Spline modelling electron insert factors using routine measurements

S Biggs[a], M Sobolewski[a], R Murry[b], J Kenny[c]

[a] *Riverina Cancer Care Centre, New South Wales 2650, Australia*
[b] *Radiation Oncology Queensland, Queensland 4350, Australia*
[c] *Epworth Radiation Oncology, Epworth HealthCare, Victoria 3121, Australia*

Abstract

There are many methods available to predict electron output factors however many centres still measure the factors for each irregular electron field. Creating an electron output factor prediction model that approaches measurement accuracy—but uses already available data and is simple to implement—would be advantageous in the clinical setting. This work presents an empirical spline model for output factor prediction that requires only the measured factors for arbitrary insert shapes. Equivalent ellipses of the insert shapes are determined and then parameterised by width and ratio of perimeter to area. This takes into account changes in lateral scatter, bremsstrahlung produced in the insert material, and scatter from the edge of the insert. Agreement between prediction and measurement for the 12 MeV validation data had an uncertainty of 0.4% (1SD). The maximum recorded deviation between measurement and prediction over the range of energies was 1.0%. The validation methodology showed that one may expect an approximate uncertainty of 0.5% (1SD) when as little as eight data points are used. The level of accuracy combined with the ease with which this model can be generated demonstrates its suitability for clinical use. Implementation of this method is freely available for download at https://github.com/SimonBiggs/electronfactors.

Keywords: electron output factor, bivariate spline modelling, data interpolation, electron therapy, insert factor
PACS: 02.60.Ed, 02.60.Pn, 87.55.Qr, 87.56.jk

1. Introduction

Electron beams are often used for the treatment of skin tumours, head and neck cancers, and breast boosts. The dose delivered by an electron beam is primarily dependent upon beam energy, patient specific field shape, SSD, and collimator design as well as patient anatomy. The specific shape used in the final aperture collimation of the treatment applicator is often unique for any given patient and will here be referred to as the insert as per AAPM TG 71 [1]. For our purposes here the portion of the output factor that is dependent on the insert will be called the insert factor. Even though there are methods in the literature for modelling the insert factor, in many centres it is often directly measured [1].

The modelling methods available range in complexity, accuracy, and resources required to implement. Historical methods such as the square root method are still in use and, when compared to measurement, can achieve uncertainties (represented to 1SD) as low as 2% [2]. Analytical modelling methods built atop of the Fermi-Eyges pencil beam model such as the lateral build up method have achieved uncertainties as low as 1% [3]. More sophisticated Monte Carlo methods have successfully predicted insert factors to an uncertainty as low as 1% [4]. Empirical models such as the sector integration method [5] have shown promise with a particular enhanced method achieving uncertainties as low as 0.6% [6] however it requires many measurements of circular fields.

To create a model that may be widely adopted in the clinical setting it must be easy to implement and have an uncertainty approaching that already available via measurement. If the use of electron fields is infrequent or high accuracy is desired then direct measurement for each irregular insert may

Email address: mail@simonbiggs.net (S Biggs)

be preferred over the implementation of a numerical method or the measurement of specific inserts for an empirical method. Most clinics have vast amounts of data currently available to them. Creating a model that builds upon this already collected data would be beneficial. An empirical model via parameterising insert shapes is a promising way of achieving these aims.

There are a number of parameters that can be constructed to characterise the dependence of delivered dose on insert shape. For the shapes where sections of the insert begin to reduce lateral scatter to the point of maximum dose, the smallest dimension of the shape becomes an important factor for prediction. Examples of models taking into account this smallest dimension are bivariate polynomial fits [7], equivalent radii methods, or building a shape out of sectors with given radii such as the sector integration method [5]. However the lateral scatter of the insert is not the only physical effect on the insert factor. Bremsstrahlung produced in the insert material, transmission, and scatter off the internal surface of the insert can have a combined effect of 2-4% on the insert factor [8]. The larger the area of the insert aperture, the lower the contribution to dose is from the insert material, however the insert scatter effect increases with collimating surface, which can be approximated as having a dependence on shape perimeter. These physical effects therefore may be aptly parameterised by the area and the perimeter of the insert shape. The work of Nair et al [9] demonstrates the feasibility of such parameterisation, and expansion on these ideas is a promising area of investigation.

