Evaluating a single treatment planning beam model for multiple beam-matched linear accelerators

by

Liebner Koen

UNIVERSITY OF THE **FREE STATE**
UNIVERSITEIT VAN DIE **VRYSTAAT**
YUNIVESITHI YA **FREISTATA**

Submitted in fulfilment of the requirements in respect of the MMedSc degree qualification in the department of Medical Physics in the Faculty of Health Sciences, at the University of the Free State, South Africa

2018

Supervisor: Dr W Shaw

Co-supervisor: Dr FCP du Plessis

SUMMARY

Key words: 2D array, diamond detector, profiles, spatial resolution, beam-matching, MLC error, Treatment Planning System, Intensity Modulated Radiation Therapy

The success of cancer treatment with radiation is highly dependent on the ability of the Treatment Planning System (TPS) to accurately calculate doses that would be delivered to the patient. The quality of TPS commissioning data, based on Linear Accelerator (Linac) measurements in water, largely determines the quality of the TPS beam model derived from this data. Modern treatment techniques such as Intensity Modulated Radiotherapy (IMRT) require highly accurate dosimetry equipment used for TPS commissioning. Once derived, the beam model should be verified with a range of tests other than commissioning procedures to test the beam model against Linac output.

The study aims to investigate equivalences and differences between 5 Siemens Artiste^{TM} Linacs of similar output, referred to as beam-matched Linacs, and how a single TPS beam model (Monaco^{TM}) can potentially be utilized for treatment planning for any of the Linacs.

Generally, dosimeters used for TPS beam data collection differ largely from those used for post-modelling verification measurements. The study investigates the correlation between a high-resolution detector (microDiamond) typically used for collecting commissioning beam data and a post-modelling verification 2D array detector (Mapcheck2^{TM}). Measurement resolution of Mapcheck2^{TM} was increased to 1 mm by repetition measurements, manually stepping the device in-between the detector-less spaces and software developed to convert data to a readable format. Dose profiles from Mapcheck2^{TM}, with increased measurement resolution, and microDiamond agreed well. This method was further used to accurately determine Multi-leaf Collimator (MLC) errors from a range of MLC stop positions across the radiation field for each Linac respectively. This allowed for quantitative comparisons that showed significant differences between the MLCs of the 5 Linacs. A new

radiological calibration curve (containing software MLC offset values) to reduce MLC errors were proposed for each Linac respectively.

Clinical IMRT treatment fields were measured with increased array resolution on each Linac and compared to dose calculations from the TPS. Gamma pass rates, from different measurement resolution and evaluation software, were above 95% for a criterion of 2%/2mm with confidence limits above 90%. Therefore, it is concluded that differences between Linacs in terms of IMRT treatment delivery were insignificant. Hence, an overall agreement in comparing the 5 beam-matched Linacs to IMRT dose calculation from a single TPS beam model respectively. Limitations of the planar IMRT QA evaluation method were discussed as well as the inability of this method to detect seemingly significant MLC errors.

TABLE OF CONTENTS   **Page**

ABBREVIATIONS

Computed Tomography (CT)

Organs at Risk (OARs)

Treatment Planning System (TPS)

Linear Accelerator (Linac)

Multi-Leaf-Collimator (MLC)

Intensity Modulated Radiation Therapy (IMRT)

Standard Deviation (SD)

Three-Dimensional (3D)

Planning Target Volume (PTV)

Dose Volume Histogram (DVH)

Sliding Window (SW)

Head and Neck (H&N)

Percentage Depth Dose (PDD)

American Association of Physics in Medicine (AAPM)

Monte Carlo (MC)

Virtual Source Model (VSM)

Quality Assurance (QA)

Mega Voltage (MV)

Gamma Index (γ)

Millimetre (mm)

Distance to Agreement (DTA)

Dose Difference (DD)

Two Dimension (2D)

Centimetre (cm)

Central Axis (CAX)

Transmission Probability Filter (TPF)

X-ray Voxel MC (XVMC)

Poly (methyl methacrylate) (PMMA)

Source-to-Surface Distance (SSD)

Tissue Scatter Correction Factor (TSCF)

International Atomic and Energy Agency (IAEA)

Monitor Units (MU)

Tissue Phantom Ratio (TPR)

Volumetric Arc Therapy (VMAT)

Relative Electron Density (RED)

Depth of Dose Maximum (dmax)

1. INTRODUCTION

1.1) Overview of Radiotherapy Process

The goal of radiotherapy is to deliver a tumoricidal radiation dose to a tumour and limit dose to the surrounding normal organs and tissues [1]. To reach this goal, diagnostic imaging modalities such as a Computed Tomography (CT) scan of the patient, are used to delineate the target (tumour) as well as the Organ at Risk (OAR) volumes (normal tissue). These data are transferred to a sophisticated computerized Treatment Planning System (TPS) for treatment planning. The TPS simulates the radiation dose that would be delivered to the tumour and OARs and the planned treatment is then transferred to and delivered on a Linear Accelerator (Linac).

The role of the TPS is to accurately represent the planned treatment execution and calculate the corresponding dose distribution as optimized by the treatment planner to achieve the goals as set out above, by monitoring doses to target volumes and OARs depending on clinical treatment intent. The TPS dose calculation engine utilizes a virtual beam model to analytically describe the radiation beam of the Linac. Such a beam model should be derived for each Linac available for treatment in the clinic, because their radiation beams differ in terms of energy spectra, energy fluence, dose rate, fluence distributions and the variations of these characteristics within the treatment beam with different treatment setups. The model thus mimics the treatment beam to the plane beneath the collimator system and subsequently calculates the distribution in dose that would be absorbed in the patient by simulating radiation transport and its interactions with tissue [2].

Generally, the Linac collimator system consists of high density shielding material to conform the radiation beam to the target volume, thus reducing dose to the surrounding OARs. Currently, most Linacs are equipped with a Multi-Leaf-Collimator (MLC) with tens of leaves of small width (e.g. 5mm) which can move independently to further shape the radiation beam to an arbitrary shaped target volume. With shaping and combining of multiple radiation beams, a prescribed radiation dose is delivered to

the target volume while monitoring, or adjusting, the predicated dose to the OARs, termed a conventional treatment technique. This technique was further refined with the introduction of Intensity Modulated Radiation Therapy (IMRT), where the intensity of the radiation beam is varied to achieve extreme dose gradients with the use of multiple, small (or large, elongated and complex) fields size MLC segments.

1.2) Accuracy Requirements of Radiotherapy

Every treatment technique has a dose delivery uncertainty and another goal of radiotherapy is to quantify and minimize this to as low as possible. The need for the standard deviation (SD) in the mean dose in the target volume to be as low as 3-5% have been widely reported [3] [4] [5] [6]. This standard deviation combines several uncertainties in the radiotherapy treatment delivery process, for example tumour localization and machine delivery variations such as calibration of the beam under clinical conditions, MLC stop positions etc. and these combines as inaccuracies in the three-dimensional (3D) dose delivery to a patient.

One of the most important components for accurate IMRT treatment delivery is, among many others, MLC positional consistency [7]. It has been reported that seemingly small errors in MLC positioning can result in large dosimetric errors. Recently a 0.5 mm leaf gap was induced in the TPS model, resulting in a dose deviation of 11% for a narrow 5 mm sliding window test beam. More clinically relevant results showed a dose difference up to 4% in the Planning Target Volume (PTV) predicted Dose Volume Histogram (DVH) for sliding window (SW) Head and Neck (H&N) IMRT [8].

1.3) Overview of Radiotherapy Beam Modelling Process

The construction of the model is performed through an analytic derivation process. Since the beam characteristics cannot be easily measured directly, the model parameters should be adjusted in such a way that it replicates all aspects of the radiation fluence of the actual treatment machine, which will be used to calculate dose distributions. These are a series of machine output measurements, referred to as Linac commissioning measurements, with an appropriate detector that are mostly

performed in a large water filled Perspex tank. The measured dose distribution dataset is characteristic of the particular machine under investigation and consist of Percentage Depth Dose (PDD) data, profiles, scatter factors, dose rate, transmission through the collimating system, etc.[9][10] and are imported into the TPS to serve as a reference set of data for beam modelling. Once the beam model parameters have been adjusted to produce dose distributions in water that are equivalent to the measured dataset, it can be said that the adjusted TPS model is representative of the actual Linac.

The characterization of treatment machines is a time-consuming process, therefore Linac manufacturers can provide customers with reference or "golden" TPS data to allow for relaxation of commissioning measurement requirements. Another method to reduce time spent on characterizing treatment units is to implement a number of Linac models with similar output, a scenario referred to as "beam-matching" [11][12]. If machine designs are equivalent with similar output, a single, reference TPS model might reflect their output (in terms of model and design) to within an acceptable level of uncertainty for treatment unit characterization. However, task group 106 of the Therapy Physics Committee of the American Association of Physicists in Medicine (AAPM)[13] raised concerns about the use of a reference data set or a single TPS model, which include: lack of evidence on manufacturing reproducibility acceptable for clinical use; although acceptable agreement with the golden or reference data set may be found in individual checks, it may be that some clinical setups will have multiple errors, which combine to produce unacceptable results.

More recently Monte Carlo (MC) based models were introduced in the commercial treatment planning environment [14]. MC dose calculation is regarded as the gold standard in terms of dose calculation accuracy and closely models the actual radiation transport in the energy deposition process. A MC model is based on computer simulations summarizing the transport of millions of radiation particles based on each particle's probability to interact with other atomic structures. MC is particularly recommended for calculation of techniques such as IMRT, as it can simulate secondary electrons explicitly in small radiation fields. Therefore in the presence of lateral electron disequilibrium it is expected to be more accurate than other patient dose calculation algorithms [15]. Ultimately a more detailed and accurate breakdown in

the transport of radiation though the Linac collimating system is performed, as well as accurate dose calculations within the patient model [16].

Potentially, these advances in dose calculation and treatment beam model construction can be used to argue that extensive machine measurements may become less important in future [15]. However, some machine measurement are still required to customize a MC virtual beam model or a virtual source model [17]. With the introduction of a Virtual Source Model (VSM), it was demonstrated that time-consuming MC simulations of the whole accelerator head are unnecessary for each radiotherapy treatment [18]. Instead, reference parameters are derived only once directly from full MC simulation per accelerator type, and then analytical functions are used to fit certain open model parameters for an individual accelerator based on water measurements [19].

Regardless of dose calculation algorithm, dose calculations from the completed model should always be verified against machine measurements or other Monte Carlo systems [16]. These comparisons include simple open beams as well as more complex beams representative of typical clinical cases used in the clinic [2, 20 21]. There exist well defined criteria that could be used to ensure accurate TPS dose calculation compared to the Linac output [22 23]. In the case of more advanced treatment techniques, such as IMRT, the derivation of optimal model parameters becomes even more important and the criterion for acceptance of the model is more strict for such treatments [24].

1.4) Overview of Beam-Matching concept

As mentioned, "beam-matching" means treatment beams of the unit being installed are modified in such a way that the dosimetric characteristics meet reference values within some specified criteria. Resultantly, treatment planning for 2 beam-matched machines can be performed using a single TPS model for treatment on either machine. This poses a clinical advantage should one machine be out of operation, patients can still be treated on another machine without a break in treatment schedule. However, even though treatment machines may be matched, per-patient Quality Assurance (QA) measurements are still performed as "per-plan" errors or deviations from the

intended plan may still occur, especially delivering IMRT treatment plans. Other potential, non-clinical advantages of beam matching include the time saved not having to perform extensive initial machine measurements on each machine as well as the cost saved in requesting a single beam model for multiple Linacs. Though this hypothesis could improve workflow in a radiotherapy clinic, it may well be seen as an increased risk for inaccurate treatment dose delivery, as multiple small differences in machine characteristics may be acceptable reviewing a treatment planning model only, but these may combine to significant differences in the delivered dose [6] [13] and this warrants a full investigation.

1.5) Fundamentals of beam matching

Vendor stated agreement between Linacs may not be sufficient in a clinical environment, especially for IMRT delivery. Vendor's acceptance criteria are usually: 1) limited to single point comparisons on PDDs and profiles, 2) based on ionization chamber measurement with no or incorrect chamber offsets, 3) has a limited set of beam geometries and 4) usually do not include output factor comparisons [11].

Previously, 8 beam matched Varian Linacs were compared and it was reported all photon and electron beams, except the 15 MegaVoltage (MV) beams in two units, could each be represented by one treatment planning model [25]. Excluded energies were due to different output factors, possibly due to different flattening filters, a scenario not identified during the initial acceptance criteria. A recommendation was including the TPS calculated dose as the reference data for the matching procedure, because if for instance the TPS predicts a slightly higher dose than delivered by the reference Linac and the other Linac delivers a dose slightly lower than the reference Linac, the two uncertainties add up. Alternatively, the TPS calculated dose could be centred around dose deviations between 2 Linacs. Therefore, to ensure a single TPS model can be used for treatment planning on multiple machines, per machine measurements should be individually verified against the said model.

Whenever a measurement is made and compared to a calculation, one can expect some difference to be seen. If the difference is within a reasonable confidence limit,

then the result can be considered acceptable. The Gamma Index (γ) evaluation criteria can be used for this purpose [22]. This criterion is based on low (percentage) and high dose (distance) region differences and can also be used for complex treatment fields, also described in AAPM TG53 [10]. In comparing data from 2 beam-matched Linacs, Hrbacek *et all* (2007) reported an overall pass rate of 70% on γ values less than 1 using a γ criterion of 1%/1mm for all scans (PDD and profile) and attributed the failed 30% to measurement inaccuracies e.g. the PDD build-up region or out of field low dose regions. Therefore they recommended a 2 millimetre (mm) Distance to Agreement (DTA) and 2% Dose Difference (DD) when performing γ analysis for scan comparisons [12], also recommended by others [26]. Venselaar, Welleweerd and Mijnheer (2001) presented a γ criterion based on the level of beam geometry complexity (shown in table 1 below)[27] and shown by the Swiss Society for Radiobiology and Medical Physics [28].

Description	Small dose gradient (%)	Large dose gradient (mm)
Reproduction of basic beam data set	1	1
Open and wedged beams	2	2
Irregular, MLC and asymmetric fields:		
Central beam axis	2	2
Off-axis (dose profiles in any direction)	3	3
IMRT fluences	3	3

Table 1: γ *criteria per beam complexity*

Confidence limits were described as a useful tool to further evaluate profile data, especially when many comparison points are available, because single points outside of the chosen tolerance does not necessarily lead to a negative overall result [27,2]. This is based on percentage difference values between different regions e.g. penumbra region or central axis region of a profile, indicating whether difference seen are real or not within a certain confidence percentage. That study also noted that points within

the penumbra region or out of the field should be excluded or handled in a separate investigation.

Further, when comparing IMRT dose distributions from one TPS measured on more than one Linac, it is important not to only evaluate overall γ analysis pass rates, but to also compare failed individual points as these may be intrinsic to all machines evaluated [29]. Failed points could indicate: 1) a particular machine output adjustment is required, 2) a particular model parameter adjustment is required, 3) the deviation cannot be corrected for in the treatment planning model or 4) it could indicate the inability of the QA measurement device to accurately represent the dose delivered. However, differences between calculations and measurements can only be meaningfully evaluated if the uncertainties are understood [24]. These include: 1.5) machine measurement uncertainty, 1.6) limitations in the dose delivery mechanisms e.g. MLC's and limitations in the accuracy of dose calculations.

1.6) Machine measurement uncertainty

Different detector types, mostly diode and ion chamber, are commercially available for Linac and beam model commissioning. Array devices are widely used as a tool to measure and compare actual machine delivered to TPS calculated doses in two (2D) or three (3D) dimensions. The main benefit of an array device is quick acquisition of many measurement points in a single measurement, especially for complex IMRT dose distributions.

Generally, beam data used for generating a TPS model is collected in a water tank with a small diode detector and/or an ionization chamber (with a larger sensitive volume), but the verification of the same model is done with film or a detector array. Therefore, these measurement devices should be compared in terms of dosimetric capabilities and it is important to understand the differences between them before meaningful interpretation of IMRT QA results can be made [30] [31].

Several factors can contribute to uncertainty in machine commissioning and model verification measurements, such as the equipment used e.g. type of phantoms and detectors and the setup thereof, data acquisition and processing e.g. scan mode and

normalization (software used), machine output variation and the individual performing the measurements [32] [33]. For example, a detector with insufficient spatial resolution used for beam data collection and subsequent model construction can lead to a 10% difference in TPS calculated profiles compared to film measurements [34]. Three uncertainties associated with radiation detectors will be discussed here: 1.6.1) resolution and size 1.6.2) data handling and 1.6.3) setup accuracy.

1.6.1. The influence of radiation detector resolution and size on machine measurements

It has been reported that diode detectors are beneficial due to its small measurement volume, but can be dose rate dependant [34] [35]. On the contrary, ionization chambers are generally dose rate independent, but have a larger size and is therefore prone to volume averaging effects: should they be partially irradiated, dose can potentially be underestimated when measuring output factors or blurring of the penumbra can occur when doing profile measurements [30] [36].

Dose deviations between output factors measured with different detector sizes showed differences of up to 35% for a 1 × 1 cm^2 field size [36]. As a guideline, a difference of more than 3% can be expected when the detector size is more than 3/4 of the field size [37]. Also, standard ionization chambers commonly used for beam data collection (with volumes in the order of 0.1 - 0.2 cm^3) can cause penumbra increases by 0.2 to 0.3 centimetre (cm), which is large enough to influence IMRT calculation accuracy [38].

The disadvantage of most arrays is the spacing distance between detectors (up to 10 mm) which may limit the ability to detect machine component variations leading to dosimetry errors. For quantitative comparisons between pre-and post-modelling detectors, the spatial resolution of the array devices must be increased, especially in the penumbra region [39]. An increased spatial resolution is depicted by a shorter distance between measurement points. Figure 1 shows a comparison between a beam data collection detector (scanned in 1 mm increments) vs an array profile, illustrating the inability of the array to accurately measure beam penumbra due to a lack of measurement points. In this example of a 5 × 5 cm^2 field size, the diode detector had 16 penumbra measurement points compared to 2 from the array. Some

vendors improved their spatial resolution by manufacturing detectors in checkerboard configuration, resulting in a detector centre to centre spacing of e.g. 7.1 mm instead of 10 mm.
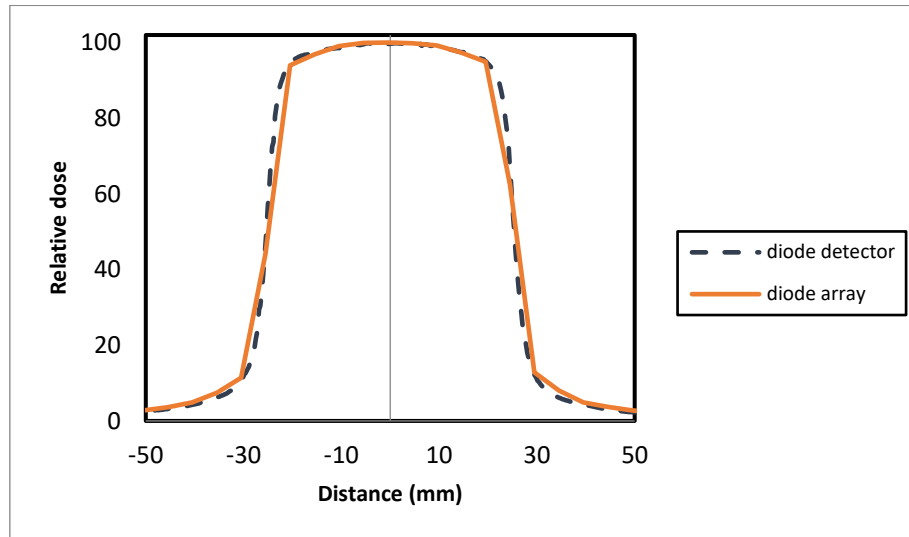

Figure 1: Example profiles from 2 detectors showing the difference in penumbra shape due to spatial resolution of measurement points

The combined effect of detector size and resolution has been described mathematically by the Nyquist sampling theorem [40], stating a sampled function is fully represented by the set of sample values if the sampling frequency is at least twice the highest spatial frequency [41].

Poppe *et al* (2007) quantified the required sampling frequency for IMRT measurements by evaluating the Fourier transform of a typical IMRT dose profile, measured with film [42]. This frequency spectra contained considerable information up to 0.1 mm^{-1}. A minimum sampling frequency of 0.1 mm^{-1} and therefore a spatial frequency or measurement resolution of 0.2 mm^{-1} (or 5 mm) was reported as an array requirement to accurately present measured IMRT dose distributions. Therefore, most arrays would require interpolation between measurement points to fulfil this requirement. This requirement was further illustrated by merging of multiple detector acquisitions e.g. using an array with detectors spaced 10 mm, one additional measurement after a 5 mm shift of the device can increase spatial resolution twofold [40], thus satisfying the Nyquist sampling theorem. It was further shown when measuring small field sizes in the order of 1×1 cm^2, a spatial frequency of at least 0.4 mm^{-1} would be required.

It has also been shown that the chosen γ criterion to evaluate two distributions is related to the amount of dose points (measured or interpolated) [43]. The resolution of the evaluated dose points should be not greater than 1/3 of the DTA criterion. For example, at least 3 points is required when evaluating a DTA of 1 mm. However, in comparing measured with planned doses from typical clinical IMRT cases, the effect of this requirement may only become noticeable when DTA becomes smaller that the distance between data points [44]. It was further reported that the effect of detector resolution and size are negligible when a γ criterion higher than 1.5% and 1.5 mm is applied for analysing the overall dose fluence of an IMRT treatment [45].

In summary, most arrays should be sufficient for pre-treatment plan QA for specialized techniques such as IMRT. However, the effect of array resolution and detector size are more important when accessing the full behaviour of an IMRT delivery system and these properties should be studied [46] [42] [47].

1.6.2. Data Handling

2D array doses generally consists of a matrix of truly measured or reference as well as interpolated dose points and it is important for users to understand the structure of dose matrixes.

In terms of measured dose distributions, ideally only truly measured data points would be beneficial in decreasing the uncertainty in measurements, especially when the aim is to detect errors in the order of 1 mm or smaller e.g. MLC positional errors. As mentioned this can be achieved by manual stepping of the device in small increments thereby increasing the dose matrix and potentially replacing interpolated with truly measured values. Unfortunately, commercial 2D array software such as Sun Nuclear SNC Patient™ (Sun Nuclear, Melbourne, FL) and VeriSoft® (PTW, Freiburg, Germany) has a limit on handling of increased measured dose matrixes and only measurements with step increments of 5 mm can be merged successfully. With the aim of verifying a beam model as accurately as possible using a 2D array, and possibly lowering the criterion to even 1%/1mm for simple measurements, it is believed a truly measured dose matrix with higher resolution, ideally in the order of 1mm, should be used [47].

In terms of reference dose distributions (typically TPS calculated), it is important to consider how commercial software handles it before comparison to measurements

through γ analysis. For example, some would interpolate TPS dose distributions to 1 mm [48]. Others, in the case of ionization chamber arrays, would account for the volume averaging effect by convolving the TPS dose distribution with the 2D lateral dose response function of a single chamber [40] [42] [49] [50], a correction supplied by the vendor, before comparison with the array measured dose distribution.

1.6.3. Device setup

Setup of radiation detectors should be done as accurately as possible [50]. Detecting small errors with a 2D array requires the user to establish the reference coordinate system which can be challenging [51]. The effect of inaccurate positioning of a diode array when obtaining an array sensitivity curve, a large open field procedure determining a calibration factor for each detector based on its response to radiation, was shown by Wang *et al* [52]. A 0.86% calibration error with a 1 mm misalignment was reported. A systematic error of < 1.5% was also shown, assuming the device should generally be aligned within 2mm of its intended location and concluded the array is not sensitive to the calibration error caused by misalignment of the device [53]. However, the focus of that work was to emphasize the importance of this phenomena when delivering flattening filter free beams. A 1 mm uncertainty in positioning the device for routine QA has been reported by others [54] [55].

1.7) Limitations in the Dose Delivery Mechanisms (MLCs)

As each MLC system, regardless of beam-matched status, may differ due to e.g. small differences in the installation and software calibrations, MLC QA tests such as the picket fence test can be performed to evaluate the accuracy and stability of each Linac's MLCs [56]. MLC QA results from 5 centres have been studied, all visually inspecting a picket fence delivered on film and reported a positional accuracy of 0.2 mm. They also reported a 1 year stability with no deviation in MLC positioning in any of the centres [57]. However, quantification of film results is not easy to perform and at the time of the publication some of the authors explored the possibility of using 2D arrays for this purpose, which nowadays is more commonly used. It has been reported that a diode array with detector spacing of 7.1 mm can detect a systematic MLC position error in the order of 1 mm in all MLCs [58]. It was further demonstrated by

shifting an array device in 1 mm increments, submillimetre MLC positional errors can be detected [59].

When doing MLC positional measurements with an array the position of the profile is very important, because a Central Axis (CAX) profile (parallel to the MLCs) may represent a profile in-between two adjacent MLCs whereas half a leaf dimension shift e.g. 2.5 mm of the profile in the perpendicular direction could be a better presentation of actual MLC positions.

Graves et al.(2001) compared the radiation MLC field size with the TPS calculated field size and found deviations up to 3 mm following a light field calibration procedure [60]. Even though this type of calibration procedure is no longer used for newer generation Linacs, a reproducible, fixed calibration procedure is still required with soft pot MLC offsets applied if necessary. The latter can sometimes be applied at the Linac itself or be incorporated into the TPS. Therefore, it is very important to understand the process and to correctly verify MLC positions in comparing measured and TPS calculated doses, and subsequently following future calibration procedures.

2. AIM

The aim of this study is to evaluate the use of a single beam model for IMRT treatment planning for multiple beam-matched linear accelerators.

3. THEORY

The reference Linac and TPS used in this study are briefly described in this section. A full description of these are beyond the scope of this document. Instead the physical characteristics and calibration method of the reference Linac MLCs (3.1) are given, because, as discussed before, MLCs can influence the accuracy of IMRT treatment delivery. This is followed by a description of the reference TPS (3.2) in an overview of the virtual source model, a MC particle generator and the transmission probability filter that calculates how these particles transcribe through the beam modifiers. Also, how the model can be further adjusted (from post modelling measurements) to achieve the best possible presentation of the Linac. Finally, to quantify the differences between Linac measured and TPS calculated doses the reference evaluation method (γ analysis) is discussed in section 3.3.

3.1) Reference Linac

Description of MLC

A Siemens® ARTISTE™ Linac (Siemens Medical Solutions, Concord, CA) has 160 MLC™ (two opposing banks of 80 MLCs each), each with a width of 5 mm and therefore a maximum fields size of 40 cm, which can overtravel by 20 cm. The two MLC banks are denoted as x1 and x2. The MLC has a single focus design, meaning the leaves do not move in an arced manner with reference to the radiation source, but moves in a single plane. The MLCs are designed with a curved end to reduce the gap between two opposing MLCs (figure 2) [61]. Also, within the same MLC bank it is arranged in an alternating pattern of upper and lower MLCs (figure 3), with a tilted MLC design to reduce interleaf leakage (figure 4). Interleaf leakage is the amount of radiation transmission between adjacent MLCs. The MLCs are tilted or focused 2.6 mm or 0.37° from the real x-ray source or focal spot position, resulting in a so-called triangular tongue and groove effect [62]. Therefore, with this design the primary x-ray beam has no direct path in between adjacent MLCs because there is a triangular overlap between them.
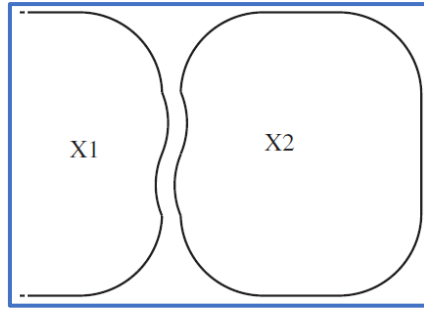
Figure2: MLC curved or S-shaped end [61]
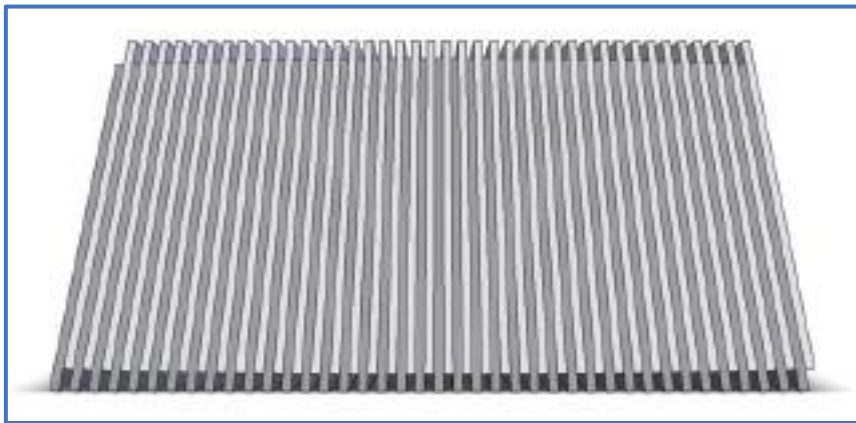
Figure 3: Pattern of upper and lower MLC's [63]
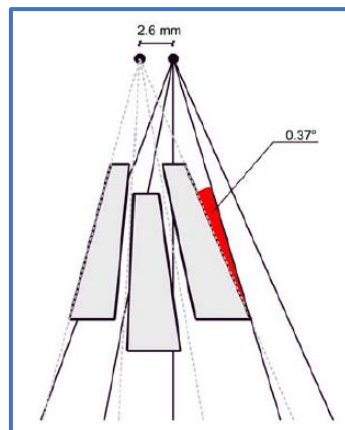
Figure 4: Tilted MLC's [64]

Siemens® ARTISTE™ 160 MLC™ are calibrated mechanically, with the use of physical blocks, as well as dosimetrically. The latter involves MLC field position measurements predominantly done with a linear array, performed in the factory during Linac assembly. The purpose of the dosimetric calibration is to establish a reference between the mechanically calibrated MLC banks and the measured radiation fields per machine energy available and to correct any difference by applying an offset to

MLC positional soft pot values [65]. Each Siemens® ARTISTE™ 160 MLC™ has unique, factory measured MLC soft pot values stored in a radiological calibration curve.

Measurements are carried out by taking dose profile measurements at a depth of 10cm in water at the isocentre plane for a set of field sizes, normalized to 100% in the centre of the field. The 50% isodose line of the profile presents the radiation field size of each MLC bank respectively. Here it is very important that the measurement system of choice must have enough spatial resolution to precisely define the 50% isodose line. Based on a series of measurements, field positions for symmetrical and asymmetrical fields are entered into the MLC control software. When performing the asymmetrical measurements, the non-analysed bank must be left at an open position. An example configuration is below (figure 5). From this, the input values are from the x1 bank of a low energy beam (example 6MV) and the position values are understood as 1/1000 fractions of cm.

It is recommended that these values be verified upon Linac commissioning as well as after any intervention that may have altered the mechanical relationship between the beam and the collimator (such as energy adjustment, beam alignment, adjustments of filters or target slides) or whenever it is deemed necessary [65].

Figure 5: Example of an input screen for the dosimetric calibration of MLC positions, referred to as the radiological calibration curve [65]

3.2) Reference Model

Description of TPS

Monaco[TM] TPS (Elekta Oncology Systems, Crawley, UK) is Monte Carlo (MC) based (Elekta CMS version 5.0). The beam model of this TPS are divided into 3 parts [19]:

1. A VSM representing the path of MC generated particles from the x-ray target to the level of the beam modifiers (3.2.1).

