Counting vs. fitting method to estimate PSF containment radius and its error.

J. Conrad
KTH-Stockholm, Particle Physics Section, AlbaNova University Center, SE-10691 Stockholm

Abstract

In preparation of the CERN beam test, the Perugia group (S. Germani et. al.) and B. Lott have used two different methods to estimate the quantiles (specifically 68 % and 95 %) of the instrumental Point Spread Function (PSF). The Perugia method is based on calculating the quantiles directly from the histogram of space angle deviations (counting method) whereas Benoit Lott's method (fitting method) is based on estimating the quantile from fitting a function (proposed by Toby Burnett) to the data and estimating the quantiles from the fit results. This short note summarizes a study of some of the statistical properties of the two methods. It is concluded that the two methods are equivalent if the data is well fitted by the fit function. Furthermore, statistical errors in both methods are estimated with reasonable precision even if the fit is worse. This implies that both methods can be used to estimate the required statistics for a given precision.

1 The two methods

1.1 The fitting method

Toby Burnett (LAT-AM-4355) showed that the PSF can be reasonably well parameterized by following function:

$$f(u) = \left(1 - \frac{1}{\gamma}\right)\left(1 + \frac{u}{\gamma}\right)^{-\gamma} \tag{1}$$

where $u = \frac{1}{2}\frac{r^2}{\sigma^2}$ and r being the angular deviation between true and reconstructed direction divided by an energy dependent scaling function. This function will be referred to as *Burnett function* in the remainder of this note.
The idea of the fitting method is to fit this function to the data and then to retrieve the containment radius (we will assume the 68 % containment quantile throughout this note and call the containment radius Θ_{68}) analytically from the fitted function. The error on this radius can be calculated from the covariance matrix of the fit parameters.

The counting method calculates the quantiles of the data distribution, which gives the containment radius at any confidence level. The error on the Θ_{68} is given by:

$$\Delta\Theta_{68} = \frac{\Theta(N_{68} + \Delta N_{68}) - \Theta(N_{68} - \Delta N_{68})}{2} \tag{2}$$

giving as measurement result $\Theta_{68} \pm \Delta\Theta_{68}$. N_{68} is the number of events which fall within Θ_{68} and $\Delta N_{68} = \sqrt{N_{tot} \cdot 0.68 \cdot (1 - 0.68)}$ (binomial error).

2 Procedure

To compare the two methods, the following procedure is used:

- Fitting method
  - Step 1: generate a data histogram by drawing random numbers from the Burnett function with $\gamma = 2$ and $\sigma = 1$.
  - Step 2: fit the Burnett function to the generated histogram and calculate the estimate of Θ_{68} and the estimate of the error on Θ_{68} from the covariance matrix.
  - repeat step 1-2 N_{exp}^{pseudo} times and store the estimates of Θ_{68} and its error.
- Counting method
  - Step 1: generate a data histogram by throwing random numbers according to a Landau function. The parameters of the Landau function are obtained from a fit to a representative data histogram provided by the Perugia group.
  - Step 2: calculate Θ_{68} from the data histogram by counting (as described above).
  - Step 3: repeat step 1-2 N_{exp}^{pseudo} times and store the estimates of Θ_{68} and its error.

Throwing according to Landau represents a histogram of scaled Θ (according to Perugia procedure), whereas throwing according to the Burnett function represents a reweighted histogram (according to Benoit Lott's procedure).

In both cases the function which generated the data is known , meaning the true value of Θ_{68}. Thus, the bias of the method can be calculated. Furthermore, the mean of the error estimates should be similar to the rms of the distribution of the estimates of Θ_{68}.

It should be noted that we use the same functional form for fitting the data histogram as was used to generate the histogram, which implies that the Burnett function yields a good fit to the data. Benoit Lott's presentation on the 2nd Pisa workshop (16/17.5 2006) shows that this is not necessarily true (the effects of having non-Burnett function contributions will be illustrated in section 4). Figure 1 illustrates a typical fit used in the presents study.

