A Core Fitting Algorithm Using the Inner Outriggers

Tyce DeYoung

November 27, 2001

Abstract

A new core fitter that uses the 66 outriggers currently installed is presented. The fitter relies on neural networks to decide whether the core is on the pond, near the pond, or distant. Appropriate algorithms are then used to determine the core location. Some improvement over the current core fitter is observed in simulations, but less than with simpler proposed algorithms using the full outrigger array. Nevertheless, the increased information may warrant implementing this fitter until the complete array is installed.

1 Introduction

A core fitting algorithm that may be appropriate for short-term use is presented here. It is intended to utilize the 66 outriggers presently installed to improve the accuracy of the core fitter (and thus hopefully the accuracy of angular reconstruction as well) while the installation of the remaining outriggers is completed. The algorithm is envisioned for use in a reconstruction chain in which γ/h separation is performed first, followed by core fitting and then angular reconstruction. The algorithm thus is optimized for γ cores, and avoids the use of any information related to the angular fit (t_χ's, for instance).

The inner outriggers inhabit a rough ring some 20–30 m thick around the edge of the pond; in any given direction there are only two or three outriggers between the pond and the fence. For most showers, therefore, there is relatively little information contributed by the outriggers; the most probable number of outrigger hits in an event is only five. More importantly, the outrigger ring is not thick enough to contain any shower core, in the sense of seeing a clear peak in amplitude vs. radius. The overall improvement in performance is thus modest, and is certainly much less than will be possible with the full outrigger array.

All results presented are based on γ Monte Carlo v. 3.0. Some 13,620 showers with a trigger threshold of 50 hits in the top layer were used. They are isotropic with a $E^{-2.4}$ (i.e., Crab-like) spectrum. The first 6,810 showers were used to train the neural networks, and the second half were used as an evaluation sample. The results shown contain only the evaluation sample, so that results are essentially free of training bias[1].

2 Description of the Algorithm

At the Santa Cruz meeting in October, results from a 'hybrid' core fitter using the full outrigger array were presented. The hybrid fitter used relatively simple parameters to decide whether a given core was on the pond, off the pond but within the outrigger field, or outside the outrigger field. Different core fitters were used for each of these classes of showers. Based on the success of that method, a similar approach was used for this fitter. However, because less information is delivered by the smaller outrigger array, it

[1] The neural nets used in the final version were the ones that produced minimal error in the evaluation sample, so there is a second- or third-order selection effect. For complete independence, a third set would be needed. C'est la vie.

was necessary to use two neural networks instead of simpler parameters to make the classifications. The neural networks provide comparable accuracy of classification at the cost of increased CPU time.

The first stage of the algorithm is a network which decides whether the core is on the pond or not. This network was kept small to reduce the required CPU time. There are three inputs: the fraction of the top layer tubes hit, the fraction of the outriggers hit, and the summed amplitude in the bottom layer. The topology of the network is 3–10–5–1, for a total of 85 connections to be computed[2]. The performance of this network is shown in Fig. 1. Cores for which the network output is greater than 0.9 are passed to the existing FFT pond core fitter, which has the best performance of the existing fitters for cores on the pond.

A second neural network classifies showers that are not sent to the FFT routine as either nearby (i.e., with the core inside the outrigger field, about 60 m from the center of the pond) or distant. This is a more difficult classification for the computer, and so a more extensive 7–10–10–1 network is used. This is a total of 180 connections, more than twice the size of the first network. The seven inputs are the three parameters used for the first network, the largest amplitude in the outrigger array, the summed amplitude in the outrigger array, the summed amplitude of hits larger than 5 photoelectrons in the bottom layer, and the radius of the center of gravity of the outrigger hits. Even so, the two classes of events (nearby and distant) are less well separated than the on- and off-pond events classified by the first network. The larger amount of CPU time required for this network is compensated by the fact that the remainder of the algorithm is quite fast.

Cores that are identified as on the outrigger field (above 0.85 in Fig. 2) are fit with the center of gravity of the outrigger hits. This position was found to correspond quite well with the core location in the case of the full outrigger array; with only 66 outriggers the correspondence is weaker, but still acceptable. Attempts to improve the quality of the fit by taking a weighted average of the outrigger and air shower centers of gravity, or by correcting the direction to the core with information from the air shower layer, were unsuccessful. Future work may produce improvements, though.