Presented here is an empirical bivariate model based upon the two parameters width and the ratio of perimeter to area. The model takes arbitrary insert shapes as input data and interpolates between them (See Figure 1). Each shape is simplified to be represented by an equivalent ellipse, determined so that the width and area is equal. Modelling of insert factors is done using a smoothing bivariate spline model [10] and the Scientific Stack for Python [11] using the Anaconda Python 3.4 distribution.

2. Methods

2.1. Insert factor measurements

Insert factors were measured in RW3 with an Advanced Markus on an Elekta Agility linac. Definitions of electron output factors given here are as per AAPM TG 25, defined as the ratio of dose per monitor unit at d_{max} [12]. When the depth of maximum dose is shifted from the reference depth this depth was searched for to a 1 mm resolution. All depth differences took into account stopping power ratio corrections as per the protocol in IAEA TRS 398 [13]. The depth searching process was aided by automated relative dose plotting by a program written in python which performed the ionization to dose conversion.

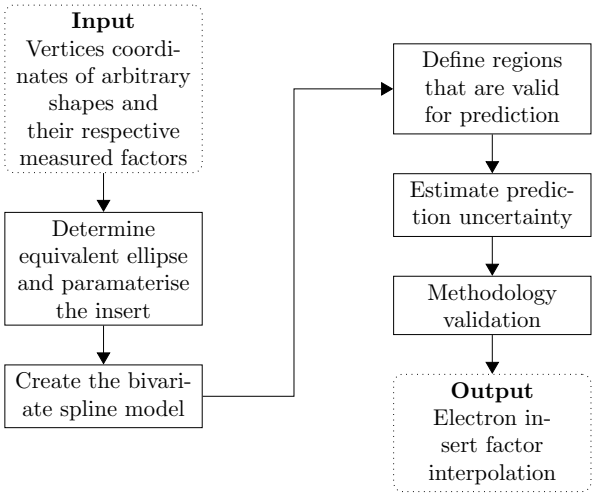

Figure 1: Flow chart of the equivalent ellipse spline model methodology.

For the purpose of methodology validation a large data set of 42 shapes was created. Of these, 24 shapes were pulled from previously treated clinical shapes stored within the treatment planning system. Supplementing the clinical shapes were 18 standard shapes, 7 circles and 11 ellipses. The circles ranged in diameter between 3.2 cm and 9.5 cm at 100 cm SSD. The aspect ratios of the ellipses varied between 1.6 and 4.3 with the highest aspect ratio ellipse being 3.2 cm × 13.6 cm. These 42 insert shapes were measured using 12 MeV electrons on the 10 cm applicator. Data was also collected to confirm the methodology for the remaining energies, 6, 9, 15, and 18 MeV with shapes over the clinical range of the 10 cm applicator. These consisted of the same set of five circles and three ellipses. The five circles ranged in diameter between 5.0 and 9.5 cm at 100 cm SSD. The highest aspect ratio ellipse chosen was 5.3 cm × 12.4 cm. This data was combined with previously measured clinical measurements ranging in number between four and seven depending on energy.

2.2. Equivalent ellipse parameterisation

The methodology proposed here for finding the equivalent ellipse is similar to the method for an equivalent rectangle as proposed by Hogstrom et al. reproduced in AAPM TG 71 [1].

The width of the ellipse was chosen to produce equivalent loss of lateral scatter effects between the insert and the ellipse. It is defined as the diameter of the largest circle fully enclosed by the insert shape. The circle meeting these conditions was found using the Basin-Hopping global optimiser [14], the BFGS local optimiser [15], and the Shapely geometry python module. This method of parameterising loss of lateral scatter assumes that the loss of scatter is primarily from the minor axis of the shape. If lateral scatter is being lost in the major axis of the shape discretion is required to determine if a similar loss occurs in the resulting equivalent ellipse.

The length of the ellipse is chosen so that there is similar bremsstrahlung production in the shielding between the ellipse and the insert. This is done by choosing the length so that the area of the ellipse is equal to that of the insert shape. This results in a similar volume of shielding for both the original insert and the equivalent ellipse. In this process the perimeter of the ellipse no longer remains equal to that of the original insert, however when comparing the inserts and their corresponding equivalent ellipses the difference in the resulting insert factor prediction was small.