2. A Transmission Probability Filter (TPF) [66] [67] that characterises the MLCs and jaws (3.2.2).

3. Another MC dose engine, a X-ray Voxel MC (XVMC) dose engine that calculates the absorbed dose in a patient 3D CT dataset, as illustrated in Figure 6 [68].
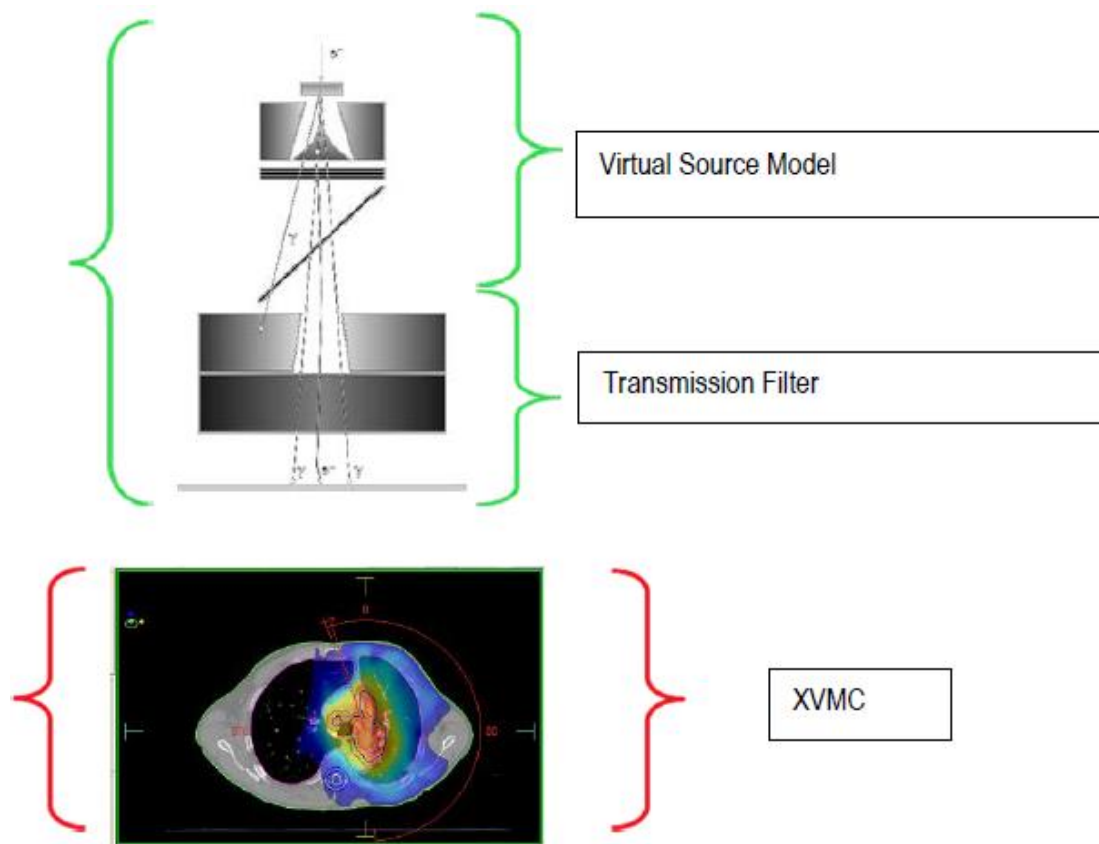
Figure 6: Three components of the Monaco [TM] Monte Carlo dose calculation

3.2.1) Virtual Source Model

The VSM represents 3 virtual sources of radiation in the head of the Linac [17] [19], each presented by a Gaussian function: primary photons, secondary photons and electron contamination:

- Primary photons originate through bremsstrahlung from the x-ray target and is regarded as particles that do not interact with other components in the Linac head.

- Secondary photons are particles that interact with components in the head and will subsequently change direction and lose energy, called scattered radiation. These originate mainly from the primary collimators and the flattening filter and include secondary bremsstrahlung and pair production photons produced elsewhere than the target.

- Electron contamination are Compton electrons produced from photon interactions in the head. These are simulated at the base of the flattening filter.

The location where these particles are recorded is the scoring plane, capturing details of all particles passing through, so called phase space files. A phase space computer file therefore contains a collection of particles with their properties that include energy, particle type, origin location, direction and statistical weight.

As elucidated before, an advantage of the VSM is that instead of calculating a detailed MC simulation of the Linac head for each radiotherapy treatment, phase space files can be studied once per Linac energy to derive analytical functions [68] to present the invariant part of the Linac head. In other words, the properties of the virtual sources are described by parameterizing the phase space information though analytical functions. These predefined analytical functions then describe the energy spectra of all 3 sources in a scoring plane presented at the top location of the beam modifier (MLC or jaw) nearest to the target, for example at a distance of 22.2 cm from the target in the case of a Siemens® ARTISTE™ Linac. Once particles are generated at this scoring plane (which remains fixed for all possible beam settings), its probability of traversing through the MLCs and jaws on a patient specific basis is determined also using an analytic method referred to as a transmission filter (discussed in section 3.2.2).

Analytical functions consist of reference parameters, based on the full MC data analysis and open parameters, based on machine measurements. Open parameters of the VSM or fitting parameters of the analytical functions are adjusted based on dosimetric measurements in water for each individual Linac [19]. This is an iteratively process to minimize the difference between calculated and measured dose. Therefore, the need for accurate measured data is crucial otherwise model parameters will be incorrectly derived. Accurate measured data is described as high resolution measurements with an appropriate detector to correctly present dose output and penumbra width, especially small field size measurements for estimation of for example the primary photon source diameter [67]. It was demonstrated that the accuracy of small field size output factors, influencing the dose calculation of for example IMRT segments, is largely dependent on this parameter.

The AAPM recommend that vendors of MC-based dose calculation systems should be responsible for providing the necessary guidance and assistance with the beam modelling to ensure that the beam model meets the required specification and benchmarking process [18]. Therefore, optimization of the reference and open parameters of the VSM discussed here is performed by the vendor until the model agrees, within specified limits, to Linac measurements. Thereafter the model, accompanied by a full validation report is delivered to the user. An advantage of this approach is, before modelling of the VSM commences, customer submitted measurements can be compared to expected measurements (from a library of same Linac machine measurements) to identify gross measurement errors.

3.2.2) Transmission Probability Filter

The TPF calculates the probability of particles, created from the virtual source model, to be attenuated by the patient dependant beam modifiers (MLCs and jaws). The amount of attenuation of these beam modifiers is transcribed in the transmission value of each. In former versions of the TPS software and still applicable to the Siemens® ARTISTE™ Linac, should the transmission value of the beam modifiers be less than or equal to 1.1%, it is modelled as fully absorbing [68]. According to Siemens manufacturer, the transmission of for example the y-jaws are less than 1%. This is particularly important when verifying the beam model against machine measurements as this may cause a dose discrepancy outside the radiation field, underneath these

jaws. Lately this total absorption is handled differently by the software as all particles passes through (regardless of transmission value), but with changes in intensity and weight. This approach ensures more particles are tracked which will contribute towards a better calculation of dose under the beam modifiers.

The TPF provides the flexibility for the user to further modify the probability of VSM particles passing through based on post-modelling Linac measurements. This allows for better characterization of especially MLCs per the actual settings of the Linac used in the clinic. More importantly the results can be used to compare and analyse similarities and possible differences between multiple Linacs when utilizing a single reference model. This single reference model will have the same VSM, but probabilistic parameters in the TPF can differ. If so, essentially a new model can be generated without having to repeat the entire modelling process.

To optimize TPF parameters, dose calculated from the TPS model in a phantom for a test beam package[69] (consisting of 8 treatment beams shown in table 2) can be compared to measurements of the same set of test beams on the Linac or any other beam-matched Linac. In this way, geometrical characteristics of the Linac e.g. leaf offset, leaf transmission, interleaf leakage, leaf tip leakage, leaf groove width, transmission of the jaws and other parameters specific to the Linac are compared to TPS calculations, as described in the table. Based on the results, TPF parameters of the TPS model can be adjusted iteratively.

As an example, the set of test beams can quantify differences between measured and TPS calculated MLC stop positions for various stop positions typically found during step and shoot IMRT delivery. Thereafter, if necessary, an adjustment can be made to the MLC major offset parameter in the TPF, fully customizable to the customer. A MLC major offset can be applied to all MLC stop positions to minimize differences found. Should the difference indicate a higher than calculated measured dose, a positive MLC major offset value can be applied to allow for overtravel of MLCs, potentially improving the agreement and vice versa. Also, the MLC transmission value, another customizable parameter, can also influence this comparison result and often the manufacturers recommend adjusting both (changing one parameter at a time). Therefore, Linacs that are beam matched in terms of basic commissioning data (e.g. PDDs and Profiles), may not require the same MLC major offset value applied in

the reference TPS model TPF. An example is shown in Figure 7. This is a comparison of an abutted field (3 consecutive segments) between 2 Linacs. The one Linac has a negative MLC major offset in terms of MLC bank, while the other has a positive MLC major offset. It is therefore important to evaluate this offset correctly [66].

Beam Name	Description & Purpose	
10×10	10×10 cm² open field. The purpose of this field was to assess the calibration of the measurement device.	
20×20	20×20 cm² open field. The purpose of this field was to assess the response of the QA device to a large field and to examine for the presence of beam asymmetry.	
3ABUT	Three 6 cm wide fields via step-and-shoot delivery. The fields were designed to abut and created two junctions. This plan was used to evaluate MLC calibration and the leaf offset parameter.	
7SEGA	Seven 2 cm wide fields via step-and-shoot delivery. The fields were designed to abut and created six junctions. This plan was used to evaluate MLC calibration and the leaf offset parameter.	
FOURL	Four L-shaped fields via step-and-shoot delivery. Each sequential L is smaller producing an overall large L-shape distribution. The plan was designed to 'abut' on each side of the L-shape. This field was used to examine the leaf position offset, the MLC transmission, and the MLC groove width settings.	
DMLC	A dynamic 10 cm sweep with a 2 cm wide field (defined by both the leaves and backup jaws). The dose distribution was sensitive to changes in the MLC position offset, transmission, and calibration.	
HIMRT	Head and neck step-and-shoot IMRT field. The intention of this field was to evaluate the impact the TPF MLC settings in a complex clinical situation.	
HDMLC	Head and neck sliding window IMRT field. The intention of this field was to evaluate the impact the TPF MLC settings in a complex clinical situation.	

Table 2: Description of test beams used to characterize MLC parameters of a Monaco model [69]
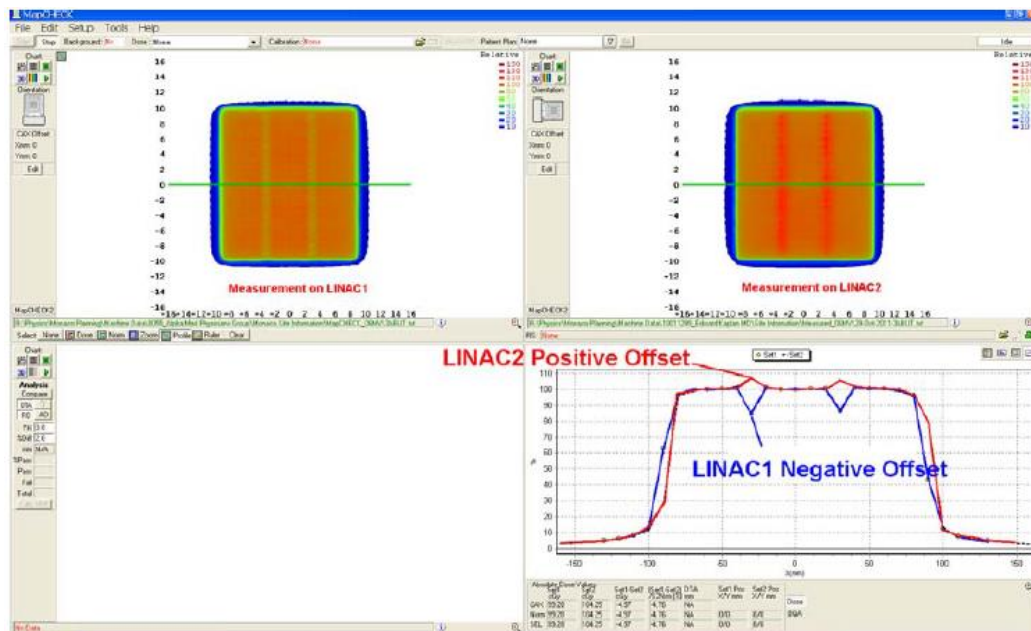
Figure 7: Measurement of the same beam on two LINACs [66]

3.3) Reference Evaluation Method

Basic principle of γ analysis

A quantitative method for comparison between measured and calculated dose distributions, termed the γ index was presented by Low et all [43]. The method is to provide a means for quantitative comparison between 2D profiles and dose distributions. γ analysis are typically based on the DD and/or the DTA between a reference and measured profile or dose distribution, as mentioned before in section 1.5. In low dose regions, γ values are dominated by dose differences. In steep dose regions dose differences are large and therefore the γ value is dominated by DTA [47]. A simplified approach to the γ index value can be applied in some instances, calculating its value as DD or DTA only [70]. For example, for a 1%/1 mm criterion at a given detector point, the DD would be determined first and if the value is less than or equal to 1%, it would be assumed DTA = 0 mm. The γ is then simply calculated as:

$$\gamma = \sqrt{\frac{\Delta DD^2}{\Delta D^2}}$$
(Eq. 1)

Where

ΔDD^2 = dose difference

ΔD^2= dose difference tolerance criteria

The value should be less than or equal to one, indicating a passing point. Should, for the same measured point, the dose difference be more than 1%, a search for the nearest point (in mm) on the reference profile with the same dose value as measured are performed. If such a point is found within the 1 mm criterion, γ is calculated as:

$$\gamma = \sqrt{\frac{\Delta r^2}{\Delta d^2}}$$
(Eq. 2)

Where

Δr^2= nearest distance to agreement

Δd^2= distance-to-agreement tolerance criterion

This was explained by an example (Figure 8 below), showing a comparison between measured and calculated dose profiles. Large dose differences can be seen in the penumbra region (a) due to a constant shift in the calculated dose of 2.5 mm. However, because the shift and therefore DTA is within the chosen criterion of 3 mm, there are no regions that fail the γ criteria. A maximum γ value of 0.83 was calculated (b).
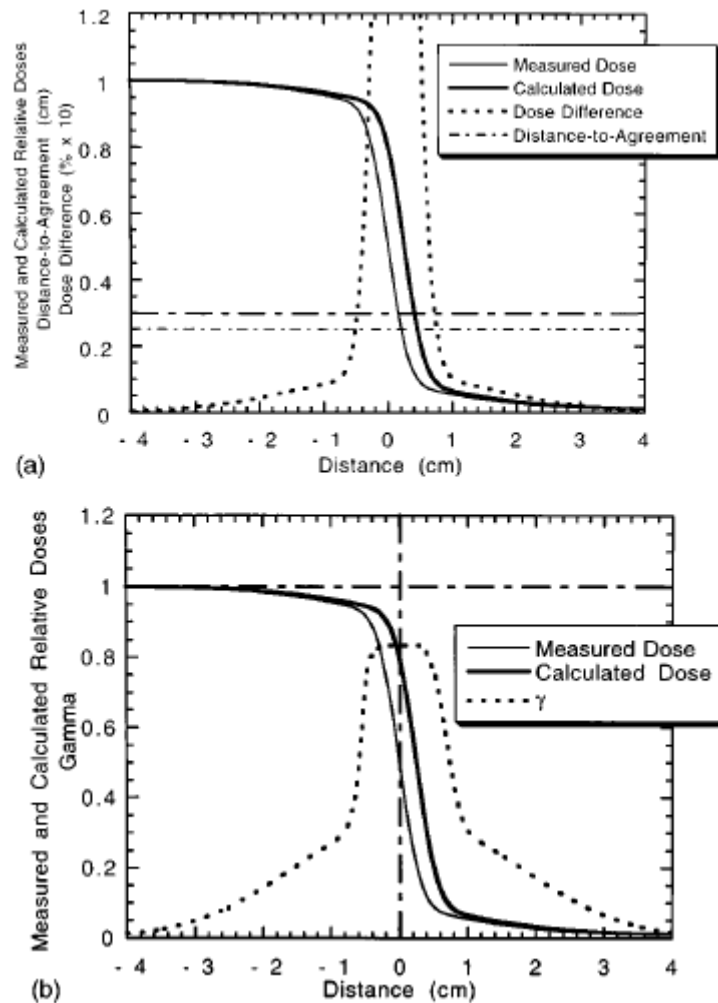
Figure 8: Example from Low et all: dose distribution comparisons [43]

If a γ value of less than 1 cannot be calculated for a specific detector point, a combination of the 2 parameters can be calculated as:

$$\gamma = \sqrt{\frac{DD^2}{\Delta d^2} + \frac{\Delta r^2}{\Delta d^2}}$$
(Eq. 3)

Depending on the type of detector used, array software handles data differently when comparing 2D measured and TPS calculated dose distributions. Example, SNC Patient^{TM} 70 investigates the dose difference between each measurement point and the corresponding TPS dose point based on the specified criterion. If the dose difference exceeds the criterion, a distance searched is performed. The distance, depending on the chosen DTA criterion, is equivalent to the radius of a circle around the measurement point within which points are examined for agreement. Measurement points are spaced 10 mm horizontally and 7.1 mm diagonally, but TPS calculations are interpolated to 1 mm by default. For example, choosing a 3mm DTA, should the %DD fail at the measurement point, that same point will be compared to 28 TPS dose points (figure 9).
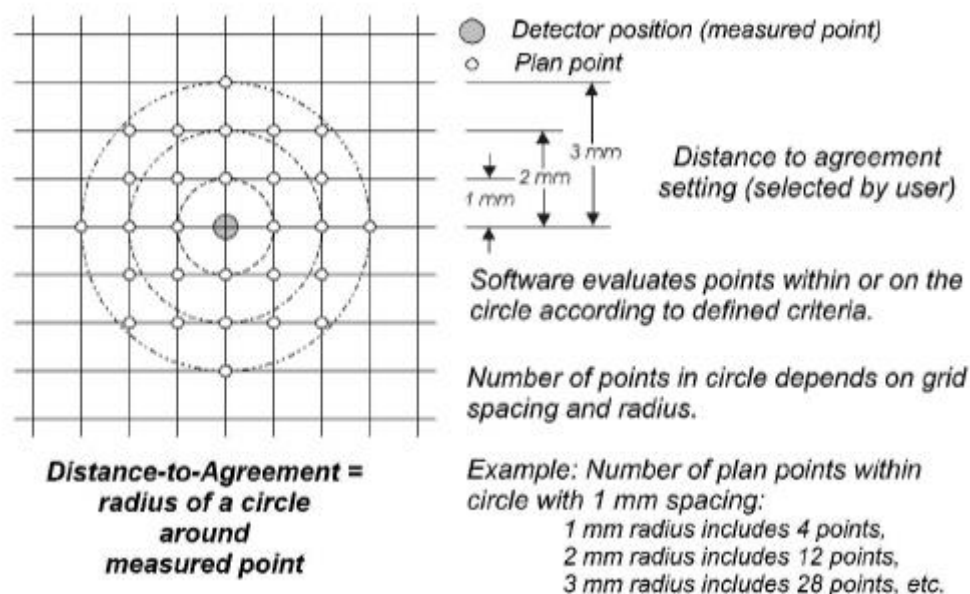
Figure 9: Example from SNC Patient software: Distance to Agreement Analysis [70]

4. MATERIAL AND METHODS

This chapter provides a description of the measurement equipment (section 4.1) and methodology used in this study for dosimetry measurements on 5 beam-matched Siemens® ARTISTE™ Linacs. These measurements are to determine the equivalence, and differences, between the Linacs based on water as well as array measurements. Water based Linac equivalence measurements, discussed in section 4.2, include absolute point doses, PDDs, profiles, scatter factors and beam quality factors performed with a diode detector. A method of increasing the spatial resolution of array measurements in terms of setup of the device, data processing and validation thereof by comparison to high resolution water-based profile measurements are discussed in section 4.3. This method was further validated by comparison to a high spatial resolution ionization chamber array. Finally, how this increased spatial resolution method was utilized for array based Linac equivalence measurements are discussed in section 4.4: to quantify MLC positional errors and compare IMRT dose distributions of each Linac respectively. IMRT dose distributions measured on each Linac was respectively compared to calculated IMRT dose distributions from the same, single reference TPS model.

4.1) Overview of Measurement Equipment

4.1.1) Mapcheck2

Mapcheck2™ with SNC Patient™ software (Sun Nuclear, Melbourne, FL) allow the user to quantitatively compare measured and calculated 2D dose distributions, utilizing the γ analysis and DTA techniques. Mapcheck2™ is shown in Figure 10 and is a 2D array consisting of 1527 diode detectors distributed in a 26×32 cm^2 (row by column) area. Diodes are situated at a depth of 1.35 cm below the surface (water-equivalent to a density of 2 g/cm^2), have a detector size of 0.8×0.8 mm^2 and a sensitive volume of 0.019mm^3. As mentioned before, the device allows for quick acquisition of many measurement points on a Linac. Diodes are spaced 1.0 cm in a parallel direction and 7.07 mm diagonally which may limit its ability to detect machine dosimetry errors

(shown before in Figure1). This checkerboard configuration of diodes is shown in Figures 11 and 12.

Figure10: Sun Nuclear Mapcheck ^{TM}
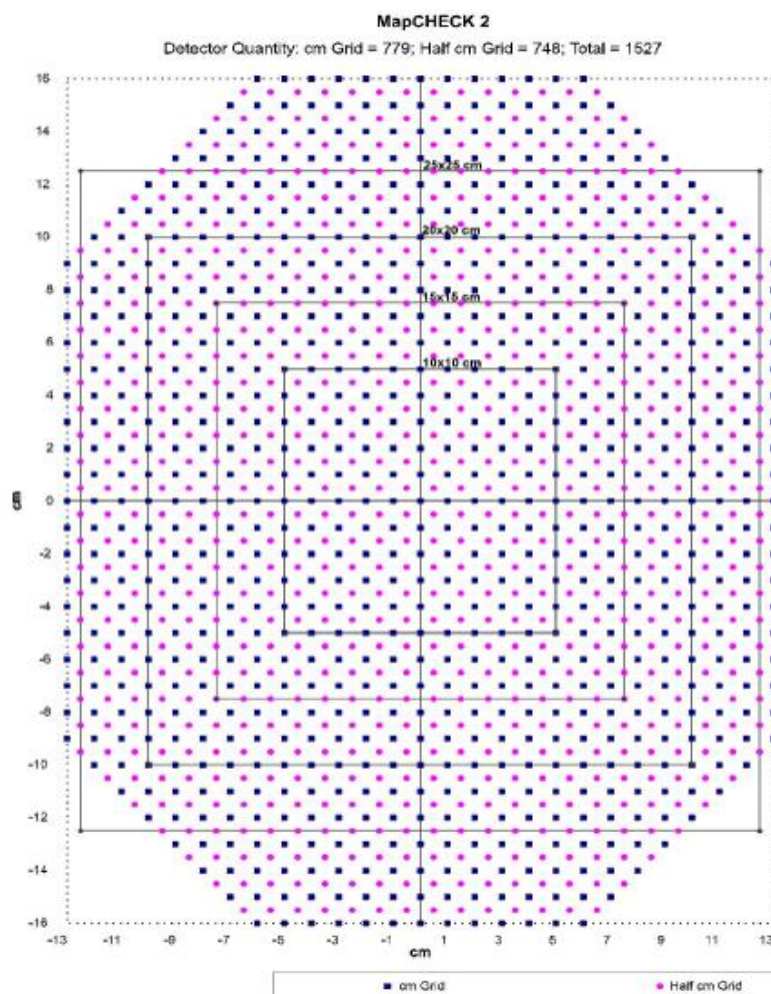
Figure 11: Mapcheck ^{TM} diode configuration indicated with dots [70]
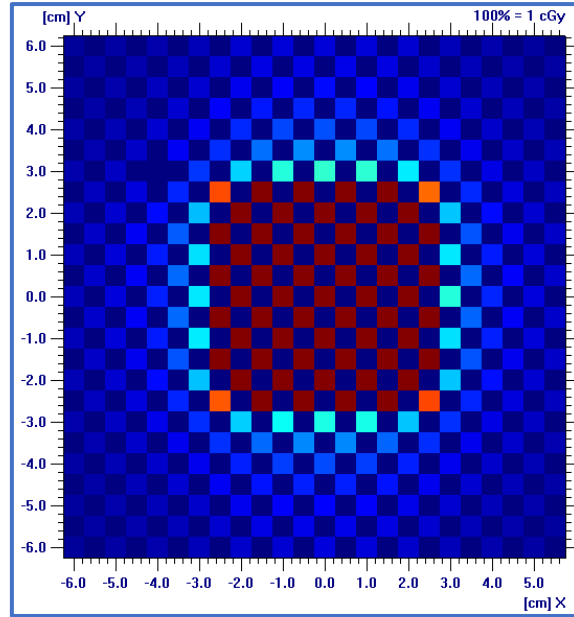
Figure 12: Mapcheck ^{TM} measurement points for a standard 5 × 5 cm^2 field size

4.1.2) 1000SRS array

The 2D-Array 1000SRS^{TM} (PTW, Freiburg, Germany) shown in Figure 13 consists of 977 liquid-filled ion chambers arranged in a 11 × 11 cm^2 outer matrix and a 5.5 × 5.5 cm^2 inner matrix area of the array. Each detector has a size of 2.3 × 2.3 mm^2 and a sensitive volume of 0.003 cm^3 or 3 mm^3. In the high resolution inner area, the centres of the adjacent chambers are placed at a distance of 2.5 mm from each other, whereas the spacing of the detectors in the low resolution outer area is 5 mm (center to center) [71]. It was shown that this detector can be used for high accuracy measurements due to its high spatial resolution [72]. Therefore, in this study the device was used to further validate high resolution small field size profile measurements obtained with Mapcheck2^{TM}.

Figure 13: The PTW 1000SRS^{TM} array
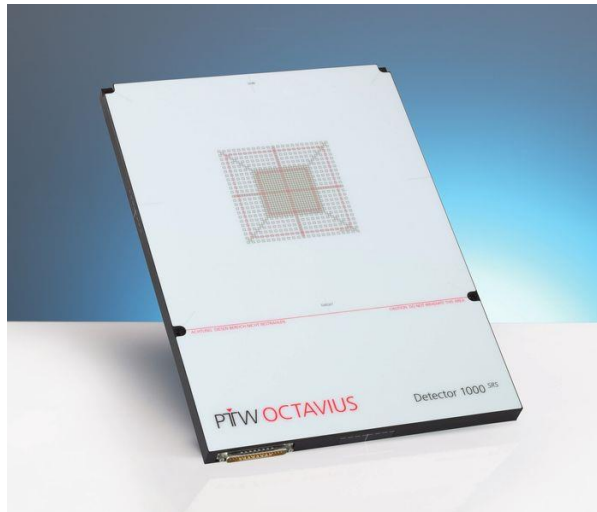
Figure 13: The PTW 1000SRS^{TM} array

4.1.3) microDiamond detector

The microDiamond detector (TM60019, PTW-Freiburg, Germany) is a synthetic single crystal diamond detector/silicon diode detector shown in figure 14. This detector has a very small sensitive volume of 0.004mm^3 within a waterproof Poly(methyl methacrylate) (PMMA) housing with a 7 mm diameter [73]. The diamond surface is 1 mm below the top surface of the housing and scanning is performed with the detector vertically in the upright orientation to minimize stem effects [74]. Both authors described the microDiamond detector (TM60019) as one of the most appropriate detectors for beam data collection and dose measurements in penumbra regions due to its finite size and subsequent minimal volume effect, which makes it ideal for small field dosimetry.

Figure 14: The PTW microDiamond detector (TM60019)

4.2) Linac Equivalence Measurements (water-based scanning)

Equivalence measurements in water on each Linac were based on beam data collected for the reference TPS model (Monaco™). Initially beam data were collected according to manufacturer's specifications with a Sun Nuclear SNC 3D Dosimetry™ (Sun Nuclear Corporation, Melbourne, FL) water phantom on one Siemens® ARTISTE™, referred to as the reference Linac. This data was used by the TPS manufacturer to produce a single 6 MV photon MC model for IMRT treatment planning, referred to as the reference model. The recommendation of including the reference model when evaluating machine equivalence (as was discussed in section 1.5) was followed, comparing measurements from each Linac to the reference model. Also, full scan comparisons were performed, not limited to a single PDD or profile point.

Therefore, water-based measurements from all 5 Linacs, summarized in table 3, were compared to dose calculations from the reference model. It should be highlighted that these 5 Linacs were historically beam- matched and installed between 2009 and 2011. Unfortunately, for profile comparison historic beam data (collected for XiO TPS from CMS (St. Louis, MO, USA)) from all 5 linacs could not be used as these scans were collected at a difference source-to-surface distance (SSD). Instead, profiles from 2 other Linacs (excluding the reference Linac) were re-measured.

For Linac output, beam quality factors and Tissue Scatter Correction Factor (TSCF) values, deviation between results of reference model calculations and Linac measurements were expressed as a percentage of the measured dose [27].

$$\% \text{ Difference} = (D_{calculated} - D_{measured})/D_{measured} \times 100\% \qquad \textit{(Eq. 4)}$$

$D_{calculated}$ = dose calculated
$D_{measured}$ = dose measured

For PDD comparison, an absolute difference of percentage dose was calculated. Profile comparison was done through γ analysis, performed in Microsoft Excel® VBA using an in-house developed code with overall criterion of 2%/2mm. The International Atomic and Energy Agency (IAEA) has set a tolerance of 2% in low dose, central profile regions and 2 mm in high dose, profile penumbra regions [2].

Measurement (in water)	Description; comparison method
Linac output	Absolute dose in water from 100 Monitor Units (MU) measured in Gy at a depth of 10 cm; % difference
PDD	Percentage depth dose for square field sizes 2 × 2, 5 × 5, 10 × 10 and 20 × 20 cm^2; absolute difference
Beam Quality (TPR$_{20,10}$)	Tissue phantom ratio (TPR) derived from the 10 × 10 cm^2 PDD at depths 10 and 20 cm in water; % difference
TSCF	Total scatter correction factors for square field sizes ranging between 2 and 40 cm^2; % difference comparison
Profiles	In-and crossplane profiles for square field sizes 5 × 5 and 20 × 20 cm^2 at depths d$_{max}$ and 10 cm; γ analysis

Table 3: A list of water-based machine equivalence measurements from each Linac, with a description of the field sizes and depths included as well as the method used for comparison to the reference TPS model

4.3) Increasing Array Spatial Resolution

For quantitative comparisons between the microDiamond detector and Mapcheck2^{TM} profiles, the spatial resolution of the array device must be increased, especially in the penumbra region as previously discussed. In this way, should the results from the two detectors be comparable, the array can be used with increased confidence for further Linac equivalence measurements (later discussed in section 4.4).

Therefore, to characterize the sensitivity of the array, Mapcheck2^{TM} dose measurements were performed with sequential shifted positions of the device in 1 mm increments across the radiation field. These were compared to beam data acquired on the reference Linac with a microDiamond detector (TM60019) for field sizes 2 × 2, 3 × 3 and 5 × 5 cm^2 at depths 5 and 10 cm in water.

