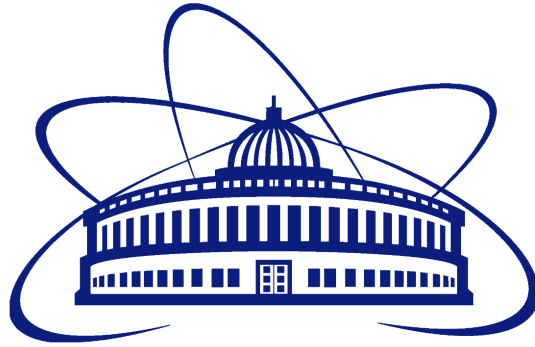JOINT INSTITUTE FOR NUCLEAR RESEARCH

Veksler and Baldin laboratory of High Energy Physics

FINAL REPORT ON THE SUMMER STUDENT PROGRAM

Optimization of tasks in the event reconstruction procedure for the BM@N experiment

Supervisor: Sergei P. Merts
Student: Sofia P. Lebedeva
Russia
St. Petersburg University
Participation period:
July 10 – August 18

Dubna, 2023

Content

Abstract

The BM@N (Baryonic Matter at the Nuclotron) is an experiment at the NICA (Nuclotron-based Ion Collider fAcility) accelerator complex. The first physics runs were carried out with the collection of experimental data in 2018. After the collection, all data is reconstructed. To improve the results of reconstruction and reduce noise signals, the problem of optimization of the threshold value was solved during this work.

Introduction

The mega-science project NICA (Nuclotron-based Ion Collider fAcility) is being created at the Joint Institute for Nuclear Research in Dubna. It is aimed at creating a new accelerator facility for conducting experiments to study the interactions of relativistic nuclei and polarized particles. The goals of NICA cover many areas of high energy physics, including the study of the properties of dense baryonic matter.
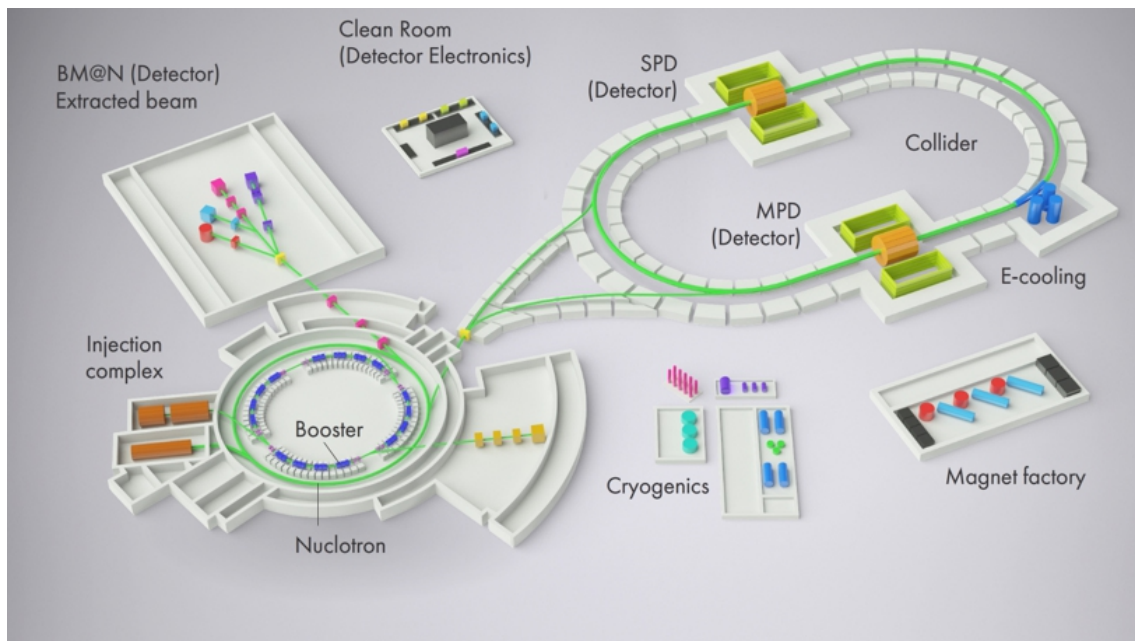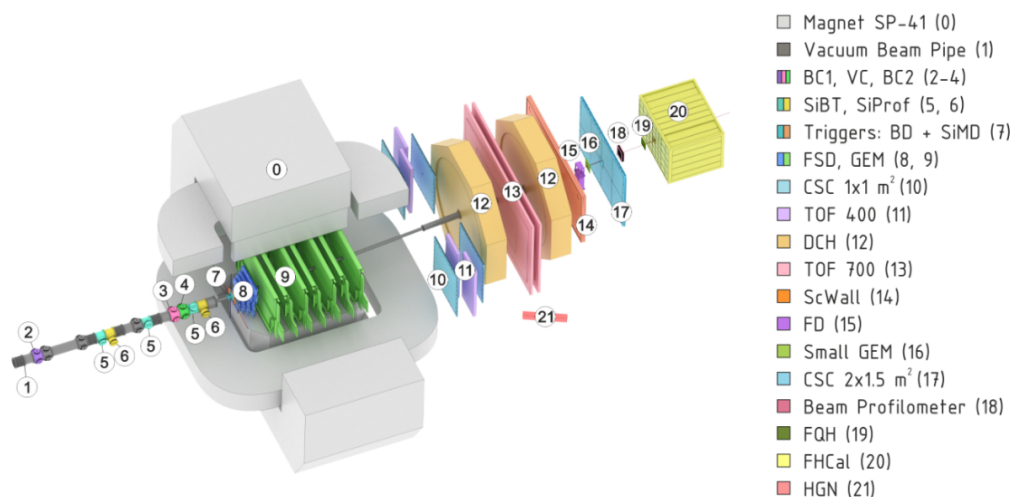