The final parameters used for modelling were width and the ratio of perimeter to area defined using the equivalent ellipse. To calculate the perimeter (P) of the ellipse given its width (w) and length (l) Ramanujan's approximation was used as given in Equation 1:

$$P \approx \frac{\pi}{2}\left[3(w+l) - \sqrt{(3w+l)(3l+w)}\right] \qquad (1)$$

2.3. Bivariate spline model

Fitting was achieved using the smoothing bivariate spline [10] class packaged within SciPy [11]. This was used to create a fit for the measured insert factors against the two parameters width, and the ratio of perimeter to area. It was found that the required spline orders were no more than two in the width dimension, and one in the ratio of perimeter to area dimension. The spline order was kept as low as possible to reduce the impact of outliers on the fit and to allow for greater coverage of the valid region. Due to this choice of spline order, combined with the requirement to remove a data point for uncertainty estimation, the minimum number of unique equivalent ellipses required to create this model is eight. The smoothing factor used within the spline class was chosen to be equal to the number of input data points as recommended by the SciPy online documentation. The spline bounding box was expanded to include the requested point of prediction, this was important when extrapolating outside the original spline bounding box. The output of the spline model is an insert factor prediction function, $f(w, {}^{P}/_{A})$, taking the inputs width (w), and ratio of perimeter to area (${}^{P}/_{A}$) and returning an insert factor prediction.

2.4. Defining valid prediction regions of the spline

The spline fit is only valid in the regions where there is a sufficient amount of data. Here sufficient data is defined specifically for the point in question, such that the model weighting of the available data is at least that which would be given to one measurement at that point. Where there is insufficient data the spline is susceptible to large differences between prediction and reality. To quantify the valid region a numerical method was devised based upon how the spline reacts should it be forced to include a simulated outlier placed at the point in question. Should the spline deform more than half way to meet the outlier this implies that the spline gave more weight to the single simulated outlier than to the available measured data. In this case the requirement for sufficient data is not met. Conversely a deform distance of less than or equal to half way implies that the weight given to the data is equivalent to at least the weight of one data point. In this case the requirement for sufficient data is met. The ratio of deform distance towards the simulated outlier is denoted as the deformability parameter (ε).

The calculation of the ε is at the point in question $(w_i, [{}^{P}/_{A}]_i)$. It is calculated by creating two deformed insert factor prediction spline functions by adding a simulated outlier either above or below the curve and including this extra point in a spline fit to give the two deformed spline functions, $f_{x^+}(w, {}^{P}/_{A})$ and $f_{x^-}(w, {}^{P}/_{A})$. These simulated outliers were assigned at $(w_i, [{}^{P}/_{A}]_i)$ with an insert factor of $f(w_i, [{}^{P}/_{A}]_i) \pm \Delta$. The magnitude of Δ was chosen to be 0.02 so that the simulated outliers were approximately five standard deviations away from the original fit. Any choice of Δ sufficiently larger

than the range of expected factor measurements at $(w_i, [P/A]_i)$ will suffice. The ε is then calculated by first finding the maximum difference between $f(w_i, [P/A]_i)$ and $f_{x^+}(w_i, [P/A]_i)$ or $f_{x^-}(w_i, [P/A]_i)$, then second by defining the ε as the ratio of the maximum difference to Δ.

Using the ε, the prediction of $f(w_i, [P/A]_i)$ is declared either valid or invalid. If the ε is greater than 0.5 this implies that the result of the deformed spline from either of the functions, $f_{x^+}(w, P/A)$ or $f_{x^-}(w, P/A)$, had a result closer to the extreme point than the original prediction. Hence the smoothing bivariate spline placed more weight to a single data point than to the information that was originally available at $(w_i, [P/A]_i)$. Therefore in that case $f(w_i, [P/A]_i)$ is declared as an invalid prediction. When the ε is less than or equal to 0.5 then the information available at $(w_i, [P/A]_i)$ is, by the same reasoning, at least equivalent to one measured data point. In this case the prediction $f(w_i, [P/A]_i)$ is declared valid. This valid region is enclosed by a deformability boundary within which ε is less than or equal to the threshold of 0.5. For an example of calculating ε see the supplementary material.

