
CHAPTER 5

X-ray Beam Analysis

5.1 Introduction

The ability to accurately calculate the dose absorbed by a sample is vital to the reproducibility of experimental results in radiation damage research and to inter-compare in MX, as well as to develop protocols aimed at advising experiments on how to spread the dose and thus optimise the use of their crystals. As alluded to in the introduction (section ??), the extent of knowledge on some of the parameters in the diffraction experiment is limited, which results in uncertainty in the calculated dose values. These uncertainties are not quantified and hence RADDOSE-3D does not explicitly calculate errors on the dose values. The uncertainty in some experimental parameters, such as the crystal volume or the unit cell volume, only result in small errors in the aggregate dose metrics (average dose whole crystal, maximum dose, DWD). On the other hand, uncertainty in the beam profile can lead to significant errors. For this reason, a module in RADDOSE-3D was implemented to allow it to simulate MX experiments using experimentally measured X-ray beam profiles, which can be read by the program. However, the measured beam profiles require preprocessing to remove systematic errors resulting from the actual measurement before they can be used in RADDOSE-3D.

The work presented in this chapter describes the preprocessing of experimentally measured beam profiles and explores the errors that can arise at this stage in the dose calculation.

5.2 Processing aperture measurements

Prior to the measurements reported here (July 2013), users at Diamond Light Source (DLS) beamline I02 were provided with the full width half maximum (FWHM) values of the X-ray beam measured by a $10\ \mu m$ aperture (see below). To obtain a more realistic value of the FWHM it was recommended that $5\ \mu m$ be subtracted from the supplied FWHM values. This recommendation however was not supported by any systematic experiments or studies.

For this reason investigations were carried out to obtain a better estimate of the true beam profile to so that a more trustworthy value of the FWHM could be provided to users. This was the initial motivation for processing beam measurements made using aperture scans.

5.2.1 Experimental methods

Aperture scans to obtain X-ray beam flux measurements were carried out in collaboration with Dr. Carina Loble on the DLS I02 beamline. For these scans, a miniap device, a piece of steel with a $10\ \mu m$ diameter circular hole, is translated across the beam by remote control. The position of the aperture is recorded along with a measure of the current detected in a silicon diode (S3590-09 model purchased from Hamamatsu) on the detector shutter (`i_pin`) (?). The detector current is proportional to the beam flux (photons/second). First the ‘centre’ x and y position is found such that the diode reading at this position gives the highest current. The aperture is then translated $120\ \mu m$ in the negative horizontal direction. Measurements of the x and y positions and the diode reading are carried out at $2\ \mu m$ intervals as the aperture is translated across in the positive horizontal direction for a total of $240\ \mu m$ ($+120\ \mu m$ from the centre). A similar scan is then carried out in the vertical direction. Examples of data from the measurements are shown in Figure 5.1.

5.2.2 Deconvoluting the X-ray beam measurements

Theoretical introduction

The area of the aperture used on I02 is $(10/2)^2 \times \pi \approx 78.54\ \mu m^2$, so each diode reading at a position (x, y) is the result of the integral of the flux from a $78.54\ \mu m^2$ area surrounding the

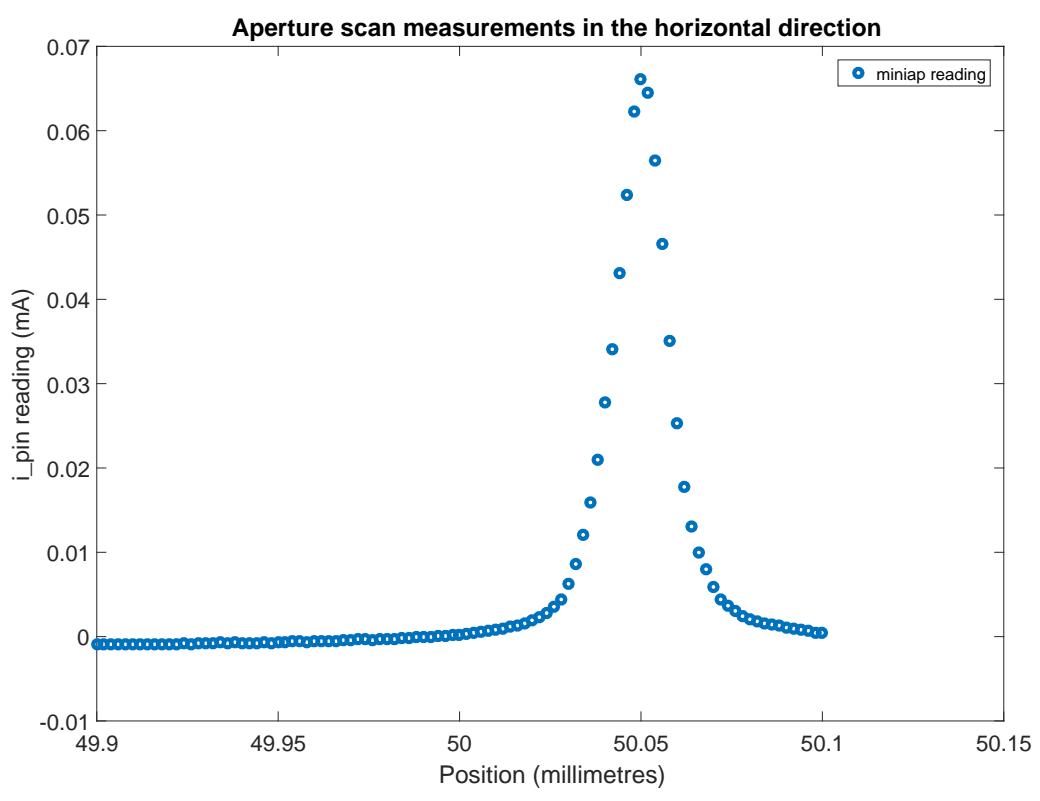
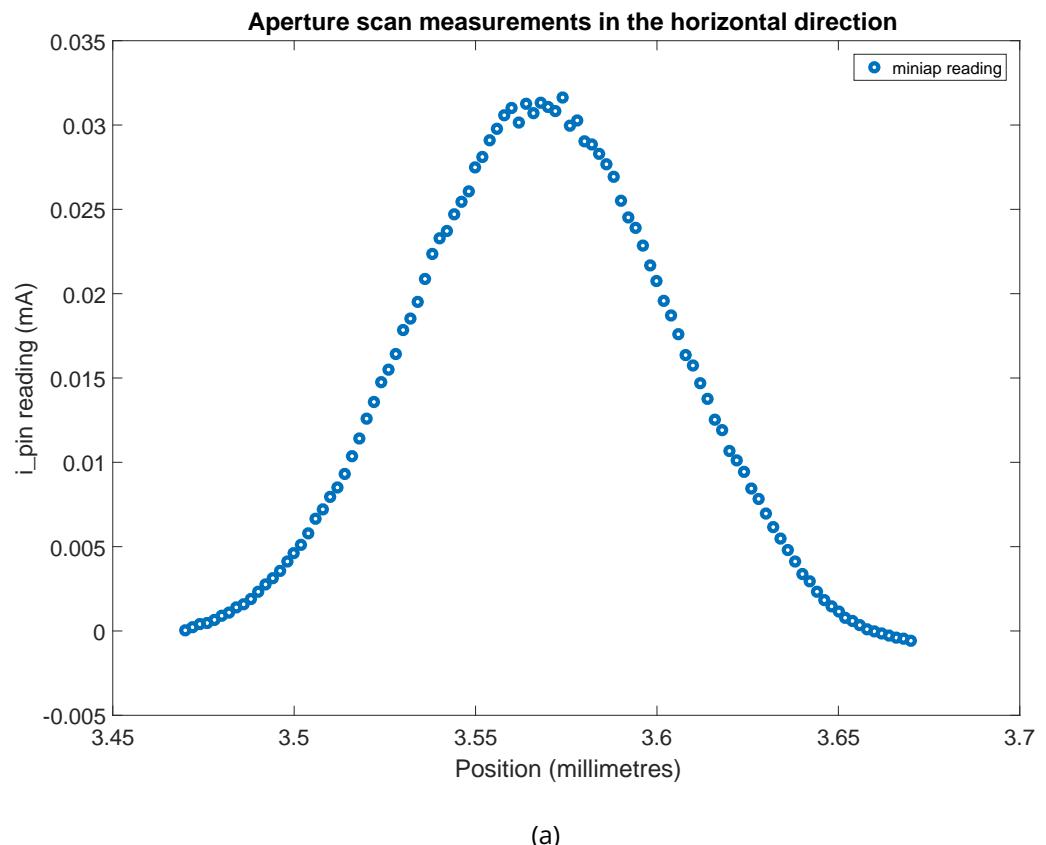


Figure 5.1: Examples of the flux measurement data collected at the Diamond Light Source synchrotron. The horizontal axis represents the aperture position in millimetres and the vertical axis represents the current in the diode (mA). (a) Data collected in the *x*-direction. (b) Data collected in the *y*-direction.

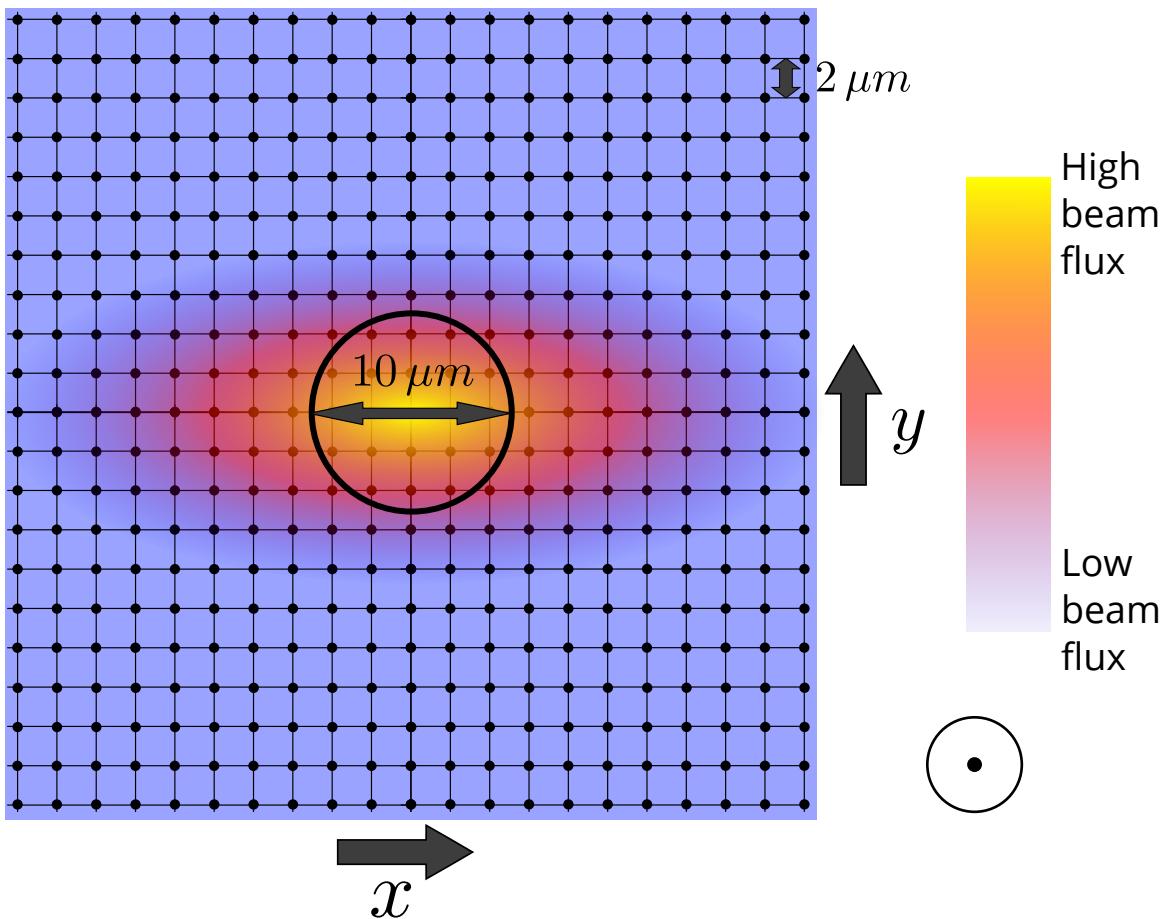


Figure 5.2: A schematic of the X-ray beam and aperture setup as viewed from the detector looking into the beam. Each point represents a spatial position where a beam measurement is taken, hence the distance between any two horizontally or vertically consecutive spots is $2 \mu m$. The circle represents the circular aperture which has a $10 \mu m$ diameter. The colour represents the beam flux intensity (not to scale). The beam is smaller in the y direction than it is in the x direction. This diagram shows that a reading at a particular point in space is the integral of all surrounding points within the aperture area. Thus the reading of the total X-ray beam profile that is measured is a convolution of readings from local surrounding space.