This section provides a description of measurement considerations when combining multiple Mapcheck2^{TM} measurements (4.3.1); the format of Mapcheck2^{TM} stored data (4.3.2) and how this data was handled to collate measurements into single file format (4.3.3)

4.3.1) Multiple Mapcheck Measurements

Prior to a series of measurements, Mapcheck2^{TM} is cross calibrated against Linac dose output to convert Mapcheck relative dose values to absolute dose and is therefore specific to each accelerator and energy. Absolute dose calibration is performed in an open 10 × 10 cm^2 square field with the array at a depth where the dose is known from an earlier measurement (through TRS398) with an ionization chamber whose calibration is traceable to an absolute dosimetry standard [75]. The required depth was achieved by adding build-up material (PMMA with a density of 1.14 g/cm^2) on top of the device. The dose calibration factor is applied to all detectors in addition to an array sensitivity correction factor, a method to compensate for the difference in each detector's response to radiation. Determining a sensitivity correction factor for each detector is generally performed annually, consisting of the average of five radiation exposures, where each exposure is identical except for the position and orientation of the array [70].

Some additional setup and measurement considerations were considered before Mapcheck2^{TM} CAX profile acquisitions in 1 mm increments:

- Interpolation and normalization:

  Only measured, non-interpolated and not normalized data points were gathered and collated. Combining 10 profiles of normalized measured fluences would result in a false, overresponse flat profile around the CAX, especially for small field sizes such as 2 × 2 cm^2, shown in Figure 15. It should be noted that stability in Linac dose output could influence these measurements, but the effect is incorporated in the reproducibility of the entire method which will be shown later.
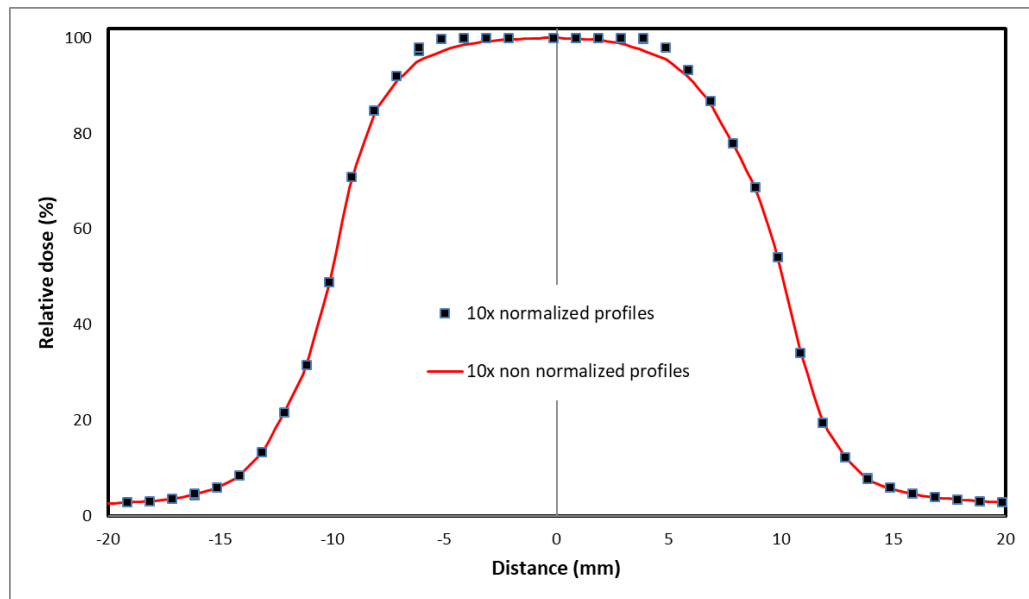
Figure 15: An example comparison of combining 10 Mapcheck measurements of non-normalized (red line) and normalized profiles (blue dots) for a 2 × 2 cm^2 field

- Shifting the array:

    A couch movement mechanism was used to shift Mapcheck2^{TM} as accurately as possible, but not relying on the couch movement readouts. Graph paper (on top of the device) was used instead. Using the treatment couch readout alone proved to be non-sufficient due to lag, specified as ±0.5 mm by the manufacturer [76].

- Device setup:

    The device was levelled to within 0.1° tilt using a digital laser level (Ryobi DLL-250). To setup the device to the radiation isocentre as accurately as possible, a visual check of the distance between isodoses of the 1st dose measurement and Mapcheck software grid lines were performed, as recommended by the manufacturer (Sun Nuclear SNC Patient™ version 6.4.1.) and illustrated in Figure 16. This figure is showing an off-centre setup in the cross-plane direction. Thereafter, Mapcheck was shifted iteratively, remeasured and evaluated until the best possible visual match could be seen, shown in Figure 17. In this example the distance difference between the sides of the square field (measured) and the grid lines are mostly equal.
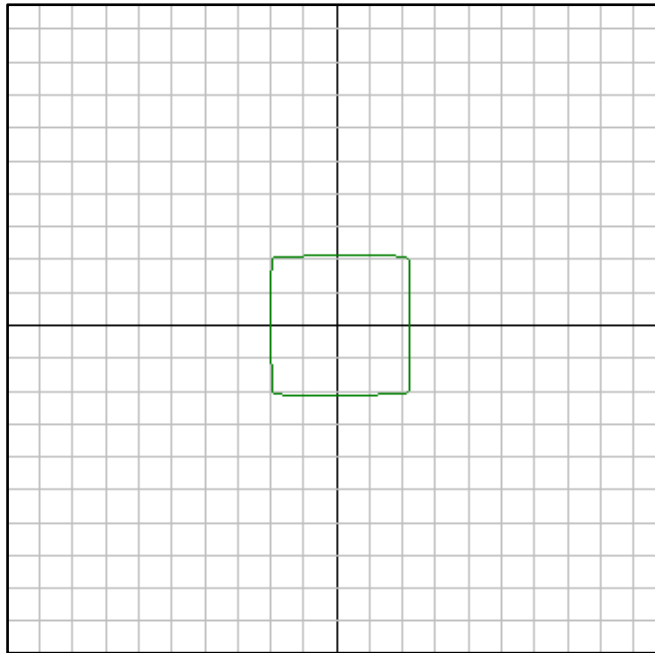
Figure 16: Mapcheck CAX setup to Linac radiation isocentre following the coincidence of a measured isodose (green line) and Mapcheck software gridlines (shown in grey): in the example an off-centre setup in the x direction can be seen.
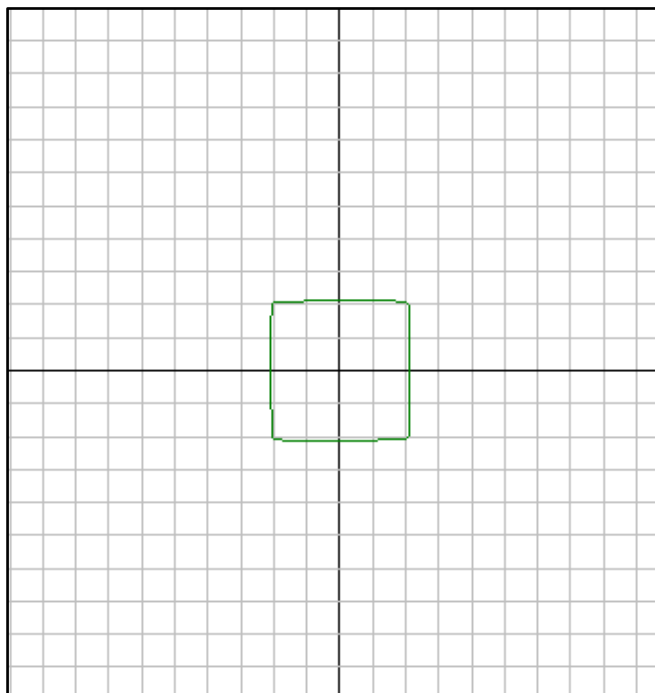
Figure 17: Mapcheck CAX setup to Linac radiation isocentre following the coincidence of a measured isodose (green line) and Mapcheck software gridlines (shown in grey): the example shows an acceptable, more equidistance between these lines

4.3.2) Mapcheck Data Format

Mapcheck2^{TM} measurements are saved in measured text file format with a ".txt" extension (Figure 18). The .txt file is a tab-delimited text file that can be viewed in a text editor such as WordPad. Among other information, the file contains matrixes of raw, corrected (background and array sensitivity compensated) and dose counts as well as interpolated dose counts. Each matrix represents diode detectors, with x and y coordinates corresponding to its physical location within the device. Each Mapcheck2^{TM} detector is connected to the input of an operational amplifier [70]. The output voltage of the operational amplifier increases as charge is collected on the feedback capacitor during radiation. After the exposure is complete, a final measurement of the output voltage is made and recorded in the raw counts matrix. This matrix is then corrected by subtracting a background voltage measurement as well as applying a sensitivity correction factor to each measurement point. Then an absolute dose correction factor is applied to obtain dose counts and finally an interpolation between these are performed.

As mentioned before, in the quest for increased resolution, non-normalized, non-interpolated matrixes were used to combine multiple measurements, in other words only truly measured values. SNC Patient™ can only handle a matrix with a fixed number of data points, therefore is not able to accommodate a matrix 10 times larger than the original.

Figure 18: An example of a Mapcheck format text file

4.3.3) Data Handling

A non-interpolated Mapcheck2^{TM} measured dose file consists of values in 1 cm increments, with a null value in-between, usually populated with an interpolated value. As SNC Patient™ cannot accommodate a larger matrix, an in-house Microsoft Visual Basic (Visual Studio (C) 2017) code in executable format was developed to combine mapheck measurements, referred to as "Mapcheck Combine" (figure 19). The code would adjust the x or y coordinates of the matrix (depending on the movement direction) by a specified amount and combine all files into a single matrix. The y direction corresponds to the inplane direction and the x direction to the crossplane direction. Also, data was converted to OmniPro I'mRT software format (IBA Dosimetry, Schwarzenbruck, Germany), as this software can accommodate larger sized matrixes or high-resolution dose planes. The code has 3 options for combining data: x-direction only, y-direction only and x and y direction simultaneously. The first 2 options would be more ideal for profile comparison, in other words for array sensitivity analysis and basic model verification measurements, especially when

evaluating the penumbra region. However, resolution in the opposing direction would remain 5 mm. Therefore, to improve this, a combination shift (x and y) was done that increases resolution diagonally. Unfortunately, due to the nature of the matrix, only measurements with a shift of 5 mm between them can be combined at a time and zero values could not be removed from the matrix without interpolation. Therefore, this option was voided.
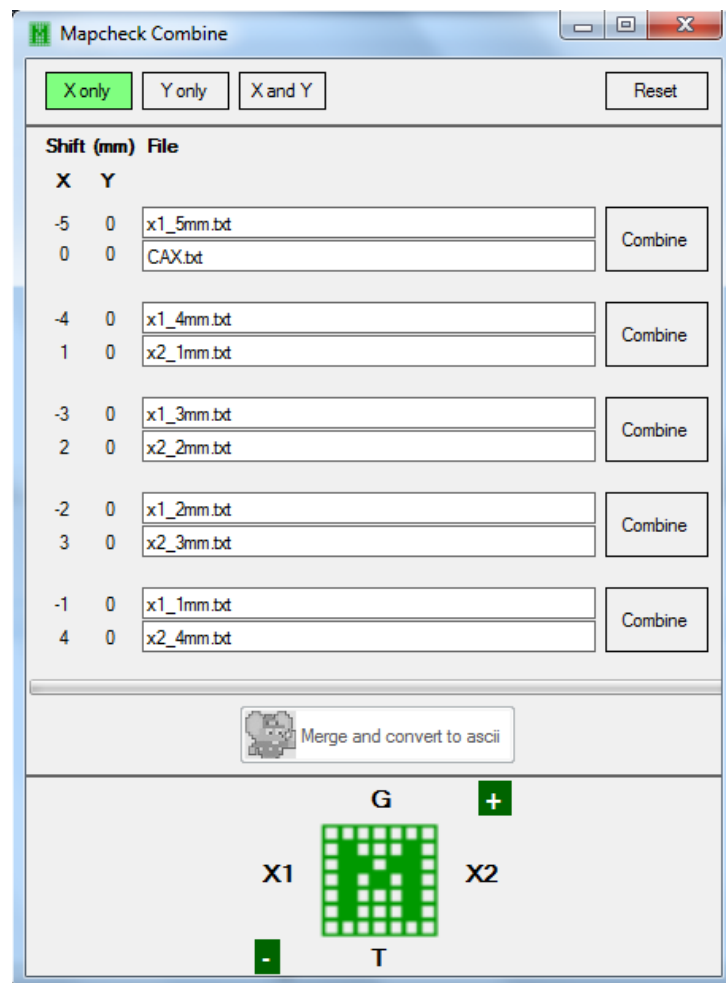
Figure 19: A screenshot of the "Mapcheck Combine" executable program

4.4) Linac Equivalence Measurements (array based)

4.4.1) Determining MLC errors

An MLC error is defined as the difference between the MLC planned position and the actual measured stop position. The MLC picket fence test [56] was measured using the proposed method of 1 mm increments with Mapcheck2™. The device was setup as described before, but also shifted by 2 mm in the y direction (non-MLC direction) using graph paper. This was done to achieve a better presentation of actual leave positions as a CAX profile presents a profile in-between two adjacent MLCs. A shift of 2.5 mm (half a MLC diameter) would have been ideal but the shift was believed not realistic.

The picket fence or no gap test consists of 12 MLC segment fields of 2 cm width delivered back to back. A maximum field size of 20 cm in the y direction and 24 cm in the x direction was chosen due to the physical limitation of the Mapcheck2™ device. Each of the 12 segments was measured 10 times, with 1 mm incremental shifts of Mapcheck2™ in between, individually. These measured doses were combined and converted into OmniPro Im'RT file format. The ascii file was imported, and multiple row analysis performed, consisting of 40 profiles each representing a MLC position of the x1 and x2 bank (figure 20) respectively. In this way, the actual stop positions of 80 MLCs (40 per bank) were measured at 12 locations, ranging across the radiation field (CAX ± 12 cm). Therefore, completing a picket fence test consisted of 120 Linac measurements and resulted in 960 MLC stop positions (480 from each MLC bank). The results were imported into Excel to determine MLC errors.

Using this methodology, the complete picket fence test was measured on each of the 5 Linacs and repeated 3 times over a period of 6 months for reproducibility. In this way, MLC errors of each Linac could be used to determine similarities and differences between their MLCs quantitatively. It was reported that Mapcheck can reliably detect MLC errors in the order of 0.1 mm using the same methodology, measuring the picket fence by shifting the device 10 times in 1 mm increments [77].

For a quantitative comparison between MLC errors from different Linacs, the Paired sample T test was used assuming the null hypothesis. The null hypothesis states that the difference of the means between two data sets should be zero. If the zero is

included in the confidence interval (chosen as 95%) then the null hypothesis is not rejected, meaning that there is not a significant difference between two data sets. This test was chosen as, in the case of beam matched Linacs, one would aim for no or at least similar differences in MLC errors between Linacs.
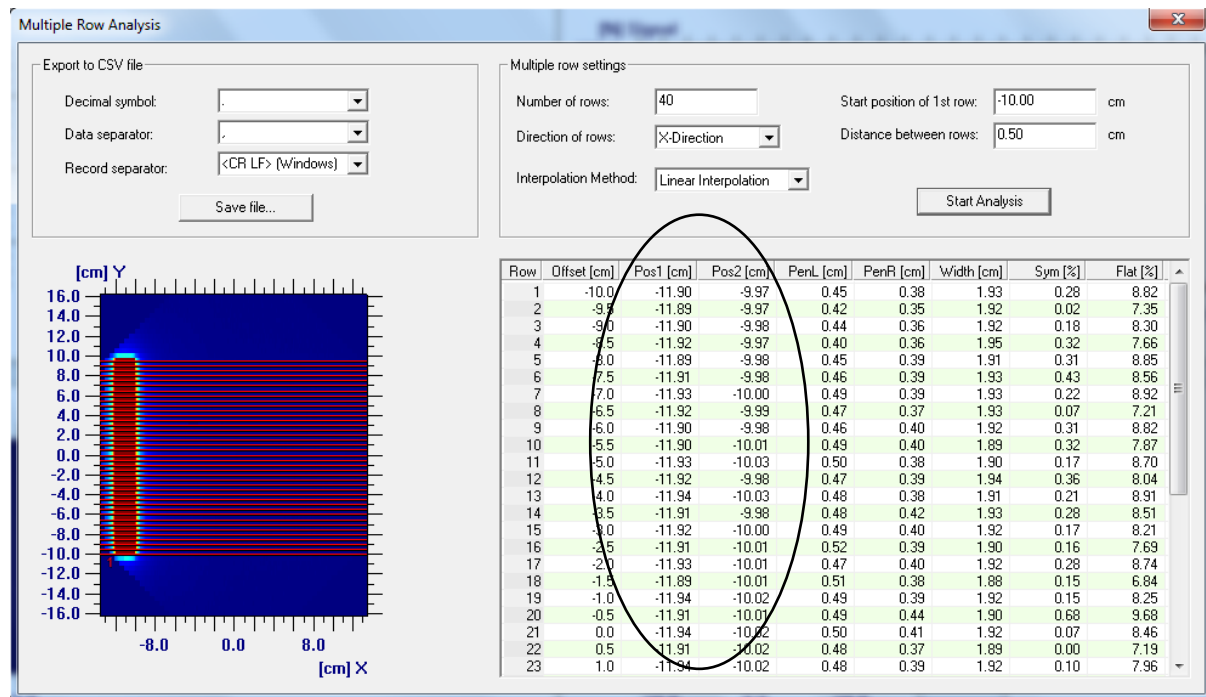
Figure 20: Multiple row analysis performed in OmniPro Im'RT on one of the 12 segments of the picket fence test. In the example the set position of the x1 bank (Pos1) is -12 cm and the x2 bank (Pos2) is -10cm and 80 actual MLC stop positions (40 from each MLC bank) are reported in cm

4.4.2) Test beam dose measurements

From the 8 test beams described in section 3.2.2, five were measured with increased resolution on all 5 Linacs and compared to calculated doses from the reference TPS model. From preliminarily results from the reference Linac, it was decided not to include all 8 beams due to the nature of the MLCs which will be discussed later. Also, the TPS manufacturer advised that these test beams should not be used for characterization of a model for a normal step-and-shoot IMRT ARTISTE™ but should rather be used for the characterization of MLC parameters in models (or Linacs) capable of delivering Volumetric Arc Therapy (VMAT) treatments. Nevertheless, these 5 beams were included for comparison between the 5 Linacs: 2 open beams (20 cm × 20 cm and 10 cm × 10 cm), MLC abutting fields (3ABUT and 7SegA) and a clinical

IMRT field. The beams were calculated with the MC reference model in an artificial phantom of 30 × 30 × 30 cm^3 with relative electron density (RED) of 1.14. The 3D dose grid size and MC statistical uncertainty were set to 2 mm and 0.5%, respectively, for all dose calculations.

4.4.3) IMRT dose distributions

2D dose distributions of IMRT fields using 6MV, predominantly used for IMRT treatment planning, were measured on each of the 5 Linacs. In order to test the capabilities of the system, each IMRT field consisted of a high number of segments (between 19 and 22) due to a low segment suppression filter applied during TPS segmentation. Each field had a minimum MU per segment of 5. The measured IMRT fields were each subsequently compared to the reference TPS model 2D dose distributions.   Figures 21 to 25 show these fields, with each field's degree of modulation shown by a dose profile through the CAX in the y direction (inplane). The IMRT fields have a high modulation degree, in other words the minimum dose within the field being approximately 10% to the maximum dose within the field. Evaluation criteria ranged between 2 to 3% DD and 2 to 3 mm DTA. This range was used to include more stringent tolerances which can potentially highlight Linac differences more easily (discussed in more detail later). Further, for each IMRT measurement, the absolute dose (in cGy) from the CAX diode (0,0 point) was compared to TPS calculated doses. In this case the measured absolute dose was from a cross-calibration to Linac output (discussed in 4.3.1), therefore the values were inclusive of any Linac dose output variations.

Gamma results obtained were from 4 evaluation methods:

- Mapcheck single measurement analysed in SNC Patient^{TM}, denoted as "Mapcheck single"
- Mapcheck double exposure (2× measurements, array shifted 5mm in crossplane direction) analysed in SNC Patient^{TM}, denoted as "Mapcheck double"
- Mapcheck single measurement converted and analysed in OmniPro I'mRT, denoted as "OmniPro single"

- Mapcheck 1 mm resolution measurements combined and analysed in OmniPro I'mRT, denoted as "OmniPro 1mm"

To summarize the results, confidence limits were calculated for each Linac, based on mean γ pass rates obtained. Some difference is to be expected, but should these be within a reasonable confidence limit, then the results can be considered acceptable. The confidence limit is the sum of the absolute values of the average differences and the standard deviation (σ) of the difference multiplied by a factor 1.96 [24].

$$\text{Confidence limit} = (100 - mean) + 1.96*\sigma \qquad \textit{(Eq. 5)}$$

In this formulation that is based on the statistics of a normal distribution, it is to be expected that 95% of the measured points (factor 1.96) will fall within the confidence limit. Confidence limit values encompasses measurement, delivery, and dose calculation uncertainty [78].

SNC Patient™ software settings were chosen similar to those commonly used as was reported in AAPM-119. These choices include 1) absolute dose (measured doses were not scaled to some normalization value), 2) 10% threshold (the region of interest was defined by the isodose line representing 10% of maximum dose), 3) Van Dyk percentage difference applied (the percentage difference in dose was with respect to the maximum point in the region and not the local point) and 4) applied measurement uncertainty, a presumed measurement error of about 1% is included in the analysis [24]. The latter accounts for example diode temperature dependence.
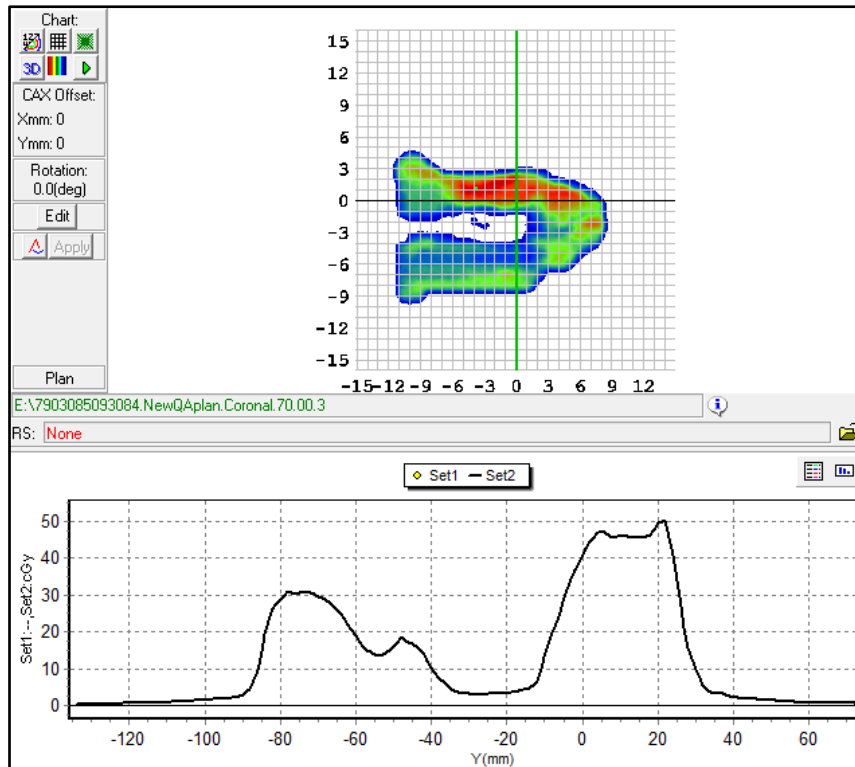
Figure 21: Dose distribution of the 1st IMRT field; the degree of modulation is shown by the dose profile (black line) through the CAX in the inplane direction
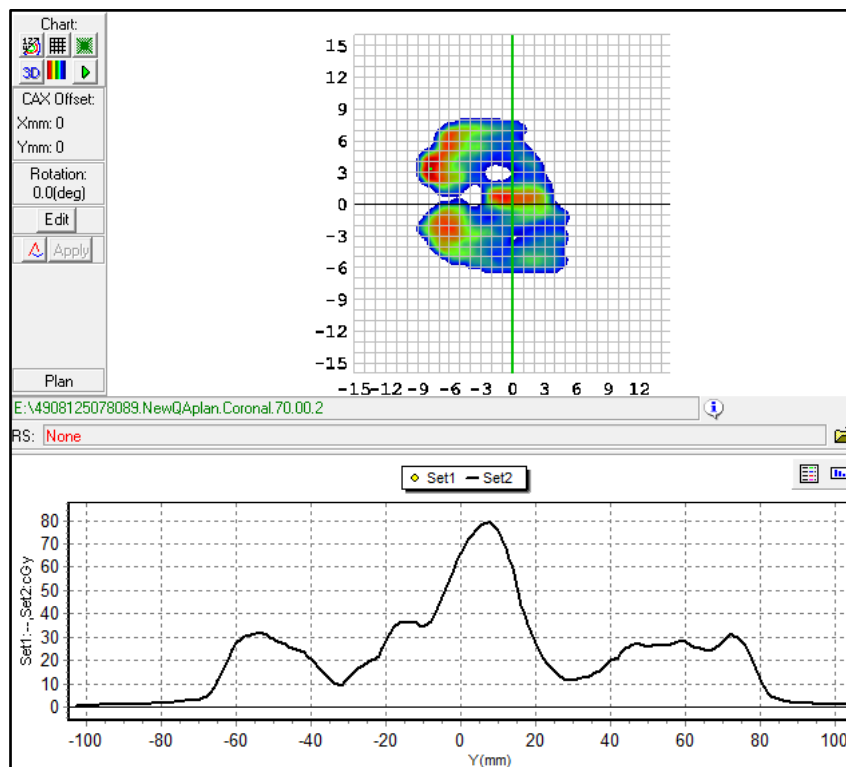
Figure 22: Dose distribution of the 2nd IMRT field; the degree of modulation is shown by the dose profile (black line) through the CAX in the inplane direction
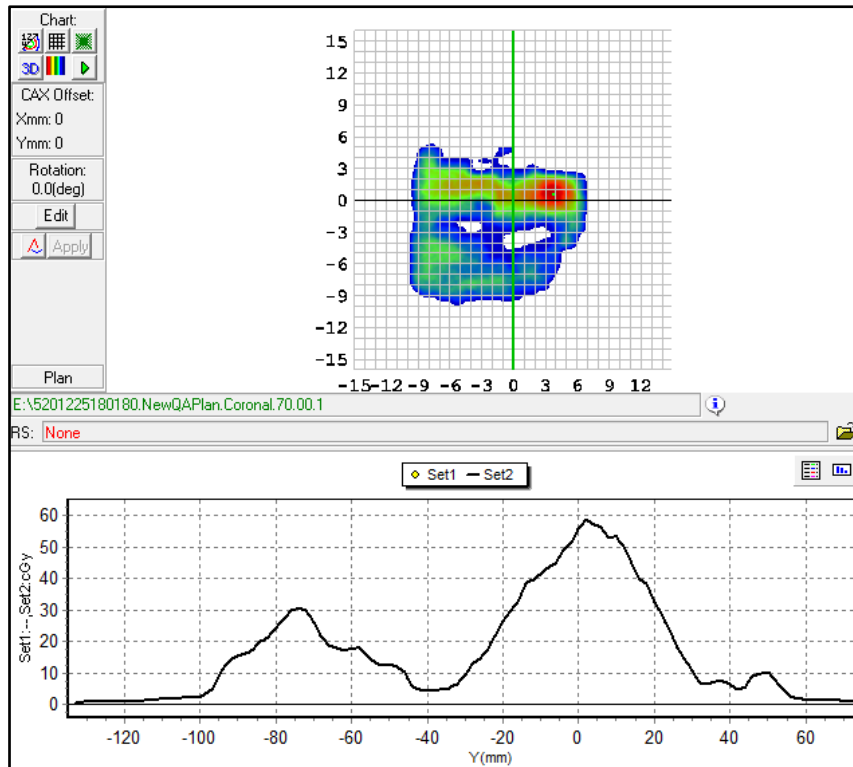
Figure 23: Dose distribution of the 3rd IMRT field; the degree of modulation is shown by the dose profile (black line) through the CAX in the inplane direction
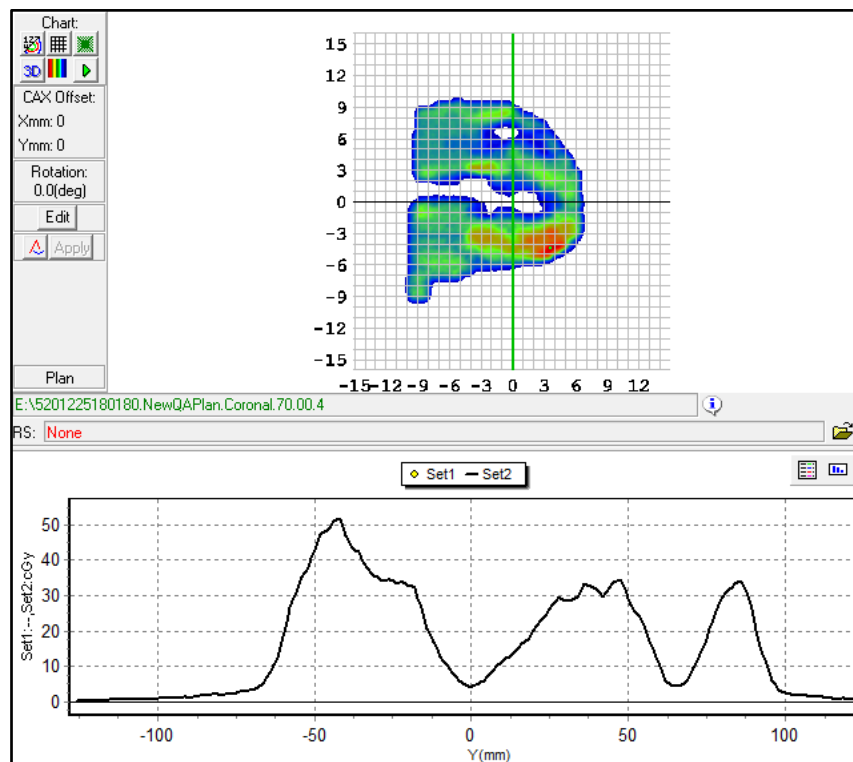
Figure 24: Dose distribution of the 4th IMRT field; the degree of modulation is shown by the dose profile (black line) through the CAX in the inplane direction
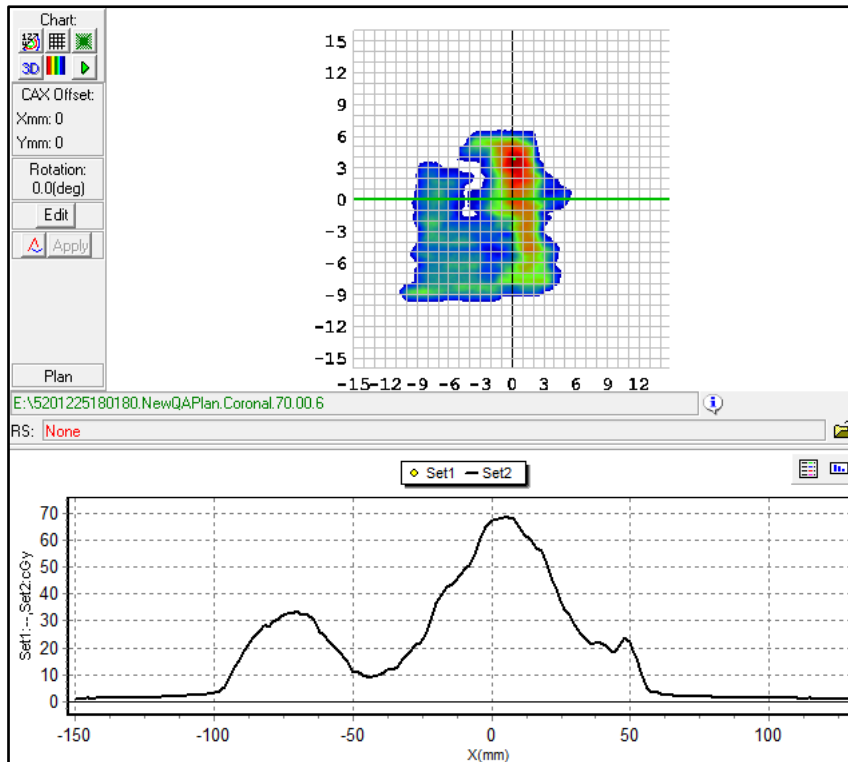
Figure 25: Dose distribution of the 5th IMRT field; the degree of modulation is shown by the dose profile (black line) through the CAX in the inplane direction

5. RESULTS AND DISCUSSION

5.1) Linac Equivalence Results (water-based scanning)

In this section, measured data scans from all 5 Linacs were used for Linac equivalence assessments by comparison to dose calculated from the reference TPS model. Published criteria of acceptable or non-acceptable results were applied throughout.