The showers remaining after the second neural net, some 65% of all triggers, are classified as distant. For these showers, the reconstructed core is projected a certain distance in the direction of the center of gravity of

[2] All networks used in this routine are fully-connected feed-forward back-propagation multilayer perceptrons. Each connection has a constant (bias) term plus a linear combination of inputs, and is activated by the sigmoid function $\frac{1}{1+e^{-x}}$.

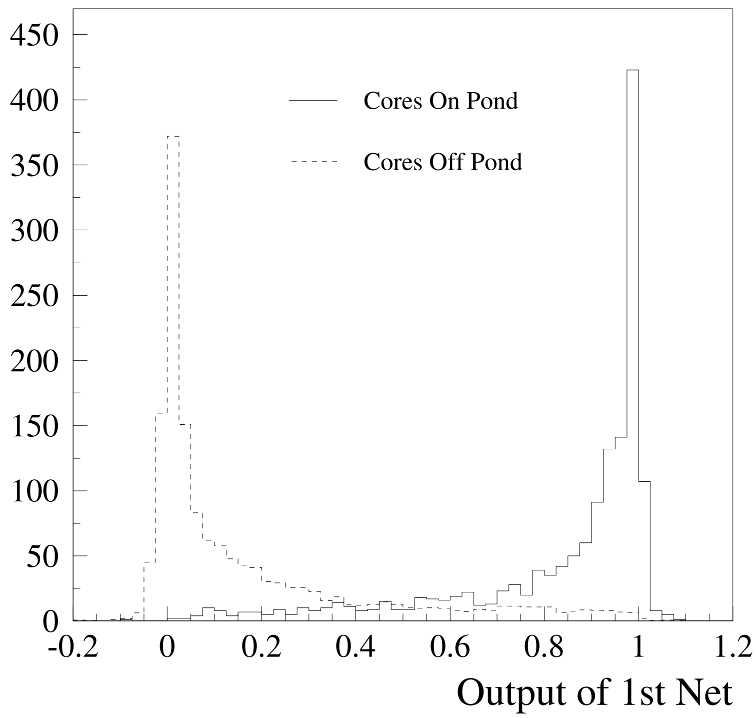


Figure 1: The output of the neural network that identifies cores on the pond. Cores truly on the pond are shown in solid, those truly off are dashed. Events with output above 0.9 are reconstructed by the FFT core fitter. The two distributions at left are normalized to an equal number of events; in absolute numbers, 15% of all cores are reconstructed by FFT.

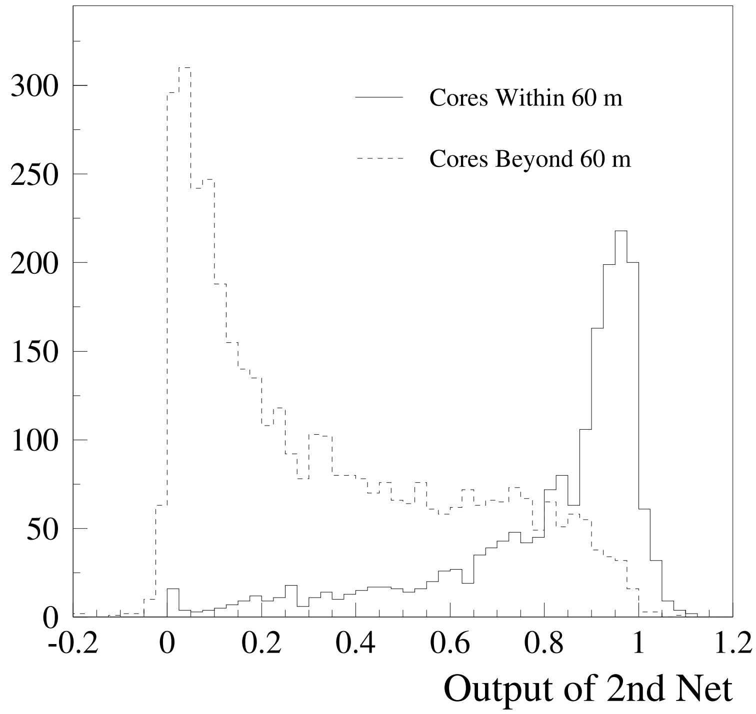


Figure 2: The second stage network classifies cores as lying within or without the area occupied by the inner 66 outriggers. Cores truly less than 60 m from the center of the pond are shown in solid, cores truly farther away are dashed. Events with output above 0.85 are classified as nearby and are fit with the center of gravity of the outrigger hits.