central point (Figure 5.2). Given that the diode measurements were taken at $2 \mu m$ intervals, it is possible to deconvolute the measured signal to obtain a truer value of the diode current to a spatial resolution of $2 \mu m$. Only this estimate of the true 2D profile of the current is necessary for the beam profile measurement. This is because RADDOSE-3D additionally requires a total flux estimate, which is distributed across the measured beam profile (a 2D array) according to the current at each spatial position.

To define the problem mathematically, the diode readings can be described as a convolution of the true diode current, f , at a point \mathbf{x} , and the area of the aperture, g . Explicitly this is

$$[f * g](\mathbf{x}) = \int_0^{\mathbf{x}} f(\mathbf{u})g(\mathbf{x} - \mathbf{u})d\mathbf{u} + n(\mathbf{x}), \quad (5.2.1)$$

where $[f * g]$ is the measured diode reading and $n(\mathbf{x})$ represents the additive noise at a par-

ticular point in space. The aim of deconvolution in this context is to find the diode current profile, $f(x)$, given knowledge of the measured current, $[f * g](x)$, and the aperture contribution, $g(x)$. Mathematically this aim can be interpreted as finding some $w(x)$ such that:

$$\hat{f}(x) = [w * [f * g]](x), \quad (5.2.2)$$

where $\hat{f}(x)$ is an estimate of $f(x)$ that minimises the mean square error. In general, deconvolution is an ill-posed problem, thus it is likely that a unique solution does not exist (?). If it is assumed that the noise has zero mean and is spatially independent (white noise assumption) then the Wiener filter (?) can be used to deconvolute the signal to find f . The Wiener Filter, W , is defined mathematically in the Fourier domain as:

$$W = \frac{1}{\mathcal{F}(g)} \left[\frac{|\mathcal{F}(g)|^2}{|\mathcal{F}(g)|^2 + \frac{\mathcal{F}(n)}{\mathcal{F}(f)}} \right], \quad (5.2.3)$$

where \mathcal{F} denotes the Fourier transform of its argument:

$$\mathcal{F}[f(x)](k) = \int_{-\infty}^{\infty} f(x) e^{ikx} dx. \quad (5.2.4)$$

The $\frac{\mathcal{F}(n)}{\mathcal{F}(f)}$ term in equation 5.2.3 is similar to the inverse of a signal to noise ratio in the Fourier domain. If this value is known exactly, then the Wiener filter often works very well. However, more commonly the value of this ratio is unknown, and it is often treated as a constant value throughout the domain.

The Wiener filter is multiplied with the convoluted function in the Fourier domain to give

$$\mathcal{F}[\hat{f}] = W \mathcal{F}[f * g], \quad (5.2.5)$$

such that $\mathcal{F}[\hat{f}]$ is the solution found with the minimum mean squared error:

$$\text{error} = E \left[(\mathcal{F}[f] - \mathcal{F}[\hat{f}])^2 \right], \quad (5.2.6)$$

where the true solution is denoted $\mathcal{F}[f]$. Multiplication in the Fourier domain is equivalent to a convolution operation in the real domain, making clear the equivalence between equation 5.2.5 and the statement of the mathematical aim (equation 5.2.2). Parseval's theorem implies that minimising the mean squared error in the Fourier domain is equivalent to min-

imising the mean squared error in the real domain. For further details on the Wiener Filter and its formal derivation the reader is referred to González and Woods (1992).

Implementing the deconvolution

The deconvolution algorithm was implemented in Matlab R2012a. First the initial diode measurements were read into Matlab and a one-dimensional Gaussian function was fitted to the data in order to obtain the parameter values in both the x and y directions (Figure 5.3). The parameter values were subsequently used for the two-dimensional Gaussian function representing the estimate of the convoluted two dimensional beam profile (Figure 5.4a).

$$g_{2D}(x, y) = A_{max} \exp \left[- \left(\frac{(x - \mu_x)^2}{2\sigma_x^2} + \frac{(y - \mu_y)^2}{2\sigma_y^2} \right) \right], \quad (5.2.7)$$

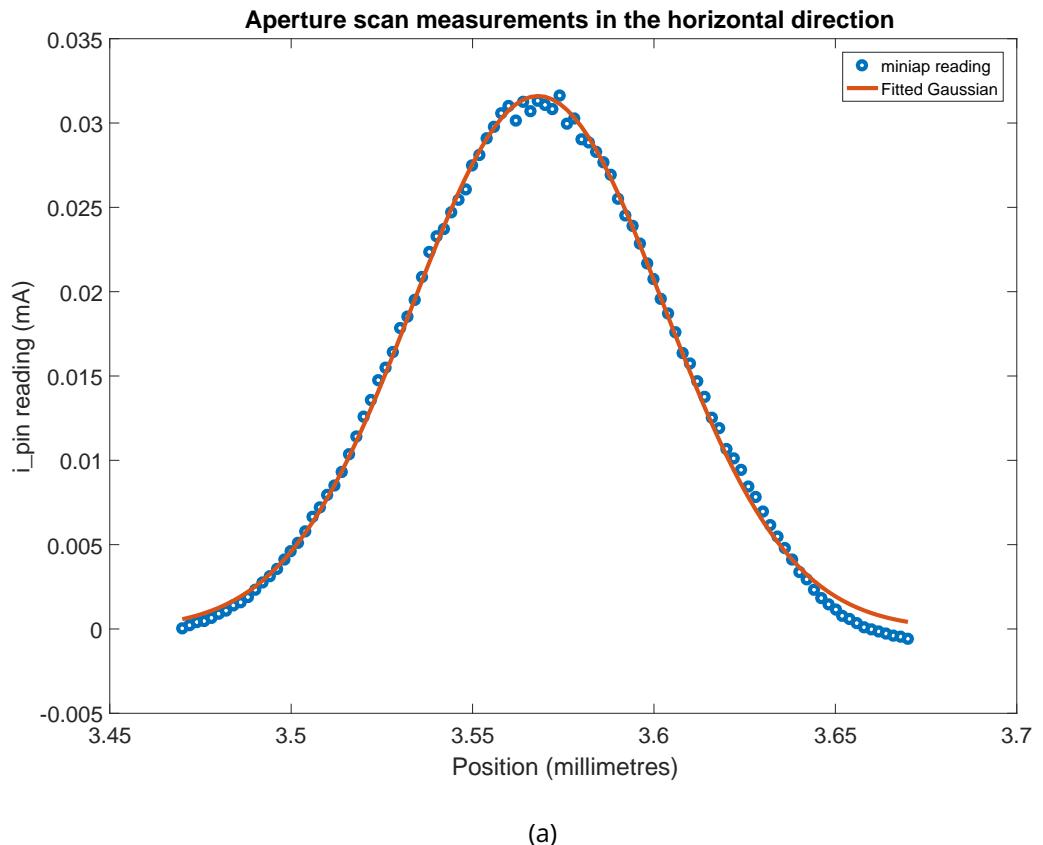
where μ_x , μ_y , σ_x and σ_y are the parameters whose values are obtained from fitting the one-dimensional Gaussian functions

$$g_h(x) = A_x \exp \left[- \frac{(x - \mu_x)^2}{2\sigma_x^2} \right], \quad (5.2.8)$$

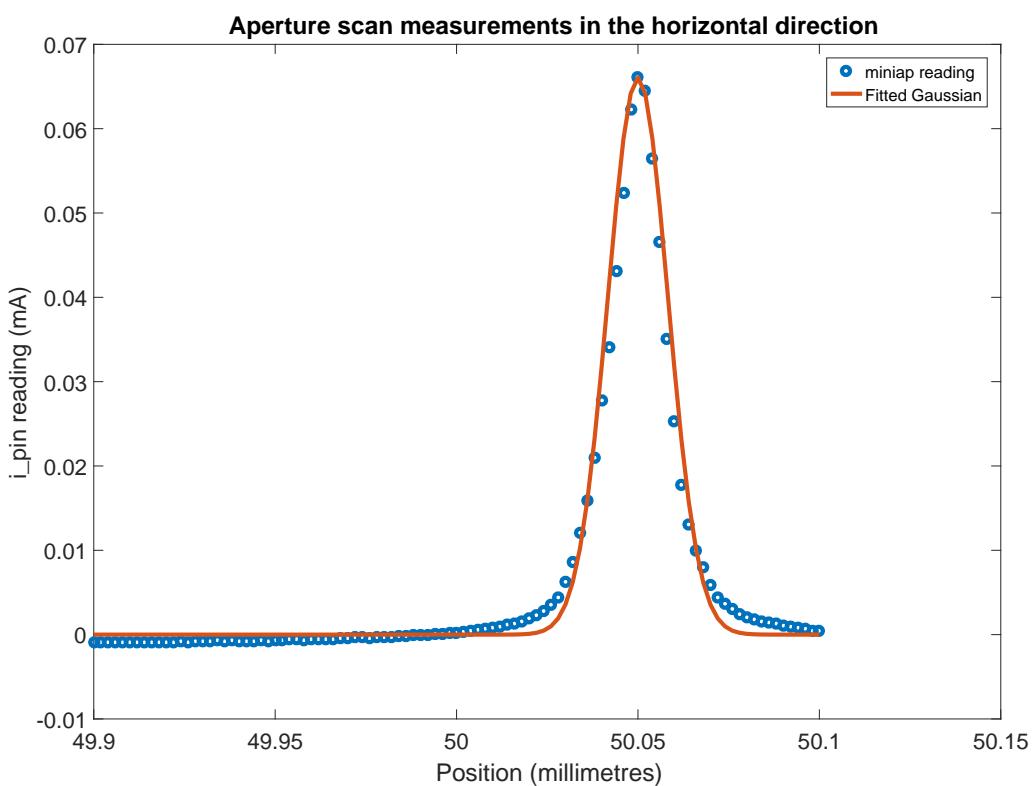
$$g_v(y) = A_y \exp \left[- \frac{(y - \mu_y)^2}{2\sigma_y^2} \right], \quad (5.2.9)$$

and A_{max} is the maximum value of A_x and A_y which are the values of the highest recorded diode current in the x and y directions respectively. If the measurements are taken perfectly (i.e. aperture is scanned exactly across the maximum flux in x and y) then $A_x = A_y$, but due to experimental error this is rarely the case.