2.5. Estimation of prediction uncertainty

To estimate the uncertainty in the model's prediction of measured insert factors the model is iteratively recreated with a point removed. Should the removed point be within the valid prediction region (see Section 2.4) of the recreated model then the difference between the recreated model's prediction and the measurement is recorded. This is repeated for all measured data points used to create the initial model. To estimate the uncertainty in prediction, the bias-corrected standard deviation is taken of the recorded differences between measurement and prediction.

2.6. Methodology validation

To validate the model, eight data points were randomly selected from the 12 MeV data set and used to predict the remaining data. This was achieved by creating a new fit with just eight data points and predicting the remaining data that fell within the new fit's valid region. The difference between prediction and measurement was recorded and then the process was repeated by selecting a new random subset of eight data points. This process was repeated in excess of one hundred thousand repetitions. The prediction differences of the results were assessed. The standard deviation of predictions from all repetitions was used as an approximation of the expected uncertainty for models with small data sets.

This validation method was extended to also test how the uncertainty in prediction varied with number of data points used, and how this was affected by the inclusion of hypothetical outlier factors. Outliers were created to be ± 0.02 from a true measurement. For any given number of measurements the approximate prediction uncertainty was calculated using a sufficient number of repetitions to minimise statistical noise. This was done for the scenarios with zero, one, and two outliers.

3. Results

The bivariate spline fit for the 12 MeV data set is provided in Figure 2. First it is represented in its native domain of width and ratio of perimeter to area, and second in the transformed domain of width and length. In this figure the deformability boundary provides the prediction boundary for the spline fit. Not all of the valid regions represent a physical insert. This is shown by drawing in the two physical bounds, the circle bound and approximate applicator bound. The circle bound is due to inserts not being physically able to have lengths smaller than their widths. The applicator bound is defined by the approximate upper limit of insert length as constrained by the applicator dimensions while taking into account the method of creating an equivalent ellipse.

The accuracy and precision of the model is demonstrated in the histogram of prediction differences for the 12 MeV data set given in Figure 3. This data set had 40 data points within it that were within the valid region for removal and prediction. The remaining two data points were not within the valid region of the recreated model that resulted after their removal. For these 40 measurements the uncertainty in differences between measurement and prediction was 0.4% (1SD). Furthermore the mean difference between measurement and prediction was 0.0%. For the remaining energies—6, 9, 15, and 18 MeV—the maximum recorded deviations between measurement and prediction were 0.8%, 0.8%, 0.9%, and 1.0% respectively.

The validation procedure which repetitively created models from random subsets of eight measurements within the 12 MeV dataset was able to predict the remaining measurements within their re-

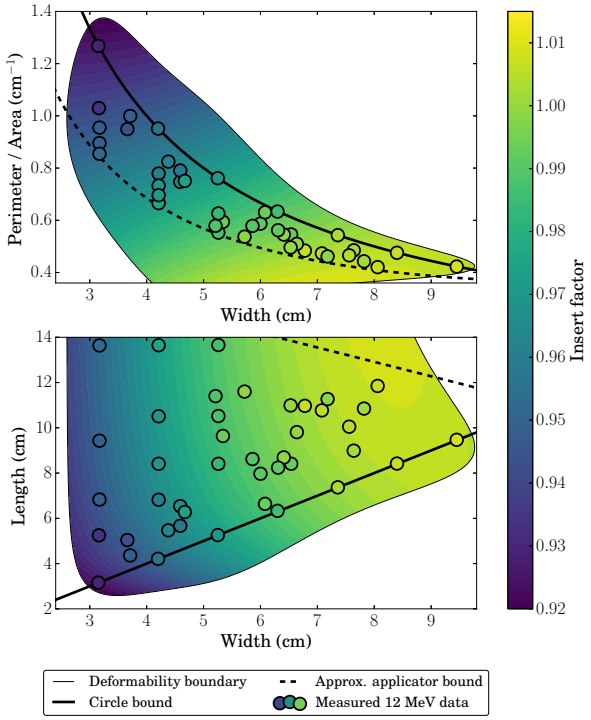

Figure 2: The bivariate spline fit for the 12 MeV 10 cm applicator set of measurements represented with respect to width and length, and width and ratio of perimeter to area. There are no inserts able to be predicted beyond the circle and applicator bounds.