Figure 1: A schematic view of the NICA accelerator complex.

The first experiment at the NICA accelerator complex is the BM@N (Baryonic Matter at the Nuclotron) experiment, which specific goal is

studying the interaction of relativistic heavy ion beams with fixed targets in the energy range from 2 to 6 GeV/nucleon. It includes collecting particle trajectory data and time-of-flight data that is required to identify particles and find their momentum, charge, and other characteristics.

Figure 2: A schematic view of BM@N experimental setup in the last run.

In the experimental setup, particles accelerated at the Nuclotron move to a fixed target through a vacuum tube. The target and various detectors are located inside the SP41 magnet. Inside the tube, the particle first encounters these detectors which notify other detectors in the system about the particle. After colliding with the target, the particle splits into many new particles that form a beam. It moves in the magnetic field through various strip detectors that measure the positions of particles and the time passing the detector. At the end of its path, the particles collide with a thick lead calorimeter that measures their energies. The data obtained after reconstructing their paths through the strip detectors and from the calorimeter is enough to know all the most important parameters of the particle.

Optimization of the threshold value through signal maximization.

Formulation of the problem

One of the external detectors outside the magnet in the setup is three Cathode Strip Chambers or CSC. There are two small CSC and a large one named LCSC. We consider the last one. Like all strip detectors, LCSC's modules consist of layers of parallel strips. Also, this chamber is divided into 8 modules, and each of them is also divided into hot and cold zones. Such separation is used to the strips inside that zones can read the signals more accurately.

The vacuum tube also passes through the entire volume of CSC and when the beam moves through it, the strips inside the chamber collect its data. After that, the received data is digitized and then it is ready for further work. There are not only signals from the beam that we need among the received data, but also random ones — or the noise. To reduce the number of such noise signals, there are set *threshold* for each strip. Below such a threshold, the signals are "cut", which means that the signal from that strip is not recorded. For example, if we set this threshold to zero we will get a very strong noise, and the necessary information will be difficult to extract from it. And on the other hand, if we raise such a threshold too high, the necessary data may be lost. Therefore, the problem is to choose the optimal threshold for each strip in the CSC.

Problem solving method

The mathematical abstraction that describes a particle in a beam at each moment is called a *track*, which means its trajectory. The track is described by the *state vector*, which contains the values of the X and Y coordinates, the values of the tangents of the slopes of the momentum components $t_x = p_x/p_z$ and $t_x = p_y/p_z$, the values of charge-to-momentum ratio or q/p, and the coordinate of the vertical plane Z implicitly. If such a vector is known at least at one point in space then by solving the equations of motion in a magnetic field, we can find a similar vector at any other point using the Kalman filter.