To obtain an estimate of the true relative X-ray beam profile, equation 5.2.7 has to be deconvoluted with a matrix that corresponds to the aperture. The aperture matrix is a square matrix created just large enough to contain the area of the circular aperture. It essentially acts as a mask, so elements of the matrix contain the value 1 if the area is inside the aperture, and 0 if the element is outside the aperture. To determine the position of a matrix element with respect to the aperture area, each point in the matrix is given the value of the Euclidean distance away from the central matrix element. This central element is assumed to be at the position of the centre of the aperture. Any points with a Euclidean distance smaller than the aperture radius are determined to lie within the aperture area and hence given a value of 1. However points with a Euclidean distance bigger than the aperture radius



(a)



(b)

Figure 5.3: The original miniat readings from translating the $10 \mu\text{m}$ diameter aperture are shown as blue circles and the solid orange line is the Gaussian fit to the data. (a) beam profile in the x -direction. (b) beam profile in the y -direction. Data collected on I02, DLS.

are outside and are set to 0.

The $\frac{\mathcal{F}(n)}{\mathcal{F}(f)}$ term in equation 5.2.3 is estimated using the `fminsearch` minimisation function in Matlab. Thus an objective function was written as follows:

1. first an estimate of the signal to noise ratio is passed to the objective function
2. using the estimates a deconvolution of the signal is performed using the two-dimensional beam profile given by equation 5.2.7 and the aperture matrix.
3. a convolution of the deconvoluted signal with the same aperture matrix is performed.
4. the Frobenius norm* of the original convoluted signal subtracted from the newly calculated convoluted signal is calculated and returned by the objective function.

If the two signals give a norm value of zero then they are exactly the same, but otherwise a positive value is returned. The `fminsearch` function returns the value of $\frac{\mathcal{F}(n)}{\mathcal{F}(f)}$ that minimises the objective function. Given the convoluted two-dimensional beam profile, the aperture matrix and $\frac{\mathcal{F}(n)}{\mathcal{F}(f)}$, the beam profile can be deconvoluted using the Matlab function `deconvwnr` which performs a Wiener deconvolution as described in section 5.2.2 (Figure 5.4b).

The deconvoluted 2D beam profile is not smooth, which is typical of deconvolution when the exact noise distribution is unknown. A further 2D Gaussian function is fitted to the deconvoluted data in order to obtain a smooth beam profile (Figure 5.5). This is carried out by taking 1D slices of the 2D beam profile in both the x and the y directions around the slice that gives the maximum readings in the centre, and then obtaining the parameters as outlined above (section 5.2.2).

Using the fitted Gaussian model, any properties of the beam can be determined, such as the Full Width at Half Maximum (FWHM) of the beam, which can then be provided to the I02 DLS users. The FWHM is defined mathematically for a Gaussian as

$$FWHM = 2\sigma\sqrt{2\ln(2)}. \quad (5.2.10)$$

*The Frobenius norm, $\|A\|_F$, of a matrix, A , is defined as $\|A\|_F \equiv \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$ where the a_{ij} represent the elements of the matrix.

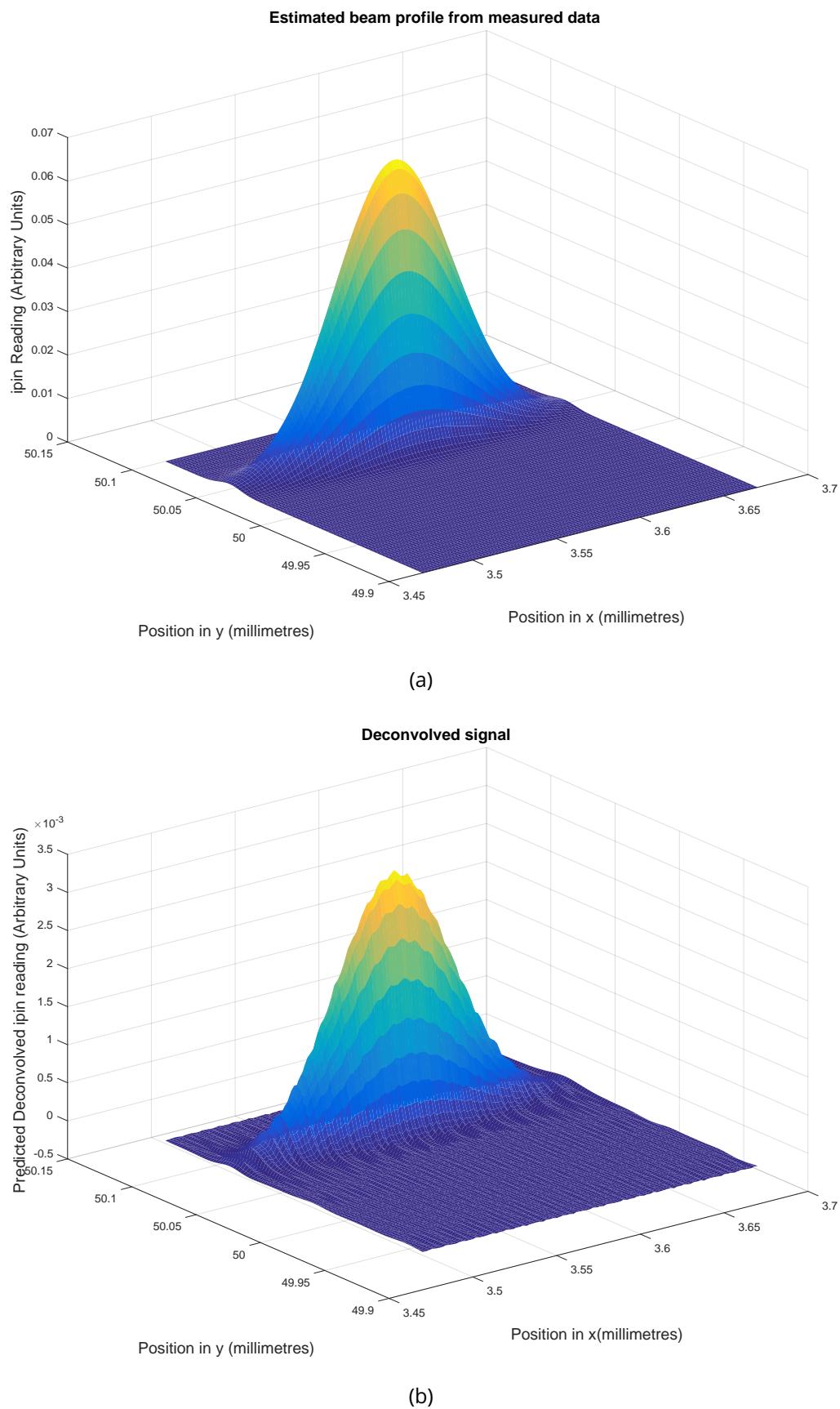


Figure 5.4: 2D X-ray beam profiles. (a) Gaussian approximation of the measured current profile given by equation 5.2.7 (b) Beam profile after Wiener deconvolution.

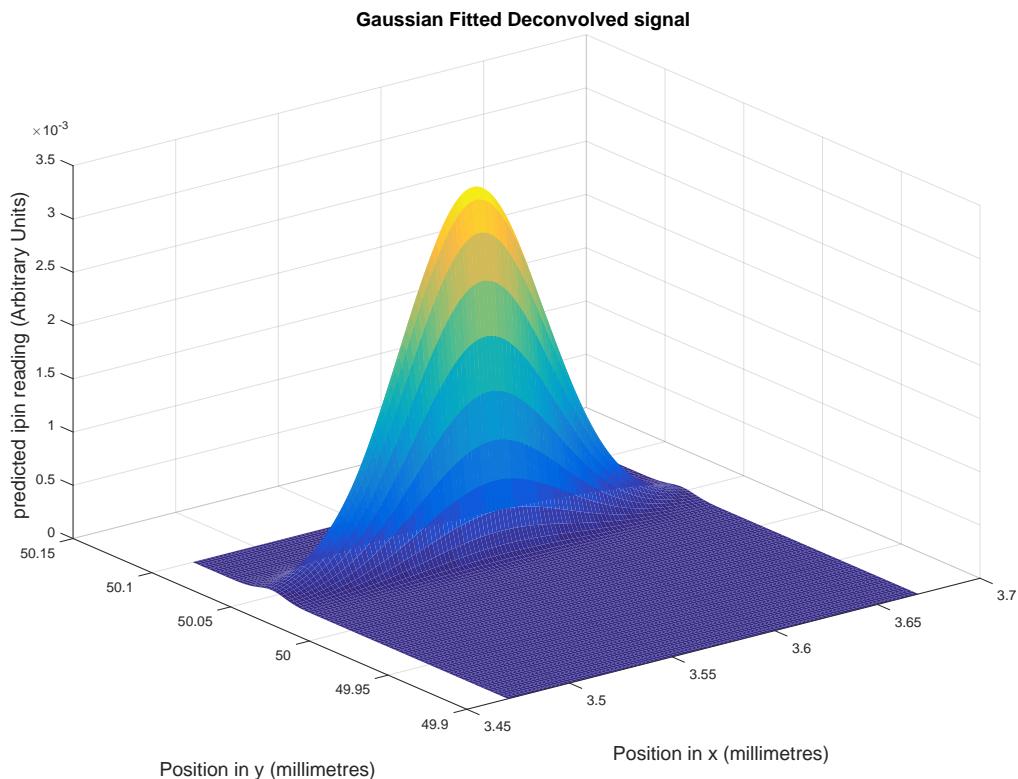


Figure 5.5: Gaussian model fitted to the deconvoluted X-ray beam profile (Figure 5.4b).

Validation of deconvolution approach

No measuring device is available to take measurements at the resolution to which the deconvolution method attempts to spatially resolve the flux readings. Therefore to validate the results, the deconvolution method was used in reverse to predict the FWHM values that would be obtained from the diode readings with a $20\ \mu m$ diameter aperture. To achieve this the 2D deconvoluted signal (Figure 5.5) was convoluted with an aperture matrix of a different size. 1D slices in the x and y directions were then found such that the flux in the central matrix element was maximum and 1D Gaussian functions were fitted to each slice. FWHM values were then calculated from the 1D Gaussians using equation 5.2.10 and these calculated values could then be compared with the readings taken from actual experimental data obtained using a $20\ \mu m$ aperture.

5.2.3 Results

FWHM values obtained from the experimentally measured (convoluted) *i*_pin readings using the $20 \mu\text{m}$ diameter aperture (referred to as Exp_{20}) were compared with calculated FWHM values derived from the numerical convolution of the deconvoluted profile with the $20 \mu\text{m}$ aperture (referred to as Conv_{20} - Table 5.1). The calculated FWHM for Conv_{20} are quite accurate.