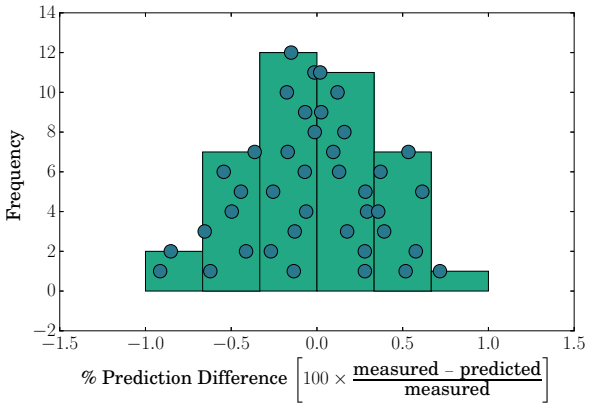

Figure 3: The percent prediction differences for the 12 MeV measured data set. Has a mean of 0.0% with an estimated population standard deviation of 0.4%. The scatter overlay is representative of the individual data points binned in the histogram.

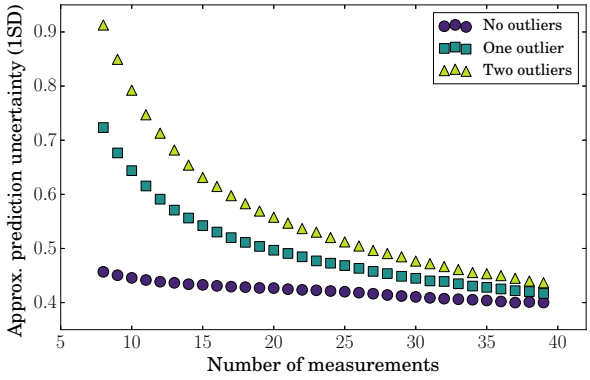

Figure 4: An approximation of the dependence of percent standard uncertainty on the amount of data collected given either zero, one, or two hypothetical outliers. Determined using the 12 MeV data set.

spective valid regions with an average uncertainty of 0.5% (1SD). In 10^6 predictions none of these random subsets of eight data points produced a prediction difference greater than 2%.

An approximation of the dependence of uncertainty on the amount of data collected—given either zero, one, or two hypothetical outliers—was produced for the 12 MeV dataset (see Figure 4). With no outliers the change in uncertainty with more measurements was minimal however, with an outlier, increasing the number of measurements by four was able to provide a 0.1% (1SD) improvement.

4. Discussion

The prediction uncertainty of 0.4% with no measurable systematic offset compares favourably to other methods available in the literature such as the improved sector integration method which achieves a prediction uncertainty of 0.4% atop a systematic offset of 0.2%.

This type of model may be introduced clinically in two ways. The first option is to measure every patient shape before treatment and make record of the electron relative output factor for each shape. Once eight or more measured data points are available for a given energy/applicator/SSD then these measurements can be used to begin predicting future factors. Should a new insert shape be required that is outside the valid prediction region, then this shape is measured and then put into the model. A second option is for the physicist to take the required measurements during commissioning. To produce a model with high coverage, with minimal

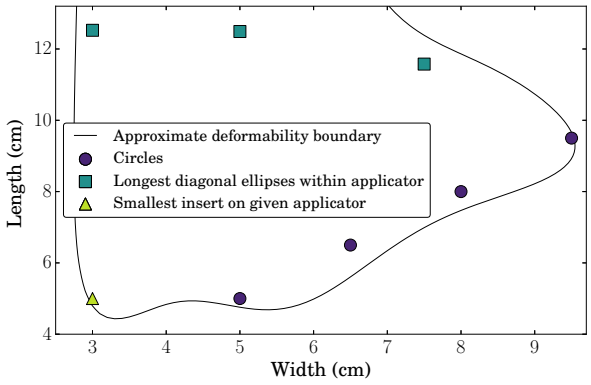

Figure 5: An example set of the minimum required number of measurements that could be taken during commissioning for a 10 × 10 cm applicator. To produce model redundancy a few clinical shapes should also be measured.

measurements, an initial bounding set of data can be collected as shown in Figure 5. However with eight measurements there is minimal redundancy in the model. As a result any outliers may heavily impact factor prediction. Therefore it is recommended that the first few times a user makes use of a given spline fit the clinical shapes should be measured, compared, and then also included within the model. Since the requirement for eight data points is due to the spline orders it is not expected that this requirement would differ with factors such as SSD, energy, or applicator size or shape.

