are applied.

for old PMTs: 0.7936 survive cuts
for new PMTs: 0.7875 survive cuts

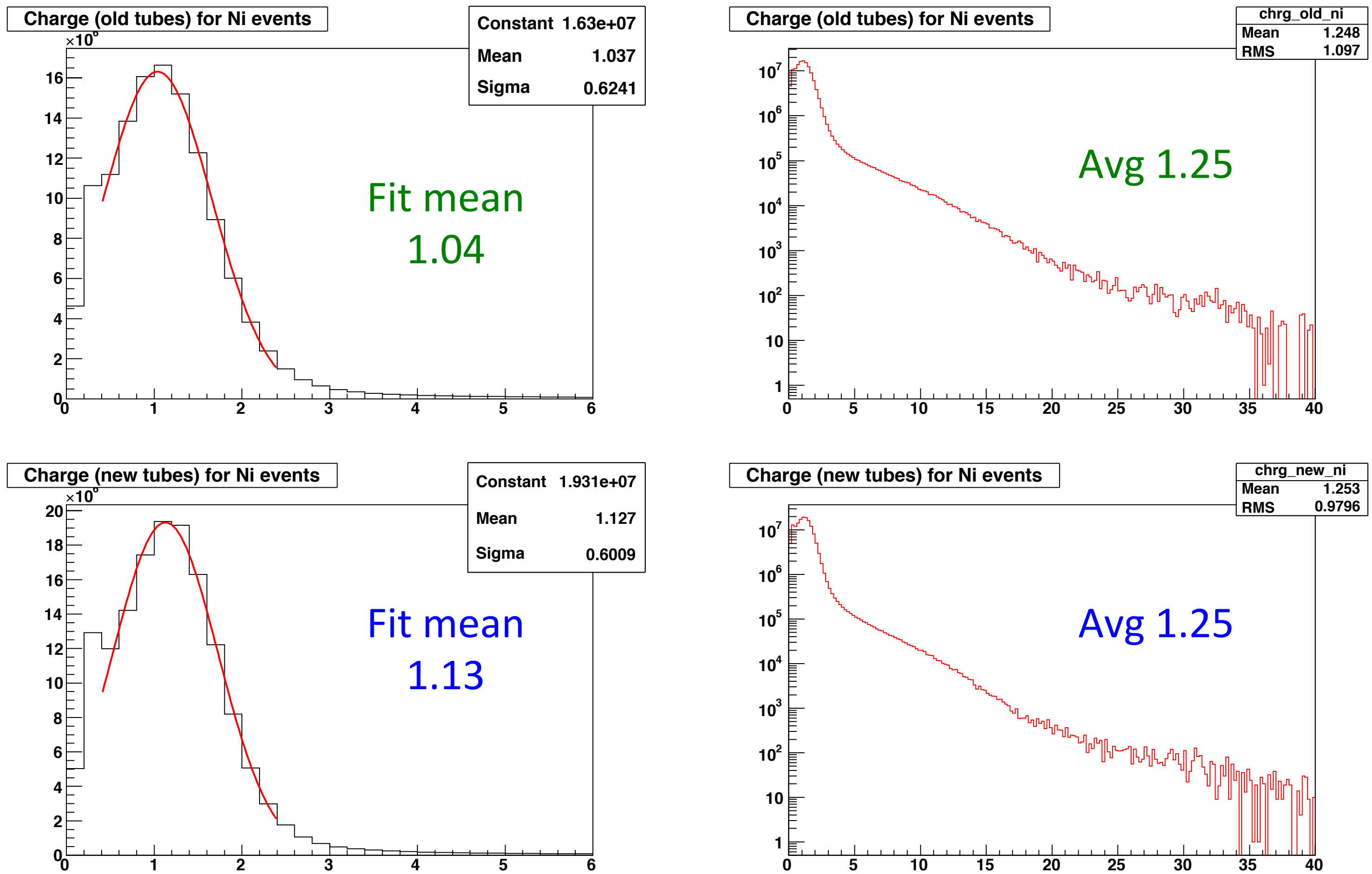
This explains the simulated old/new ratio of 1.008