Table 5.1: Comparison of calculated FWHM values with experimentally observed FWHM values. The calculated FWHM values using the $10 \mu\text{m}$ aperture (in italics) are calculated from the deconvoluted profile. Thus they were not expected to match the experimental values.

	Aperture diameter (μm)	FWHM in x (μm)	% Error in x	FWHM in y (μm)	% Error in y
Exp_{10}	10	81.6	0.49	19.5	4.84
Deconv_{10}	<i>10</i>	<i>81.2</i>		<i>18.6</i>	
Exp_{20}	20	81.8	2.57	21.9	1.83
Conv_{20}	20	83.9		22.3	

rate, showing less than a 3% error in both the x and y directions when compared to Exp_{20} . This suggests that the deconvolution method used provides reasonable estimates of the 2D X-ray beam profile.

FWHM values obtained from the experimentally measured (convoluted) *i*_pin readings using the $10 \mu\text{m}$ diameter aperture (referred to as Exp_{10}) were compared with calculated FWHM values derived from the deconvoluted profile with the $10 \mu\text{m}$ aperture (referred to as Deconv_{10} - Table 5.1). The errors of the FWHM values between Exp_{10} and Deconv_{10} are also quite small. In theory however, these values do not necessarily have to be close because the FWHM from Deconv_{10} are calculated from the deconvoluted profile, whereas the FWHM from Exp_{10} are calculated from the convoluted beam profile. The fact that these values are close provide evidence that users should not subtract $5 \mu\text{m}$ from the quoted FWHM values. Furthermore it indicates that deconvoluting the X-ray beam does not significantly change the X-ray beam profile. This is important because it suggests that the measured 2D current profile does not have to be further altered once the profile has been generated from the 1D aperture scans.

5.3 2D X-ray beam profile measurements

The previous section presented work on beam processing performed on measurements of the beam profile obtained from 1D aperture scans. In this section, the analysis is described of measurements of the X-ray beam profile carried out at PETRA III synchrotron, beamline P14, Hamburg, as detailed in section ???. Specifically, the measurement was made using a scintillator combined with an Allied Vision GC1350C CCD camera, which resulted in a 2D image of the X-ray beam. The processing of these images presents different challenges to those that arise from 1D aperture scan measurements.

The 2D image of the X-ray beam profile was exported as a portable graymap (pgm) file where each pixel contains an integer value between 0 and 255 inclusive (Figure ??). These values were determined by the size of the signal from the scintillator. To spatially resolve the flux over the beam image, RADDOSE-3D calculates the flux, F , for a particular pixel as:

$$F = \frac{p_{ij} T_F}{A_p \sum_{ij} p_{ij}}, \quad (5.3.1)$$

where p_{ij} is the pixel value in the image, T_F is the total measured flux, and A_p is spatial area covered by the pixel. This means that the pixel values can be regarded as weights that correspond to the relative photon flux at that spatial position. In general the pgm file will always contain non-zero values that correspond to background signal. This can be seen in Figures ?? and ?? where the pixel values never decay to zero at any point in the slices through the beam. RADDOSE-3D interpreted these non-zero values as beam intensity, so some of the flux that should have been in the beam was instead allocated to the background pixels, which resulted in an underestimate of the absorbed dose in the crystal. Therefore it was important to identify the background and remove it from the beam profile measurements before the dose simulation was performed. Several methods for removing the background were explored and the results were compared in order to investigate the effects on the RADDOSE-3D simulation output.

5.3.1 PGM file preprocessing

Original beam profile

The first profile that will be considered is the original pgm file. No preprocessing is performed on the file so that the results from other processed beams can be compared with the unprocessed case. Figure 5.9a shows the profile of this raw beam.

Deconvolution of beam profile

The beam profile measured with the scintillator and CCD camera as described in section ?? can smear as a result of charge diffusion into adjacent pixels due to CCD pixel subvariations. This response is known as the point spread function (PSF). Thus the image formed in the pgm file is a convolution of the actual X-ray beam profile and the PSF. To deconvolve the beam image, the following procedure was undertaken:

1. An initial blind deconvolution (using the MATLAB `deconvblind` function) was performed to determine an estimate of the PSF (Figure 5.6a). The blind deconvolution requires an initial guess for the PSF, but the PSF restoration in the blind deconvolution is heavily affected by the grid size of the PSF rather than the values of the grid elements. With no knowledge of the PSF (or of the signal to noise power spectrum) an initial guess for the PSF was made as a 7×7 grid of 1's.
2. The restored PSF was then modelled using a Laplacian function of the form

$$f_{lap} = \exp\left(-\frac{|x| + |y|}{b}\right) \quad (5.3.2)$$

where b is a parameter whose value is to be determined (Figure 5.6b). The Laplacian function was chosen because it modelled the shape of the recovered PSF function very well.

3. The pgm file was then segmented using the MATLAB `activecontour` function, which uses the Chan and Vese region based energy model (?), to find the spatial area of the beam. Additionally the centroid of the resulting image (the centre of the beam) was determined (Figure 5.7a).

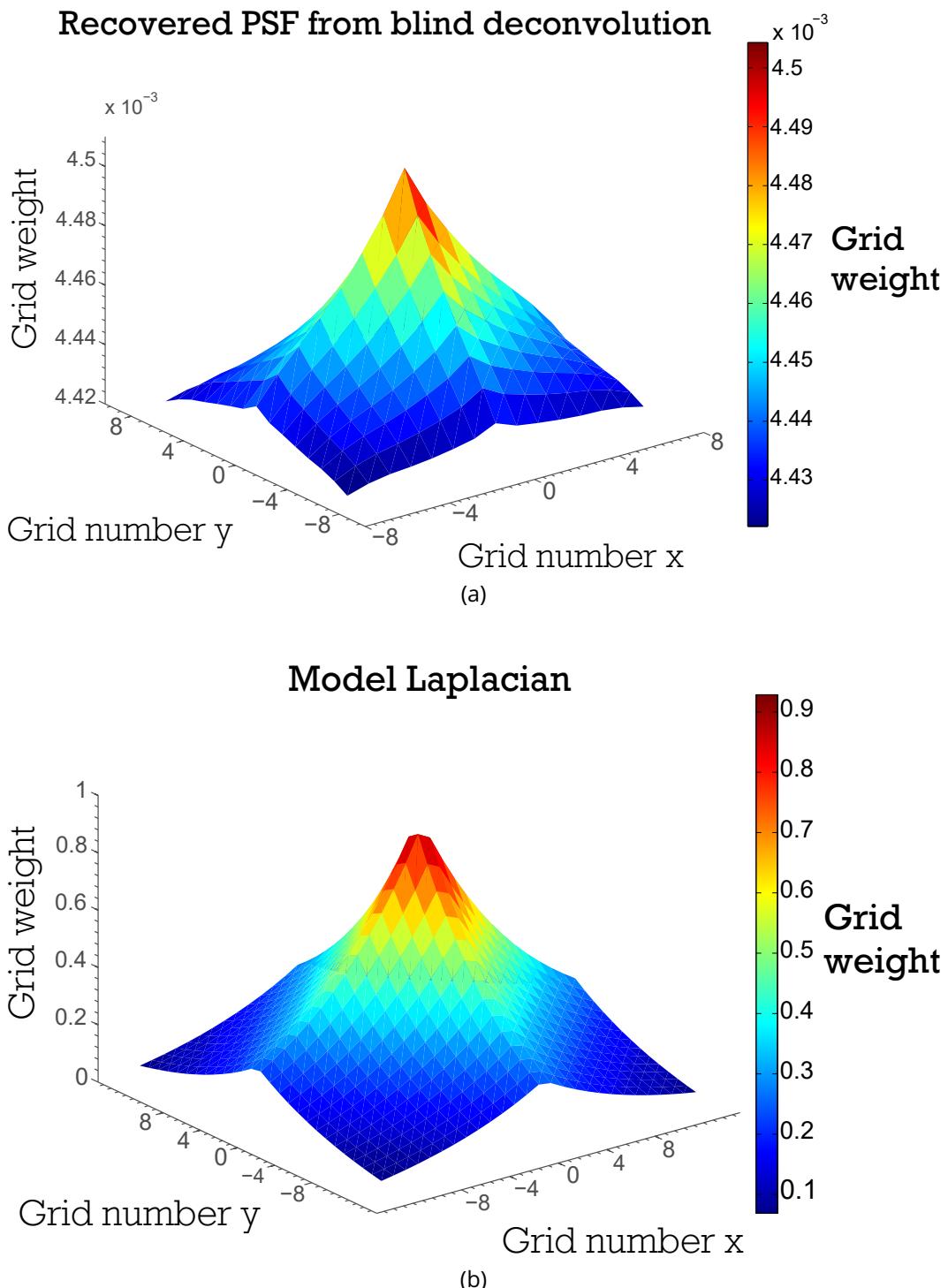


Figure 5.6: (a) PSF recovered from performing a blind deconvolution. (b) Laplacian function of the form given in equation (5.3.2) with a 7×7 grid and $b = 10$.

4. A rectangle that bounded the beam (inner red rectangle in Figure 5.7b) was chosen as $145 \mu\text{m} \times 140 \mu\text{m}$ because this contained the beam whilst allowing a buffer for beam area that may have been missed by the segmentation algorithm. These dimensions were stored.
5. The (square) grid size for the PSF, b in equation 5.3.2 and the rectangular dimensions were then used as inputs for an objective function in a minimisation process to determine optimal values for these parameters. The objective function was constructed as follows:
 - A perfect (hypothetical) top hat beam was created inside the inner box (Figure 5.7b).
 - The Laplacian PSF was created using the grid size and b .
 - The top hat beam was convolved with the PSF to create a theoretical (convolved) beam image.
 - the positions of zero pixel values from the theoretical beam image were also set to zero in the original beam image.
 - The height of the theoretical beam image was scaled to the height of the original beam.
 - The matrix 2-norm of the difference between the original beam image with zeroed outer pixel values and the theoretical beam image was calculated and used as the output of the objective function to be minimised (the matrix 2-norm is equal to the square root of the maximum eigenvalue of the matrix multiplied by its conjugate transpose matrix).
6. Another blind deconvolution was performed as before but this time with the returned grid size from the minimisation procedure as the input. The resulting beam profile was then returned for input into RADDOSE-3D (Figure 5.9b).