5.1.1) Linac output and Beam Quality

Table 4 shows Linac output (Gy/100MU), measured at 10 cm depth under reference conditions for calibrations, and beam quality parameters (PDD_{10} and $TPR_{20,10}$) compared between each Linac and the reference model (TPS in table 4) respectively. Linac output dose differences are less than 1.0% and beam quality parameters differ by less than 0.5% which was deemed acceptable.

A criterion of 0.5% on TPS dose verification has been previously set [10]. However, it was noted that this does not include uncertainties associated with determining Linac output with an ionization chamber, as an estimated relative standard uncertainty at the reference depth in water of up to 1.5% exists [75]. Furthermore, the MC calculation was performed with a statistical uncertainty of at least 0.5%. For the same setup conditions, a percentage difference of 1.84% for 95% of absolute dose measurements have been reported for the same TPS but different Linac [79]. Another study reported point doses within 2% [80]. However, both these studies included point dose measurements from various depths and off-axis fields. In the current study only Linac output, measured at a single dose point is reported. Further, that studies included a more comprehensive comparison between measured and calculated dose points for validation of heterogeneity corrections (not only water), for example the dose calculation in lung. It is recommended to compare measured and calculated doses above and below a heterogeneity material, outside of the build-up region [2 23 81].

	Linac Output (Gy/100MU) (%Diff from TPS)	PDD_{10} (%) (%Diff from TPS)	$TPR_{20,10}$ (%) (%Diff from TPS)
TPS	0.783	65.10	0.659
Linac1	0.788 (0.63)	65.14 (0.06)	0.657 (0.30)
Linac2	0.785 (0.25)	65.07 (0.04)	0.658 (0.15)
Linac3	0.780 (0.38)	65.10 (0.00)	0.656 (0.46
Linac4	0.788 (0.63)	65.20 (0.15)	0.657 (0.30)
Linac5	0.786 (0.38)	65.10 (0.00)	0.656 (0.46)

Table 4: Linac output and beam quality parameters (PDD_{10} *and* $TPR_{20,10}$*) of the model (calculated) and each of the 5 Linacs (measured). Shown in brackets are the percentage difference between values from each Linac and the TPS model respectively.*

5.1.2) Percentage Depth Doses

Figures 26,28, 30 and 32 show the PDD curves from the 5 Linacs including those calculated from the reference TPS model. Excellent agreement between these can be seen. To interpret and quantify differences, figures 27,29,31 and 33 show the respective subtraction of each measured PDD from that calculated. Differences are within 1%, except for the build-up regions (up to a depth of dose maximum (d_{max})) and a couple of points on the 20 × 20 cm^2 fields. For the 2 × 2 cm^2 field much larger than 1% discrepancies up to 13 mm exists. Data of the 2 × 2 cm^2 PDDs are only available for Linacs 1 to 3 as shown in figures 32 and 33.

Similar results, PDD comparisons within 1%, for the same TPS have been reported previously [82]. The difference in the build-up region is mostly due to different surface or entrance doses (at depths 0.0 and 0.1 mm), excluded from the absolute dose difference graphs. Surface dose differences range between 17% (5 × 5 cm^2) and 33% (2 × 2 cm^2) and is comparable to previously reported results [83] [84]. Surface dose and build-up region differences could be measurement or dose calculation related which include but is not limited to: type of detector used, the amount of charge particle disequilibrium and the user setup thereof, TPS calculation grid size and the algorithm used [6] [85].
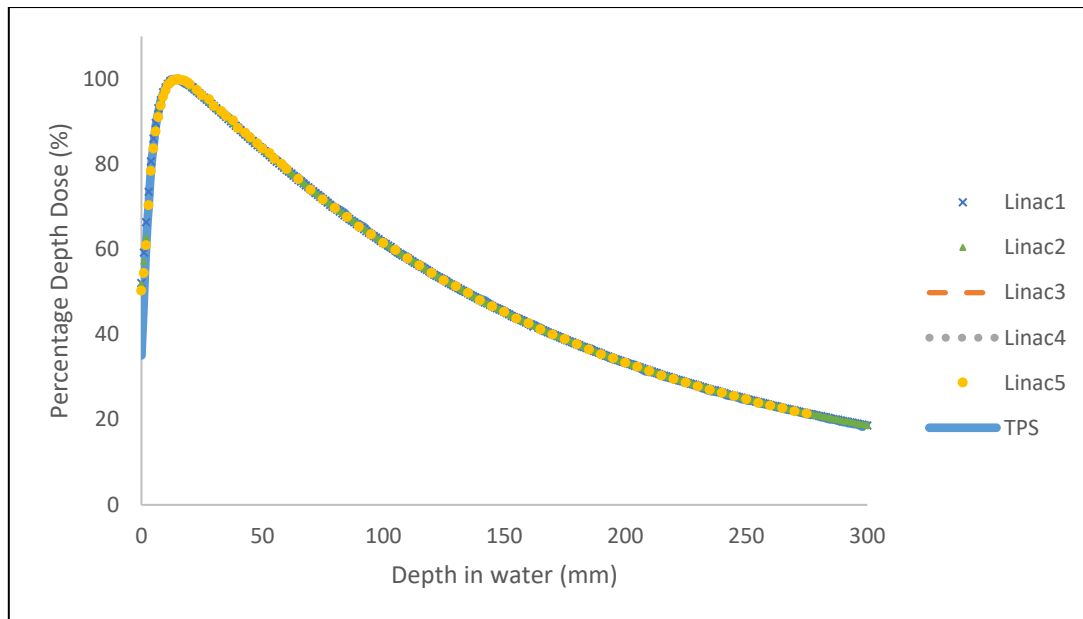
Figure 26: 5 × 5 cm^2 square field PDDs (normalized to the absorbed dose at d$_{max}$) *of 5 Linacs (measured) including the TPS (dose calculated) at SSD = 90 cm*
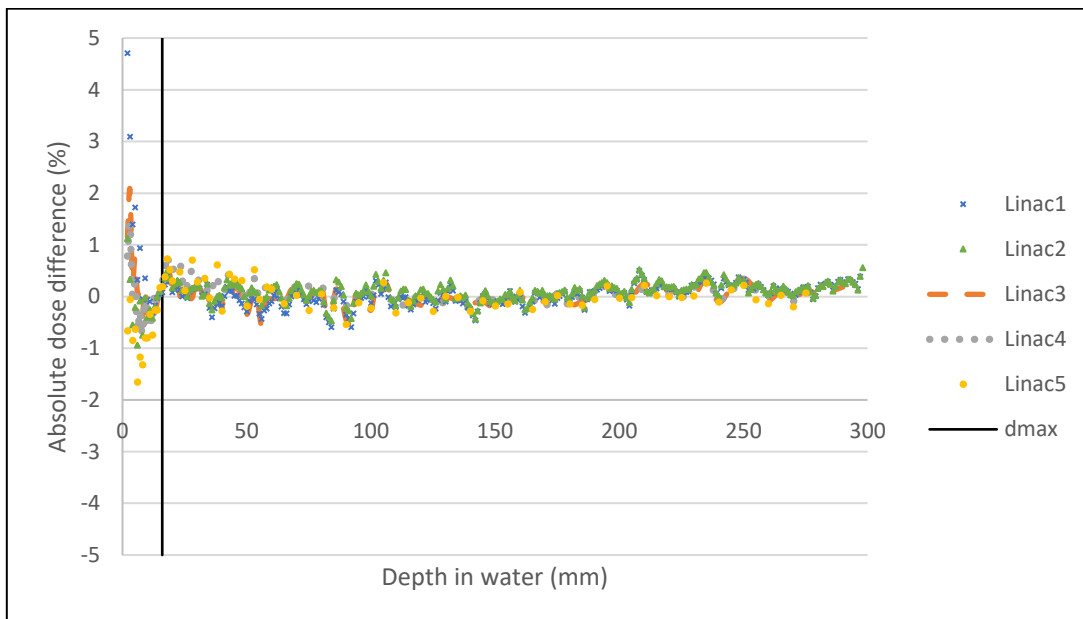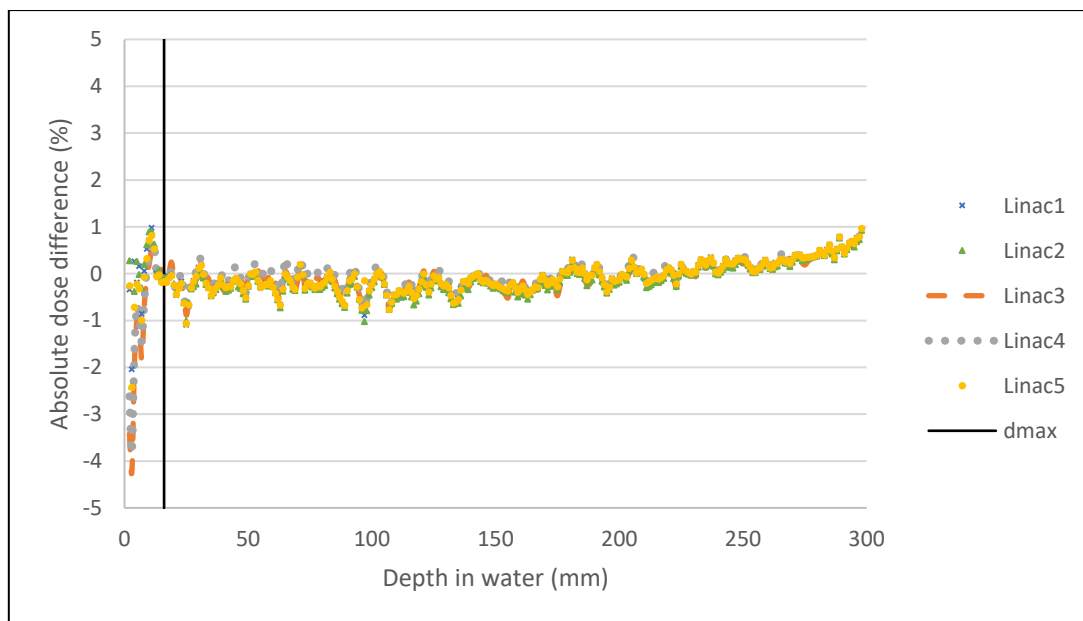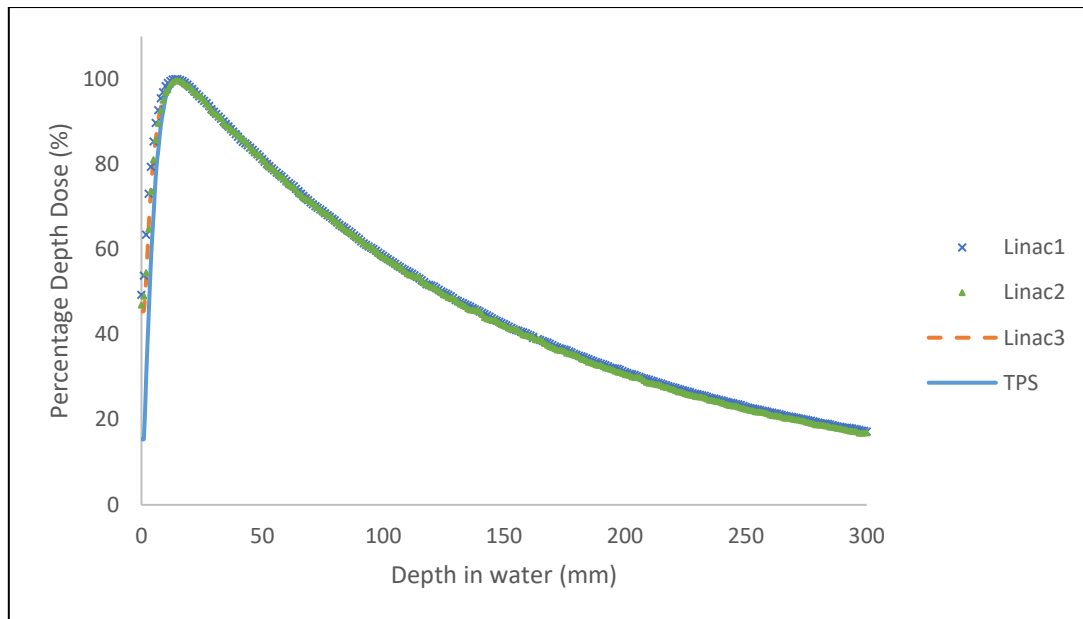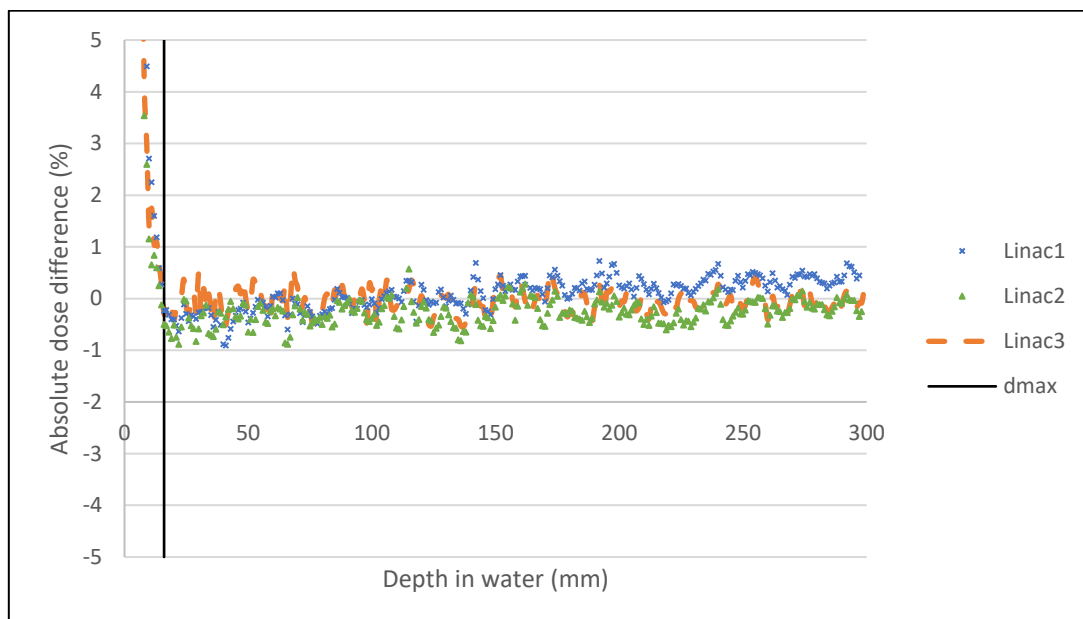

Figure 27: This graph shows the difference between the measured PDD's (from all 5 Linacs) and the calculated PDD respectively for a 5 × 5 cm^2 square field. Results are within 1%, except in the build-up region (up to the black vertical line that represents d$_{max}$). For a more comprehensive, visual interpretation of the results, large surface dose differences up to 17% at 0 mm are not shown in the graph.
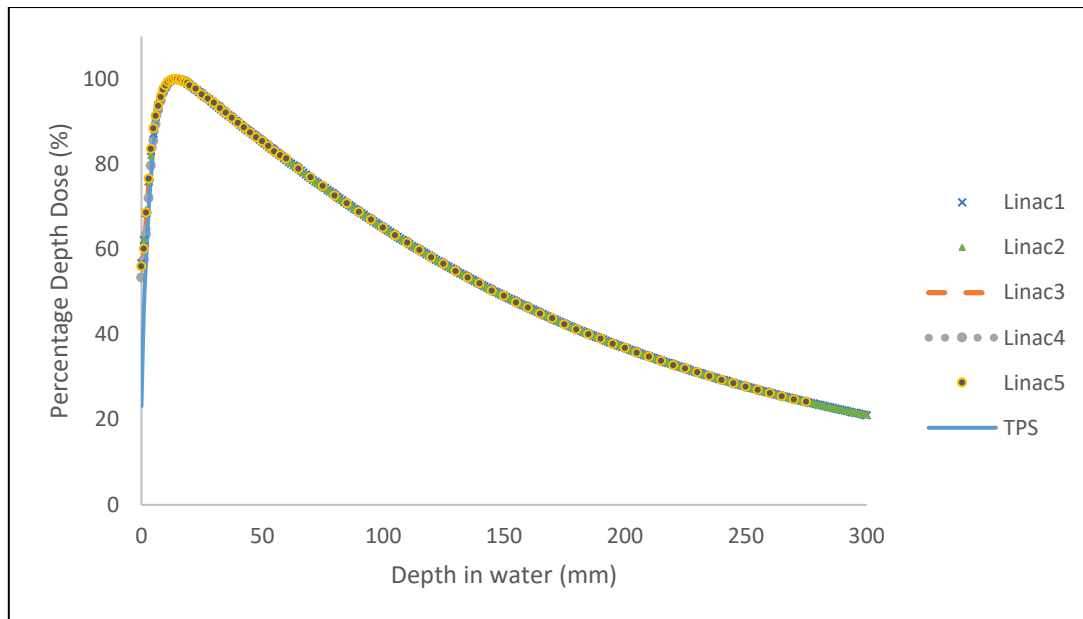
Figure 28: 10 × 10 cm² square field PDDs (normalized to the absorbed dose at d_{max}) *of 5 Linacs (measured) including the TPS (dose calculated) at SSD = 90 cm*
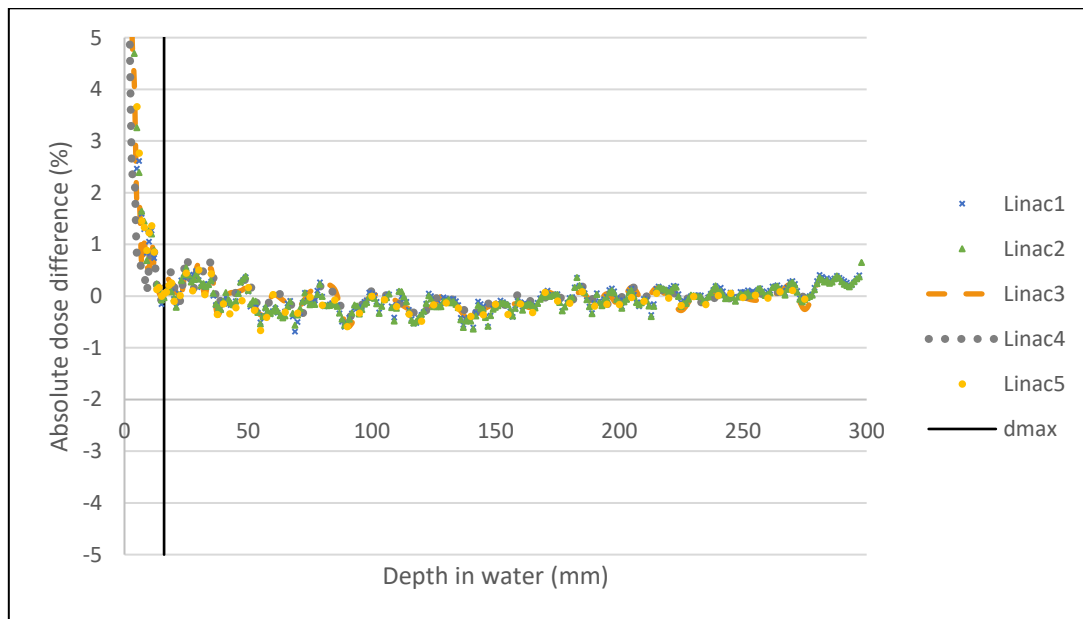
Figure 29: This graph shows the difference between the measured PDD's (from all 5 Linacs) and the calculated PDD respectively for a 10 × 10 cm² square field. For a more comprehensive, visual interpretation of the results, large surface dose differences up to 30% at 0mm are not shown in the graph. The black vertical line represents d_{max}.
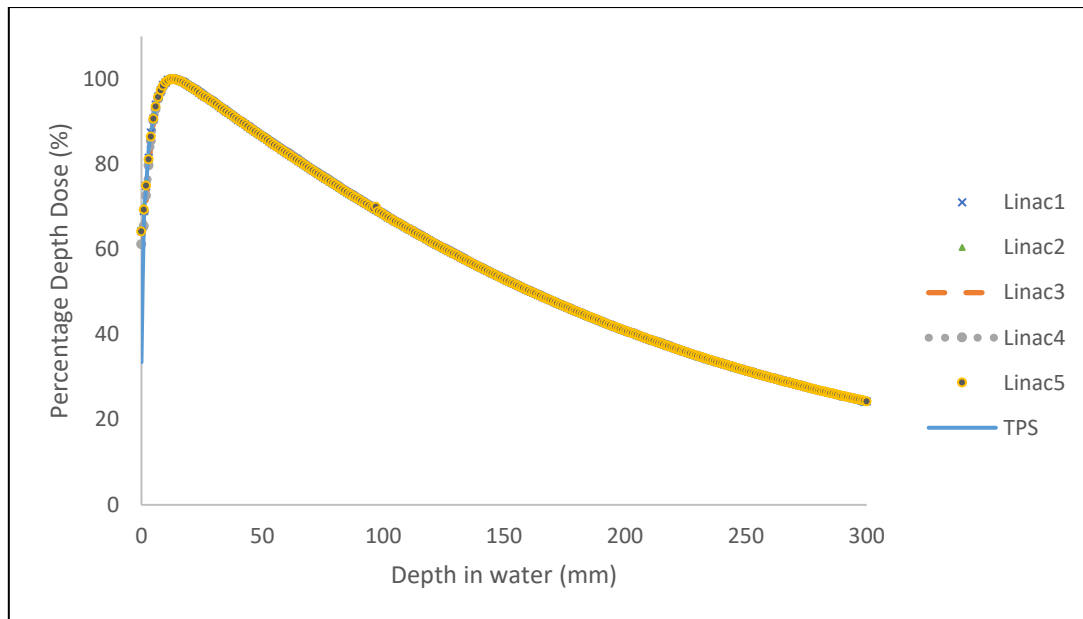
Figure 30: 20 × 20 cm² square field PDDs (normalized to the absorbed dose at dmax) of 5 Linacs (measured) including the TPS (dose calculated) at SSD = 90 cm

Figure 31: This graph shows the difference between the measured PDDs (from all 5 Linacs) and the calculated PDD respectively for a 20 × 20 cm² square field. For a more comprehensive, visual interpretation of the results, large surface dose differences up to 30% at 0mm are not shown in the graph. Slightly above 1% differences can be seen at depths 25 and 97 mm. The black vertical line represents dmax.

Figure 32: 2 × 2 cm² square field PDDs (normalized to the absorbed dose at d_{max}) of 3 Linacs (measured) including the TPS (dose calculated) at SSD = 90 cm

Figure 33: This graph shows the difference between measured PDDs (from 3 Linacs) and the calculated PDD respectively for a 2 × 2 cm² square field. Differences within 1% can be seen from a depth of 13 mm and deeper. The black vertical line represents d_{max}.

5.1.3) Total Scatter Correction Factors

TSCFs are shown in Figure 34. The TSCFs are normalized to the corresponding 10 × 10 cm^2 field. Apart from a few exceptions, the agreement between measured and calculated TSCFs for field sizes up to 20 × 20 cm^2 is better than 1% (shown in table 5) for all Linacs. Exceptions include the 2 × 2 cm^2 field of Linac 5 (1.81%) and the 3 × 3 cm^2 field of Linac 3 (1.40%). Overall the agreement was better than 2% for all fields sizes and Linacs, with one exception seen for Linac4 at field size 40 × 40 cm^2 (4.04%). The fact that different operators, using the same detector (microDiamond), were responsible for these measurements, could have affected the result. Different tolerance values have been recommended, ranging between 1% and 2% [86] [87], and results obtained are satisfactory. An average output factor agreement of 0.4 ± 1.1% have been reported [82]. These included off-axis output factor measurements, not included in the current study.

Others reported differences less than 2% between TSCF measured with a diamond detector and calculated with the Monaco TPS for small fields. It has been reported that the disagreement between the measured and calculated values can be minimized by iteratively reducing the size of the primary photon source diameter in the Monaco beam model [67]. This iterative process, until agreement between measured and calculated TSCF is achieved, was optimized by the TPS modelling team. Therefore a 0% difference exists between calculated and reference Linac measured (Linac1) TSCF of the 2 × 2 cm^2 field.
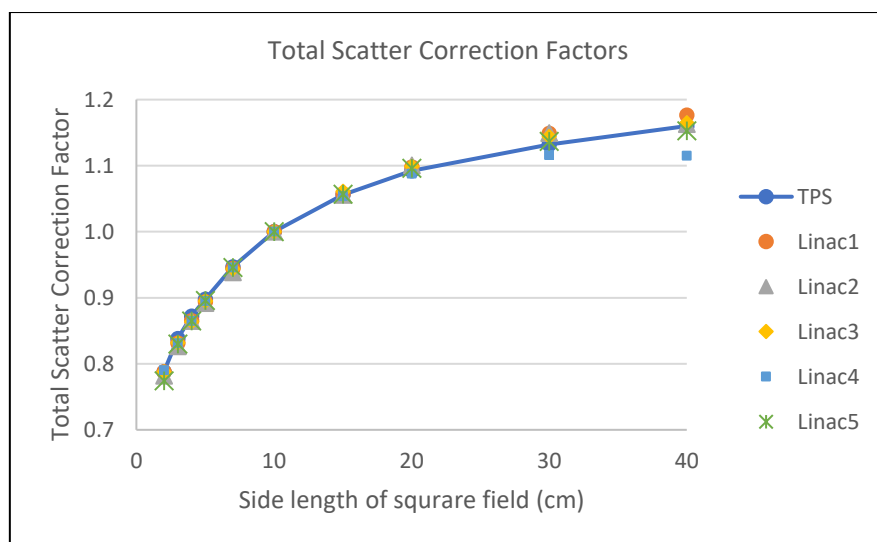
Figure 34: Total Scatter Correction factors from each Linac and the reference model

	Square field size (cm2)								
	2×2	3×3	4×4	5×5	7×7	15×15	20×20	30×30	40×40
Linac1	0.00	0.78	0.79	0.44	0.30	0.11	0.56	1.46	1.42
Linac2	0.79	1.40	0.81	0.69	0.89	0.00	0.75	1.54	0.24
Linac3	0.00	0.72	0.94	0.34	0.32	0.49	0.64	1.05	0.43
Linac4	0.25	0.84	0.81	0.34	0.11	0.19	0.37	1.43	4.04
Linac5	1.81	0.96	0.81	0.22	0.11	0.09	0.36	0.44	0.61

Table 5: Percentage difference between TPS calculated and Linac measured TSCF

5.1.4) Profiles

Figures 35 to 42 shows profile comparison between calculated and measured profiles from 3 Linacs (Linacs 1 to 3). The figures include calculated γ results for each Linac, field size and depth in water. Of the 24 profiles evaluated, 17 had overall γ values of 1 or less using a γ criterion of 2%/2mm. The results show a good agreement between Linac measurements with very similar disagreements to calculations from the model.

From the 7 profiles that have failed γ points, one failed the criterion in the out of field region (figure 38(a)). This discrepancy is visible, to variable extent, in most of the in-plane scans at a depth of d_{max}, with the measured values being higher than the calculated ones. This does not indicate differences between the Linacs as the same effect was seen on all 3 of them, but rather a difference between the Monaco model and a Siemens Artiste Linac. This could be due to inaccuracies in the Monaco beam model as it handles transmission values through the beam modifiers of a Siemens Artiste as total absorbing [68]. Total absorption was introduced in earlier versions of XVMC to limit the size of the scoring plane. This means should transmission values be less than 1.1%, which is true for the jaw transmission of a Siemens Artiste, no particles are generated in the scoring plane and the jaw is modelled as fully absorbing which could influence dose calculations under the modifiers.

The other 6 fields that have failed γ points were all inplane profiles failing in the open beam area (one-sided) for the 20 × 20 cm^2 at both depths for all 3 Linacs, including the reference Linac. For these fields, the model calculated a sharper horn profile causing approximately 1.0 - 2.0% of points to be out of tolerance. This effect is more

prominent on one side of the profiles, which could indicate asymmetric beams or a systemic inaccuracy in the scanning equipment or setup thereof. Nonetheless, the model cannot account for asymmetric beams.