MC

Have a closer look at the simulated charge distributions. Can characterize them in two ways: using fit peak positions, which are different for old and new PMTs; or using average values, which are the same for both. **Look at data on next slides.**

Old PMTs

New PMTs

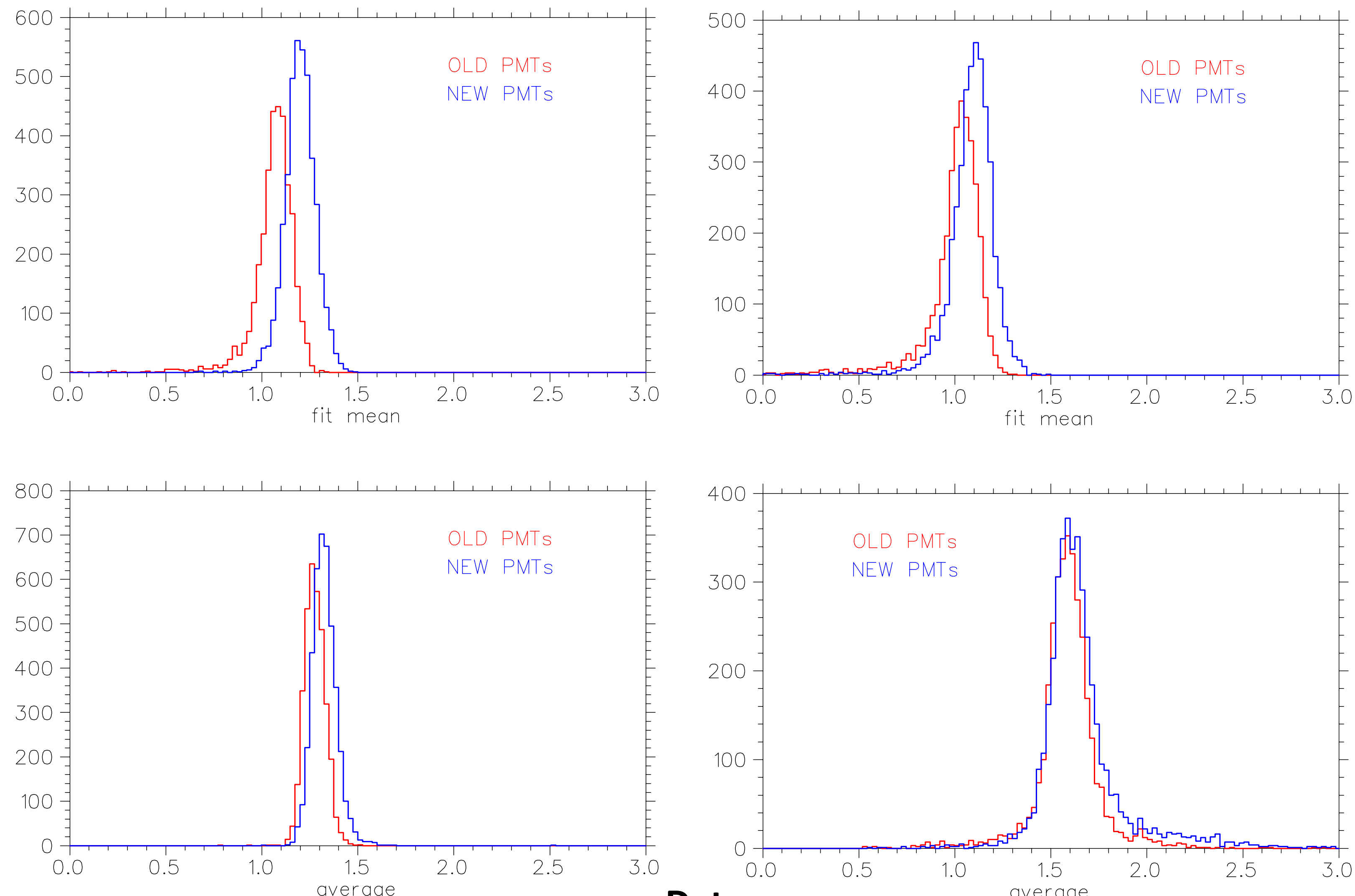


MC

One difference between data and MC is that in data, not all PMTs have the same distribution. Plots show the spread in values, for old and new PMTs, for physics (on-time – off-time) and dark noise (off-time) events, for fit means and averages. (Only consider barrel PMTs.)