Removal of background using deconvolution results

Since the rectangular dimensions of the hypothetical top-hat beam were given as outputs of the minimisation procedure, this shape can be used to distinguish between the X-ray beam

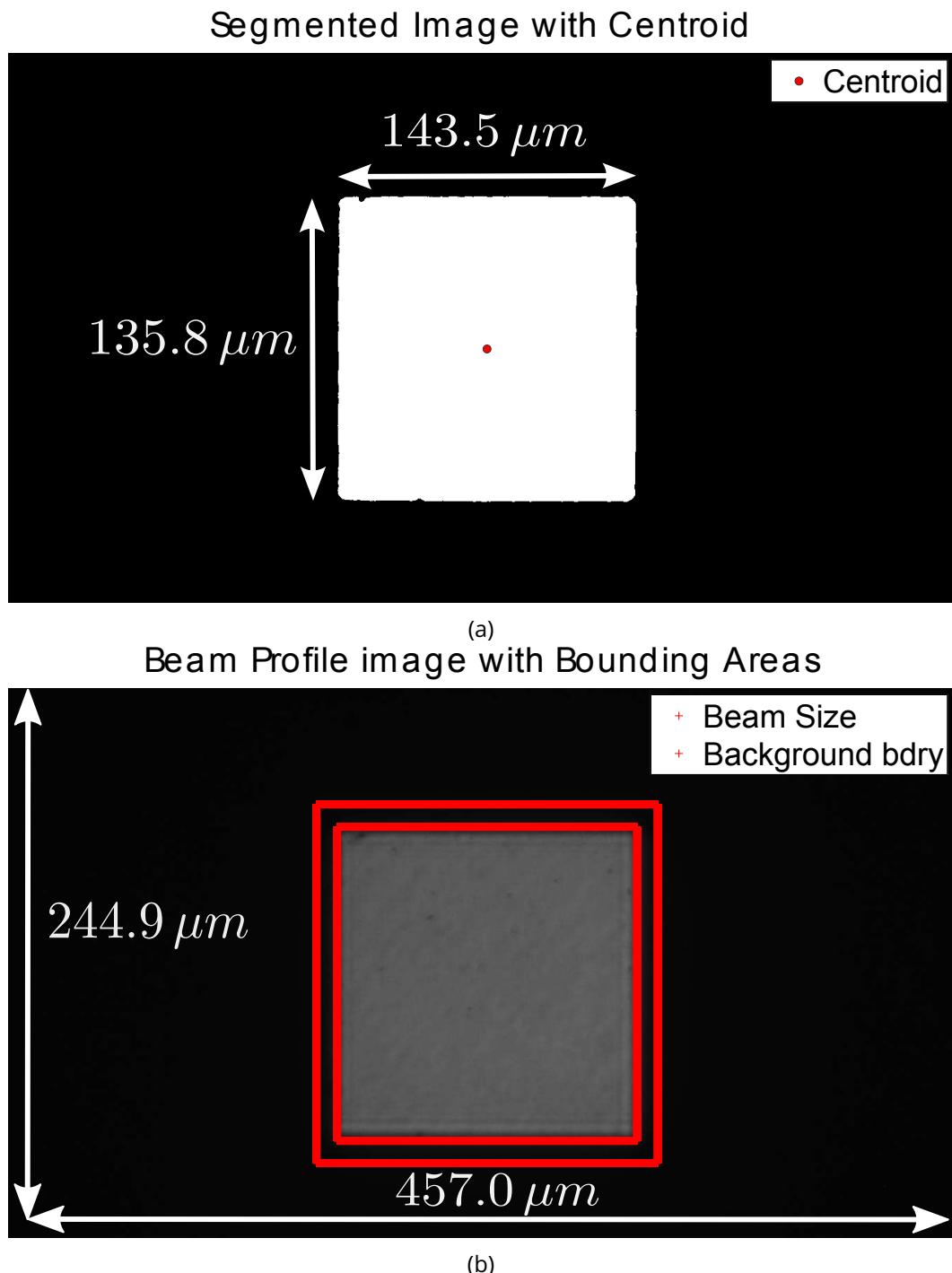


Figure 5.7: (a) Segmented image of the beam with its centroid marked as a red circle in the middle. (b) Beam image with the centroid and two bounding boxes overlaid. The inner box is an estimate of the beam size with dimensions $145 \mu m \times 140 \mu m$. These dimensions were chosen according to the coverage of the beam area in the image, as opposed to the size of the aperture from the slit separation (slits were set at $140 \mu m \times 140 \mu m$). The outer box is the boundary, $170 \mu m \times 170 \mu m$, where the pixels outside the box are considered part of the background (section 5.3.1).

and background. The two beams described above were also further manipulated by setting any values outside of the rectangle dimensions (inner box) to zero (Figures 5.9c and 5.9d).

Perfect top hat beam

A perfect top hat beam was created within the inner boundary of Figure 5.7b i.e. the value 100 was inserted for pixels that lay inside the inner rectangle, otherwise the pixel values were set to zero (Figure 5.9k).

Beam thresholding

In his D. Phil. thesis, Dr. Oliver Zeldin preprocessed pgm files by setting a threshold value (?). The threshold (or dark current) value was determined by taking the average of the pixel values that were “far away” from the main beam centre. However, to my knowledge, there was no systematic way to determine “far away”. In the PETRA III MX experiment described here, the slits were adjusted to give an aperture of $140 \times 140 \mu\text{m}^2$. To determine how the choice of “far away” affects the threshold to be subtracted, it was decided that background would be considered as any pixel value beyond $30 \mu\text{m}$ away from the slit edge (Bkgd_{30}). The $170 \times 170 \mu\text{m}^2$ outer red box in Figure 5.7b is the Bkgd_{30} boundary. Another two beams were created assuming that the background could be considered as any pixel beyond $20 \mu\text{m}$ from the slit edge (Bkgd_{20} with a $160 \times 160 \mu\text{m}^2$ boundary) and $85 \mu\text{m}$ from the slit edge (Bkgd_{85} with a $225 \times 225 \mu\text{m}^2$ boundary). These boundaries are shown in Figure 5.8. The threshold value for subtraction was determined by taking a mean average of the background pixel values and rounding it to the nearest integer (Figures 5.9e, 5.9f and 5.9i). Another way to determine the background was to take the maximum value of the background values and subtract that value from every pixel in the image (Figures 5.9g, 5.9h and 5.9j).

5.3.2 RADDSE-3D simulation results

All of the beams illustrated in Figure 5.9 were used as input for RADDSE-3D to simulate the experiment described in section ?? for insulin crystal ID 0259. Each RADDSE-3D run produced a set of DWD values; each value corresponded to one of the 271 datasets produced using the rolling window processing method described in section ?? . These doses

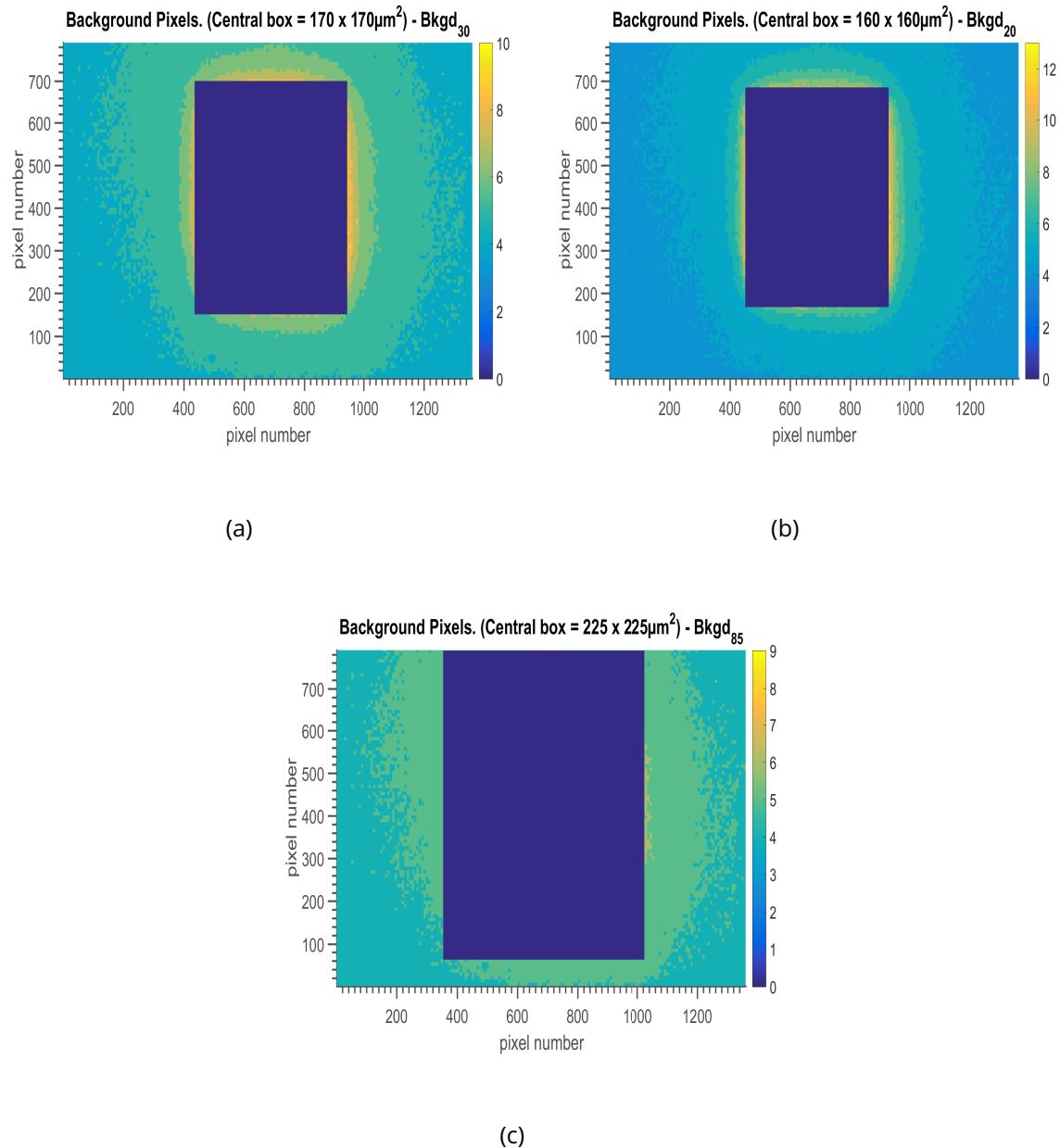
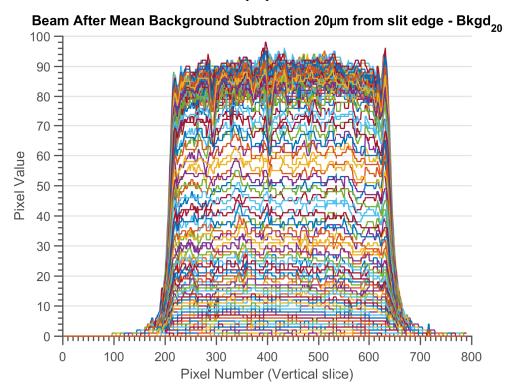
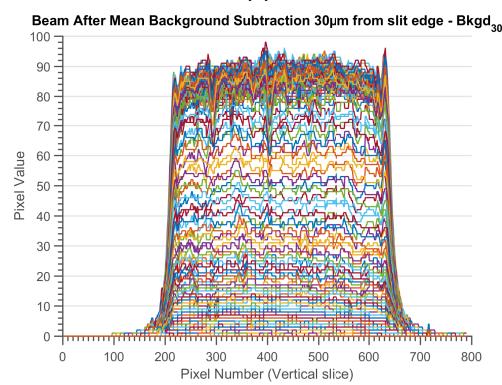
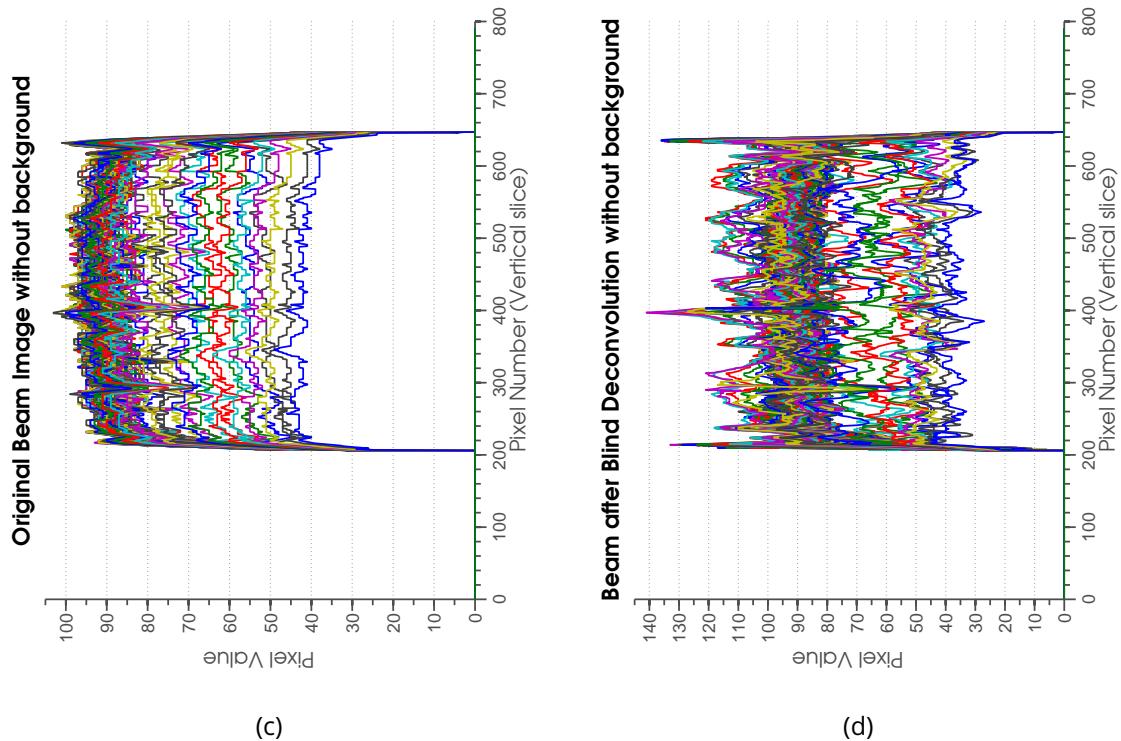
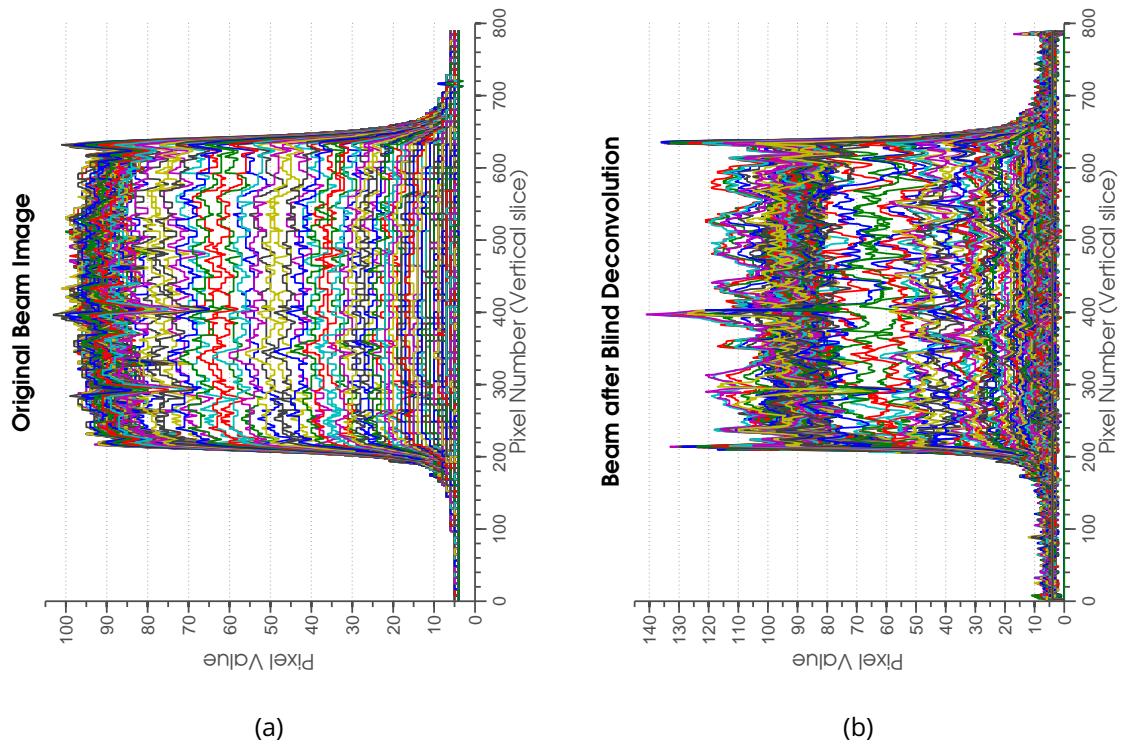