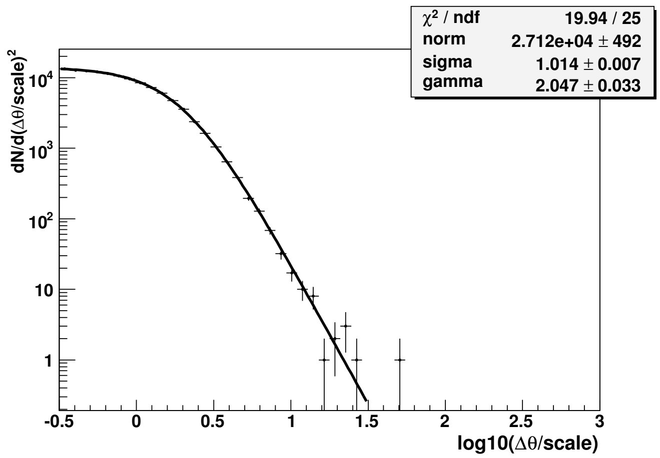

Fig. 1. A typical distribution with the fit of the Burnett function

3 Results

Table 1 summarizes the results on estimate, true and expected error as well as bias. These results are obtained using 10000 pseudo-experiments with 1000, 10000 and 100000 events in each PSF.

Both the counting method and the fitting method give consistent results. The relative error scales as expected as $1/\sqrt{counts}$ and the relative error presented here is comparable to what has been presented by the Perugia group and by Benoit Lott. Both methods give reasonably exact estimates of the statistical error. The bias of both methods is smaller than 1 %. The bias of the Perugia method (.5 %) seems to be independent of statistics, indicating a systematic effect (a guess could be the binning).

Examples of the distribution of error estimates and the difference between the true and estimated values of Θ_{68} can be found in figures 2 and 3 for the fitting method and in figures 4 and 5 for the counting method. The histograms are based on 1000 pseudo-experiments with 10000 events in each PSF.

	σ_{est}	σ_{true}	rel. error	rel. bias
Fit				
1k	0.14	0.13	0.05	0.01
10k	0.04	0.04	0.01	0.0007
100k	0.013	0.013	0.004	0.005
Count				
1k	0.0289	0.0288	0.049	0.005
10k	0.0089	0.0088	0.015	0.005
100k	0.003	0.0028	0.005	0.005

Table 1
Results of the study on the fitting (Fit) and counting (Count) method. The first column contains the number of events in the histogram, the second column the rms of the distribution of estimates, σ_{true}, the third column is the relative error and the last column is the relative bias. This results were obtained using 10000 pseudo-experiments.

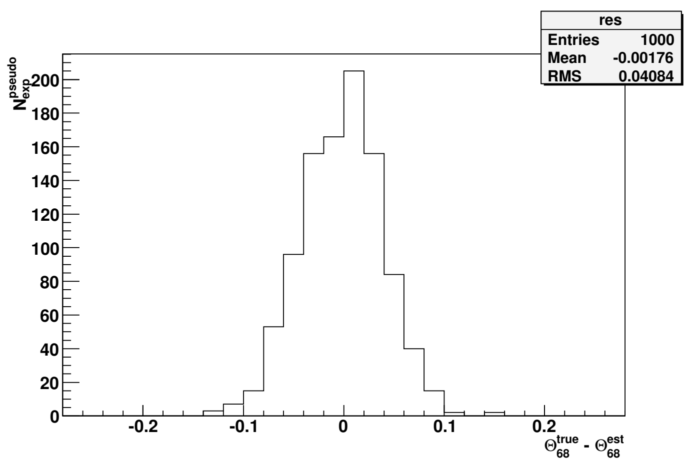

Fig. 2. Example of the distribution of the difference between the true and the estimated Θ_{68} for the fitting method (1 k pseudo-experiments, 10 k events in each PSF histogram)

4 What if a bad fit is obtained ?

In order to illustrate the effect of obtaining a bad fit with the Burnett function, in the generation step an additional Landau-shaped contribution is added, which will not be accounted for in the fitting. This mimics the excess of

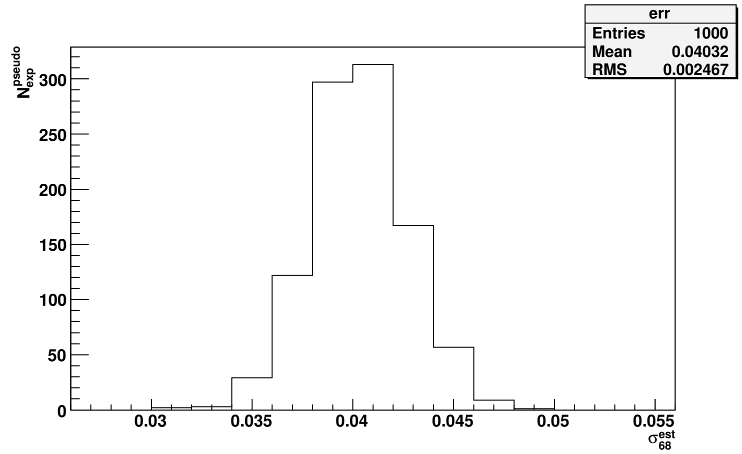

Fig. 3. Example of the distribution of the estimated errors for the fitting method (1k pseudo-experiments, 10k events in each PSF histogram.)