Fit mean

Avg

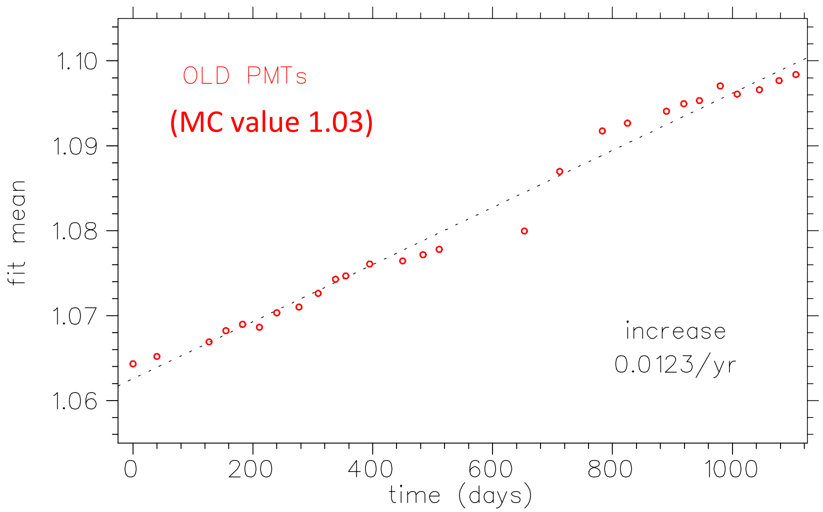


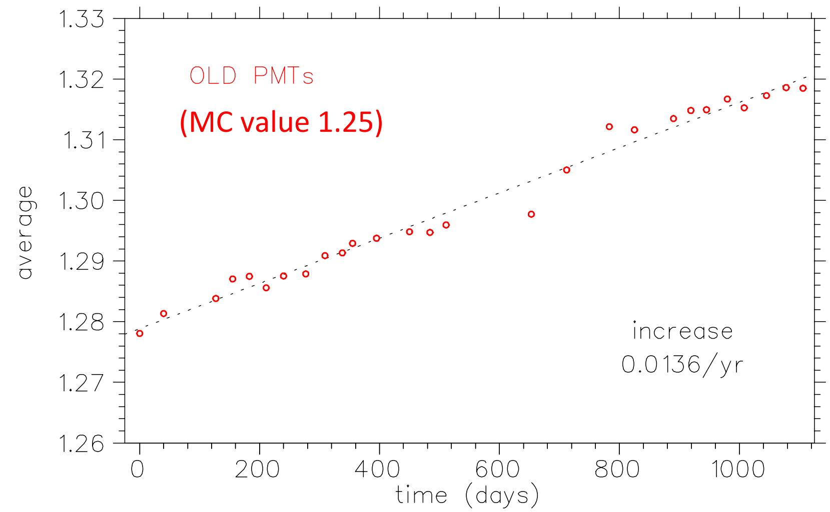
Physics Hits

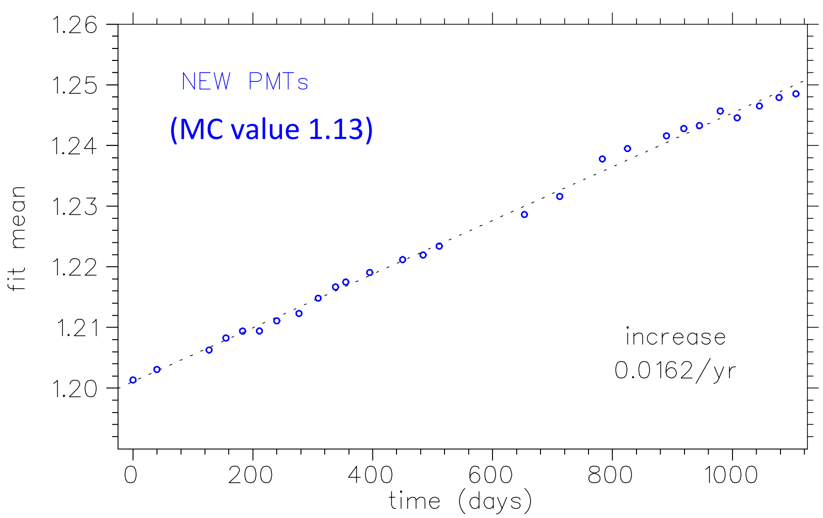
Data

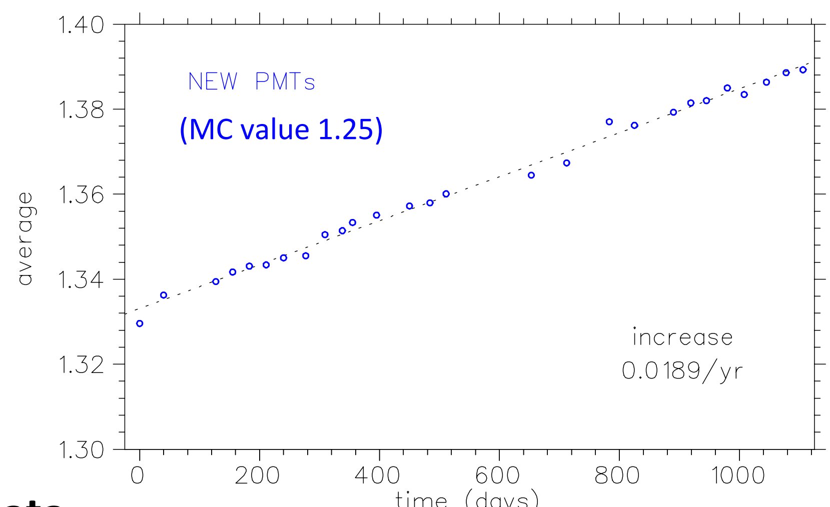
Dark Noise Hits

Another difference is that in data, charges (both fit means and averages) are increasing with time. Plots are **for physics events**, with distributions of individual PMTs summed before analysis. Values for new PMTs are increasing faster than for old PMTs.