Figure 5.8: Background pixel values for the beam image (beam image dimensions are $457.0 \mu\text{m} \times 244.9 \mu\text{m}$ for all images). The pixel values in the central box for all images are set to zero and the mean and maximum values of the non-zero pixel values are taken for each of the backgrounds. (a) Bkgd₃₀: Rounded mean pixel value = 5, maximum pixel value = 10. (b) Bkgd₂₀: Rounded mean pixel value = 5, maximum pixel value = 13. (c) Bkgd₈₅: Rounded mean pixel value = 4, maximum pixel value = 9.



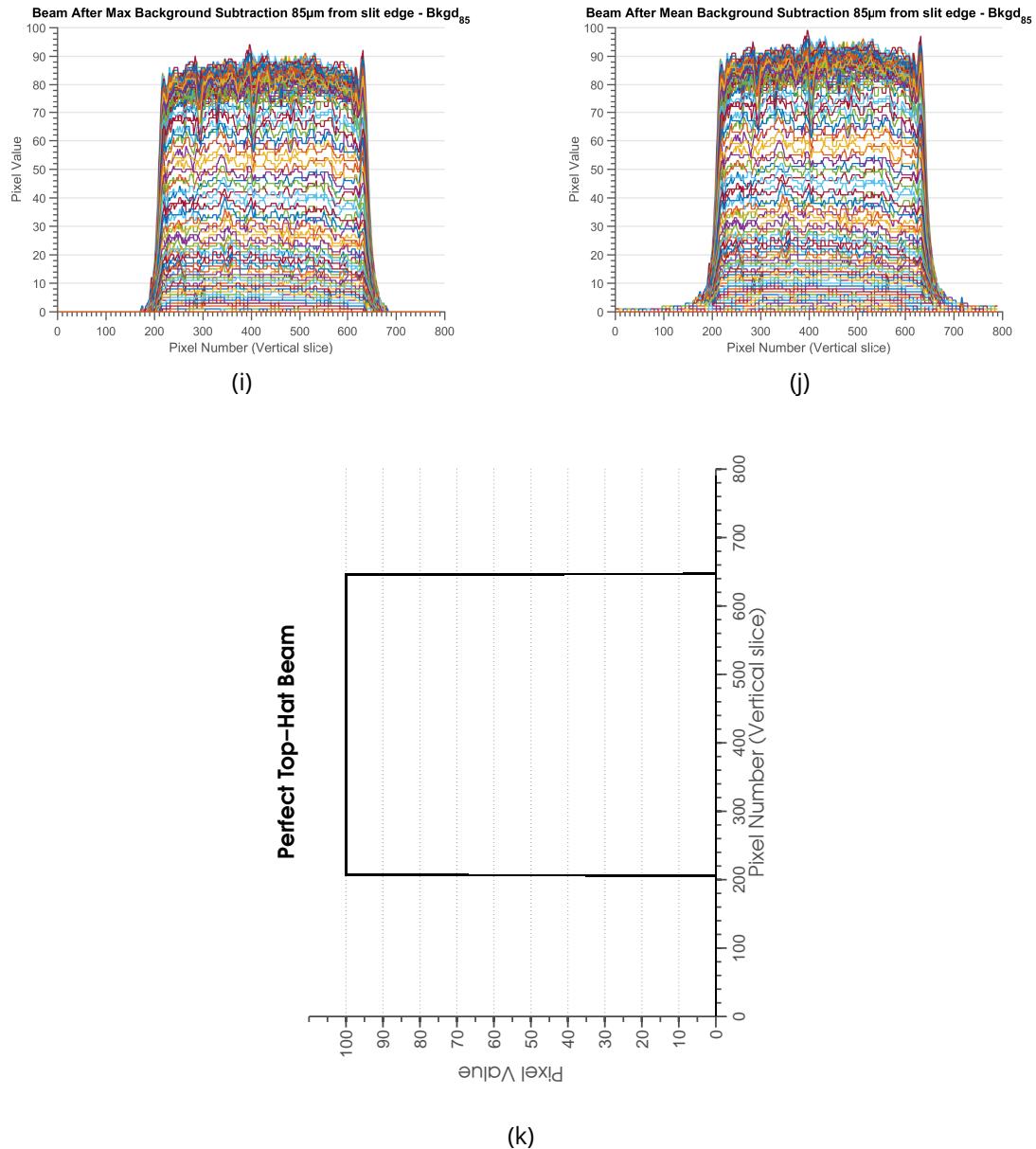


Figure 5.9: Processed beam profiles used for simulations in RADDOS-3D. Each figure is a plot with all vertical slices across the corresponding pgm images overlaid on a graph. They are essentially the same as Figure ?? except in these images all vertical slices are shown, as opposed to just the central slice.

were plotted against the relative intensities (I_n/I_1) for each dataset and a line of best fit was determined using data where $I_n/I_1 > 0.4$. $D_{1/2}$ values were then obtained from the line of best fit for each processed beam and plotted against the percentage of pixel values in the image that are equal to 0 as shown in Figure 5.10. The results show that the higher the threshold value (i.e. the more pixels that are considered as background and are set to zero) the higher the $D_{1/2}$ value. This occurs because RADDOSE-3D is run independently from the scaling of the diffraction data and therefore the I_n/I_1 values are the same for each run. As alluded to above, more zero pixel values means that the same total photon flux is distributed over a smaller area (which always contained the crystal), and hence the crystal absorbs more energy which contributes to higher doses. This gives the appearance that the crystal is less radiation sensitive (because the $D_{1/2}$ is higher).

The other major factor in the calculation of the dose values is the calculation of the absorption coefficients. In Figure 5.10, the blue markers correspond to simulations where the absorption coefficients were calculated with RADDOSE version 2 (?) (denoted RDV2 in Figure 5.10) which gave an absorption coefficient of $4 \times 10^{-4} \mu\text{m}^{-1}$ (solvent content supplied as 64%). The red markers represent simulations where the absorption coefficients were calculated for an “average protein”, $\text{H}_{49.8}\text{C}_{31.8}\text{N}_{8.56}\text{O}_{9.54}\text{S}_{0.249}$ (denoted “*default*” in Figure 5.10). The absorption coefficient, assuming 50% solvent, is $2.37 \times 10^{-4} \mu\text{m}^{-1}$ (details of the calculation can be found in Holton and Frankel (2010)). The absorption coefficient of insulin is bigger than that of the average protein because it contains more sulphur than average, and it also contains zinc, which has a relatively high absorption cross section. The difference between the absorption coefficients leads to differences in calculated doses of about a factor of 2.