For our task, we will look for the state vector of each track in the plane, which contains *hits*. The hit is the reconstructed three-dimensional point at which the particle crossed the detector. Each track will be propagated to this plane named Z_0 using the Kalman filter. The analyzed values will be the distances dx and dy from the position of each track in this Z_0 plane to each hit on it along the x and y axes, and all of this for each module.

After collecting this data, we will build histograms on them named "All-to-All" - or all tracks to all hits. For a Large Cathode Strip Chamber, we will get 16 histograms with the number of distances dx and dy for each of the eight modules. Obviously, in that case, distances for most of the track-hit pairs will be false and the hit does not really fit the track in question. The number of such distances will be almost the same for all tracks, so this data form the *background*. The true distances from each track to its real hit will have a Gaussian distribution, and we can see it on the resulting histograms. The resulting histogram will look like the Fig.3:
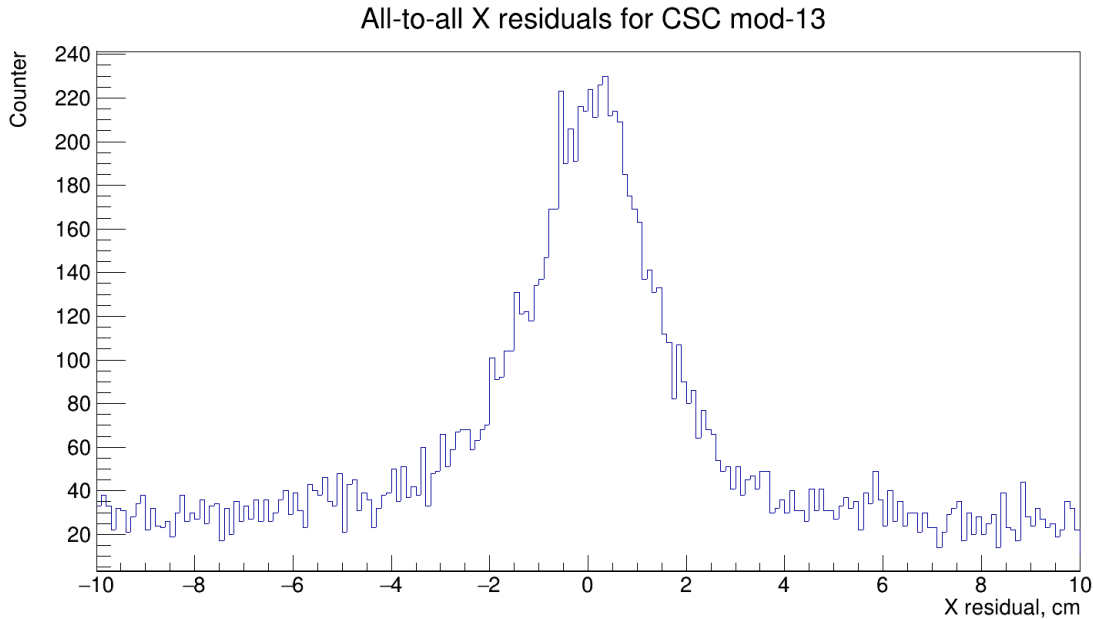
Figure 3: All-to-all histogram for the 13th module.

The background is approximated by a simple polynomial of the second degree using the special macro. And the peak which can be seen in the picture, is fitted by the Gaussian distribution. This method was originally used for the alignment procedure since in this way we can find the distance to the center and move the detectors this distance to get the center at zero according to the theory. But after the alignment procedure is over, the method can be used for selecting thresholds.

For example, for a low threshold, we expect that there will be many false data and the background will be high, and the Gaussian peak opposite will be small. For a high threshold, on the contrary, there will be a low background and a high peak, but the signal itself will be weak. In order to evaluate which threshold will be the best, we will consider three values. The first one is *Signal* (S), or the area under the graph for the Gaussian peak, excluding the background. The second one is *Signal-to-Background ratio* (S/B), which is calculated by dividing the integrals of the two fitting functions of Signal and Background, respectively. And the third one is the *Significance* of the signal against a given background, which is calculated as: $\frac{S}{\sqrt{S+B}}$. We will look for the maximum for each of these values, depending on the value of the set threshold.