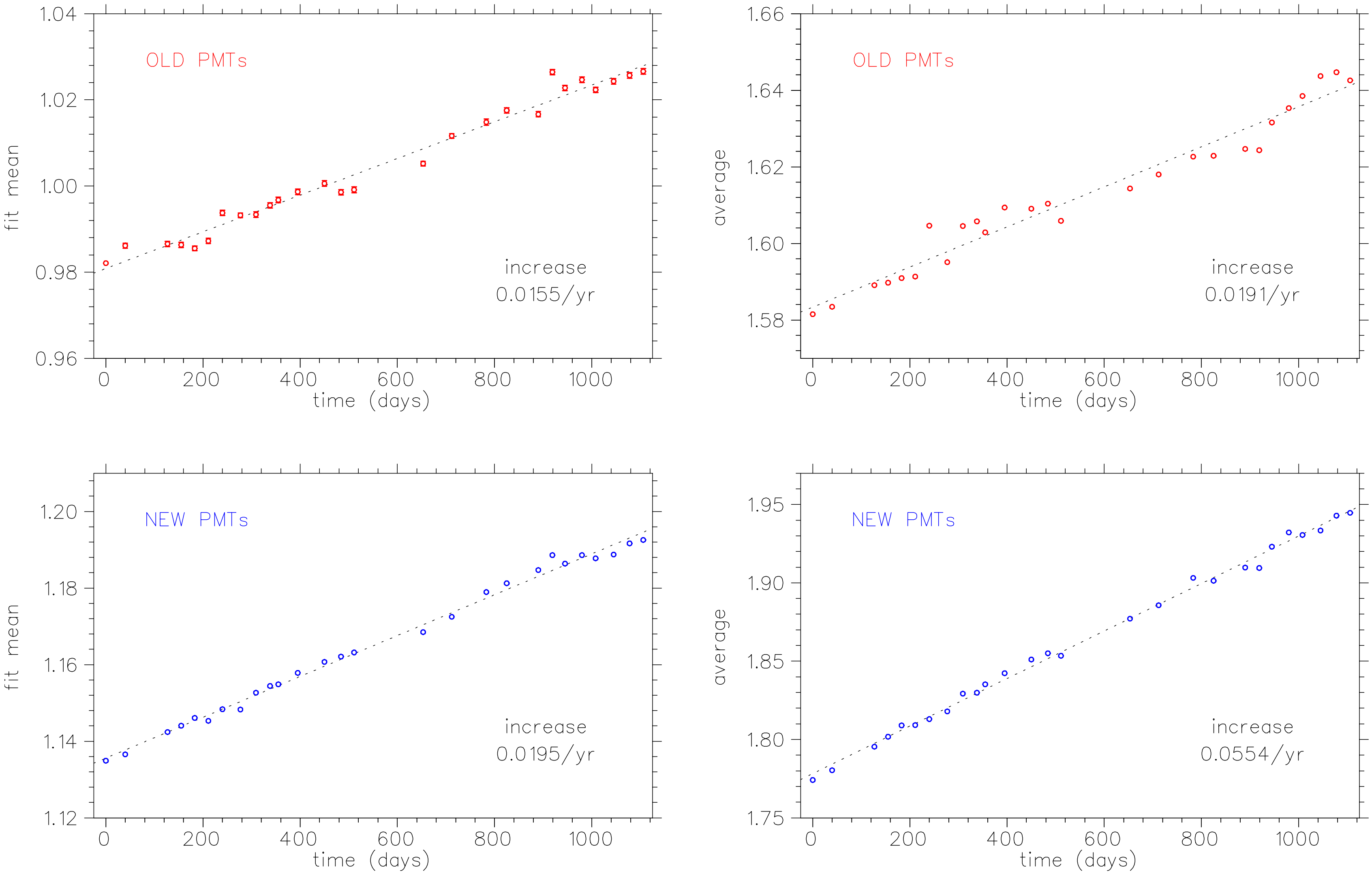


Data

fits

averages

Similar to previous plot, but **for dark noise events**. Note that the increase in the average value for new PMTs (lower right) is anomalously large. For more plots, see Tacik's presentation to the Calibration Group on October 1, 2014.