For profile comparisons from measurements on a different Linac but planned on the same TPS, results less than 2%/2mm have been reported by others [67] [82]. Sikora (2011) further reported similar results and recommended an overall profile comparison criteria of 3%/2mm inside the field and 5%/2mm outside the field [19]. However, that work included very small field sizes (down to 0.8×0.8 cm^2), not included in the current study.
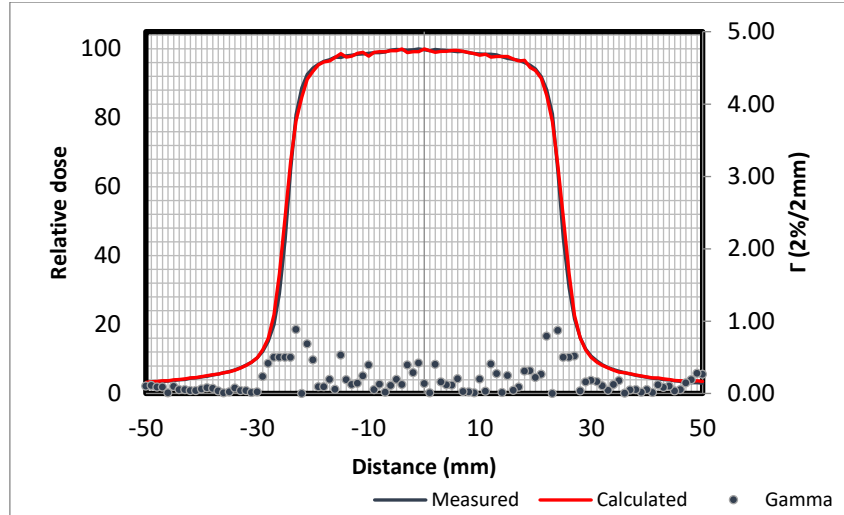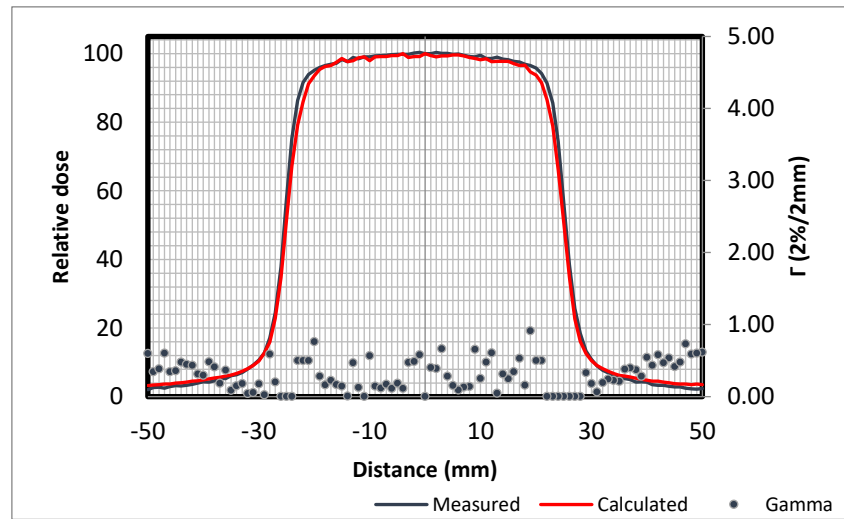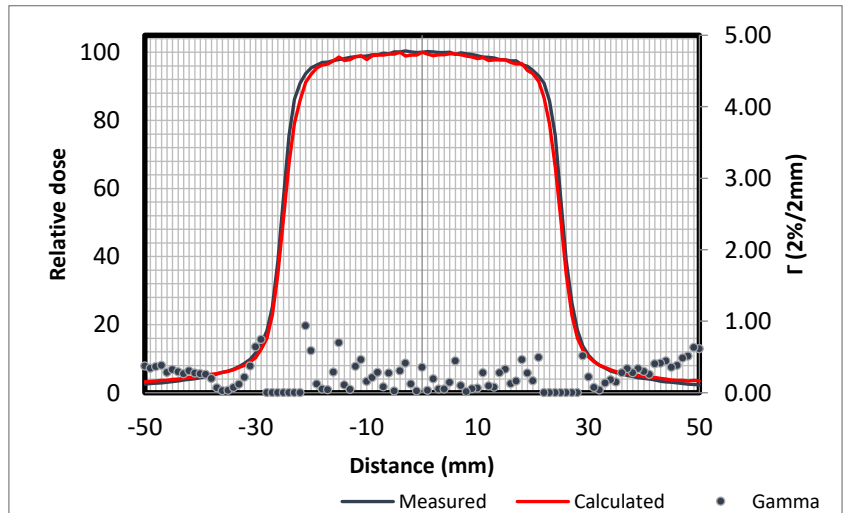
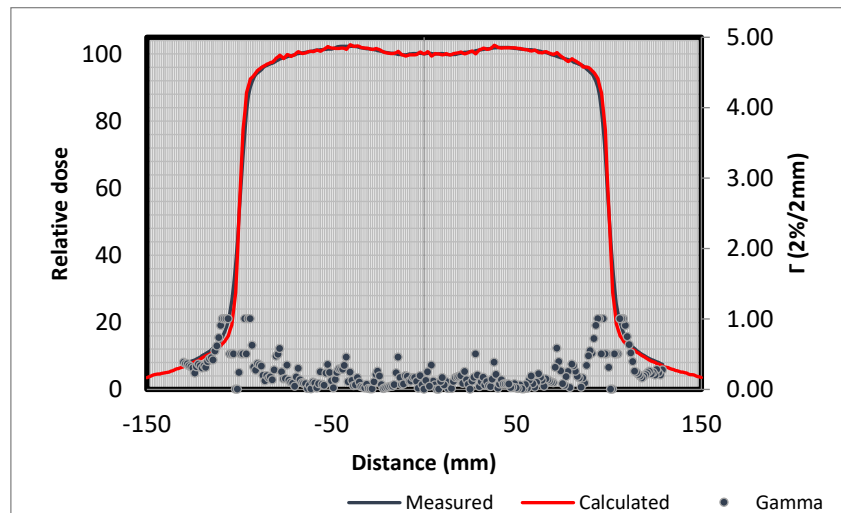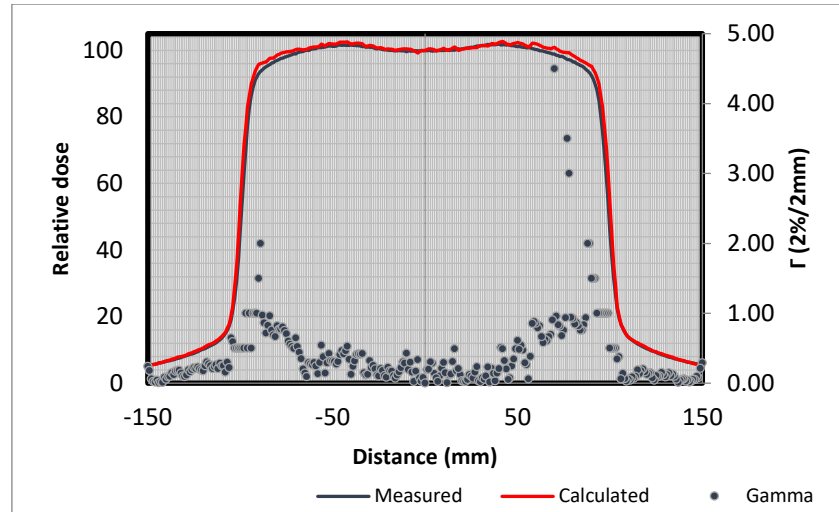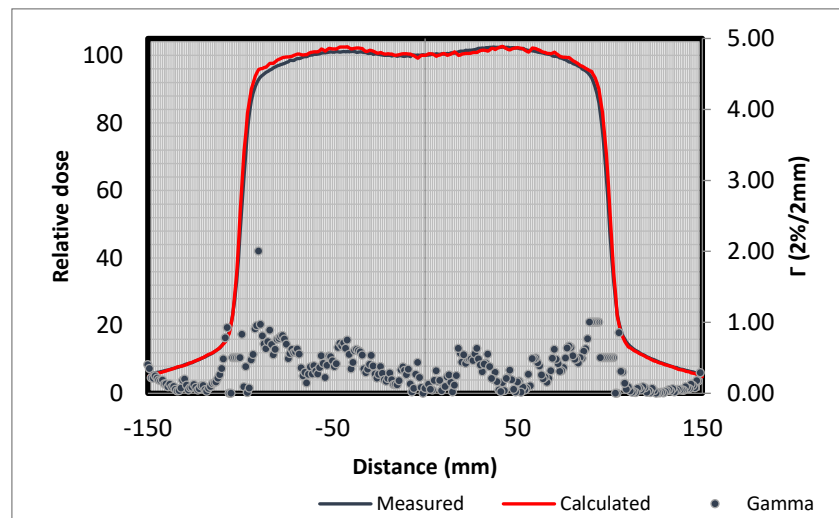*Figure 35 (a-c): Measured profiles (5 cm × 5 cm @ 10 cm depth and crossplane direction)*
from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively.
A very good agreement with all γ values less than 1 are shown.
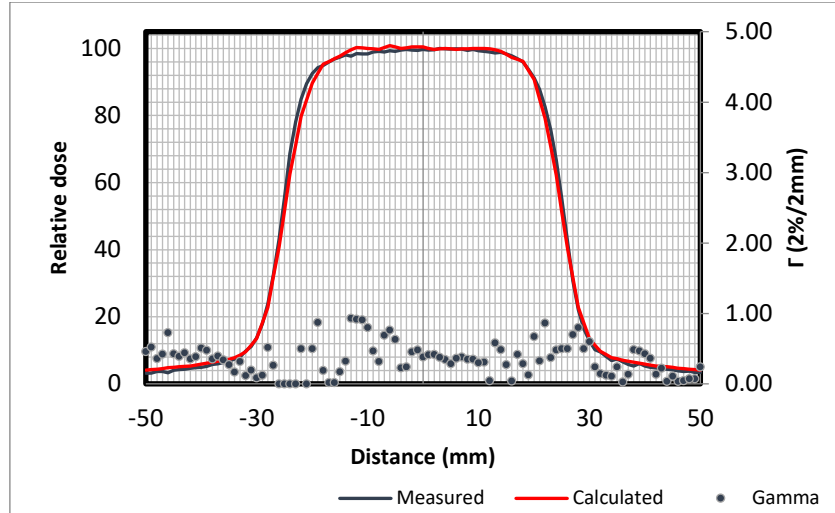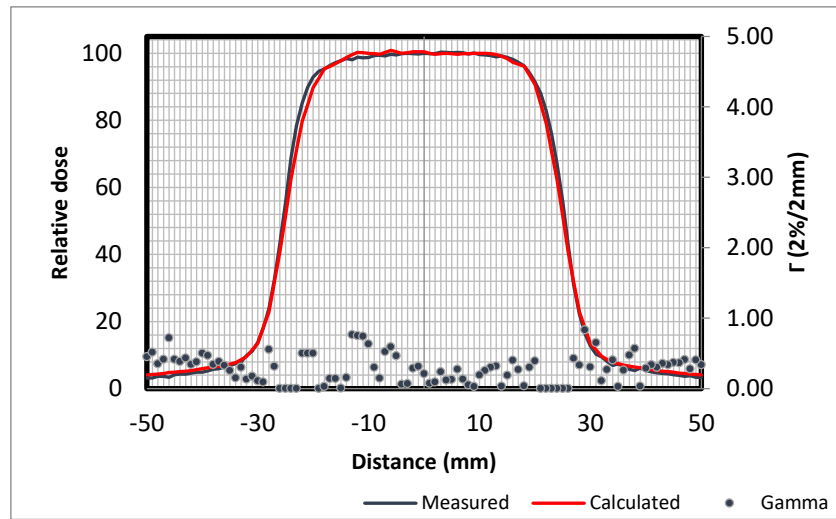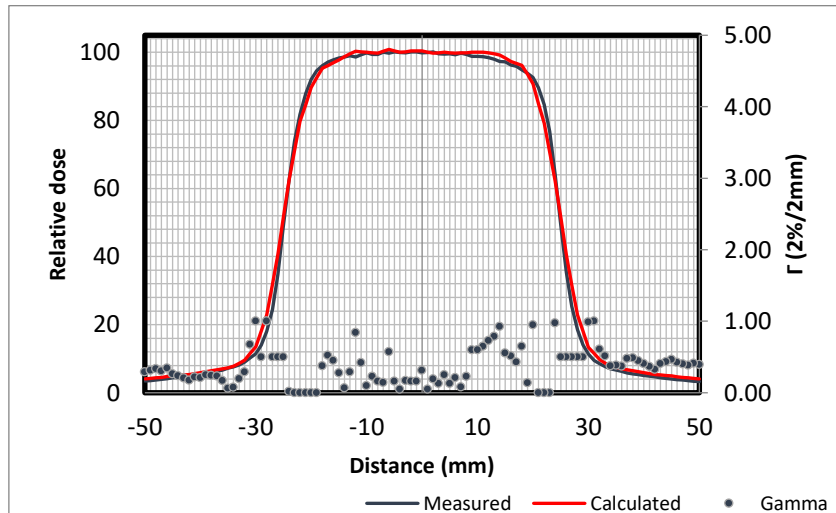
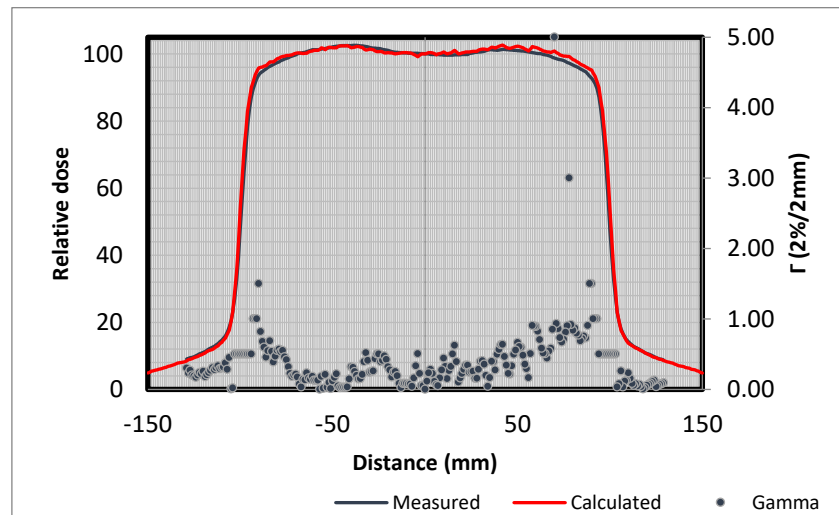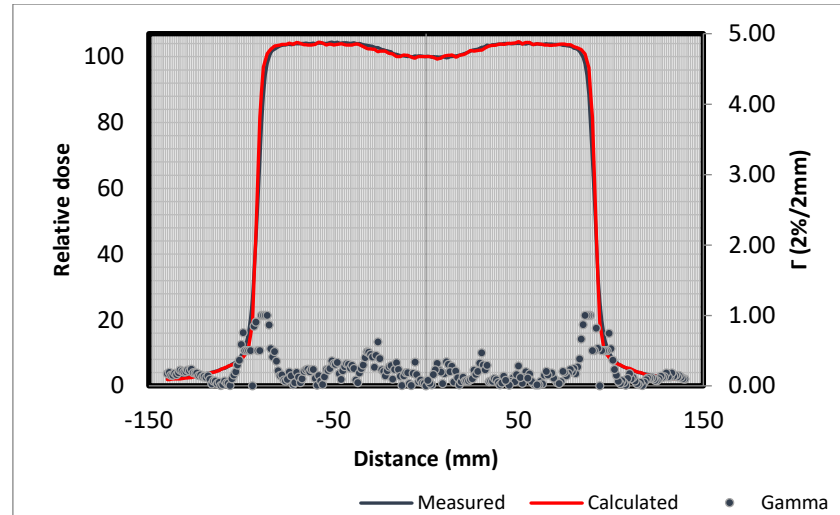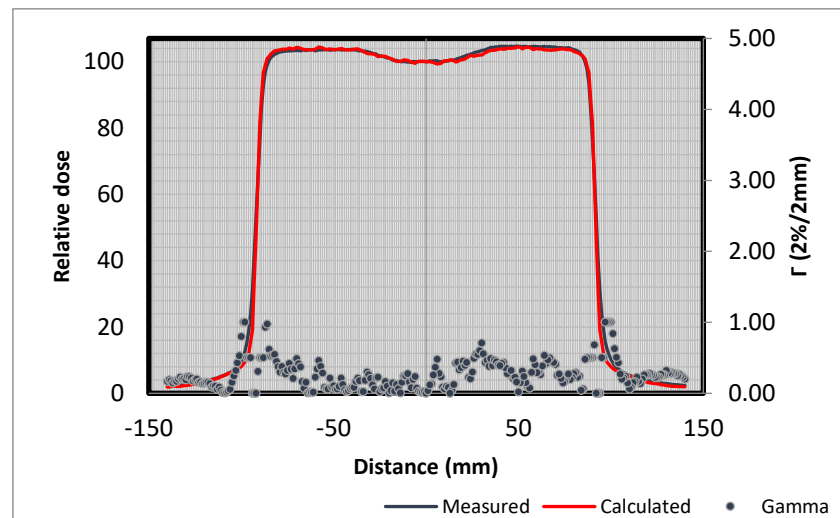*Figure 36 (a-c): Measured profiles (5 cm ✕ 5 cm @ 10 cm depth and inplane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. A very good agreement with all γ values less than 1 are shown.*
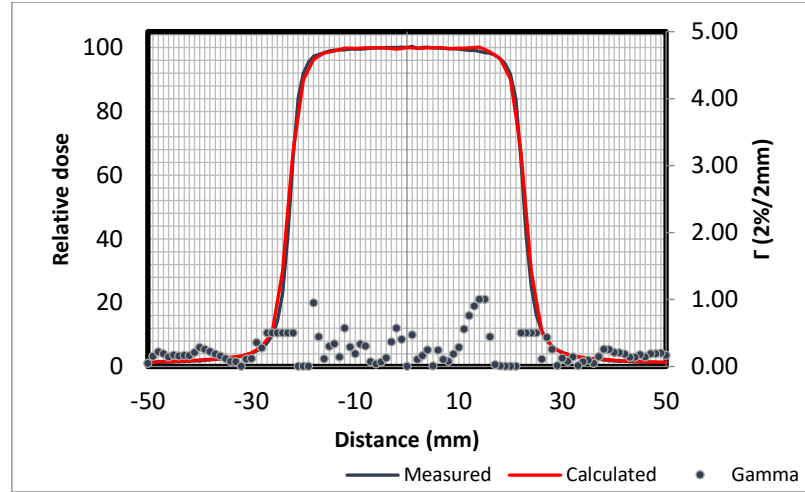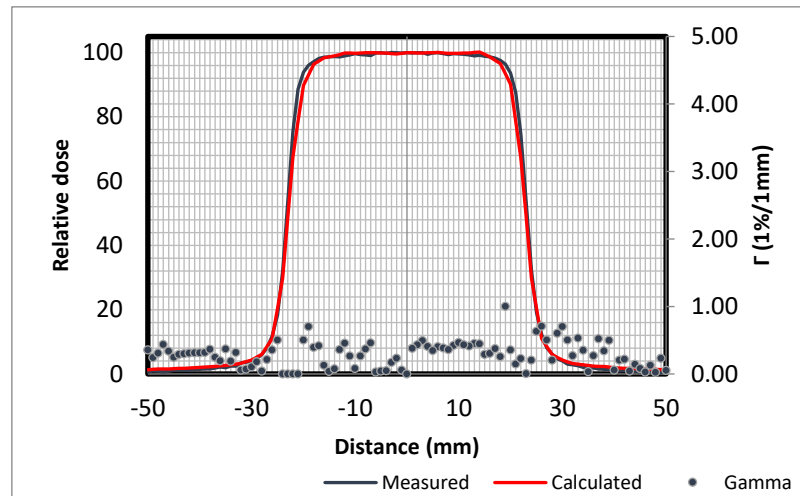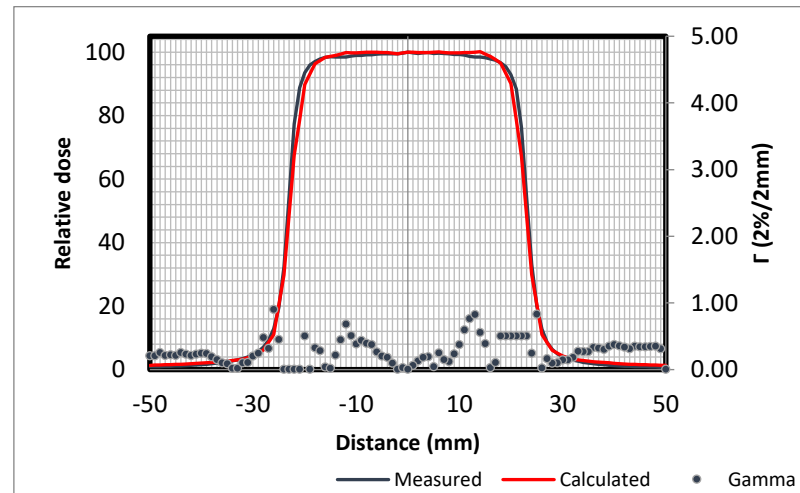
Figure 37 (a-c): Measured profiles (5 cm × 5 cm @ dmax depth and cross plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. A very good agreement with all γ values ≤ 1 are shown.
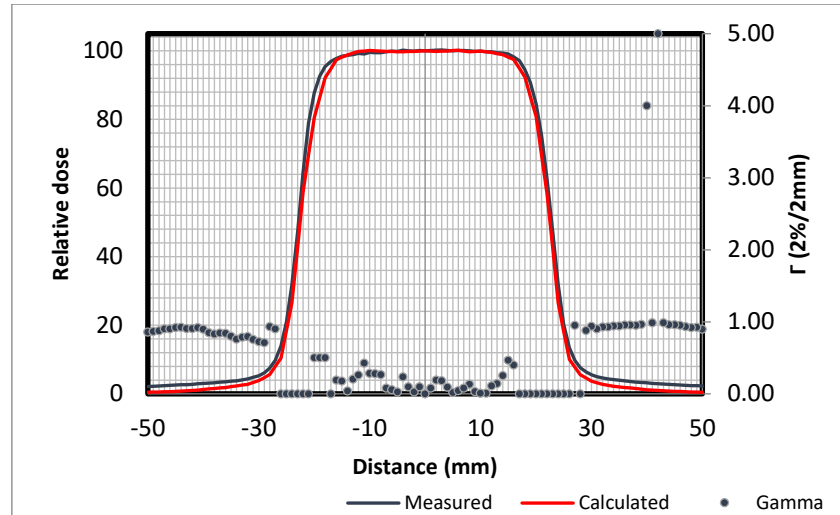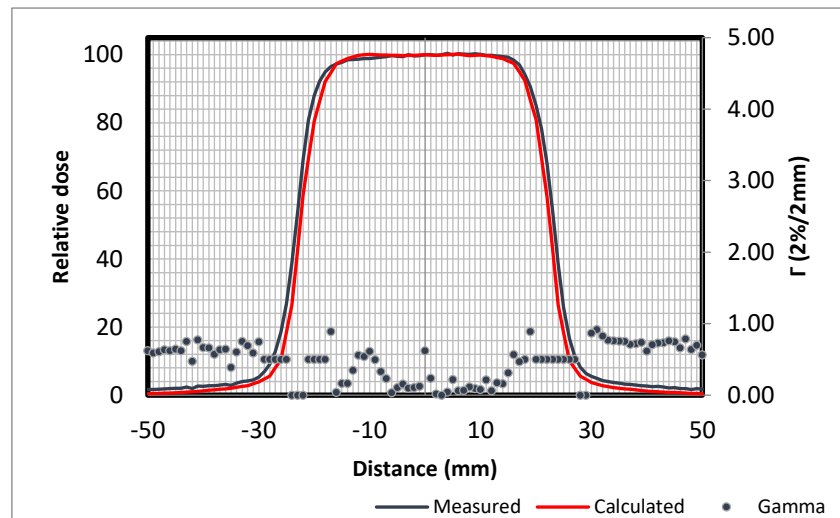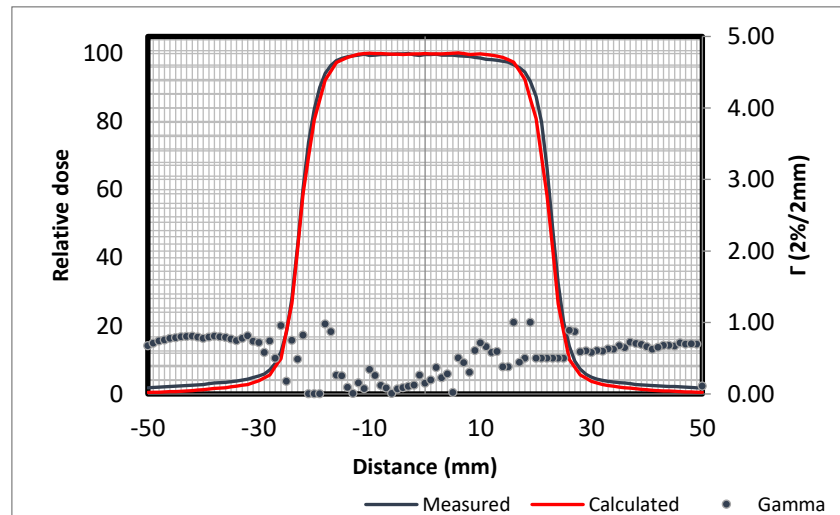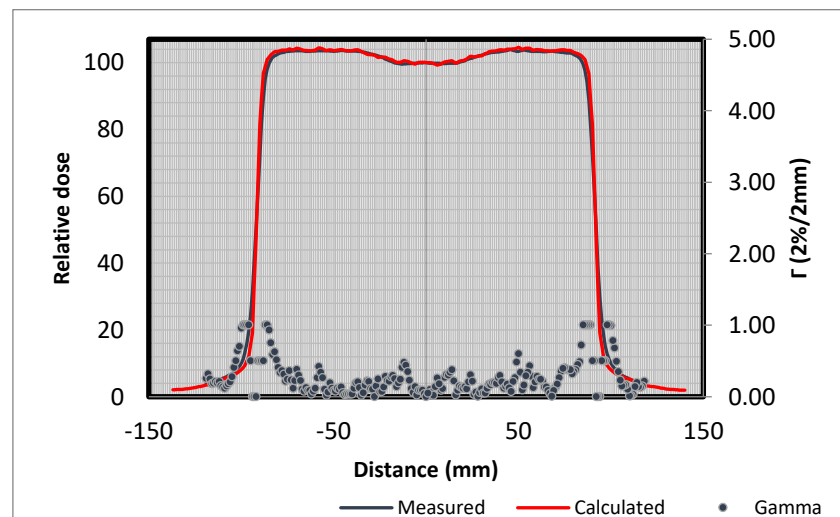
Figure 38 (a-c): Measured profiles (5 cm × 5 cm @ dmax depth and in plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. All scans show a dose discrepancy, mostly within tolerance, under the beam modifiers (y jaws)
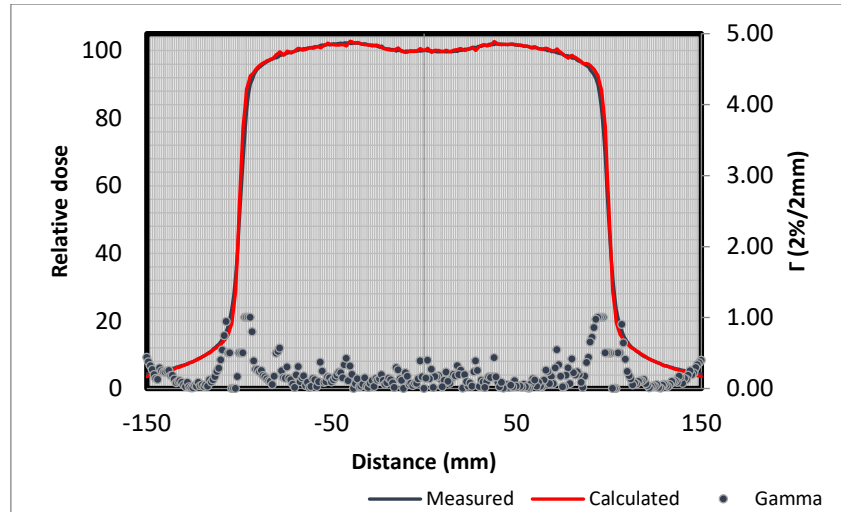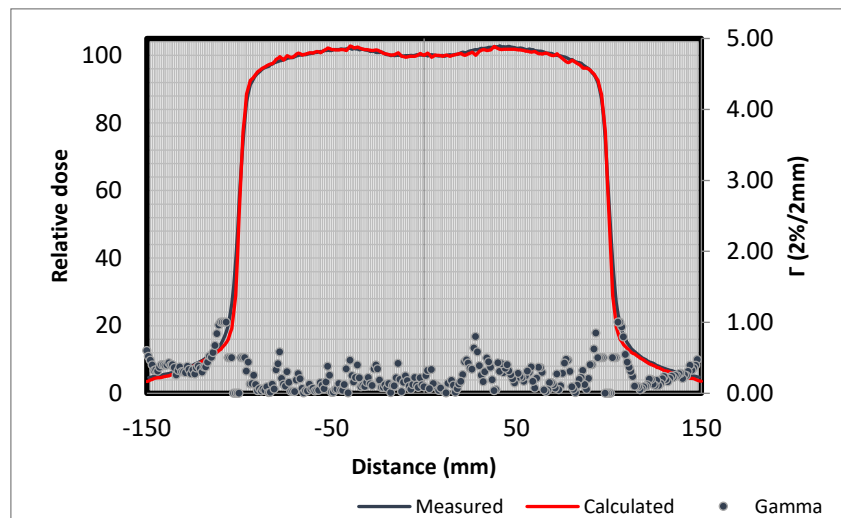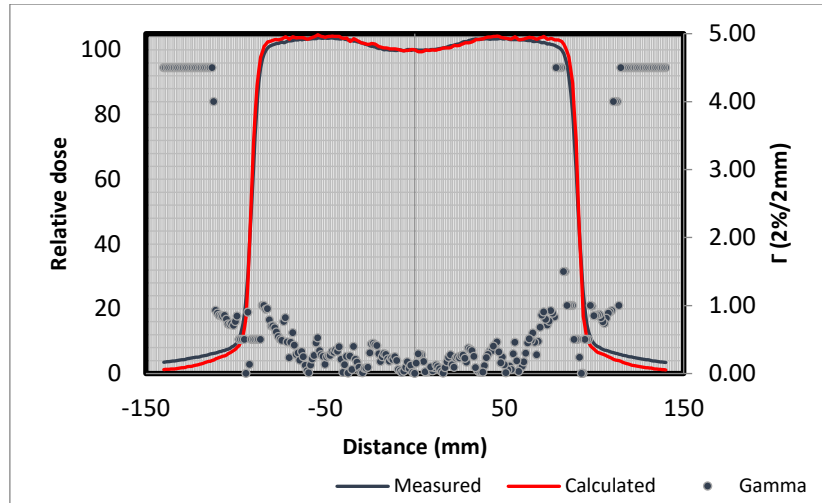
Figure 39 (a-c): Measured profiles (20 cm ✗ 20 cm @ 10 cm depth and cross plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. All scans show a very good agreement with γ values ≤ 1.

Figure 40 (a-c): Measured profiles (20 cm × 20 cm @ 10 cm depth and in plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. For all scans the model calculated a sharper horn profile.

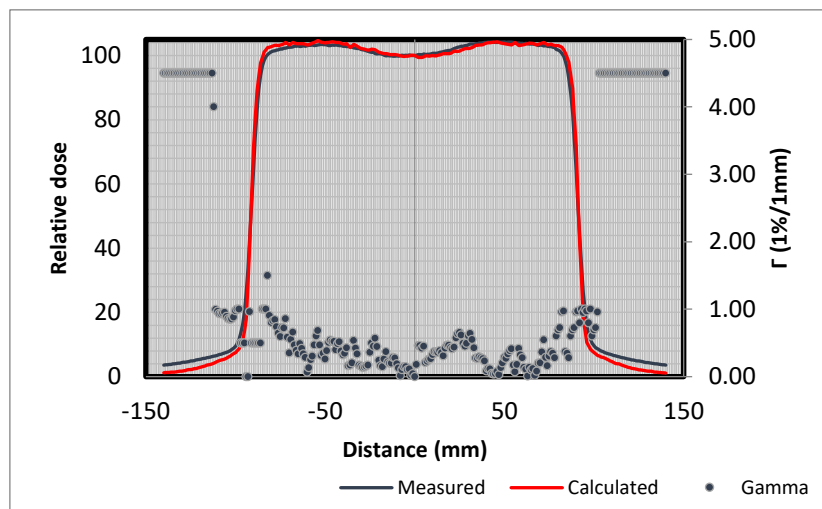*Figure 41 (a-c): Measured profiles (20 cm × 20 cm @ d_{max} depth and cross plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. All scans show a very good agreement with γ values ≤ 1.*
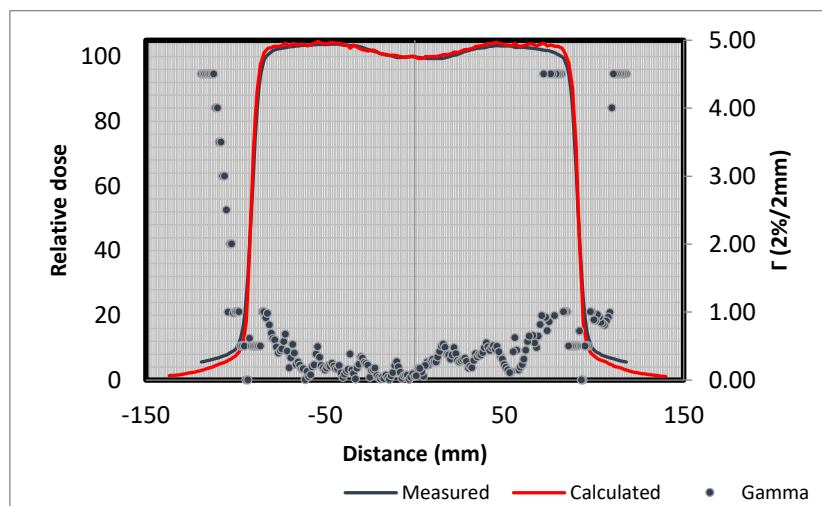
Figure 42 (a-c): Measured profiles (20 cm ✗ 20 cm @ d_{max} depth and in plane direction) from Linac 1(a), Linac 2 (b) and Linac 3 (c) each compared to the TPS calculation respectively. All scans show out-and in-field discrepancies.