Results

As a result, data was obtained for several threshold values, and All-to-all graphs were drawn from them for each module of a Large CSC (№8-15). There are such graphs for module number 13 and for the X coordinate in Fig.4, but similar pictures were obtained for other modules too.
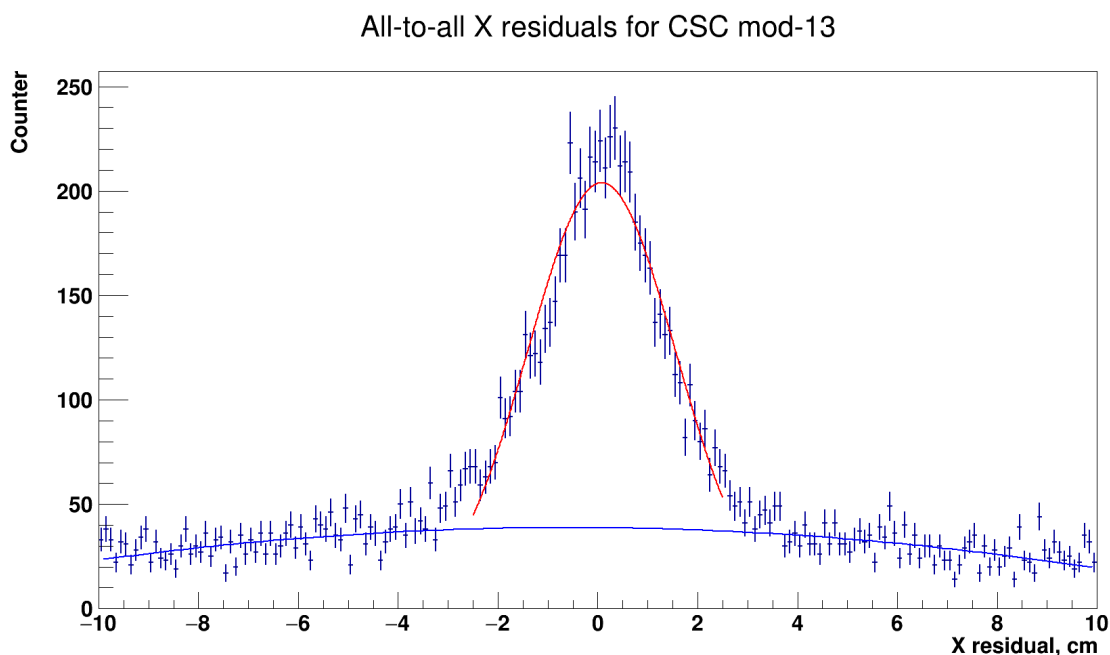

Figure 4: Fit of All-to-all histogram for the 13th module.

Using the above method, the values of S, S/B, and Significance can be founded for each threshold. The next step was to plot a new graph for each value versus the threshold and find its maximum. Thus, we have 3 graphs for each of the X and Y coordinates, and for each of the 8 modules: a total of 48 diagrams are obtained. Fig.5 shows several of them.

Now the standard value of the total threshold set is 30. At a threshold equal to zero, the signal also tends to zero, since the Gaussian peak of the signals is almost invisible against such a background. With a strong increase in the threshold, the peak is always clearly visible, but the signal itself becomes a little weaker. All this can be seen in the graphs.