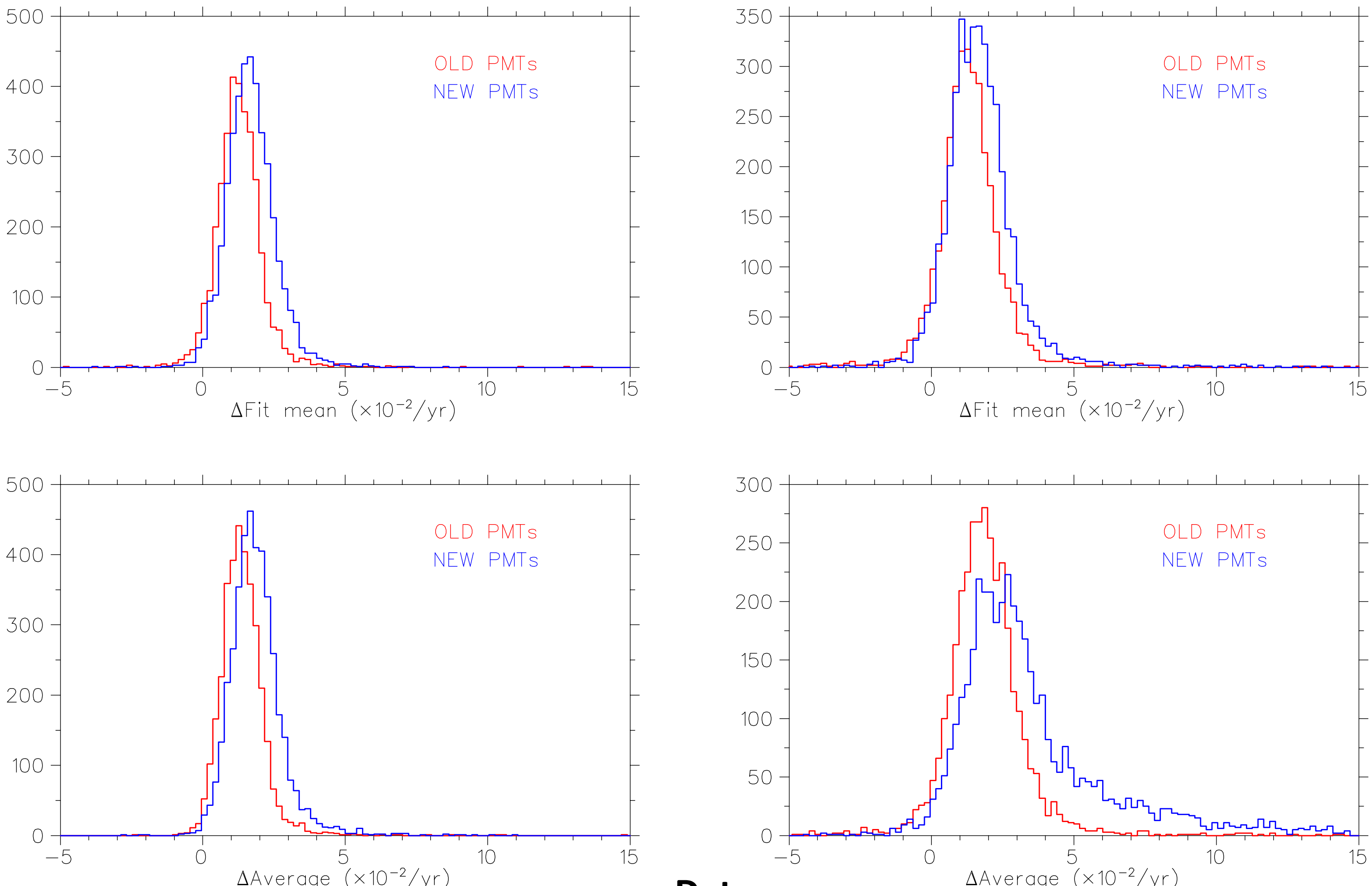


Data

fit

average

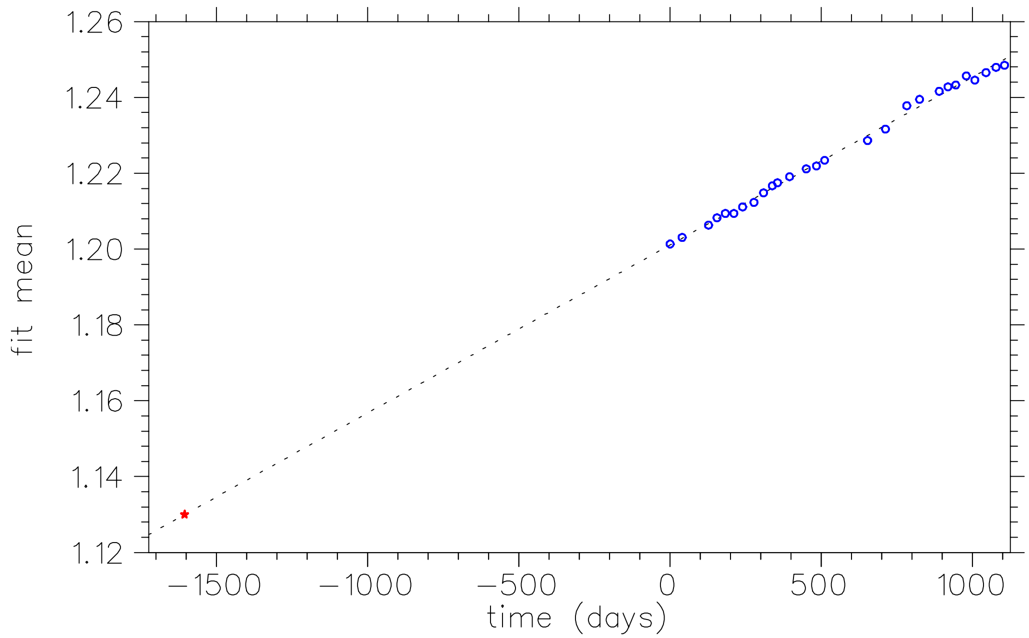
Can find the rate of change of each PMT separately, and plot distributions. The spread in values is larger for new PMTs than for old ones. The spread is larger for dark noise than for physics events, especially in average for new PMTs.