The results seem to suggest that the two factors investigated here: beam profile and absorption coefficient, are the most important for dose calculations. However, the differences between the profiles of the original beam image, the top hat beam and the deconvoluted beam have a much less pronounced affect on the dose calculations compared to the absorption coefficient. For example, the percentage difference of $D_{1/2}$ between the original beam with no background (denoted $\text{Original}_{\text{NB}}$ in the figure legend of Figure 5.10) and the deconvoluted beam with no background is 0.8% (using coefficients calculated using RADDOSE version 2). This shows that deconvolution of the image will not significantly affect the dose estimate. This is because the differences in the pixel values are relatively small when

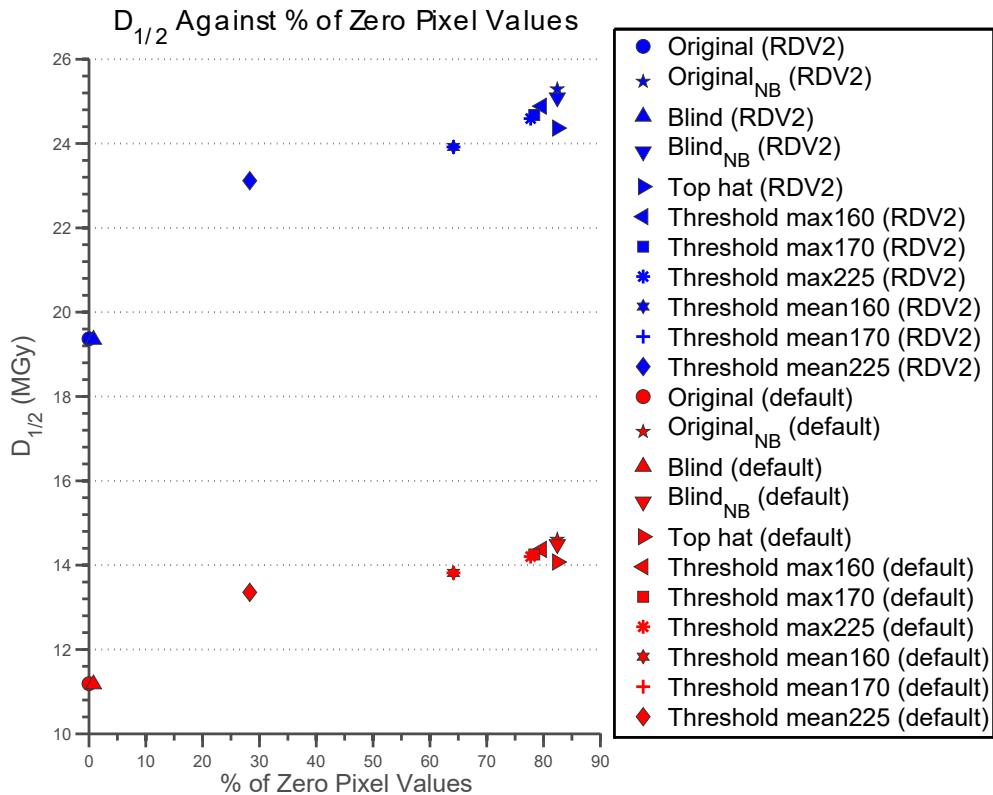


Figure 5.10: $D_{1/2}$ values plotted against the percentage of zero pixel values in the processed pgm beam images. Blue markers correspond to simulations where the absorption coefficients were calculated with RADDOSE version 2 (denoted 'RDV2' in figure legend) and red markers represent simulations where the absorption coefficients were calculated for an "average protein" (denoted 'default' in figure legend). The subscript 'NB' denotes beam profiles where the "background" was removed. The figure shows a trend that the higher the proportion of zero pixel values, the higher the $D_{1/2}$ value. It also shows that the method used to calculate the absorption coefficients can significantly affect the estimated doses, as expected.

compared to the differences one may expect if a Gaussian beam profile is compared with a tophat profile.

Using the threshold method there is a difference in $D_{1/2}$ values ($\approx 3.9\% - 6.0\%$ depending on the background threshold value) if the maximum pixel value of the background is subtracted from the image or the mean pixel value is subtracted from the image. This is consistent with the observation that the higher the threshold value, the higher the $D_{1/2}$ value. The difference between the maximum threshold value from the $160 \times 160 \mu\text{m}^2$ rectangle (Bkgd_{20}) and the mean value is 3.89% (using RDV2 calculated absorption coefficients). This result demonstrates that even if the area that is considered background is the same (Figure 5.8), deciding whether to subtract the mean background pixel value or the maximum pixel value will also have a small affect on the resulting $D_{1/2}$ value.

5.4 Processing multiple 1D aperture scan measurements

In section 5.2, measurements of the X-ray beam were described which were taken by scanning a $10\ \mu m$ aperture across the beam horizontally and vertically through the centre of the beam. A 2D beam profile was calculated by fitting a Gaussian function to the separate profiles and substituting those parameter values in the equation of a 2D Gaussian. The implicit assumption made is that the beam profile is similar in every region of space, including the central slices. This is in addition to the explicit assumption that the beam profile is Gaussian. X-ray beam profiles can vary drastically from beamline to beamline even at the same synchrotron, so the assumptions made by the previous method are not generally applicable to all beam profiles. Therefore, a method to generate 2D X-ray beam profiles from multiple 1D aperture scans across the X-ray beam was developed to provide a more generally applicable solution to this problem.

5.4.1 Acquiring the 1D aperture scans

In order to estimate the dose for the SAXS experiments described in chapter ??, aperture scans to obtain X-ray beam flux measurements were carried out in collaboration with Dr. Adam Round and Dr. Martha Brennich on the ESRF BM29 (SAXS) beamline. A $100\ \mu m$ diameter circular aperture was used to scan across the X-ray beam area and measurements were taken at $10\ \mu m$ intervals with an OSD1-0 photodiode purchased from Optoelectronics. The scanning was performed 6 times with the collection of 3 horizontal and 3 vertical scans (Figure 5.11).

5.4.2 Creating a 2D beam profile

It is clear from the data shown in Figure 5.11 that fitting a Gaussian shape will not model this beam profile well. Instead a rectangular grid is set up with the edges of the measurement positions in Figures 5.11a and 5.11b used as the boundaries of the grid. The flux at and beyond the grid boundaries is assumed to be zero. The diode measurements from the vertical aperture scans were placed in their corresponding positions on the grid and interpolation between these values was performed using the `RectBivariateSpline` function

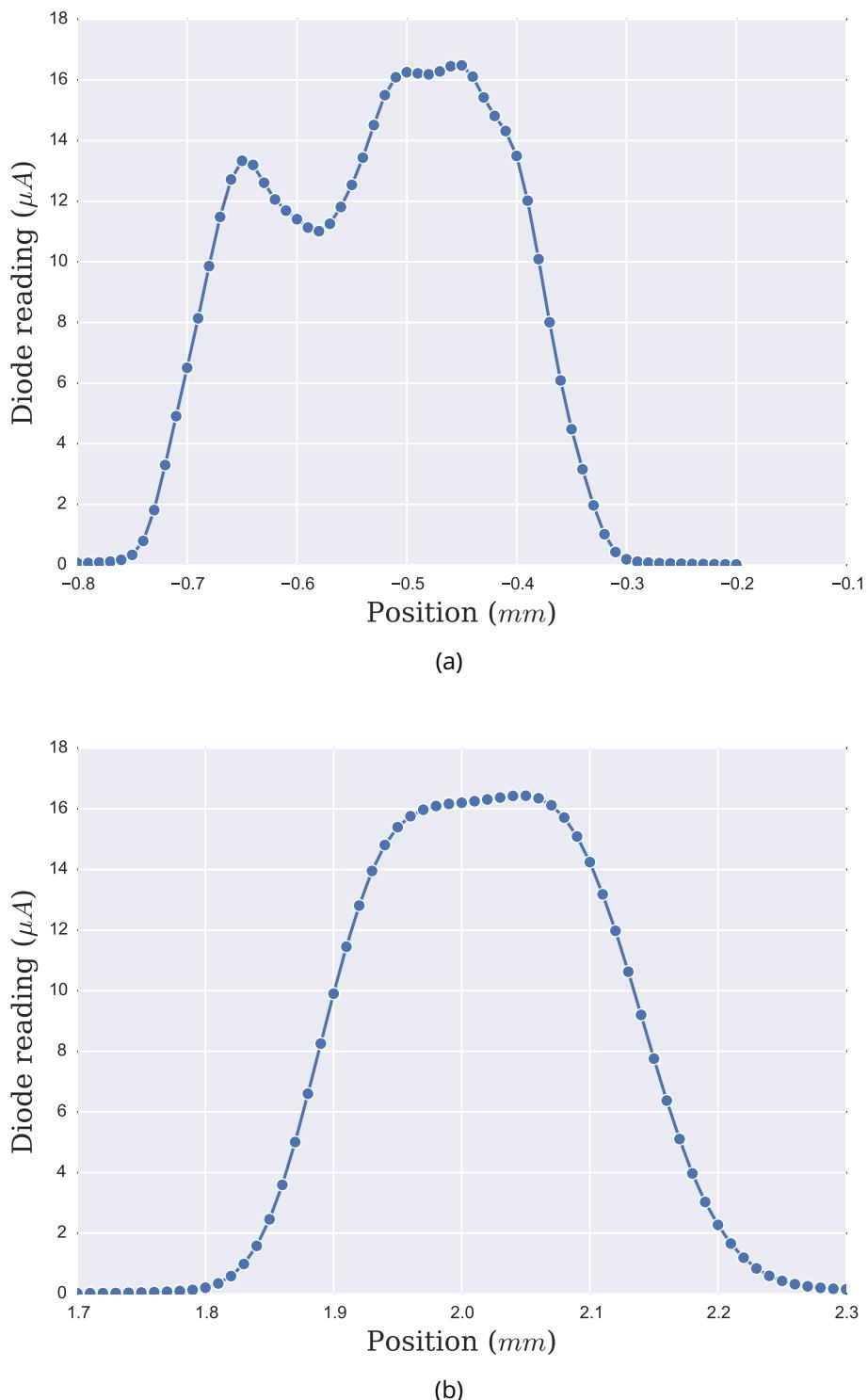


Figure 5.11: Diode readings from aperture scans across the beam collected at beamline BM29, ESRF, displayed in Figure ???. (a) Vertical scan through the centre of the beam. (b) Horizontal scan through the centre of the beam. 4 other scans (2 vertical, 2 horizontal) were also taken at other locations across the beam.

in SciPy package in the Python programming language (?). The same procedure was performed for the data in the horizontal direction. The results are shown in Figure 5.12. The two interpolated arrays from Figure 5.12 are then averaged to obtain the final beam profile (Figure 5.13).

The final X-ray beam profile in this case is a much truer representation of the overall beam profile than that generated by fitting a Gaussian profile. The main reason is because the measurements are explicitly used as points on the interpolated beam profile grid as opposed to fitting a mathematical function. Secondly, the method uses data from scans taken anywhere in the X-ray beam profile rather than just the central slices. However, the method is not perfect because the final averaging causes a loss of some features. Figure 5.12a exhibits a ‘valley’ in the beam profile; however after averaging, the valley resembles a ‘shoulder’ (Figure 5.13a). This could be rectified by applying a weighted averaging scheme where positions that contain measured data in one beam profile array and not the other are up weighted. A more desirable approach would be to use an interpolation routine that could be applied to all measured data in both directions, thus avoiding the need to average at all. It is also important to collect as many aperture scans as possible because the resulting beam profile will be more representative of the true X-ray beam profile, as then more actual values are measured.