For all 48 graphs, the maximum of each value was reached at different thresholds. But most often it was at a value between 15 and
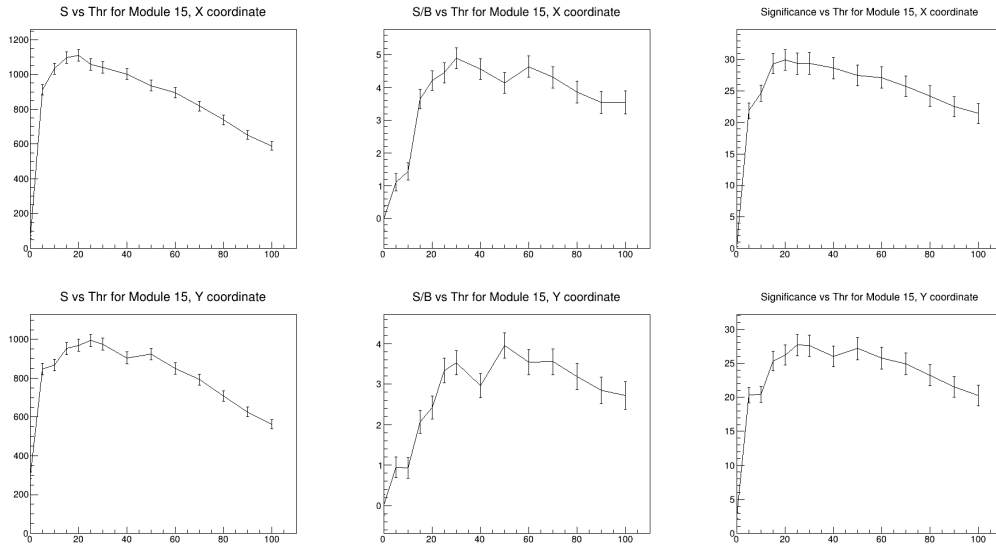
Figure 5: Plots for maximization the signal, module 15

25, which is quite close to the current threshold. For this reason, the points between these were additionally considered (Fig.6), and finally, we find that it would be best to set the threshold to 20 at this stage.
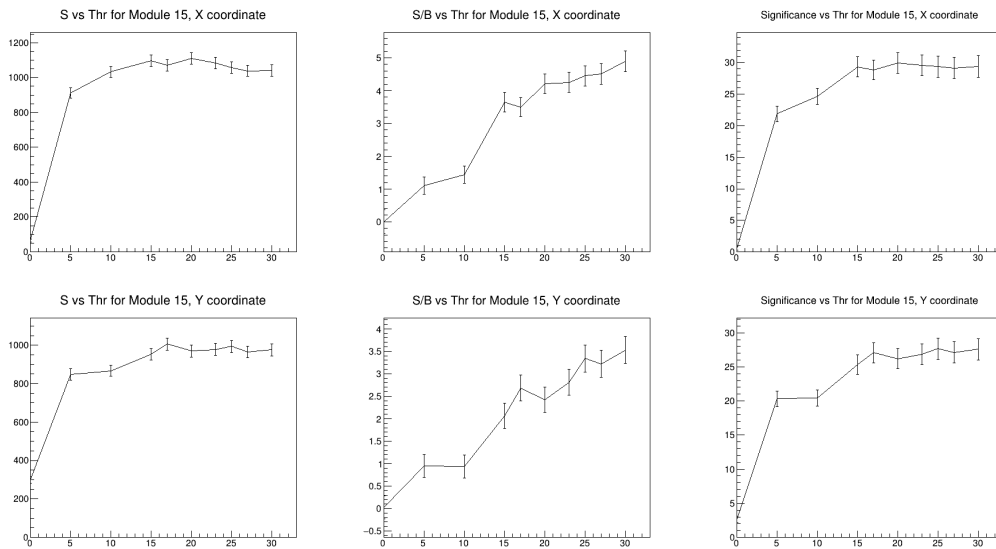