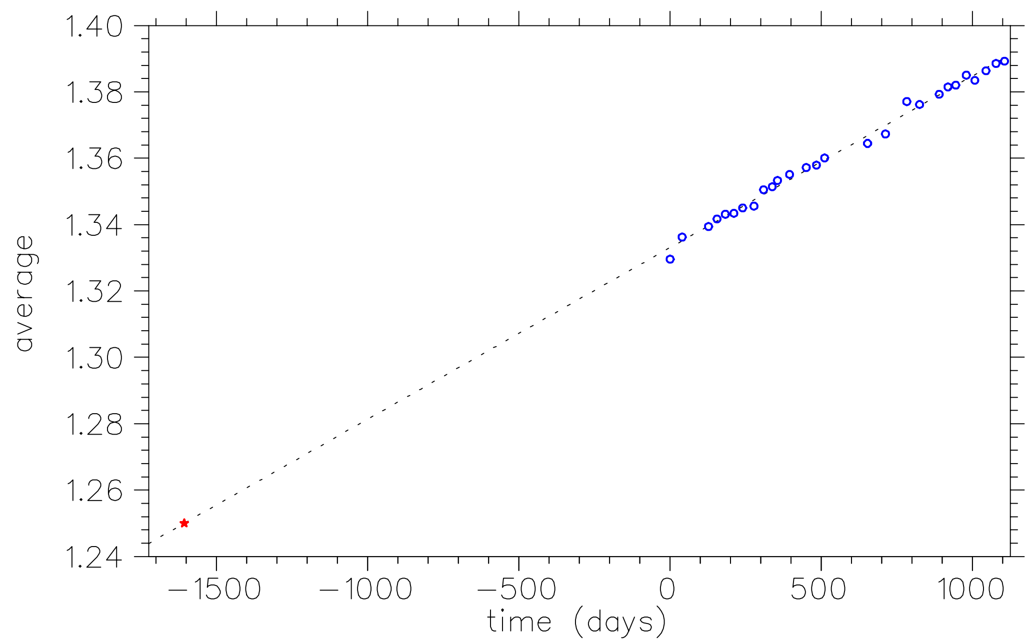


Data

physics

dark noise





This shows what happens when one extrapolates the fits for new tubes.