5.2 Increased array spatial resolution

5.2.1) Increased Resolution Method Reproducibility

Considering the measurement methodology discussed in 4.3.1, reproducibility of the method (Mapcheck2™ setup and subsequent measurements in 1 mm increments) measured 5 times over a period of 3 months are shown in Figure 43. The reproducibility is 0.1 mm or better, as calculated by OmniPro I'mRT, which is satisfactory.
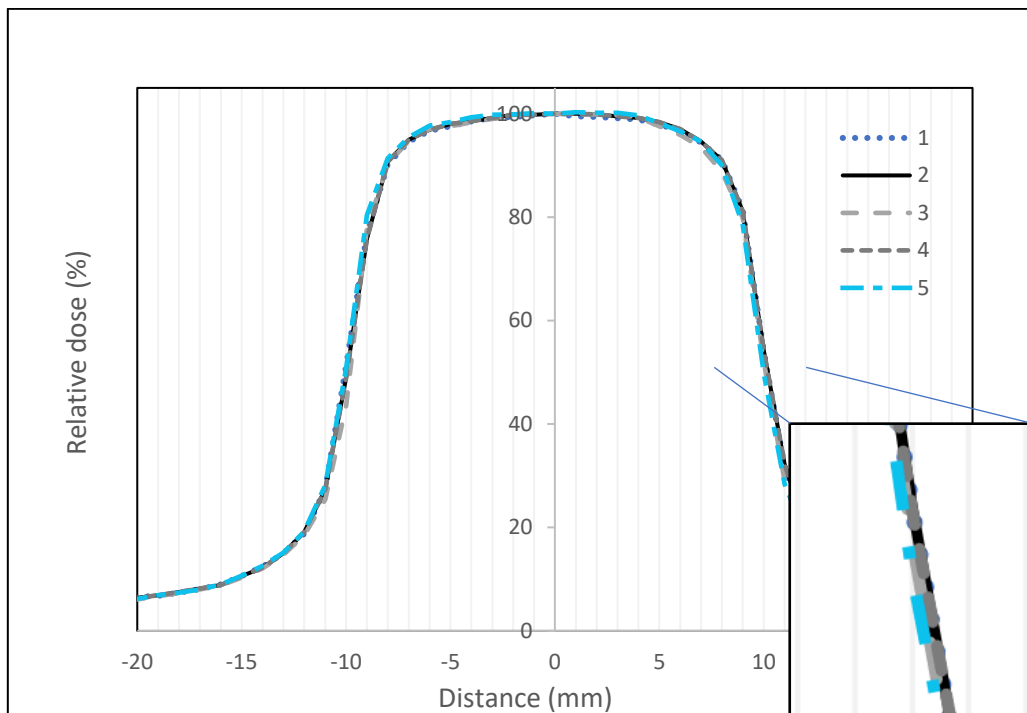

Figure 43: Setup reproducibility for a 2 × 2 cm^2 crossplane profile, measured 5 times. The right-side penumbra is zoomed to show the effect. The gridlines in the figure are spaced 1 mm apart.

5.2.2) Array sensitivity

Figures 44 to 55 show the comparison between measurement devices: individual scans from the microDiamond (TM60019) detector measured in water compared to array Mapcheck2™ measurements in 1 mm increments. The graphs contain a 1 mm grid for visual interpretation of the differences found.

For the criterion of 1 mm DTA and 1% DD, generally the agreement between the diamond and Mapcheck detectors was very good and within a γ value of 1. An average γ value of 0.41 ± 0.15 was calculated within the field and 0.27 ± 0.07 outside the field.

The two 5 cm × 5 cm crossplane comparisons had out of tolerance values in the tail end of the profiles (beyond 50% isodose). In evaluating all the crossplane scans, Mapcheck in general slightly overestimates the dose beyond the penumbra region with 0.5 to 1.0%, as indicated by higher γ values in these areas. A detector's, specifically a diode's, overresponse to low energy scattered photons in the low-dose tail of profiles (at a distance of 1 cm from the field edge) have been reported [88] [81]. This overresponse is because the average beam energy is lower outside the treatment field when in fact the diode was cross-calibrated to a harder, primary photon beam. However, only a dosimeter with a higher atomic number than tissue, such as a Mapcheck's Si diode, will overrespond in this way. The effect is not visible for the diamond detector with an effective atomic number close to tissue. This result should be taken into consideration when using Mapcheck to characterize model transmission and leakage values.

Four of the 6 inplane comparisons had out of tolerance values, all from the penumbra high dose region on the gun side of the Linac. As previously mentioned, the absolute γ value is of little value and could be truncated, especially by slightly out of tolerance DTA results in the penumbra region usually translated into high dose differences. The DTA for these out of tolerance profiles was 1.2 mm on average. Mapcheck yielded a sharper penumbra of 0.6 mm ± 0.17 (scan left) and 0.36 mm ± 0.14 (scan right), possibly due to the small volume of its diodes. The broader penumbra on one side of inplane profiles have been reported in a study of Mapcheck's predecessor, the SNC Profiler[TM] [89]. The positive y-axis has an increased amount of side-scatter relative to the negative y-axis, which is adjacent to the panel's electronics. They reported the effect of additional side-scatter on calibration factors was an asymmetric response up to 2.21%, but it should be highlighted that the effect became only significant for field sizes larger than $10 \times 10 \ cm^2$.

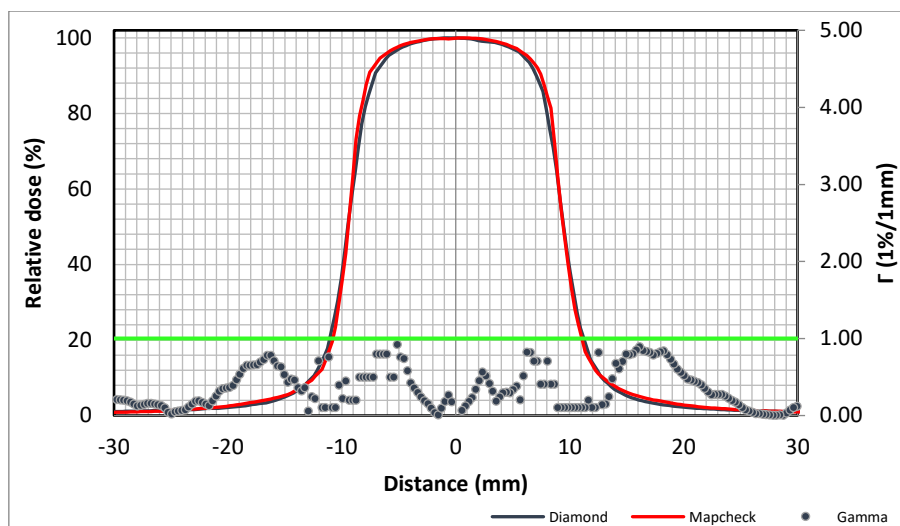*Figure 44: Comparison between microDiamond and Mapcheck measured CAX profiles: 2 cm × 2 cm @ 5 cm depth in the crossplane direction. The figure show a very good agreement between these detectors with all γ values below 1.*
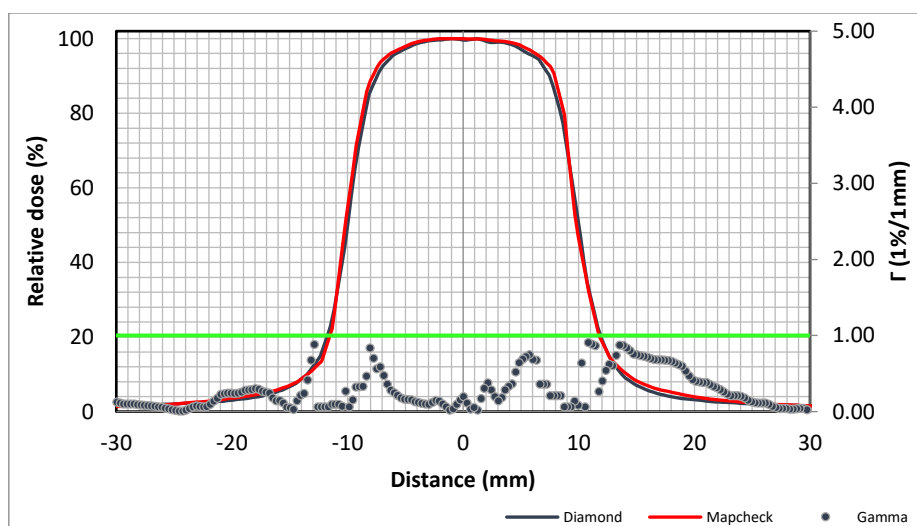

Figure 45: Comparison between microDiamond and Mapcheck measured CAX profiles: 2 cm × 2 cm @ 10 cm depth in the crossplane direction. The figure show a very good agreement between these detectors with all γ values below 1.
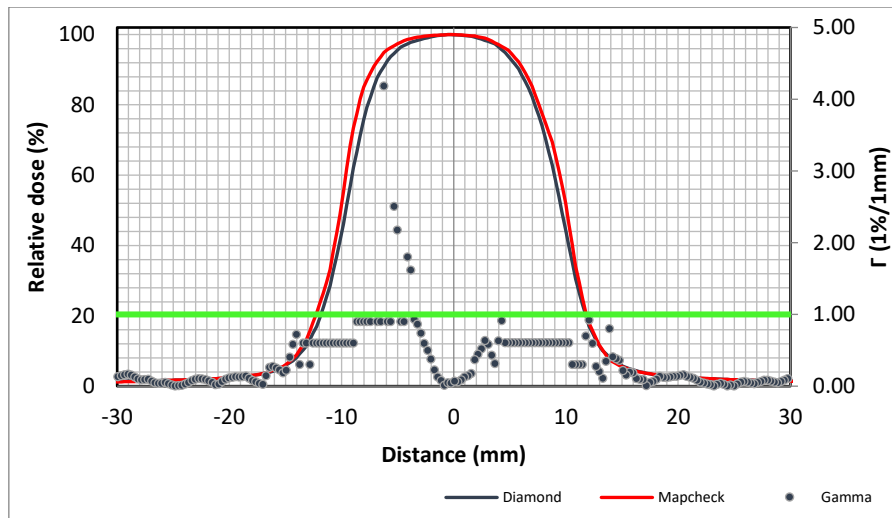
Figure 46: Comparison between microDiamond and Mapcheck measured CAX profiles: 2 cm × 2 cm @ 5 cm depth in the inplane direction. Out of tolerance values on the gun side of the profile are indicated with γ values up to 4.00.

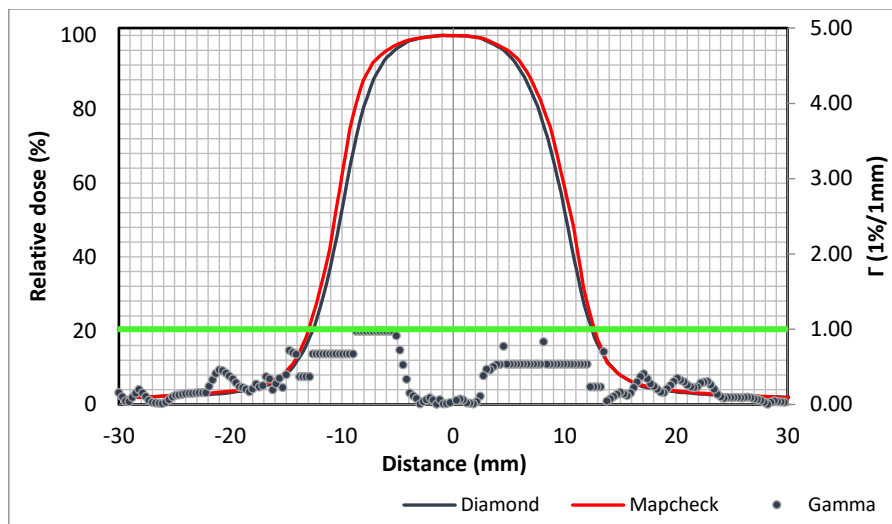*Figure 47: Comparison between microDiamond and Mapcheck measured CAX profiles: 2 cm × 2 cm @ 10 cm depth in the inplane direction. The same effect as Figure 46 are shown: Mapcheck measuring a broader penumbra on one side. However, in this case all γ values are below 1 or within 1%/1mm.*
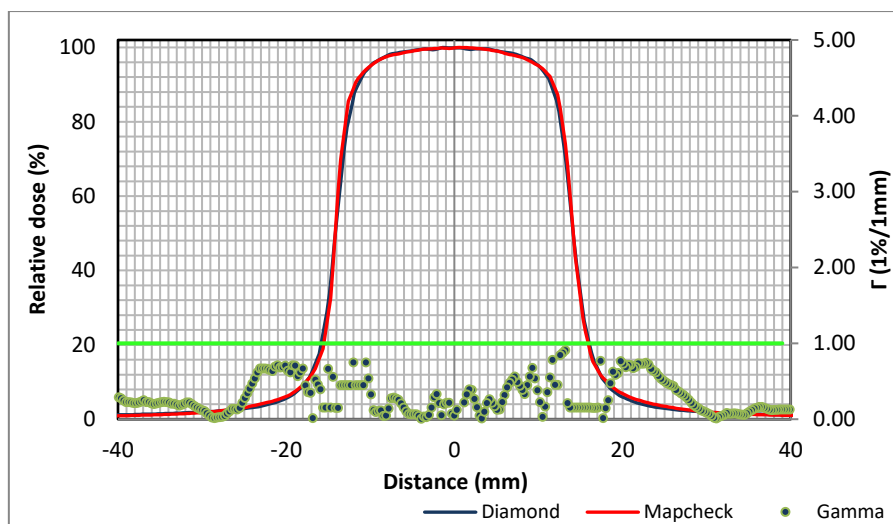
Figure 48: Comparison between microDiamond and Mapcheck measured CAX profiles: 3 cm × 3 cm @ 5 cm depth in the crossplane direction. The figure show a very good agreement between these detectors with all γ values below 1.

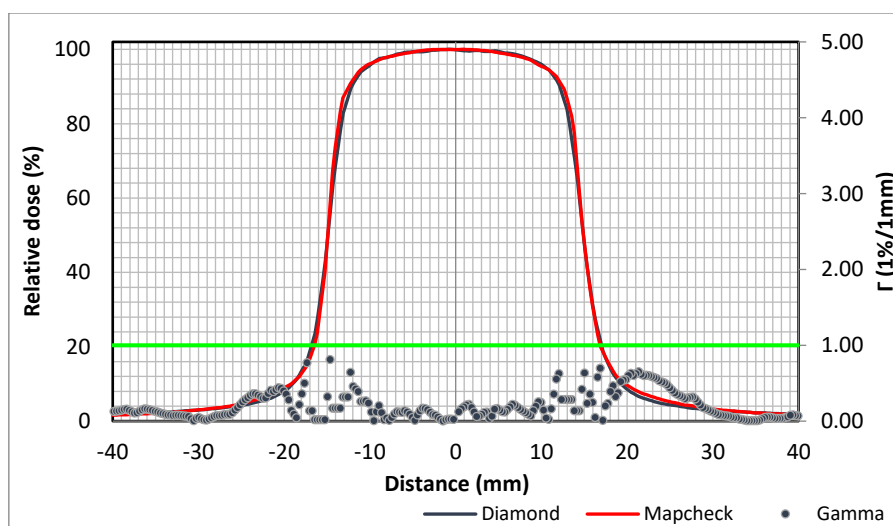*Figure 49: Comparison between microDiamond and Mapcheck measured CAX profiles: 3 cm × 3 cm @ 10 cm depth in the crossplane direction, showing a very good agreement between the detectors.*
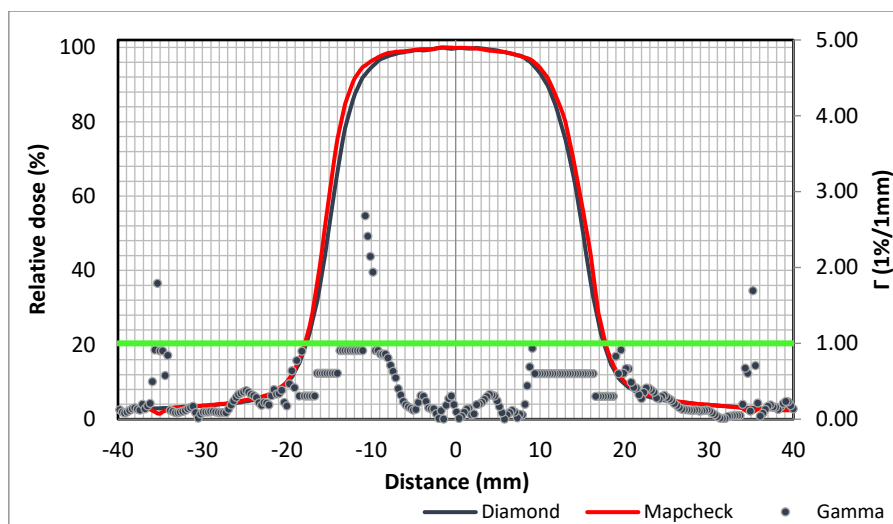
Figure 50: Comparison between microDiamond and Mapcheck measured CAX profiles: 3 cm × 3 cm @ 5 cm depth in the inplane direction. The same effect as before can be seen: Mapcheck measuring a broader penumbra on one side.
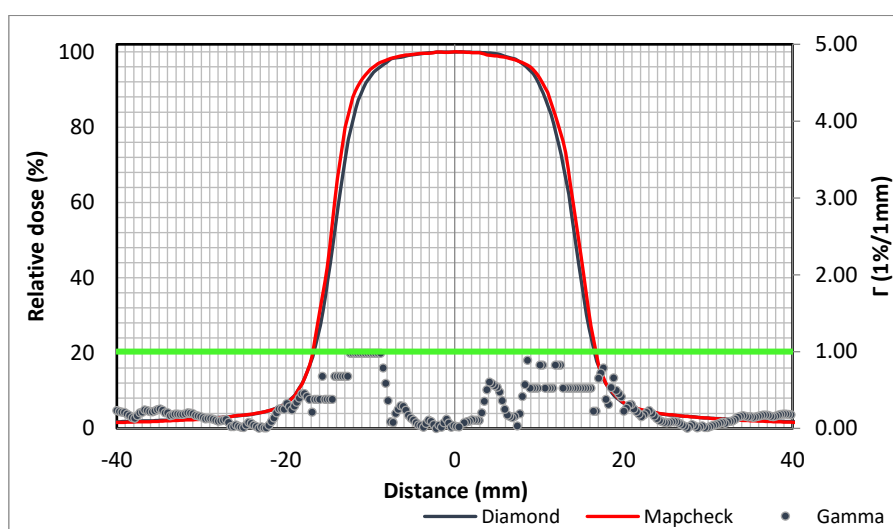

Figure 51: Comparison between microDiamond and Mapcheck measured CAX profiles: 3 cm × 3 cm @ 10 cm depth in the inplane direction. All γ values are ≤ 1.
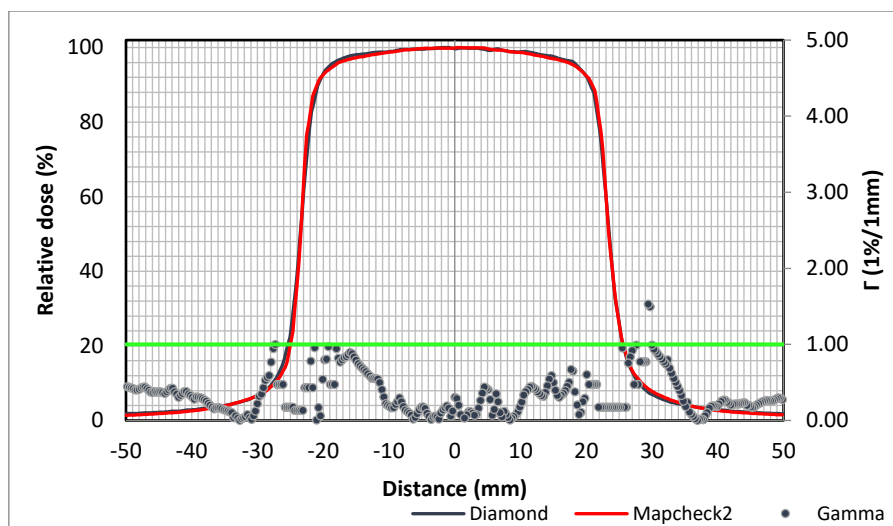
Figure 52: Comparison between microDiamond and Mapcheck measured CAX profiles: 5 cm × 5 cm @ 5 cm depth in the crossplane direction. Most of the γ values are ≤ 1
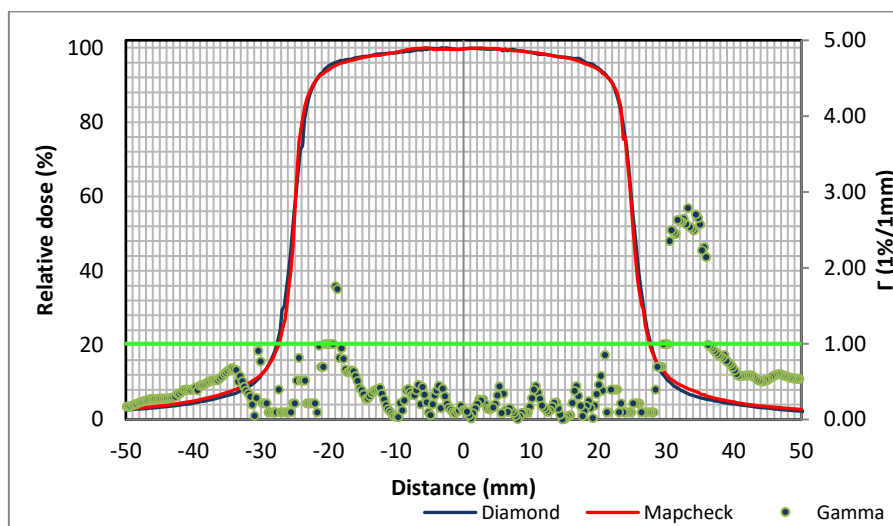

Figure 53: Comparison between microDiamond and Mapcheck measured CAX profiles: 5 cm × 5 cm @ 10 cm depth in the crossplane direction. Mapcheck overestimates the dose beyond the penumbra region with 0.5 to 1.0%, possibly due to the response of the detectors to lower energy photons.
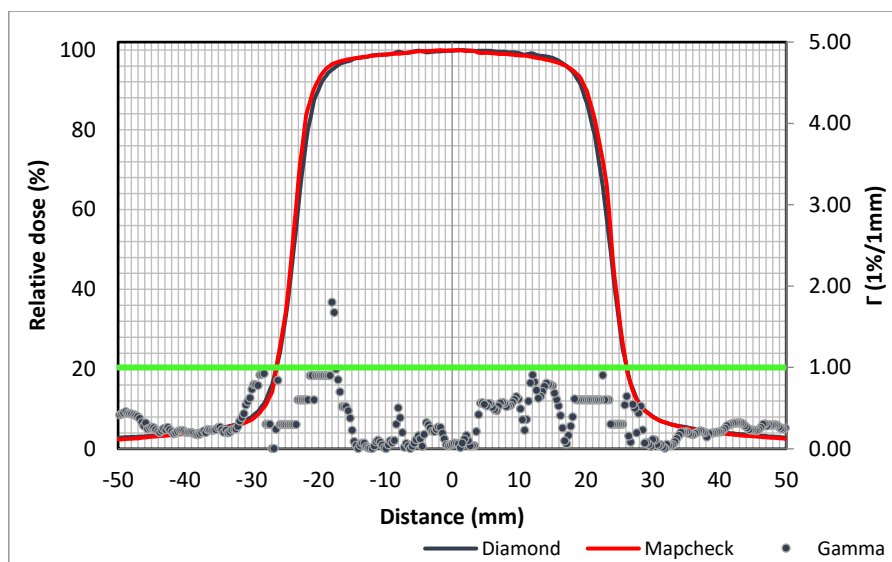
Figure 54: Comparison between microDiamond and Mapcheck measured CAX profiles: 5 cm × 5 cm @ 5 cm depth in the inplane direction. The figure shows 2 out of tolerance γ values.
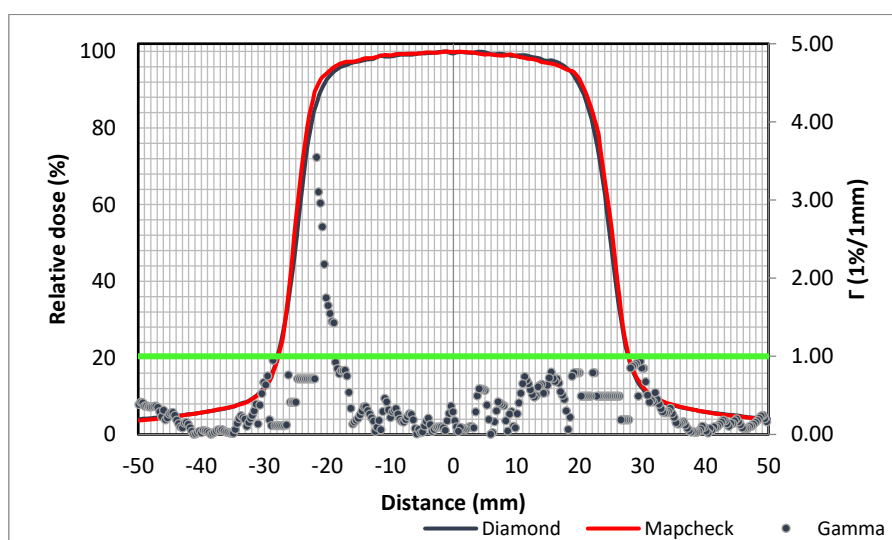

Figure 55: Comparison between microDiamond and Mapcheck measured CAX profiles: 5 cm × 5 cm @ 10 cm depth in the inplane direction. Out of tolerance γ values due to a assymetric Mapcheck profile, possibly due to added side scatter from its electronics.

Figure 56 summarizes the result of 12 profile scans evaluated, plotted as a normal boxplot with the top and bottom of the rectangle representing the 75th (Q3) and 25th (Q1) percentile values respectively (0.89 and 0.32) and the median (0.52) as the line within the box. Ideally this range should be as small as possible. The upper whisker corresponds to Q3 + 1.5*(Q3-Q1) and lower whisker to Q1 – 1.5*(Q3 – Q1) or to the nearest data point thereof. Values above or below the respective whiskers are plotted

as outliers. The minimum outlier in figure 56 is within the lower whisker range and is therefore not shown. Most of the values are within the specified criterion which indicate that the Mapcheck device can be used, with caution, for accurate comparison between Linac measurements and TPS calculations. Cautions to consider are Mapcheck's response to radiation: an asymmetric response in the in-plane direction by approximately 1.2mm and an out of field overresponse in the cross-plane direction up to 1%.
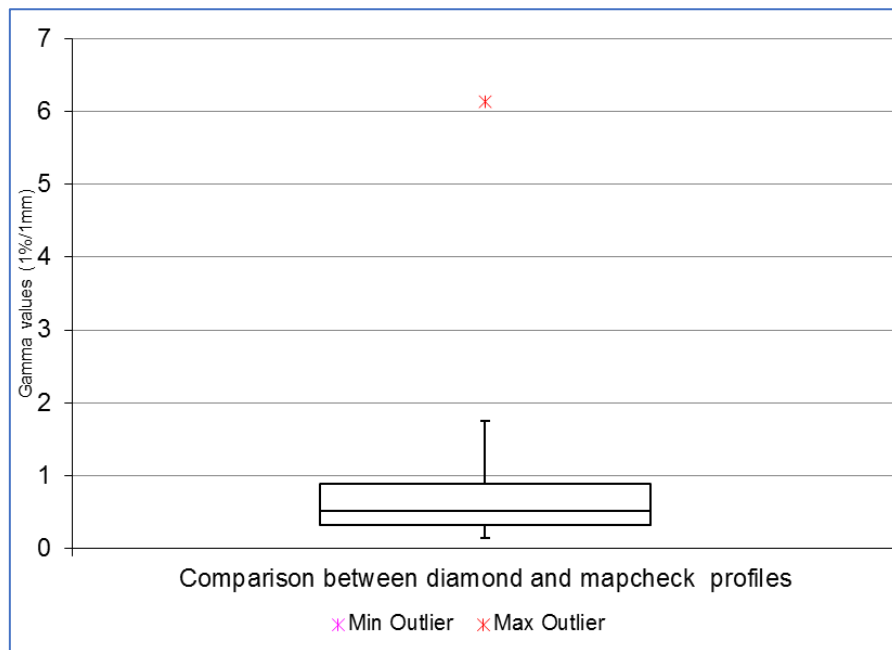

Figure 56: Boxplot of individual γ values obtained from profile comparisons between microDiamond and Mapcheck machine measurements. The majority of values are within the specified criterion (γ ≤ 1)

Figure 57 shows a comparison of the two detectors for a complete set of CAX asymmetric 2 × 2 cm^2 fields, the picket fence, a routine MLC QA test. Once again, a very good agreement was seen between these. Here the γ comparison was only performed inside the field (down to 20% isodose) as overlapping fields complicated the calculation. Overall 98.85% of comparison points (347 in total) evaluated had a γ value of 1 (criterion 1%/1mm) or below with an average of 0.64 ± 0.57. Several segments have out of tolerance values in the open part of the beam due to erratic scans. This could be measurement related but unfortunately reproducibility was not tested. However, as the penumbra agreement between detectors are very good, the result is particularly helpful to characterize correlation or differences between different

Linac MLC's (discussed in 4.4.1). This shows Mapcheck, with increased resolution that removes under sampling effects can be used to determine MLC positional accuracy.
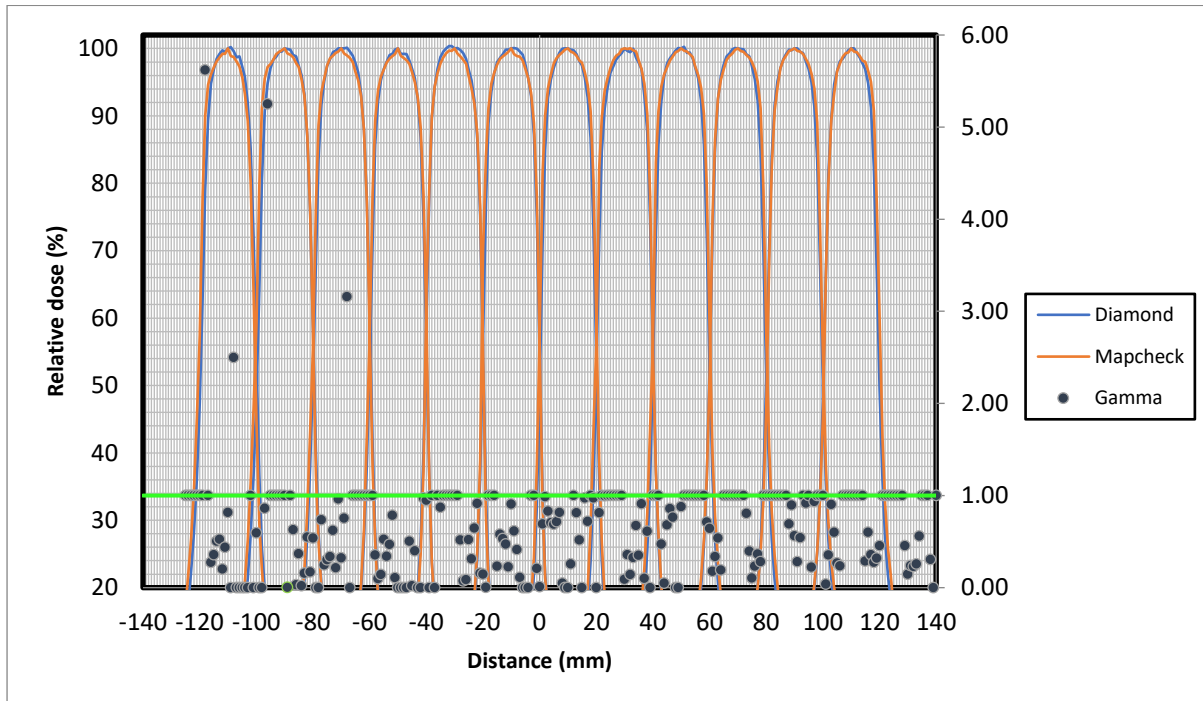

Figure57: MLC picket fence comparison between microDiamond and Mapcheck detectors that shows a very good agreement. especially around the 50% relative dose values.