Figure 6: Plots for maximization the signal, module 15

Now we can compare X residuals for a certain module with the default threshold value and for a new one found during the work. This was done in Fig.7.
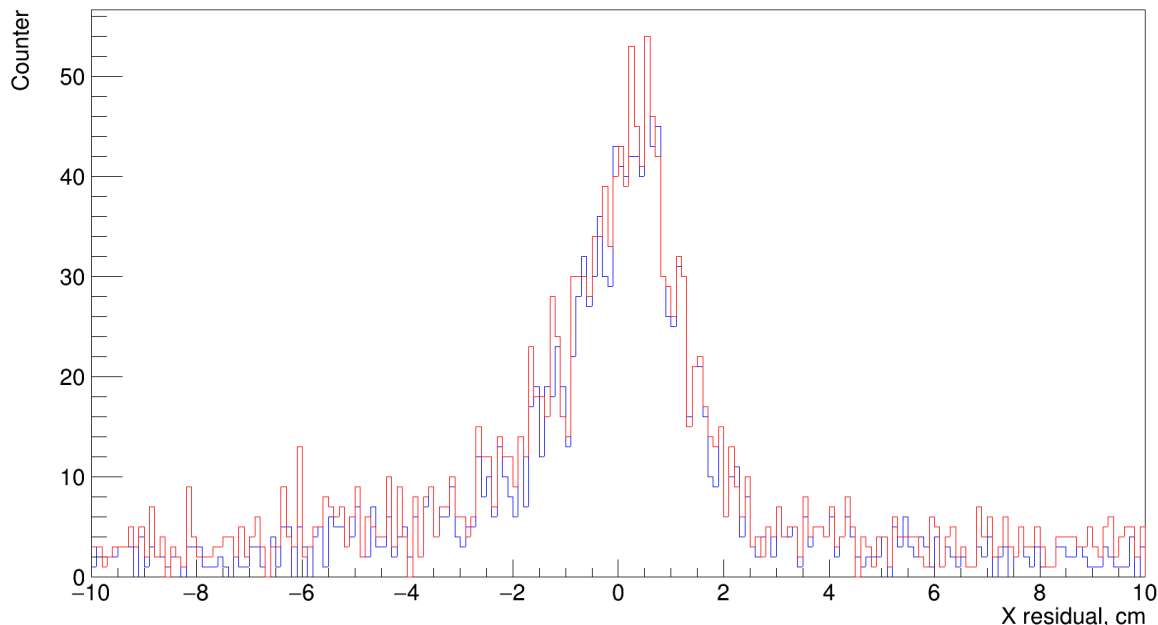
Figure 7: ompare old (blue line) and new (red line) all-to-all histogram for X residuals, module 15

The signal on the new histogram is indeed slightly larger than on the old one, but the difference is not that big. This happens because the optimal threshold was quite close to the default one.

Optimizing the value of the threshold common to the entire camera through signal maximization is only the first stage of the study. In the future, it is planned to select the optimal threshold value separately for each CSC module and each strip inside them.

Optimization of the Magnetic Field Map.

During the participation period, in parallel with the main task, two small additional tasks were also completed.

Theory overview

As mentioned above, one of the main components of the BM@N experimental setup is the SP-41 large-aperture dipole analyzing magnet. Before conducting a direct experiment, a *Magnetic Field Map* was measured inside the magnet. The Field map is a six-dimensional array that includes three components of the magnetic induction B_x, B_y, B_z, and three-dimensional coordinates: X, Y, Z. For such an object, its optimization is also relevant. Previously, I have already improved the Field map by the correcting wrong areas on it, which were caused by a technical error during the recording of data for the map. Now the task of optimizing the field map has appeared again.

Optimization process

Despite the fact that the map contains spatial coordinates directly, in reality, they are not used anywhere. Instead, they are calculated from the maximum and minimum values of each of the coordinates and their number, which are known in advance. Therefore, to reduce the size of the field map, it was proposed to remove spatial coordinates from it, leaving only the magnetic field components.

For this purpose, the classes *BmnFieldPoint* and *BmnFieldMap* was rewritten. The object of the first one contains all the points from the map, and in the second one, the Field map is read. This optimization resulted in a map size reduction of about 8%. To further reduce this size, the data type for its points has been changed from *Double* to *Float*. Together with the previous improvement, this resulted in a size reduction of almost 40%.

Another optimization task was the smoothing of the Field map, especially of its By component. Since the *Smooth* method from ROOT is only available for 1D or 2D histograms, and not available for 3D, which we have, the map was smoothed using a *moving average.*

The way the map was smoothed was simple. For each i-th point in space where the field components were measured, the new value of $B_y^{(i)}$ was calculated as:

$$B_{y,1}^{(i)} = \frac{B_y^{(i-1)} + B_y^{(i)} + B_y^{(i+1)}}{3} \tag{1}$$

After that, this value was recorded in place of the old point in place of the i-th point. This procedure was done for all three components of B. As a result, we got a new Magnetic field map, and it is noticeably smoother. A comparison of the old and new maps can be seen in Fig.7. An old map is drawn on it in blue, and a new improved one is drawn in red.
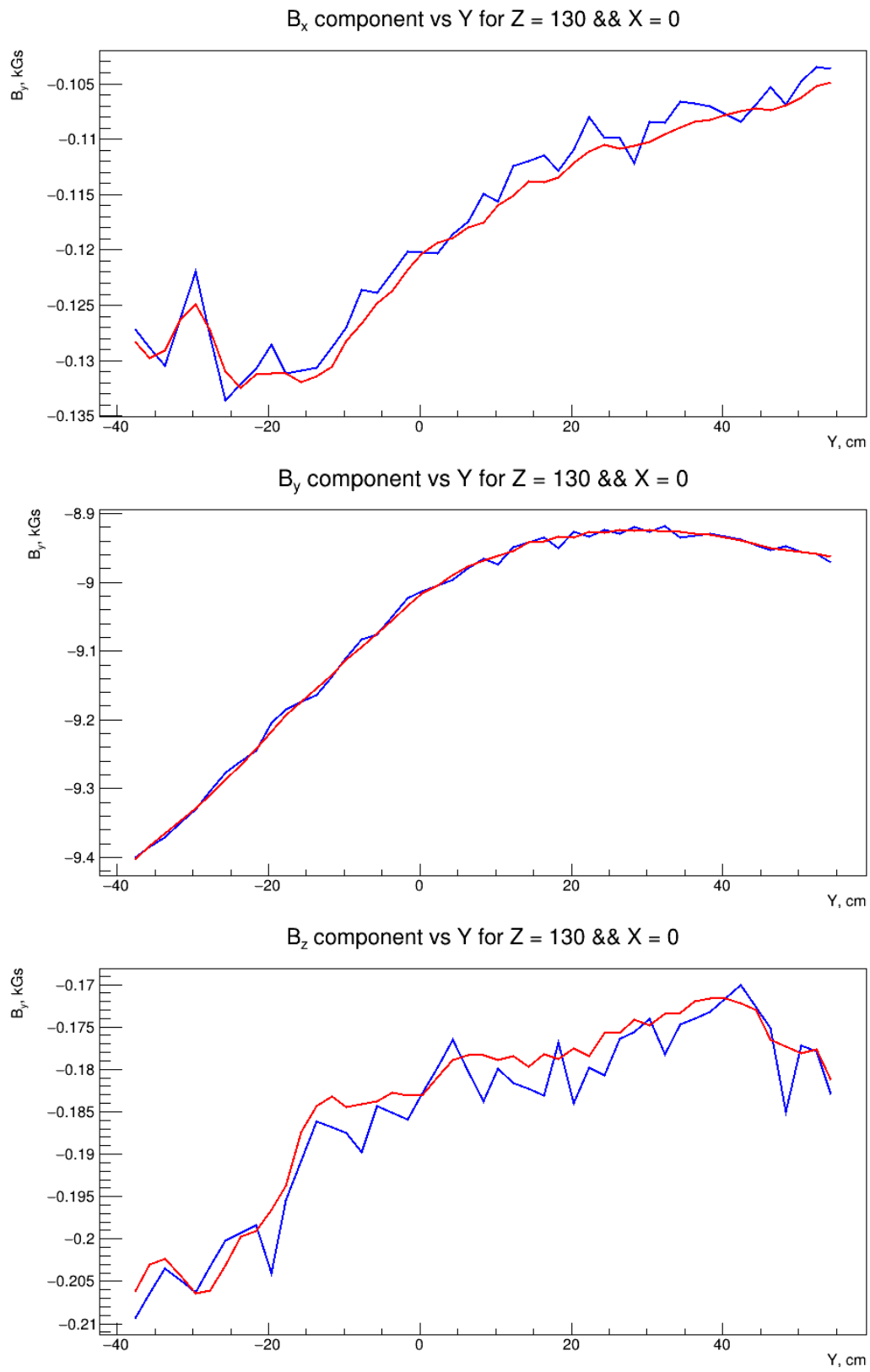
Figure 8: Comparison of the old and new Field maps by the B_y component

Summary and future

To sum up all the topics mentioned, a few conclusions might be done. For the threshold optimization problem, the main of them is that the search is not over. For different modules, as well as for strips within them, the optimal threshold value can be different. In the future, it will be necessary to find each of them, possibly using other methods.

The method of smoothing the map does indeed give a visible result, but there are doubts about the correctness of its use. In the future, it will be possible to develop a new method, which will be more similar to the classic *Smooth*.