The use of the residuals from this model equips the physicist in highlighting outliers in measured data. For the purpose of checking prior clinical measurements alone this model is useful. Furthermore, any future measurements are able to checked for consistency. When using data available in the clinic it is important to detect and intelligently rectify outliers. The resulting model will only be as good as the original data set.

This methodology has had no validation for electron field sizes smaller than 3.2 cm in width for 12 MeV. It is suspected that the model may begin to break down when the insert shape begins to lose a sufficient portion of lateral scatter to d_{max} along the long axis of the shape. This would be less of an issue if only circles are used for these small field sizes as opposed to ellipses or irregular shapes. Should modelling of these non-circular small shapes be required a parameter representing loss of lateral scatter could be used instead of the width parameter used within this model. An example of such a potential parameter which would simulate the loss of lateral scatter could be the volume under a bivariate Gaussian bounded by the insert shape. The spread for the Gaussian would need to be determined empirically. It is not expected that larger applicators, either square or rectangular, would result in higher levels of uncertainty. This is founded upon the greatest change in insert factor, with insert shape and size, being change in lateral scatter contribution which is more pronounced in smaller field sizes.

All code used to reproduce the results of this paper is available online and may be freely used and distributed according to the Affero General Public Licence version 3+ licence. Attempts have been made to make the code transparent and it is up to the user to confirm that the code is fit for use.

5. Conclusion

A method that is easy to implement, has an uncertainty approaching direct measurement, and builds upon the vast amounts of data already available in most clinics has been proposed. It was achieved by the parameterisation of insert shapes into a bivariate spline model. The 12 MeV validation data set had a prediction uncertainty of 0.4% (1SD). For the remaining energies the maximum recorded deviation between measurement and prediction was 1.0%. The validation methodology showed that one may expect an approximate uncertainty of 0.5% (1SD) when as little as eight data points are used. The level of accuracy combined with the ease with which this model can be generated demonstrates its suitability for clinical use.

[1] Gibbons JP, Antolak JA, Followill DS, Huq MS, Klein EE, Lam KL, et al. Monitor unit calculations for external photon and electron beams: Report of the AAPM therapy physics committee Task Group No 71. Med Phys 2014;41(3):031501. doi:10.1118/1.4864244.

[2] McParland BL. A parameterization of the electron beam output factors of a 25 MeV linear accelerator. Med Phys 1987;14(4):665–9. doi:10.1118/1.596145.

[3] Gebreamlak WT, Tedeschi DJ, Alkhatib HA. Dose calculation for electron therapy using an improved LBR method. Med Phys 2013;40(7):071717. doi:10.1118/1.4810938.

[4] Verhaegen F, Mubata C, Pettingell J, Bidmead AM, Rosenberg I, Mockridge D, et al. Monte Carlo calculation of output factors for circular, rectangular, and square fields of electron accelerators (6–20 MeV). Med Phys 2001;28(6):938–49. doi:10.1118/1.1373402.

[5] Jursinic PA, Mueller R. A sector-integration method for calculating the output factors of irregularly shaped electron fields. Med Phys 1997;24(11):1765–9. doi:10.1118/1.597962.