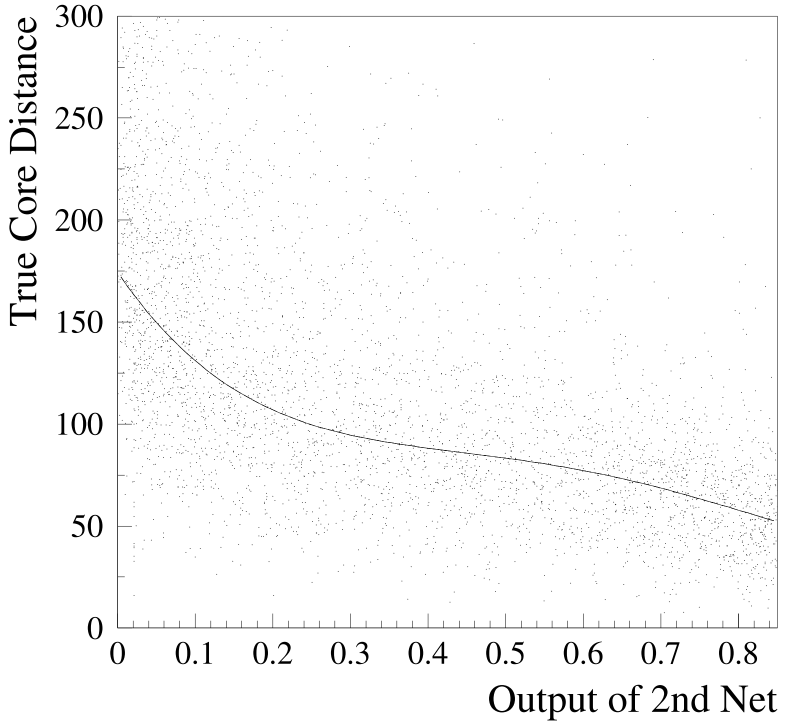
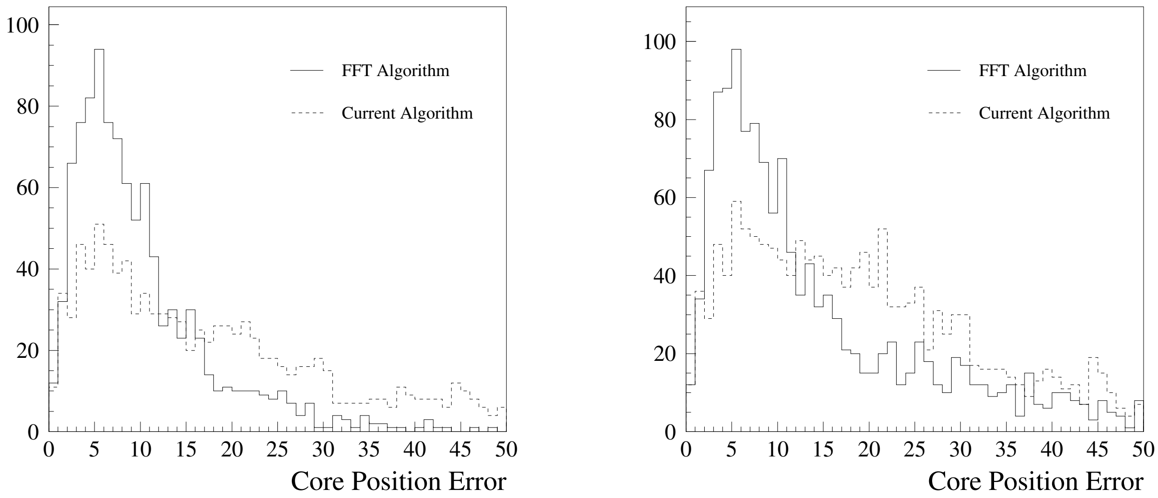


Figure 3: The polynomial function used to estimate the distance to the core based on the output of the second neural network. Dots represent individual γ showers, and the line is the polynomial estimate of the distance.

the outrigger hits from the center of the pond. The distance the core is projected is based on a 4th-degree polynomial fit to the most probable core distance as a function of the output of the second network. The fit is shown in Fig. 3, where the considerable spread in the distances clearly illustrates the difficulty of the problem.