Top plot: using charge means. Line fit to data equals MC value of 1.13 at

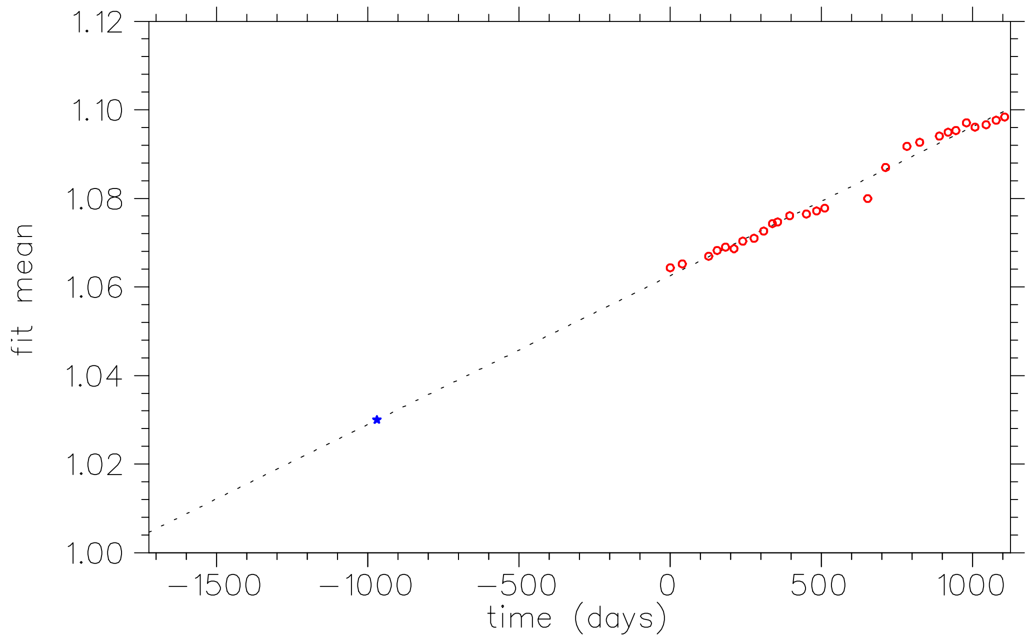
t = −1606 ± 21 days

Bottom plot: using average charge. Line fit to data equals MC value of 1.25 at

t = −1607 ± 29 days

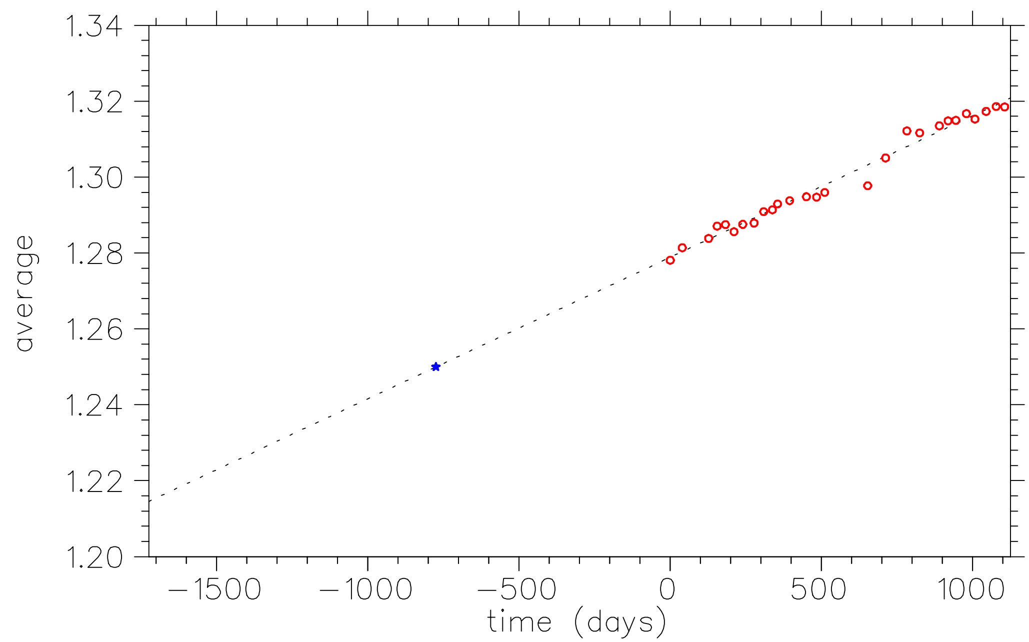
Extraordinarily good agreement!

Date in the skdetsim MC routine skrn1pe_sk3.F is: Jan 22, 2007, which corresponds to day -1675 when day 0 is Aug 24, 2011.



This shows what happens when one extrapolates the fits for old tubes.

Top plot: using charge means.
Line fit to data equals MC value of 1.03 at
t = −970 ± 30 days



Bottom plot: using average charge.
Line fit to data equals MC value of 1.25 at
t = −775 ± 28 days

Poor agreement.