[6] Gajewski R. An enhanced sector integration model for output and dose distribution calculation of irregular concave shaped electron beams. Med Phys 2009;36(7):2966–75. doi:10.1118/1.3148583.
[7] Turian JV, Smith BD, Bernard DA, Griem KL, Chu JC. Monte Carlo calculations of output factors for clinically shaped electron fields. J Appl Clin Med Phys 2004;5(2):42–63. doi:10.1120/jacmp.v5i2.1976.
[8] Faddegon BA, Villarreal-Barajas JE. Final Aperture Superposition Technique applied to fast calculation of electron output factors and depth dose curve. Med Phys 2005;32(11):3286–94. doi:10.1118/1.2068947.
[9] Nair RP, Nair TKM, Wrede DE. Shaped field electron dosimetry for Phillips SL75/10 linear accelerato. Med Phys 1983;10(3):356–60. doi:10.1118/1.595282.
[10] Dierckx P. An algorithm for surface-fitting with spline functions. IMA J Numer Anal 1981;1(3):267–83. doi:10.1093/imanum/1.3.267.
[11] Oliphant TE. Python for scientific computing. IEEE Comput Sci Eng 2007;9(3):10–20. doi:10.1109/MCSE.2007.58.
[12] Khan FM, Doppke KP, Hogstrom KR, Kutcher GJ, Nath R, Prasad SC, et al. Clinical electron-beam dosimetry: Report of AAPM radiation therapy committee Task Group No 25. Med Phys 1991;18(1):73–109. doi:10.1118/1.596695.
[13] Andreo P, Burns DT, Hohlfeld K, Huq MS, Kanai T, Laitano F, et al. Absorbed dose determination in external beam radiotherapy: An international code of practice for dosimetry based on standards of absorbed dose to water. Vienna: International Atomic Energy Agency; 2000.
[14] Wales DJ, Doye JPK. Global optimization by Basin-Hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. J Phys Chem A 1997;101(28):5111–6. doi:10.1021/jp970984n.
[15] Nocedal J, Wright S. Numerical optimization. New York: Springer Science+Business Media; 2006.

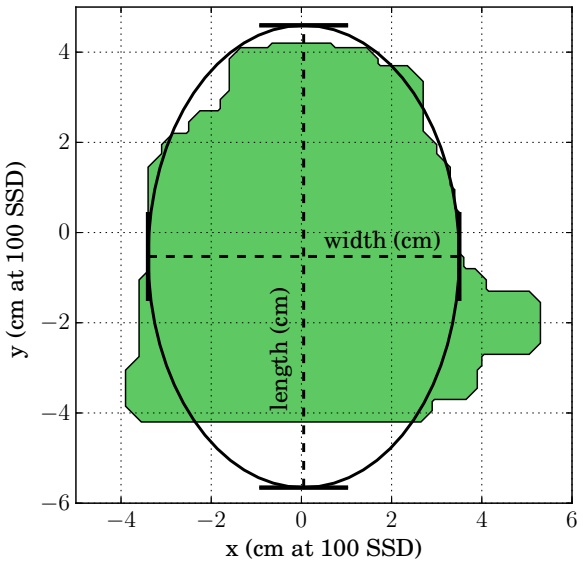

Figure 6: Example equivalent ellipse with the resulting width and length labelled.

6. Supplementary material

The following figures (Figures 6, 7, and 8) were submitted as supplementary material.

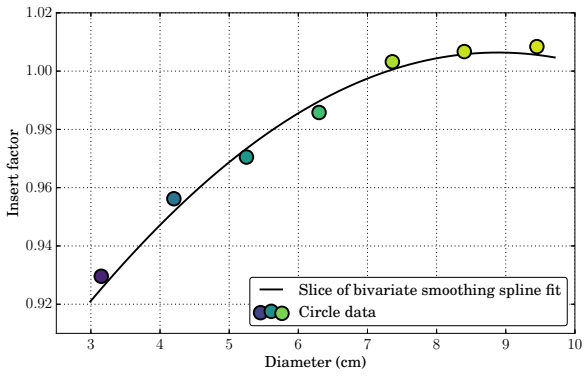

Figure 7: Slice of the bivariate smoothing spline fit giving a fit for the circle insert factors