3 Performance

This core reconstruction scheme makes use of the existing Fast Fourier Transform algorithm for cores tagged as "on-pond." This algorithm is extremely accurate, as shown in Fig. 4, with a most-probable error of 5 to 6 m and 80% of cores reconstructed to within 15 m of the true position. The



(a) Cores tagged as on the pond.

(b) Cores truly on the pond.

Figure 4: Accuracy of the FFT fitter, which is called for cores that are identified as landing on the pond. The dashed lines show the accuracy of the current core fitter on the same sets of events, for comparison. On the left the classification algorithm of the new fitter is used to identify on-pond cores; this introduces a small selection bias. Results for all cores truly on the pond are shown at right.

performance of the current core fitter on the same events is also shown in Fig. 4.

For cores that are not on-pond but that are identified as nearby, the center of gravity of the outrigger hits is used as a core location. The accuracy for these events is shown in Fig. 5. While the improvement is not so great as with on-pond cores, the information provided by the outriggers does improve the overall accuracy.

Fig. 6 shows the accuracy of reconstruction for cores identified as distant, for which a distance is estimated and the core projected in the direction of the outrigger COG. The overall accuracy is quite poor (though as shown in Fig. 6(a) there is a slight increase in cores fit with reasonable accuracy) so that good angular reconstructions of these cores may not be possible. In

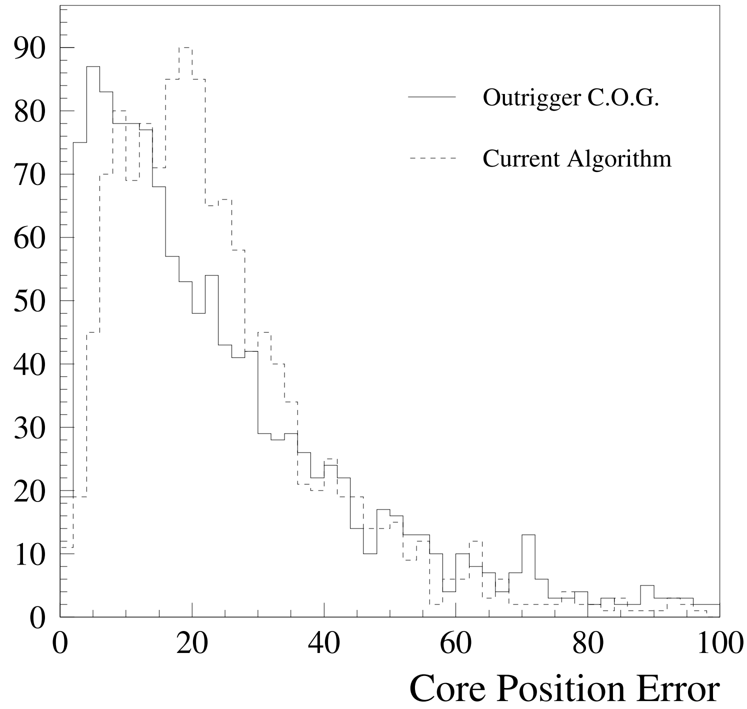


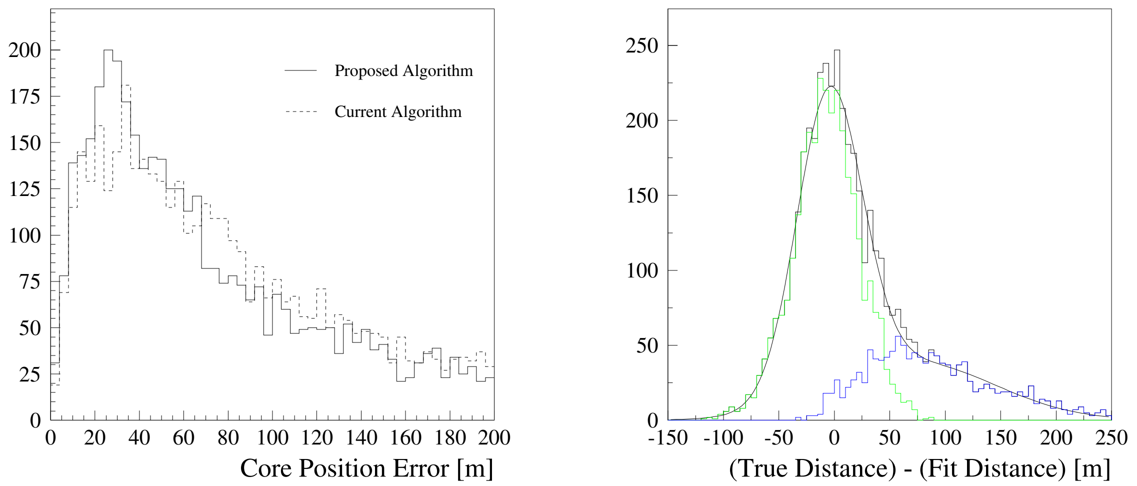
Figure 5: Accuracy of the outrigger center of gravity, which is used for cores that are identified as landing near but not on the pond. The dashed line shows the accuracy of the current core fitter on the same set of events, for comparison.