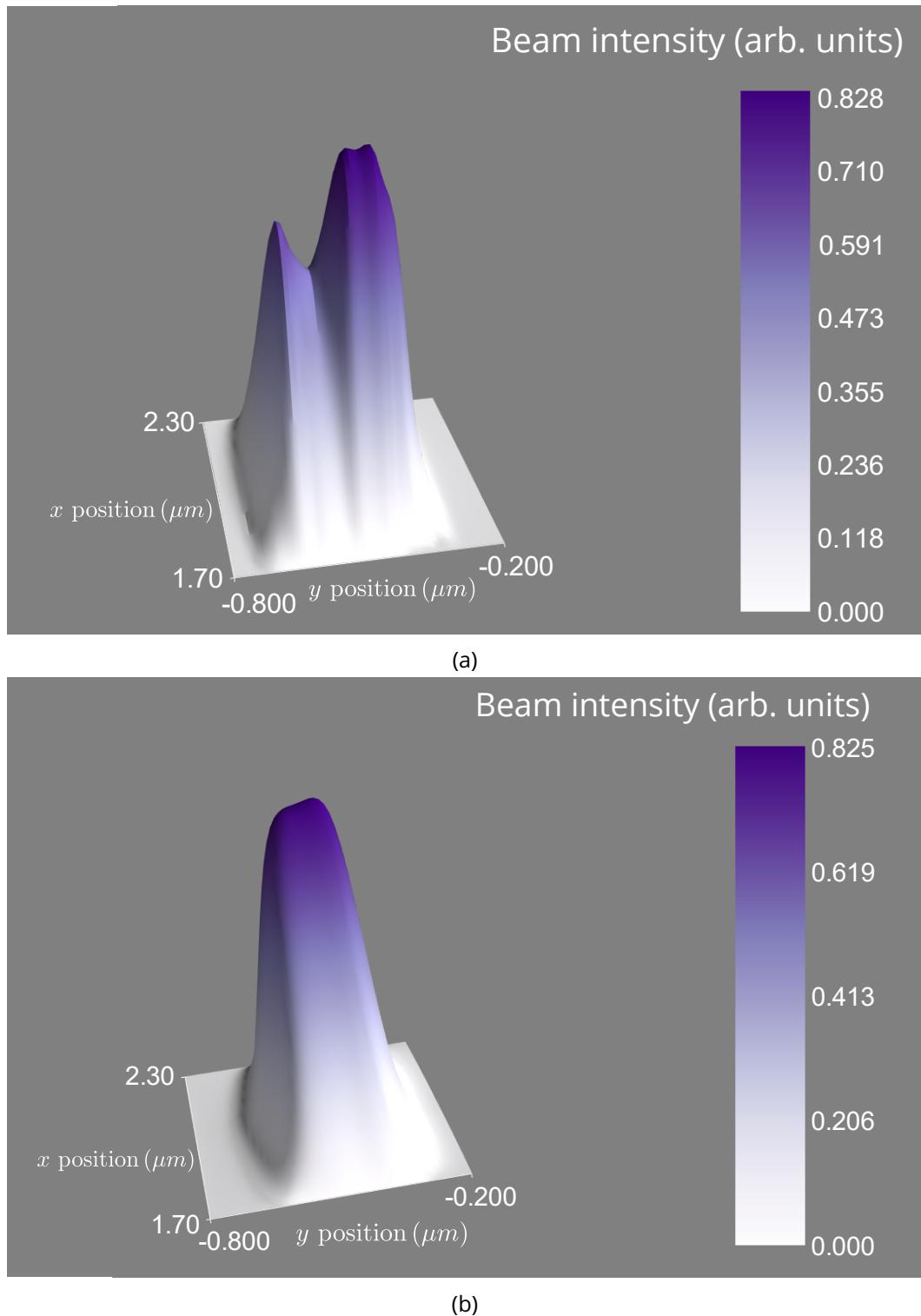


Figure 5.12: Interpolation of diode readings from the aperture scans. (a) Interpolation of vertical scans. (b) Interpolation of horizontal scans.

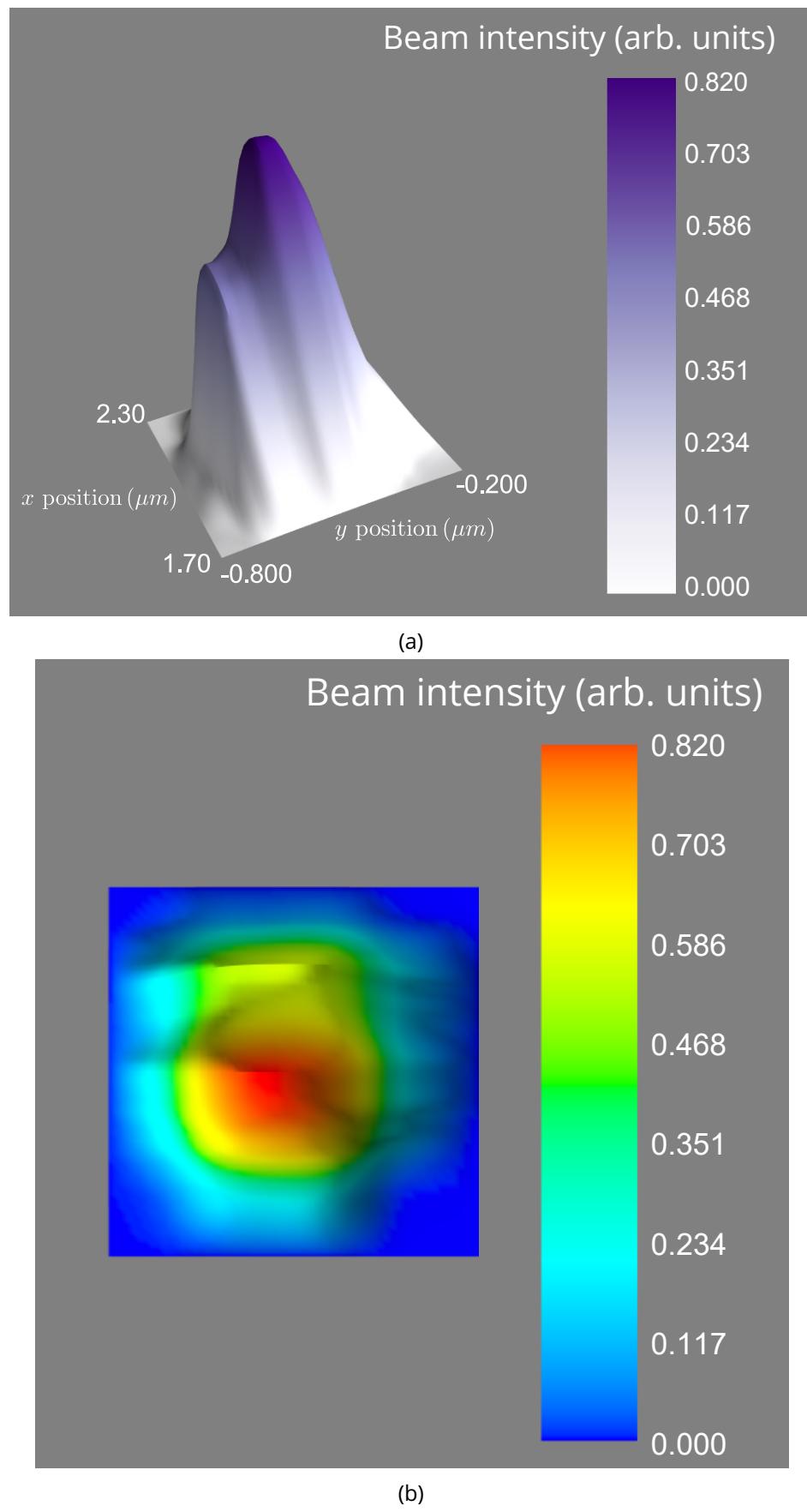


Figure 5.13: (a) Final X-ray beam profile generated from the average of the two beams in Figure 5.12. (b) Overhead view with a rainbow colour scheme.

5.5 Discussion

The ability to calculate the dose absorbed by a crystal is vital for the comparison of results from different radiation damage studies and to give guidance to experimenters so MX data collections can be optimised. To ensure that the dose estimates are reliable, it is necessary to accurately parameterise the diffraction experiment. A custom module was written into RADDOSSE-3D by Dr. Oliver Zeldin to allow the experiment to be simulated with experimentally measured X-ray beam profiles (?). However the experimentally measured beam profiles often need some form of preprocessing before they can be used in the simulation. The type of preprocessing largely depends on the experimental method used to measure the beam profile.

One method to measure the beam profile is to use an aperture to scan across the X-ray beam and measure the current in a silicon diode produced by the X-ray photons at each position. Often a single aperture scan is performed horizontally and again vertically. If the scans resemble Gaussian profiles (Figure 5.1) then it is possible to use the parameter values obtained from 1D Gaussian fits to the data in a 2D Gaussian model representation of the beam. This approach enabled accurate predictions of the data that would be generated using apertures of various sizes, hence validating the method. The model also highlighted another important aspect. It showed that the FWHM values of the true beam are very likely to be close to the FWHMs measured from the diode readings of the aperture scans. The implication of this result for modelling is that deconvoluting (with subsequent smoothing of the noise) the X-ray beam profile from the aperture size does not lead to a significantly different beam profile. However, more experiments with different aperture and beam sizes may need to be performed to verify the generality of this result.

A different method of measuring the beam profile involves taking a 2D image of the beam using a scintillator and a camera. The advantage of this method is that there is no need to explicitly model the 2D beam profile. However, the drawback is that regions of the image that represent background (i.e. areas of zero flux from the X-ray beam) have non-zero pixel values. Several beam processing methods including deconvolution, image segmentation and standard average subtraction, were applied to remove background. It was found that the various methods gave similar results in terms of the calculated half dose ($D_{1/2}$) values

i.e. the resulting beam profiles following the processing did not affect the dose values. The biggest factor affecting the dose values is the proportion of the image area that corresponds to background. This suggests that to get an accurate beam profile from a 2D image it is necessary to know the collimation of the X-ray beam so that a suitable background region can be masked. It is important that the scintillator is placed at the sample position because beam divergence will result in a slightly different beam profile if the scintillator is placed in a different plane.

A more reliable (and thus recommended) method to measure the experimental beam profile is to take a series of 1D aperture scans to sample as much of the beam profile region as possible. This bypasses the need to subtract background from an image because flux is measured via the current it produces on the silicon diode. Sampling the 2D space allows the beam to be interpolated between data points. This negates the requirement to explicitly model the beam profile with a 2D Gaussian function for example. In the work presented here, the interpolation was performed separately on horizontal scans and vertical scans and then averaged. This method still results in a loss of some features as evidenced by lack of the “valley” in the final averaged beam profile, even though it was clearly present in the vertical scans. To avoid the loss of features, in the future it would be desirable to perform the spline interpolation on all data, as opposed to the vertical and horizontal data separately.