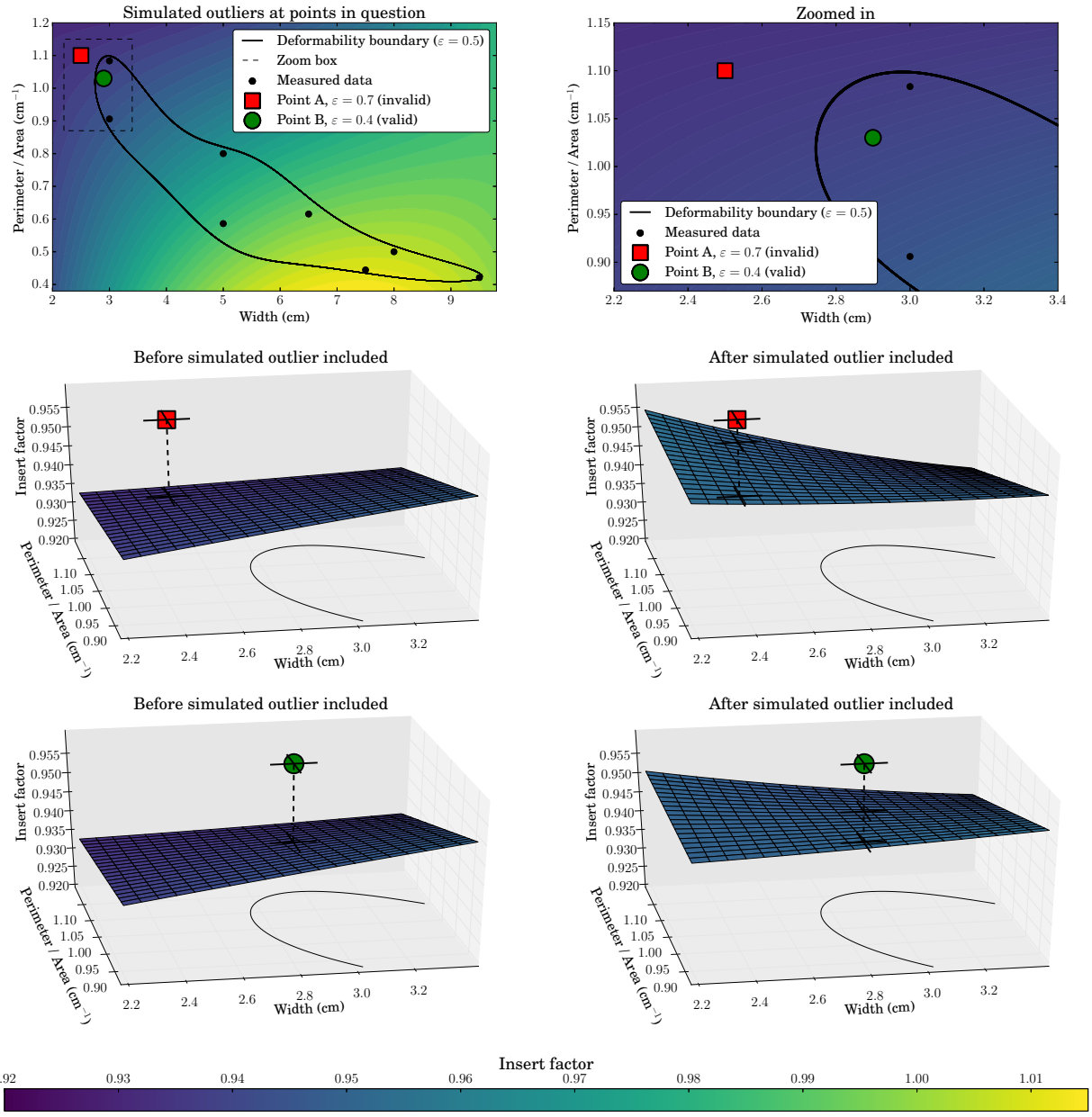

Figure 8: A demonstration of the method used to determine whether or not a point of interest is valid for prediction by testing how the model reacts to the inclusion of a simulated outlier at the point in question. The comparison is between a point of insufficient data (point A) and a point of sufficient data (point B). At point A a simulated outlier is placed denoted by a square. At point B a simulated outlier is placed denoted by a circle. At point A the ε is 0.7 which implies the model reacts to a simulated outlier at this position by weighting it more than all the available data. Therefore a prediction at point A is declared invalid. At point B the ε is 0.4. This implies that at least one measured data point worth of weight from the measured data exists at this point in question. Therefore a prediction at point B would be declared valid.