fact, the RMS error increases with the new scheme from 82 m to 123 m; this is due to the fact that the new scheme, by projecting cores farther from the pond than the maximal 50 m under the current fitter, can produce larger total errors if the direction to the core is incorrect. Once the error is off the scale of Fig. 6(a), however, it makes little difference what the actual error is.

However, the estimate of distance from the pond to the shower core (that is, the radial component of the fit) is good to about 30 m for cores in the vicinity of the pond (less than 150 m away), as shown in Fig. 6(b). This is because at the extreme edge of a very high energy shower, the hits are very splotchy — the irregularity of the pattern signals that the core is very distant, but there is no preferred direction. The simple fact that the core is distant, though, may be good enough for use in γ/hadron separation algorithms or energy fits.

4 Conclusions

Overall, this hybrid scheme does give some improvement in the accuracy of core reconstruction, as shown in Fig. 7. The improvement is not as large as is expected from the full outrigger array. However, even with only a few outriggers in place, we are able to do a much better job of distinguishing cores on the pond from those near the pond and those far away. Even without extremely accurate core positions for distant events, this information may allow us to do a much better job of energy reconstruction and γ/hadron separation.



(a) Total error in core position.

(b) Radial component of the core position error.

Figure 6: Reconstruction accuracy for distant cores. The overall accuracy is poor, but the estimated radius is good to about 30 m. The radial error is well fit by a double Gaussian, the larger of which (in light green) represents the precision for cores in the general vicinity of the pond ($R_{true} < 150$ m), and the smaller (in blue), offset from zero, contains extremely distant cores whose radius is strongly underestimated.

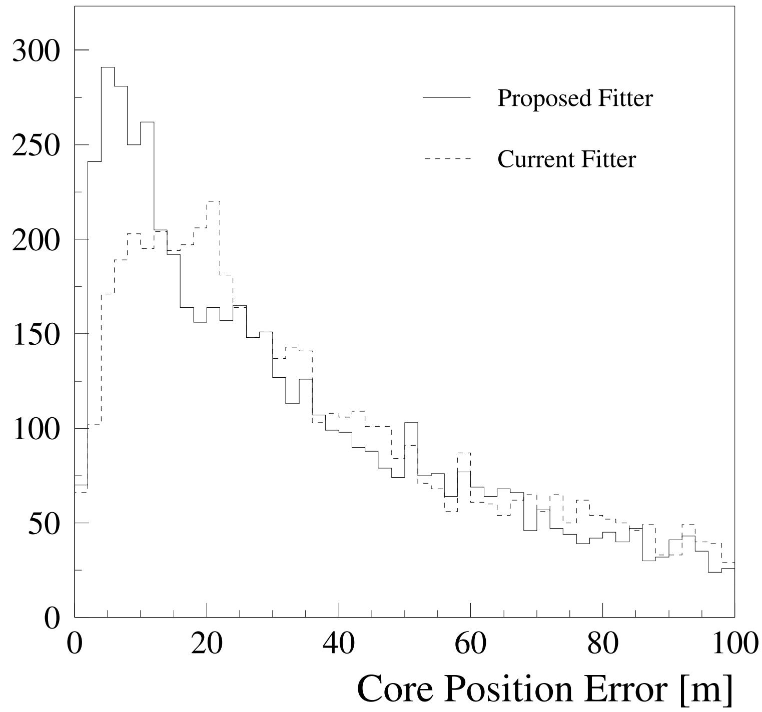


Figure 7: Accuracy of the proposed core reconstruction scheme (solid) compared to that of the current core fitter (dashed).