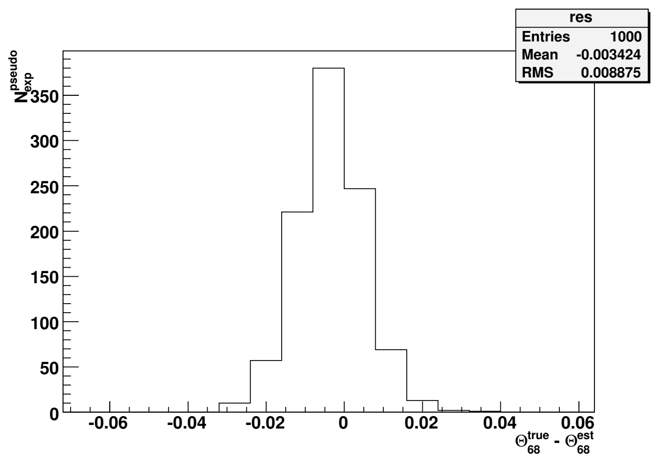

Fig. 4. Example of the distribution of the difference between the true and the estimated Θ_{68} for the counting method (1 k pseudo-experiments, 10 k events in each PSF histogram)

large deviations seen in the data. The total integral for this Landau contribution is adjusted to roughly yield a χ^2 comparable to what was presented by Benoit Lott (1 % of the contribution of the Burnett function, a MPV of 0.4 in $\log_{10} \Delta\Theta/scale$ was used). For this quick illustration we simulate 1000 pseudo-experiments with 10000 events in the PSF. An example of a resulting fit is given in figure 6.

Figure 7 shows the distribution of the difference between true and estimated Θ_{68}. It can be seen that the use of the Burnett function as a fitting function in this case causes a $\sim$ 10 % bias on the estimate. However, the rms of the distribution and the mean error estimate agree to within 6 % , which is probably good enough for the purpose of estimating required statistics. The relative error (1.7 %) can be compared with the one found for a perfect fit (1.4 %, obtained from 10000 pseudo-experiments), which seems reasonably consistent.

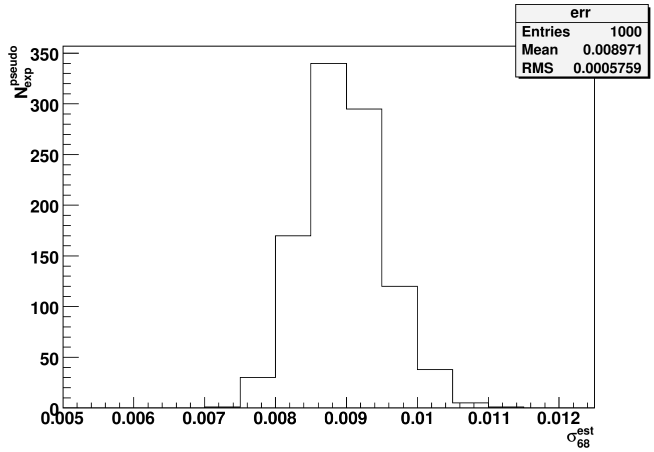

Fig. 5. Example of the distribution for the estimated errors for the fitting method (1k pseudo-experiments, 10k events in each PSF histogram.)

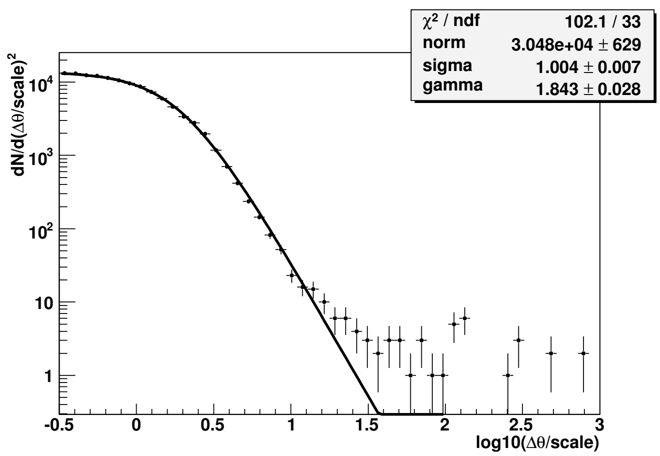

Fig. 6. A histogram generated by a Burnett function plus a Landau with MPV= 0.4. The fit of the Burnett function is also shown.

5 The 95 % containment radius

For the 95 % containment radius, the same arguments apply as for the 68 % containment radius. However, if (as seems to be the case in data) the Burnett function gives a relatively bad fit (due to a contribution at higher deviations which is not accounted for in the fit, see previous section), the relative bias increases to $\sim$ 30 %. Figure 10 shows the absolute bias for the 95 % radius for both a perfect and a bad fit.

For the estimate of required statistics this does not matter. The mean error estimate and the rms of the distribution still agree quite well in both cases. The relative error is a little larger than in the 68 % case but increases only negligibly due to the bad fit (from 4 % in the perfect fit case to about 5 % in

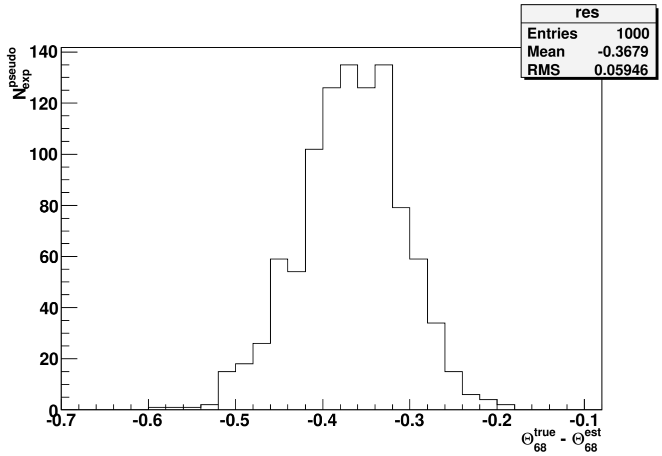

Fig. 7. Example of the distribution of the differences between the true and the estimated Θ_{68} for the fitting method (1 k pseudo-experiments, 10 k events in each PSF histogram) in the case that the Burnett function gives a relatively bad fit to the data.

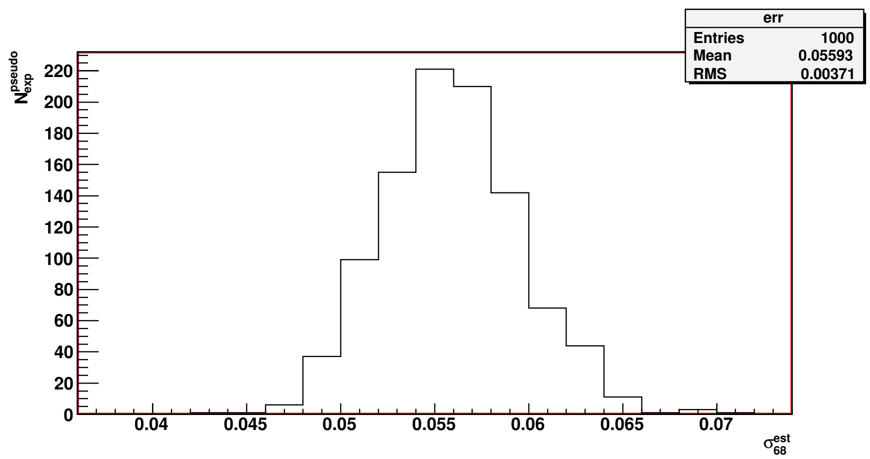

Fig. 8. Example of distribution for the estimated errors for the fitting method (1k pseudo-experiments, 10k events in each PSF histogram) in the case that the Burnett function gives a relatively bad fit to the data.

the bad fit case).

6 Conclusions

From the presented results it is concluded that the two methods are statistically equivalent. Both methods give reasonably correct estimates of the error of Θ_{68}, and can thus be used to estimate the required statistics for a given precision. In the case of the fitting method, this is true even if the fit is relatively bad.

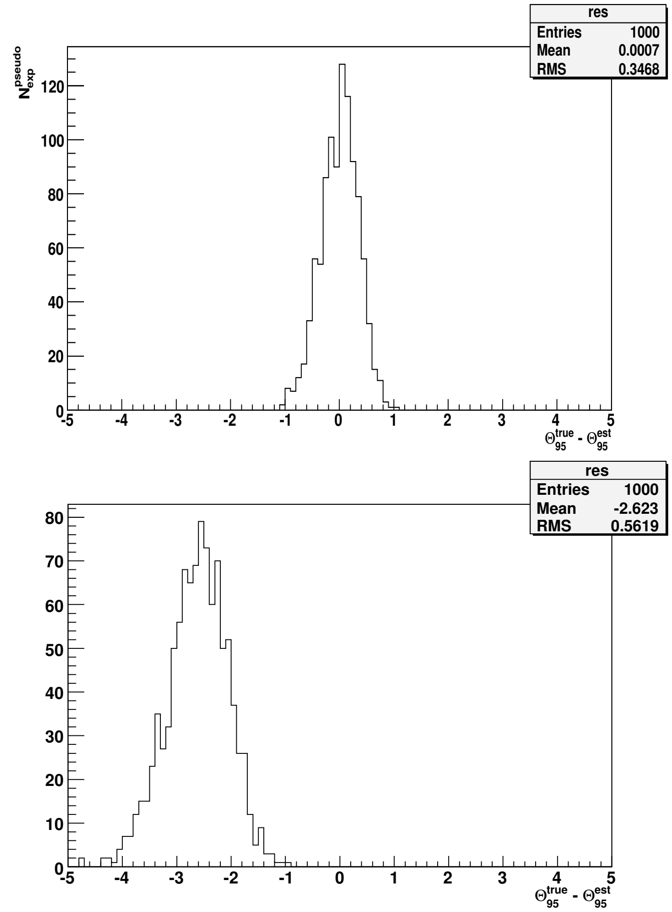

Fig. 9. Example of the distribution of the differences between the true and the estimated Θ_{95} for the fitting method (1 k pseudo-experiments, 10 k events in each PSF histogram) in the case that the Burnett function gives a good fit to the data(upper panel) and in the case that the Burnett function gives a bad fit to the data (lower panel)

The bias of the counting method is neglible, probably only due to binning effects. If the correct fitting function is used also the fitting methods yields very small bias. However, not surprisingly, bad fits can cause significant bias in the estimate. If (as seems indicated in the data) an excess of large deviations (w.r.t the Burnett function) are present, the bias increases for higher quantiles of the containment radius.

7 Acknowledgments

Thanks to S. Germani and B. Lott for providing explanations and access to their macros.

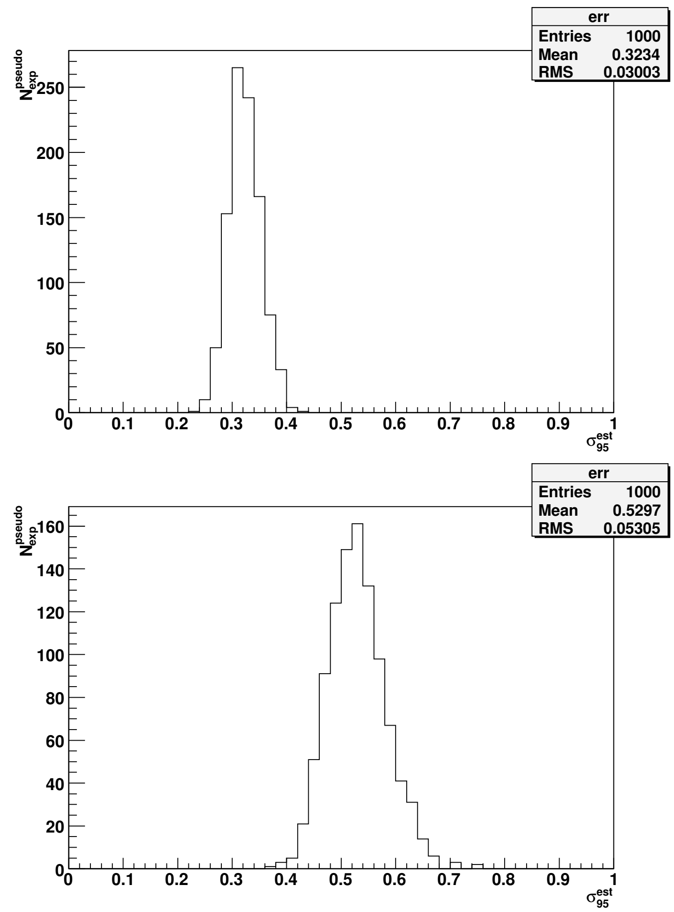

Fig. 10. Example of distribution for the estimated errors (95 % containment) for the fitting method (1k pseudo-experiments, 10k events in each PSF histogram) in the case that the Burnett function gives a good fit to the data (upper panel), and in the case that the Burnett function gives a bad fit to the data (lower panel)