For further validation of Mapcheck2™ profile measurements, measurement with an additional device (2D-Array 1000SRS™) was done. Due to the high spatial resolution of this array, only 2 shifts were required to match a microDiamond detector (TM60019) scan within the specified criteria as shown in Figures 58 and 59. 1000SRS™ profiles showed a shallower penumbra of 0.15 mm ± 0.10 mm compared to microDiamond which could be due to the size of the ionization chambers causing a dose volume averaging effect. Recently the 1000SRS™ array was validated, a device specifically designed for stereotactic small field measurements [71]. The authors also compared profiles from this array with the microDiamond detector (TM60019) and reported a close match, with γ index pass rates of 91.9% (x-axis) and 97.0% (y-axis) for a 0.5% DD / 0.5 mm DTA criterion.

Figures 60 and 61 show the comparison between Mapcheck2™ and 1000SRS™. Mapcheck2™ has a sharper (again probably due to detector design) but comparable penumbra. Further, this comparison shows above tolerance results outside the

radiation field in the crossplane direction and on one side of the inplane profile. These results are equivalent to earlier comparisons between microDiamond and Mapcheck and confirms the observations on Mapcheck's response to radiation.

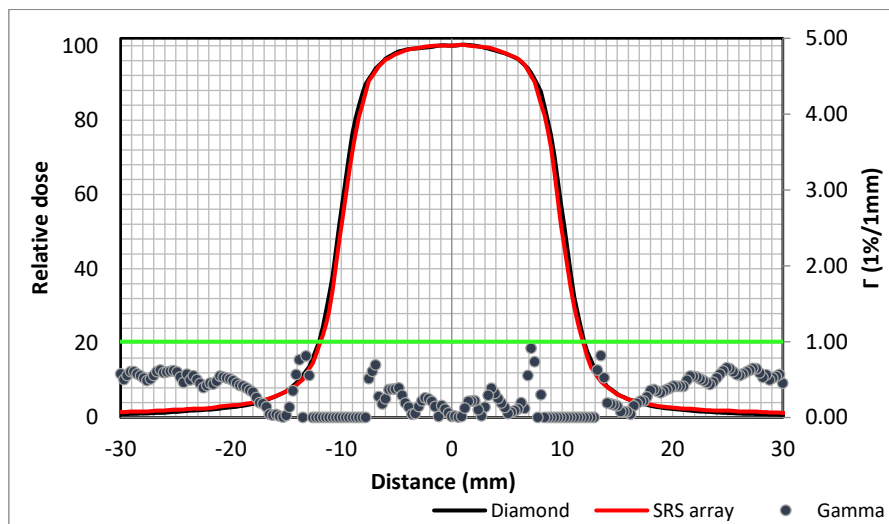*Figure 58: Comparison between a microDiamond and SRS CAX profile: 2 × 2 cm² @ 10 cm depth crossplane, showing a very good agreement between the detectors as γ values are < 1*
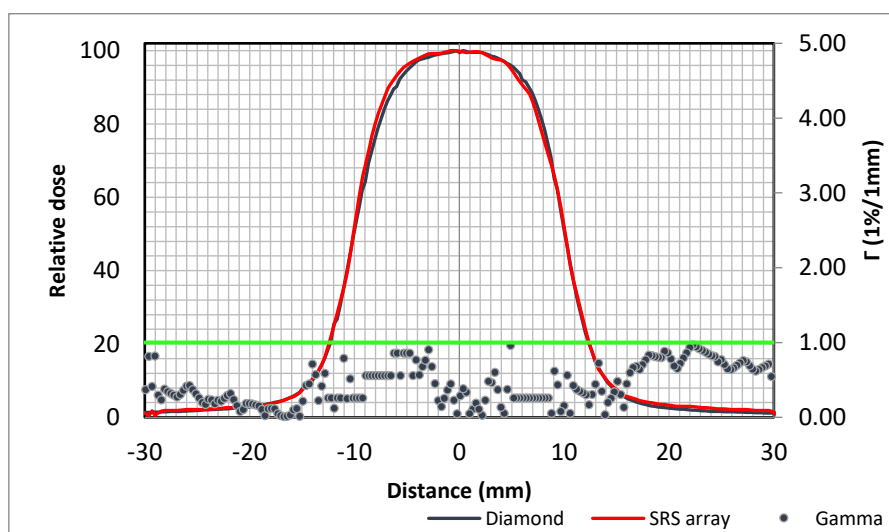

Figure 59: Comparison between a microDiamond and SRS CAX profile: 2 × 2 cm² @ 10 cm depth inplane, showing a very good agreement between the detectors as all γ values are < 1
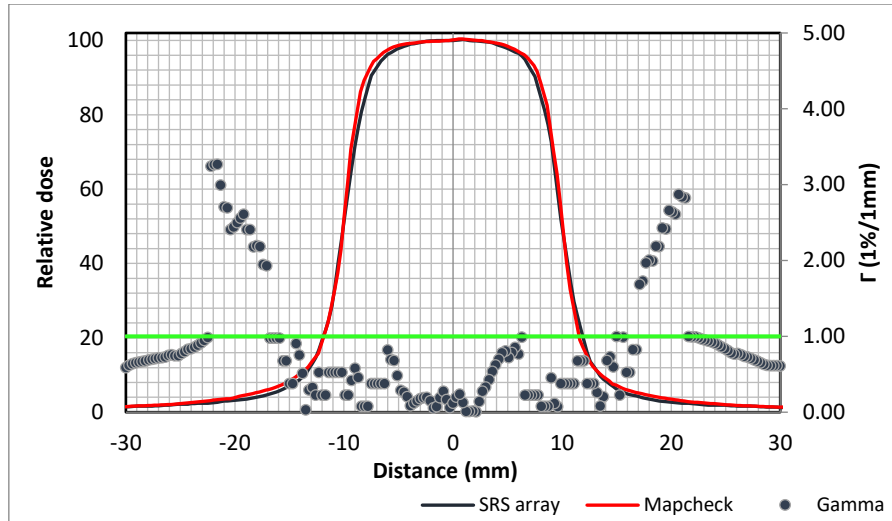
Figure60: Comparison between a SRS and Mapcheck CAX profile: 2 × 2 cm² @ 10cm depth in the crossplane direction. Out of field discrepancies are present, similar to Mapcheck and microDiamond comparisons.
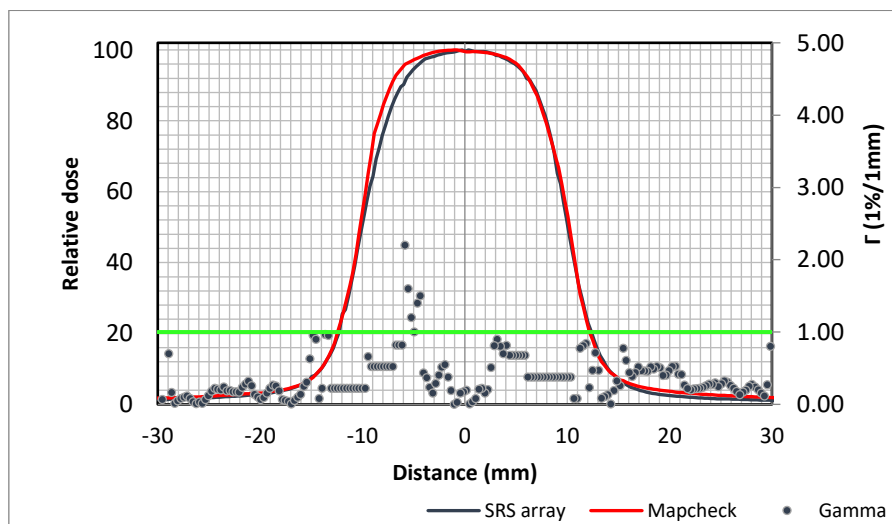
Figure61: Comparison between a SRS and Mapcheck CAX profile: 2 × 2 cm² @ 10cm depth in the inplane direction. The profiles agree fairly well, but the same trend as before is visible: Mapcheck overresponding on one-side of the profile. This effect is larger than before as the SRS has an even shallower penumbra compared to microDiamond.

Improving the measurement resolution in the crossplane or MLC direction could be beneficial for IMRT dose measurements. Figure 62 highlights the effect, in comparing a measurement of a standard IMRT field with multiple exposure of the same field, of increasing measurement resolution. However, the effect of increased resolution on clinically relevant γ pass rates (criteria 2 to 3% DD and DTA 2 to 3 mm) will be discussed later.

Figure62: Comparison between an IMRT field measured with a single Mapcheck exposure without any interpolation and multiple exposures in 1mm increments combined. The top pane of the figure shows CAX profiles of the singl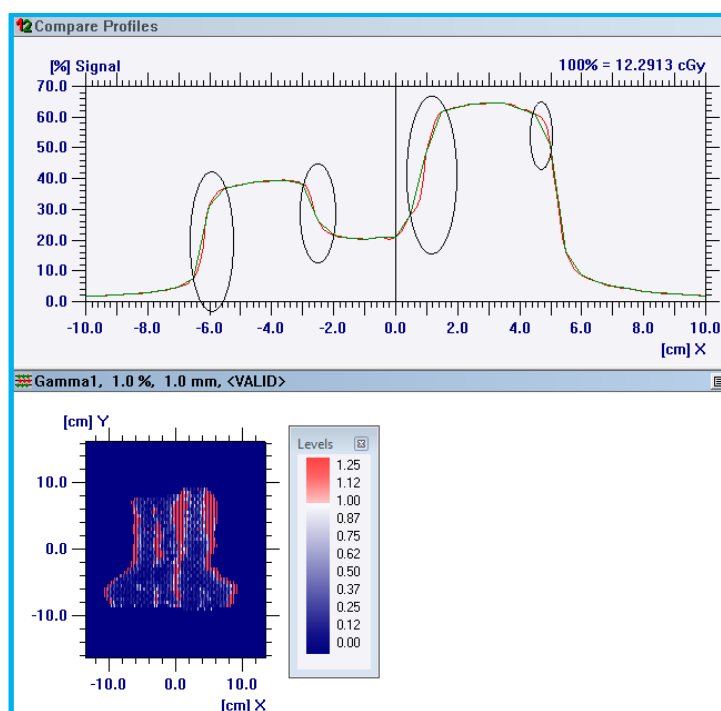e (green line) and multiple (red line) exposures. The effect of removing interpolation points are highlighted in the profile penumbra regions and is also visible in the γ analysis results (bottom pane). It should be noted that the green profile is a non-interpolated profile as interpolation points were removed and will display a different profile had these been included.

The dosimetric accomplishment of manually increasing array measurement resolution may be weakened by acquisition and processing workload. Therefore, workload reduction was investigated by comparing profiles from less array shifts example only 2 shifts, 4 shifts, etc. (figure 63). Only a 20% reduction in the number of shifts, in other words combining 8 Mapcheck2^{TM} incremental shift measurements, gave profile comparison results within a γ value of 1 compared to a microDiamond detector profile. Six measurements gave slightly out of tolerance results, with a 1.2 mm and 1.2% dose difference at a particular point in the penumbra region. Therefore, a significant reduction in the number of incremental shifts required to produce a profile comparable to the microDiamond detector could not be determined and 10 measurements yielded optimal results.

Apart from this amount of measurements being time consuming and non-practical in a clinical environment, it is believed the benefit of the method for model commissioning, especially quantifying differences between Linacs, justifies the means.
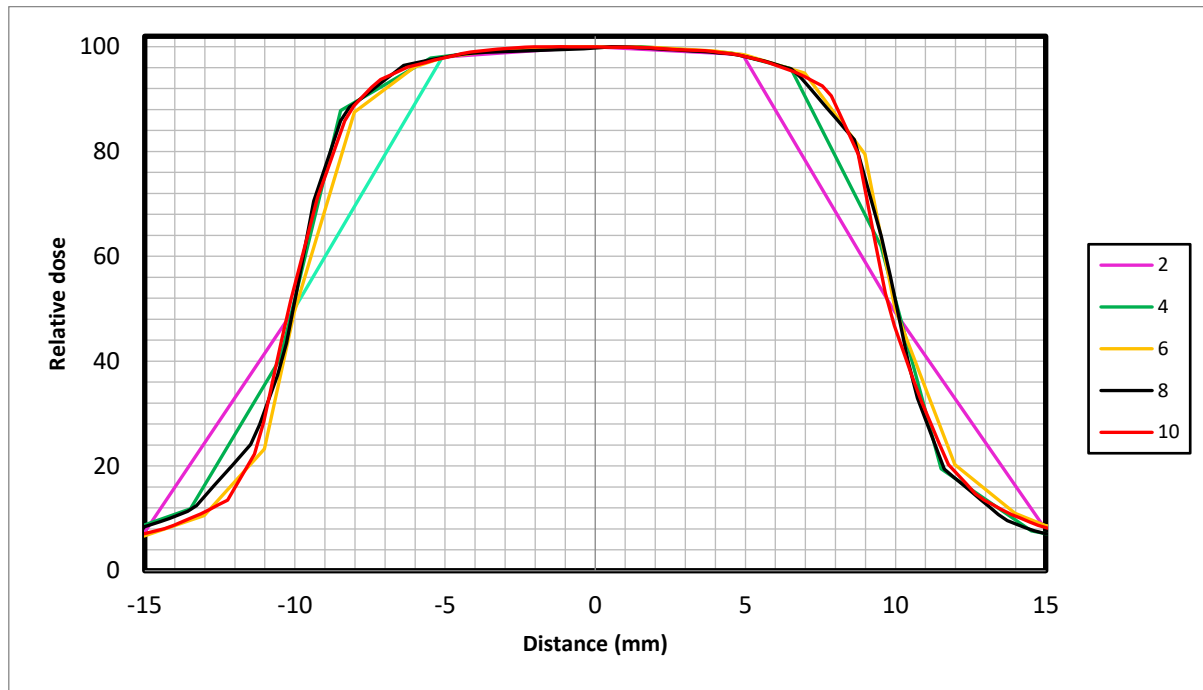

Figure63: Effect of number of Mapcheck measurements on profile penumbra. A significant reduction in the number of shifts required to produce a profile comparable to microDiamond measurements could not be shown. 8 shifts (black line) also yielded γ values ≤ 1 (1mm/1% criterion)

5.3   Linac Equivalence Results (array-based measurements)

5.3.1)   Multileaf Collimators

5.3.1.1)   MLC error quantification

Figures 64 shows a visual presentation of the complete MLC picket fence measured on each Linac respectively, typically seen when exposing a film with this beam sequence. In the figure white strips present MLC overtravel e.g. x-axis position +10 cm of Linac 1 (figure 64(a)) and darker strips present MLC under travel e.g. x-axis position -2cm of Linac 4 (figure 64(d)). A picket fence picture with black and white strips, visible for all 5 Linacs to variable extent, indicates that a non-linearity in MLC positions across the radiation field exists. However, visual interpretation of MLC errors is not quantitative, and it is difficult to relate these to absolute MLC tolerance values. Furthermore, it does not distinguish between MLC banks. For example, it cannot be said that the overtravel on CAX of Linac 5 (figure 64(e)) is due to one specific MLC bank alone or due to a combination of overtravel from both banks.
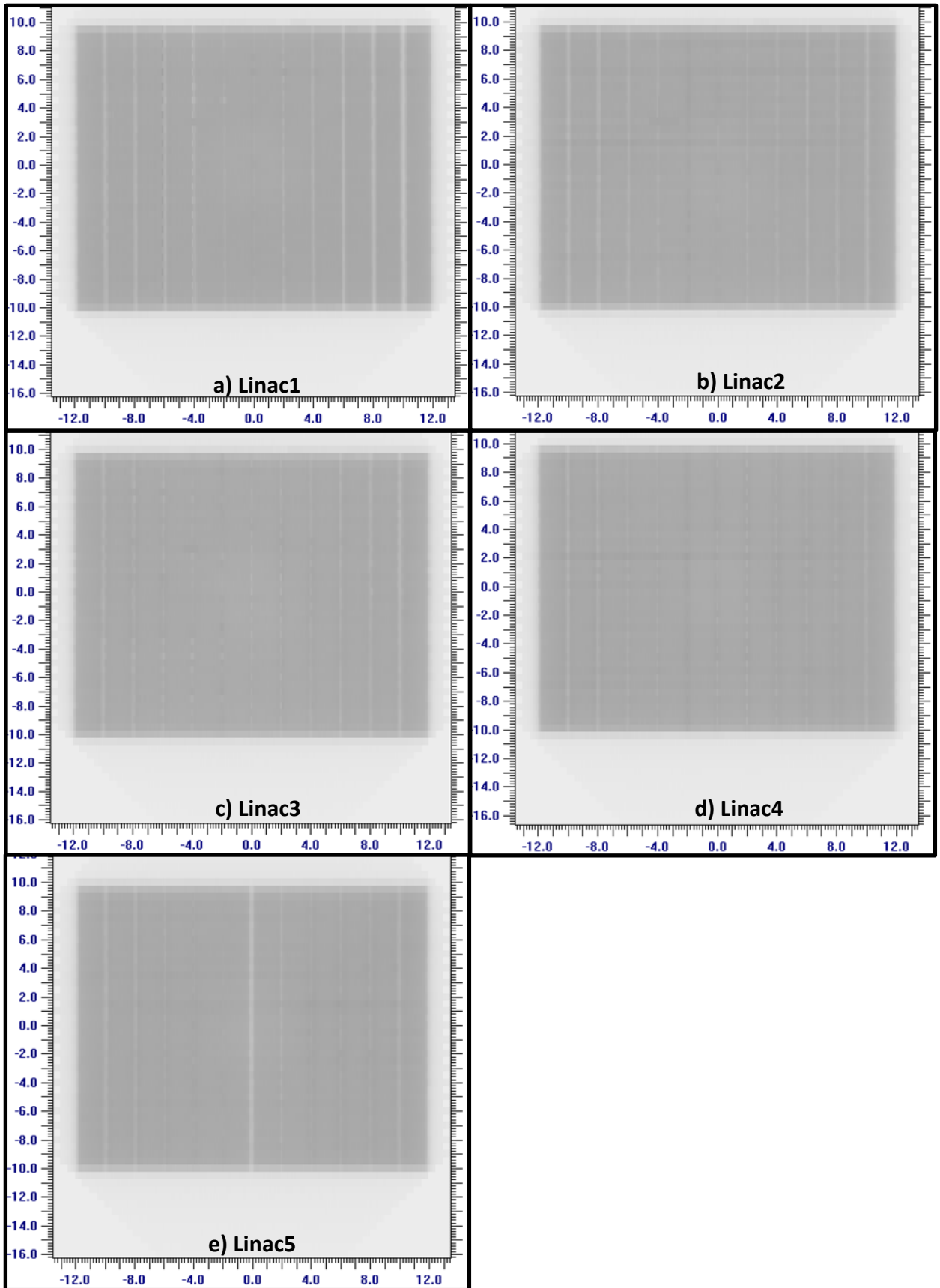
Figure64 (a-e): Picket fence patterns from (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac 4 and (e) Linac 5. X-axis show the MLC bank distance from CAX (0) in cm. White and dark stripes present MLC over-and under travel respectively and a close to perfect MLC bank(s) position(s) visible at position -2 cm of Linac5 (e).

Figures 65 to 75 shows individual MLC positioning errors form each MLC bank for all 5 Linacs at positions ranging between -10 cm and +10 cm in 2 cm intervals. Applicable to all these figures are:

- each bank evaluated consist of 40 MLCs, up to 10 cm from CAX in the non-MLC direction (inplane).
- a positive MLC error presents overtravel and a negative MLC error under travel, for each MLC bank individually.

MLC over travel errors of more than 1mm can be seen in figures: 65(a), 66(a), 70(e), 74(a) and 75(a). Most of these are from Linac 1 and quantifies the more prominent white strips of its picket fence (figure 64 (a)). At positions CAX -10 cm and CAX +10 cm, an over travel can be seen on all the Linacs, gradually decreasing to more accurate results around CAX for all the Linacs (figure 70). This trend appears to be systematic, with varying extent. From the results it can be said that:

- overtravel is largely attributed to a specific MLC bank only, for example figure 65(e) and figure 70(e).
- same direction errors can potentially propagate into larger errors and different direction errors can reduce the overall MLC error at that position, for example a smaller (illustrated in figures 66(e) and figure 75(d)) or a larger (illustrated in figure 75(e)) MLC error.
- large variation within one MLC bank are seen, for example the difference between the minimum and maximum MLC error of the x1 bank of Linac 2 at stop position -10cm (figure 65(b)) was 1.05 mm. Similarly, for Linac 1 at position x1 = +4 cm (figure 72 (a)), MLC errors of numbers 1 and 38 was +0.5 mm and -0.5 mm respectively.
- for simplicity, reproducibility of MLC error measurements is not shown on these graphs. Measurements were repeated at least twice on different calendar days on each Linac (5 times for Linac 3) and the average MLC error values are shown. A consistent standard deviation of ± 0.2 mm was calculated throughout, on all the Linacs, which is in line with the reproducibility previously shown.
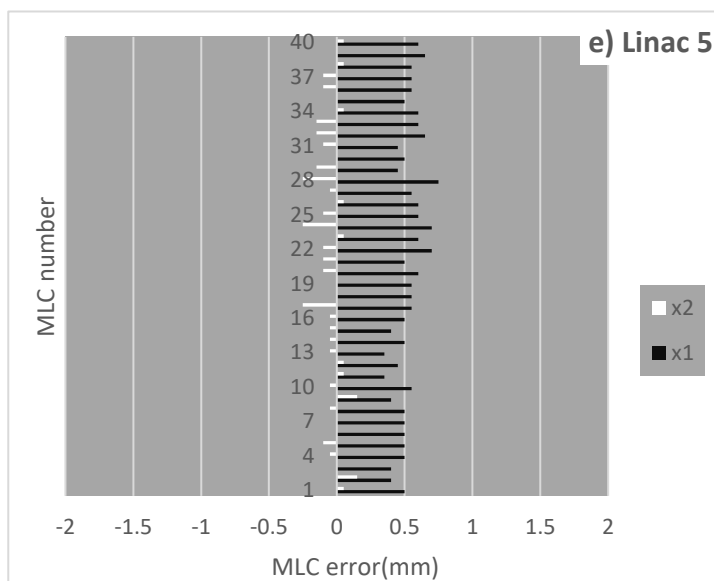
Figure 65 (a-e): MLC errors at set position -10 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. A positive MLC error presents overtravel and a negative MLC error under travel. Greater than 1 mm errors can be seen for a) Linac1 at this position, predominantly from the x1 bank. From e) Linac5, overtravel is from a single MLC bank, x1.
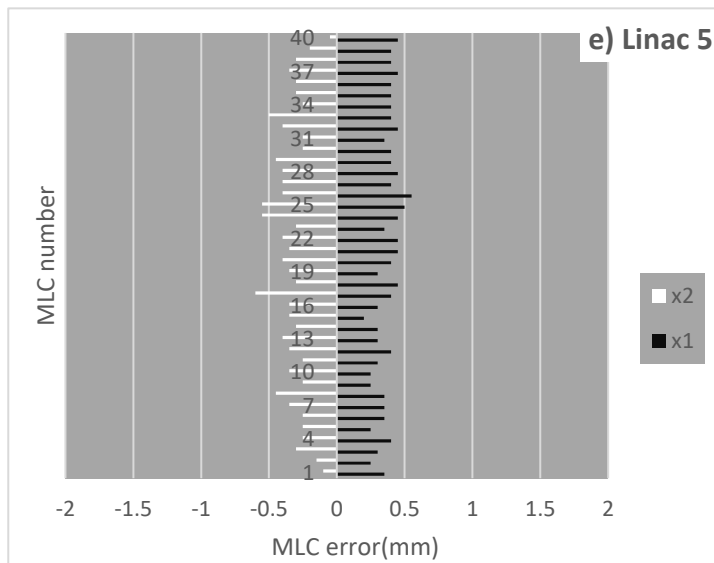
Figure 66 (a-e): MLC errors at set position -8 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. Large differences between same number MLC's between Linacs (>1 mm) are shown, for example, MLC number 1 of the x1 bank between Linacs 1 and 4.
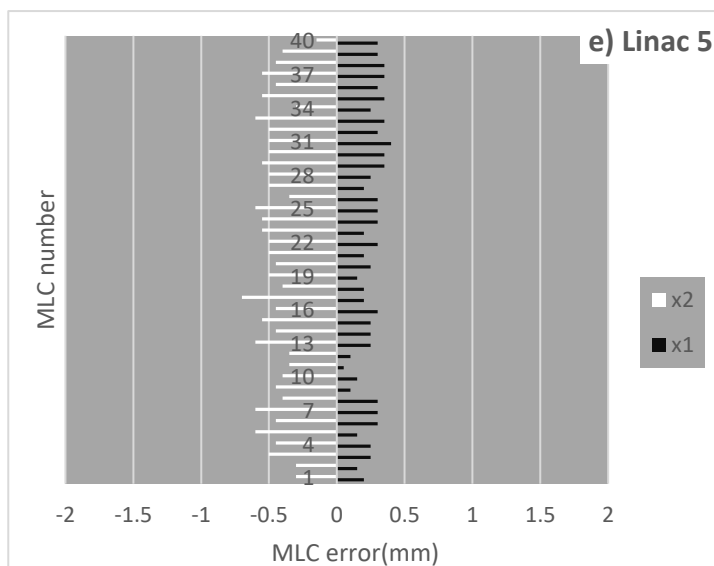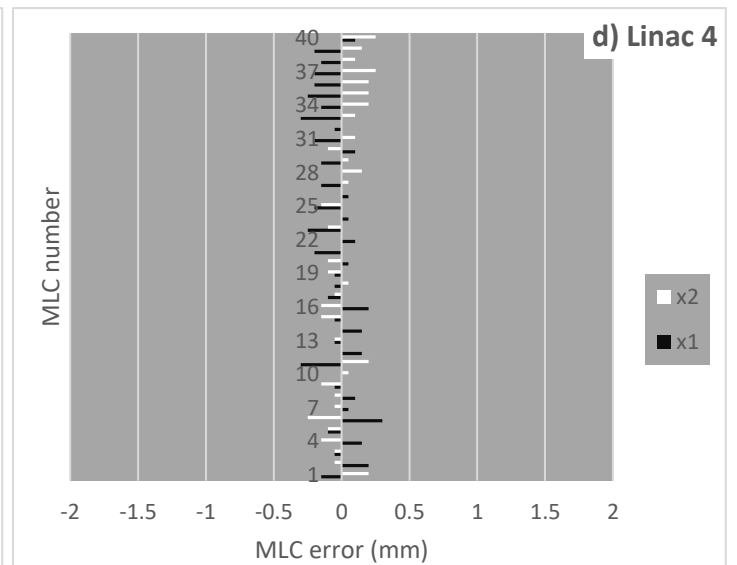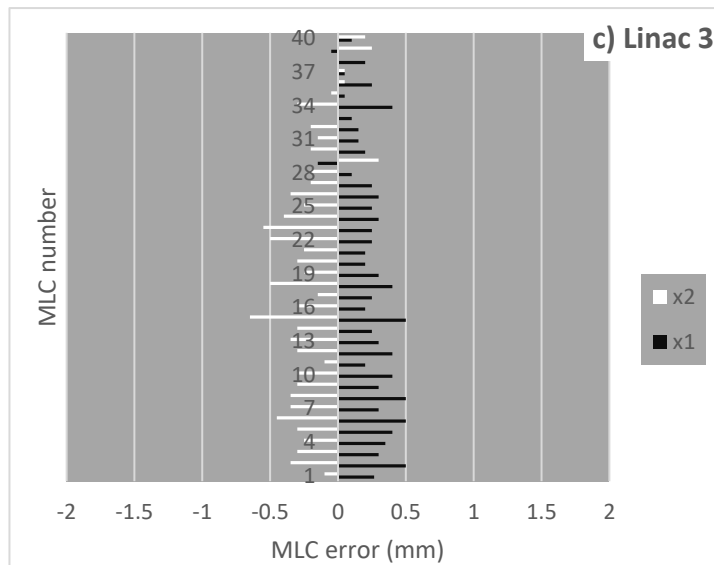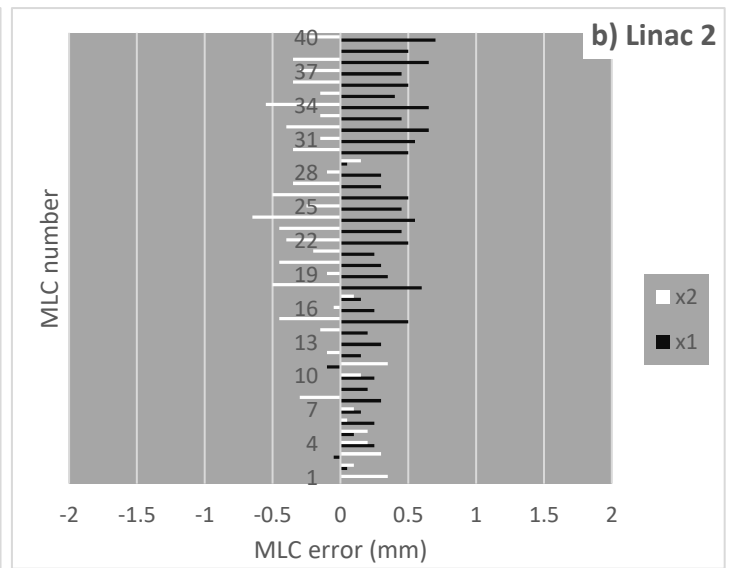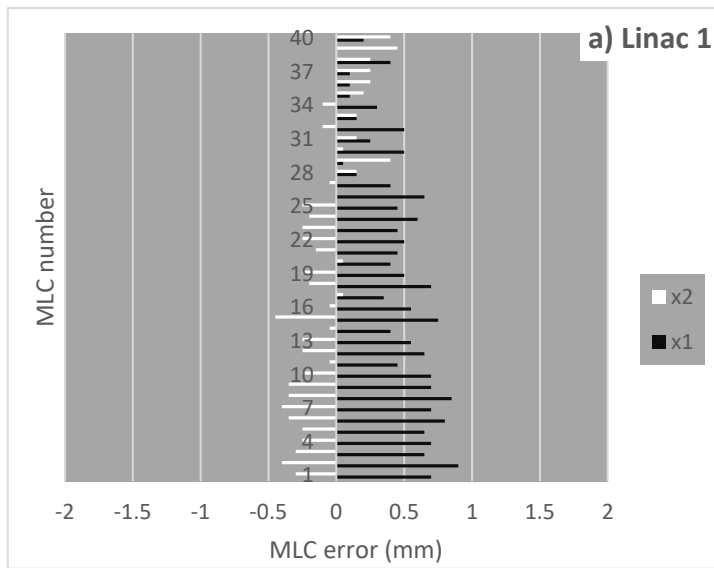
Figure 67 (a-e): MLC errors at set position -6 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5.
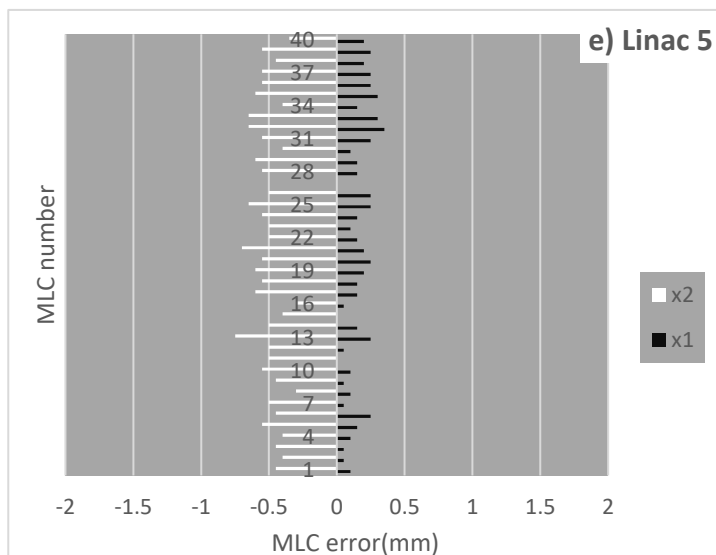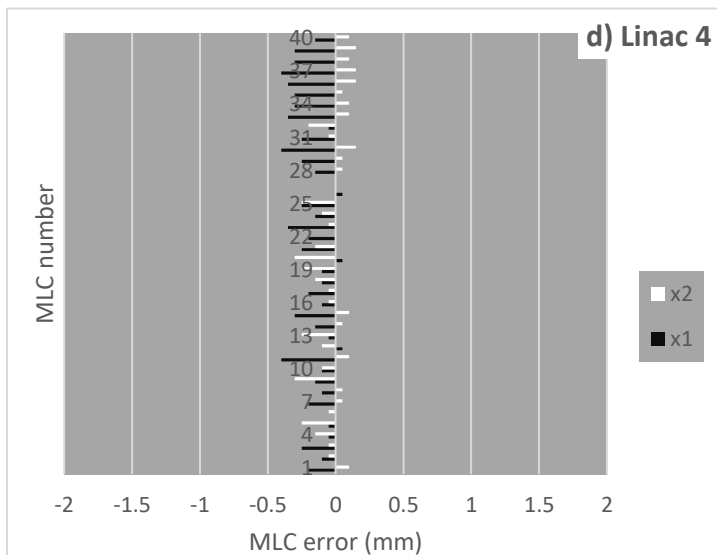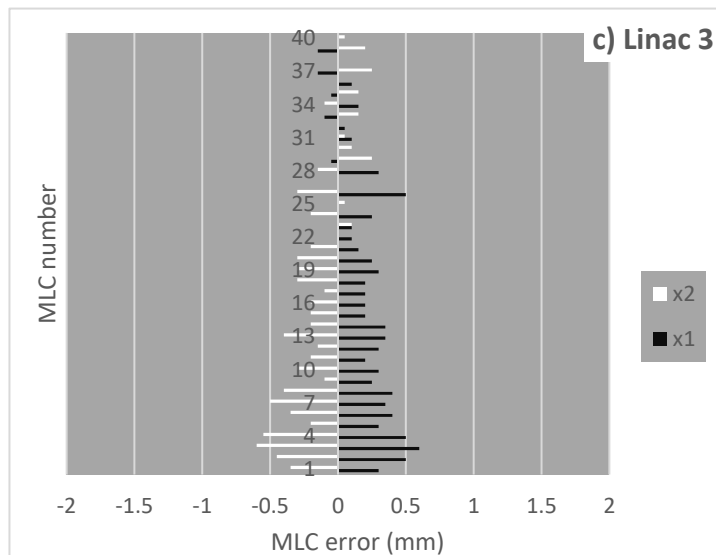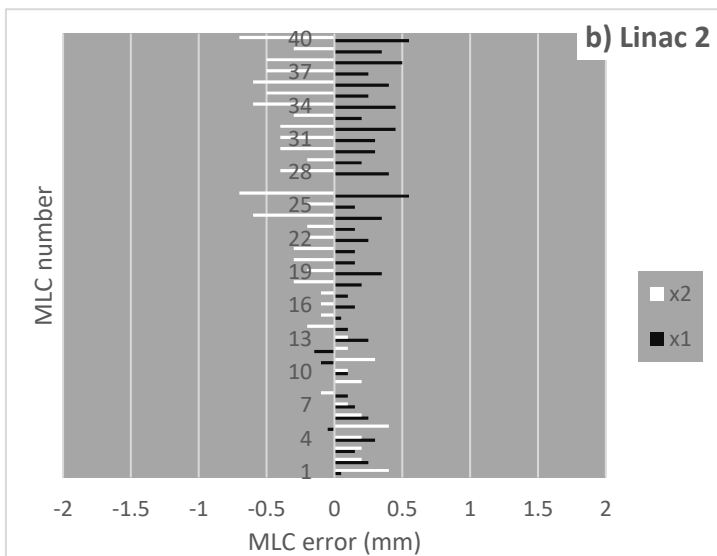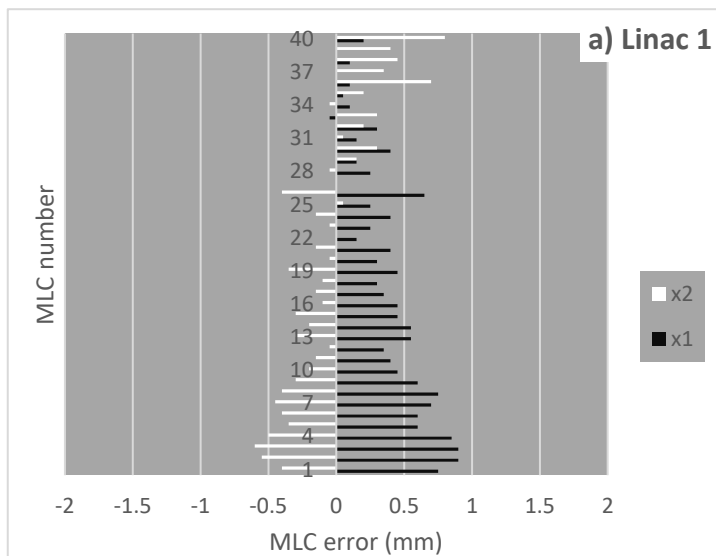
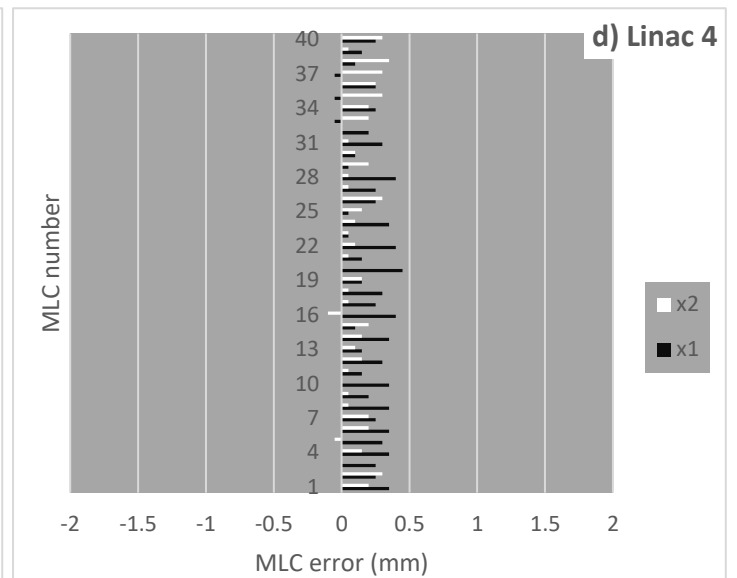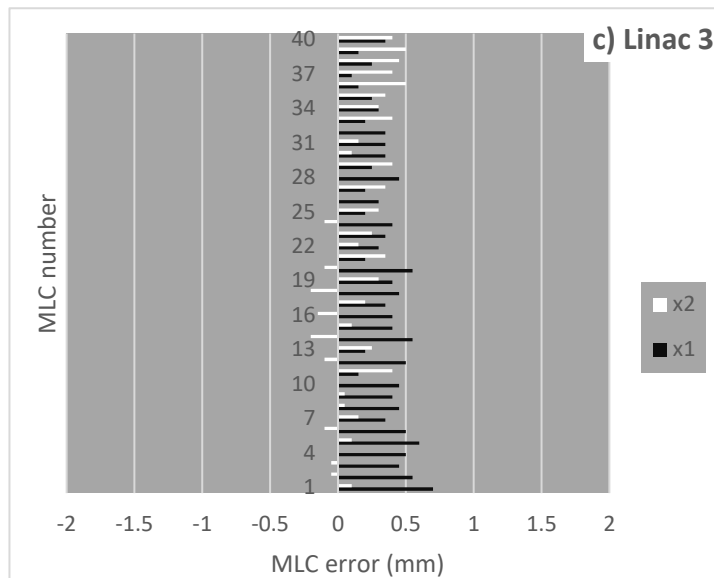Figure 68 (a-e): MLC errors at set position -4 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5.
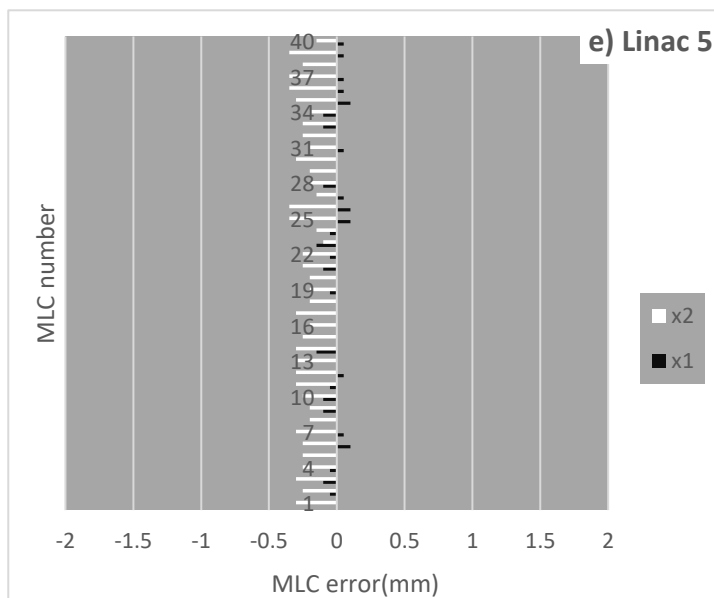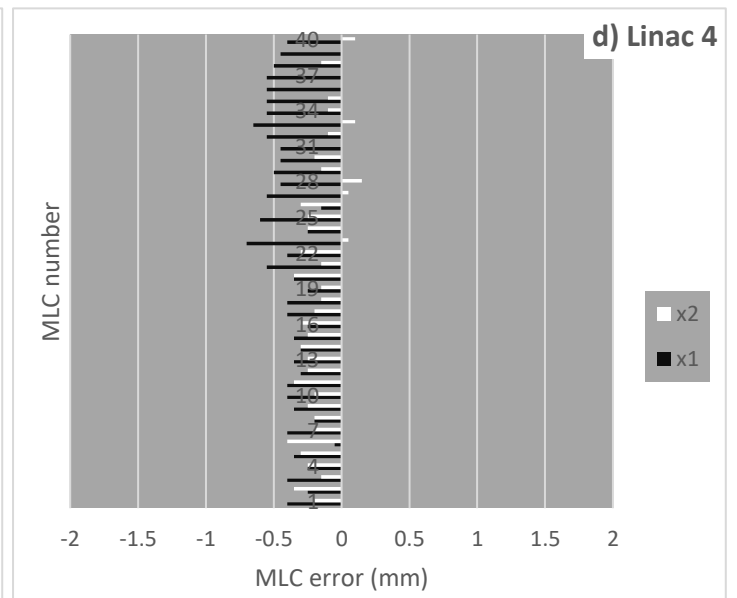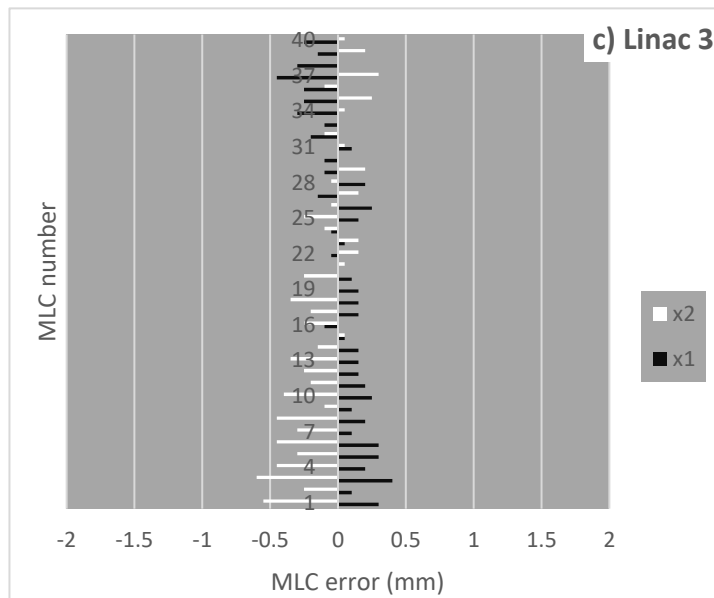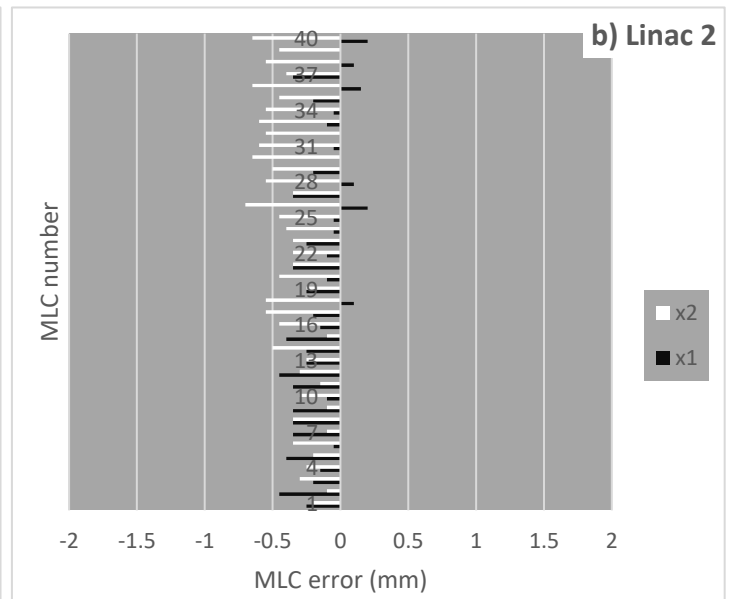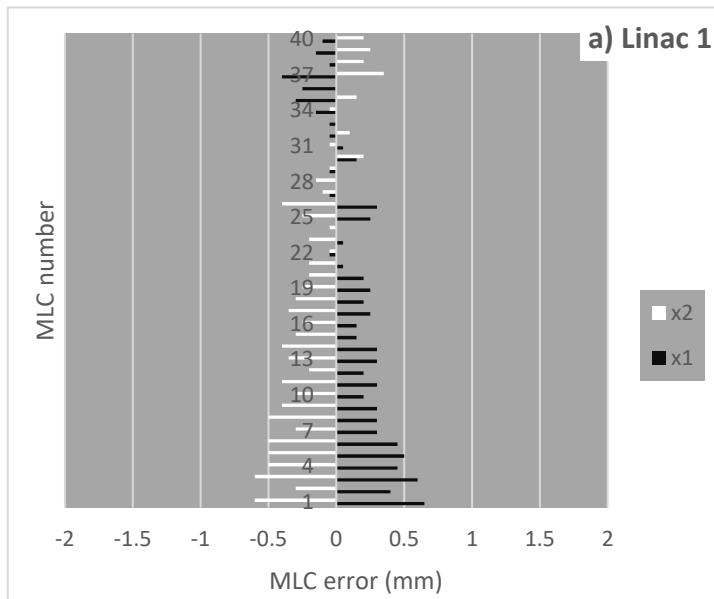
Figure 69 (a-e): MLC errors at set position -2 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5.
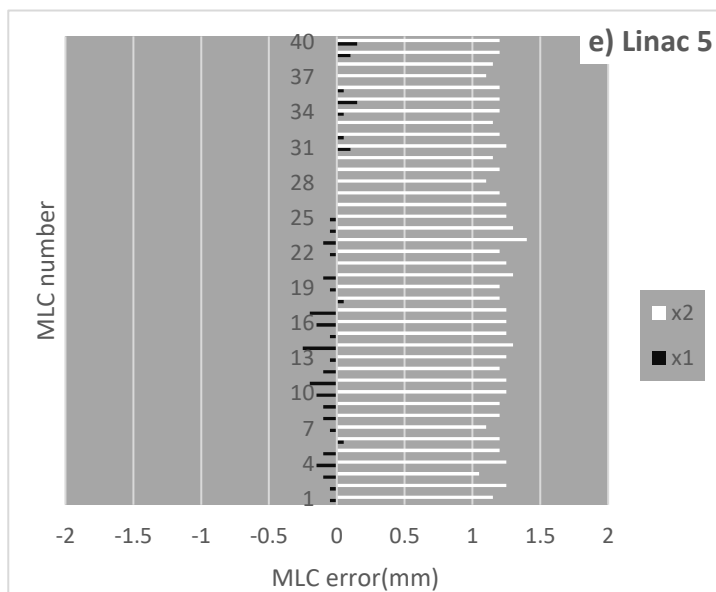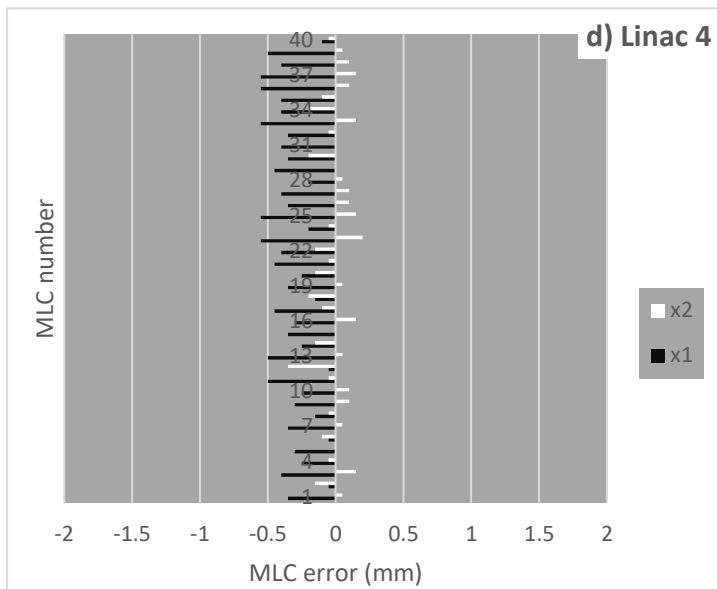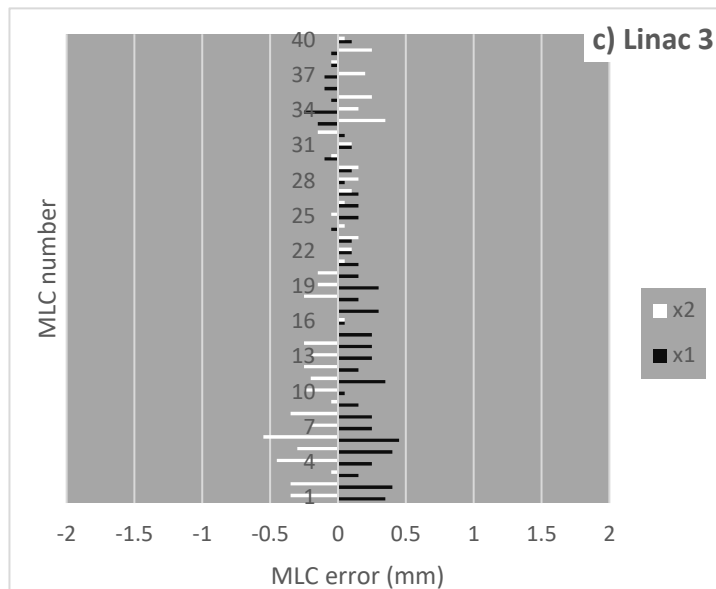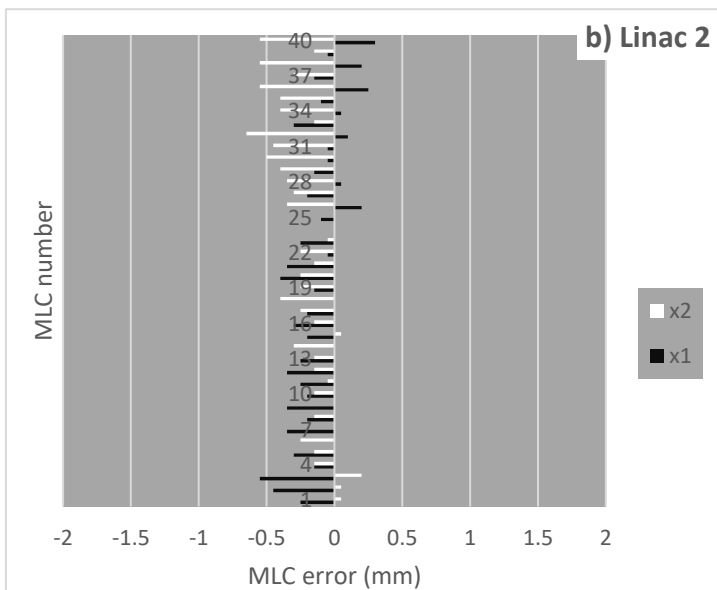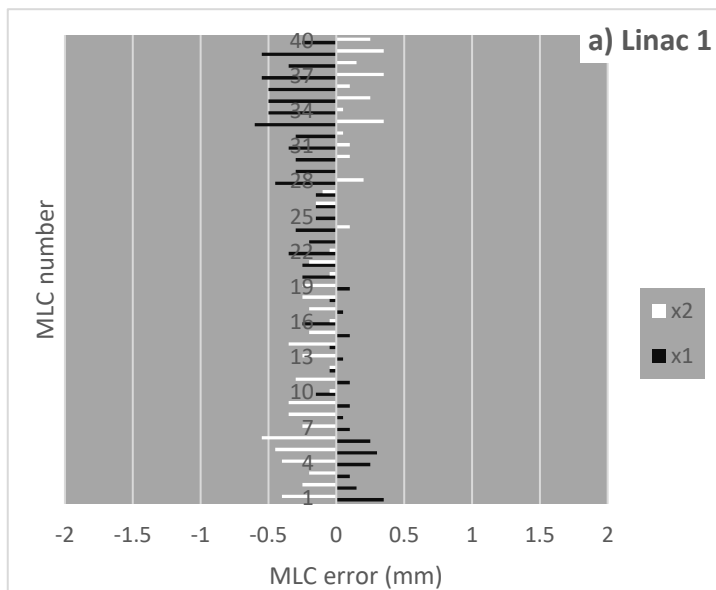
Figure 70 (a-e): MLC errors at set position CAX for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. x2 bank of Linac5 overtravel's by more than 1 mm, whereas MLC errors from the x1 bank at the same position are minimal.
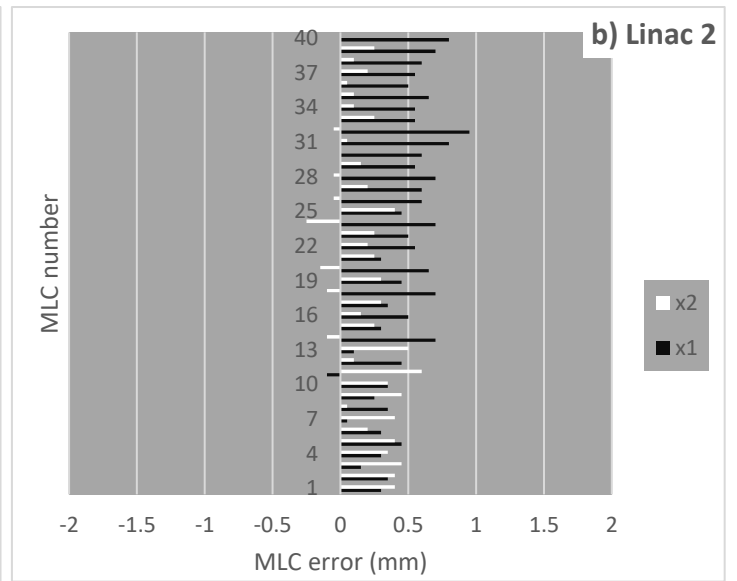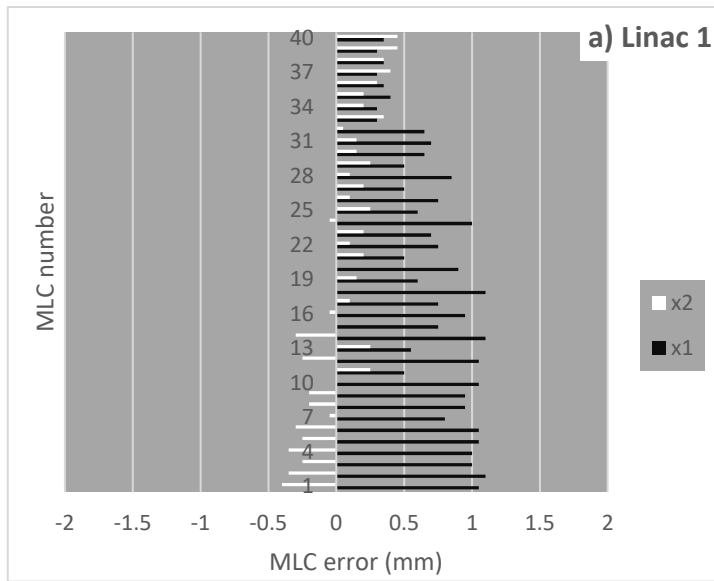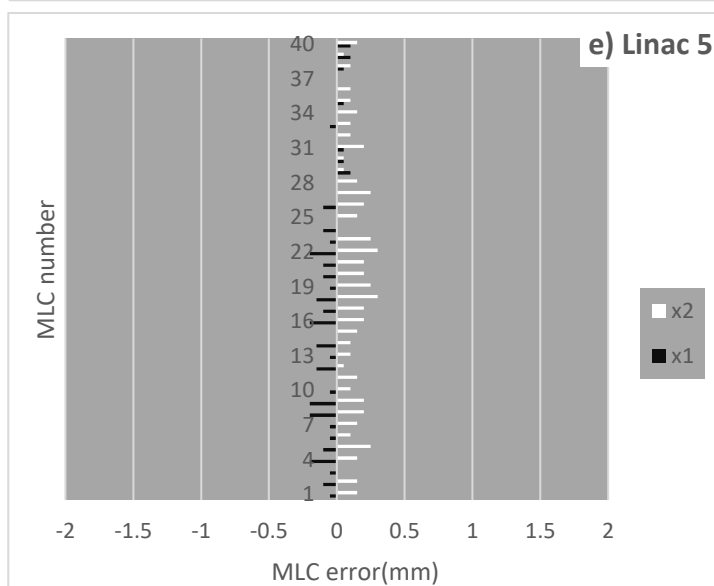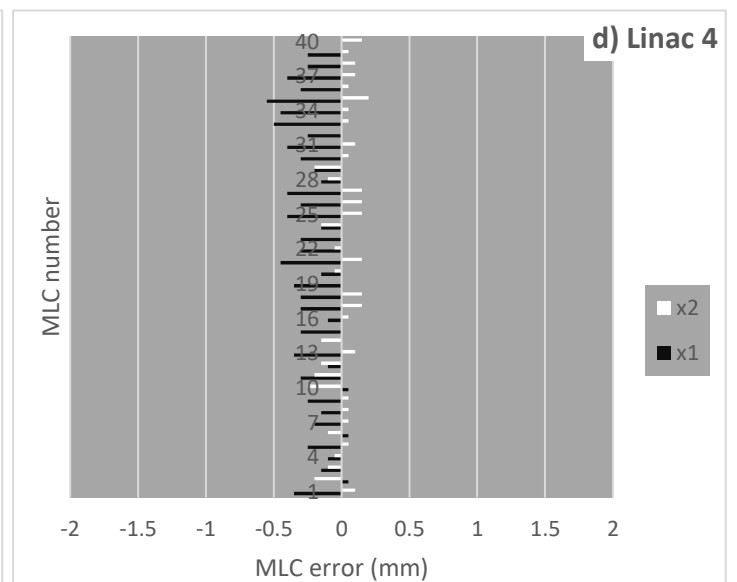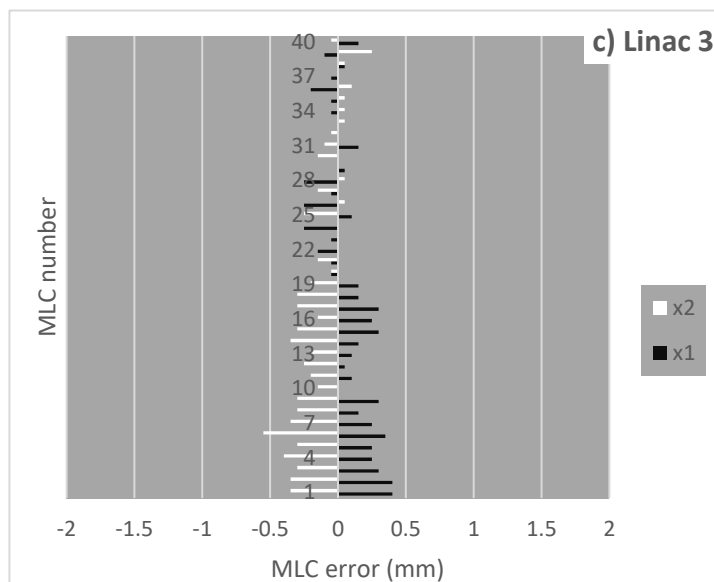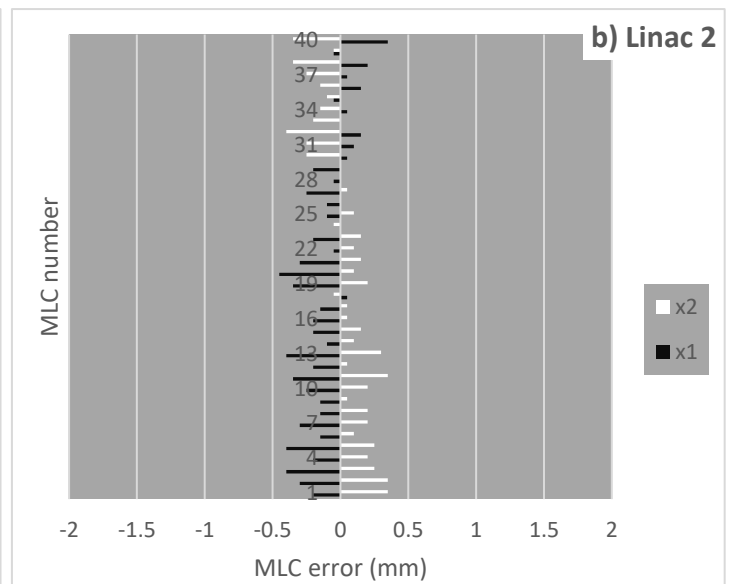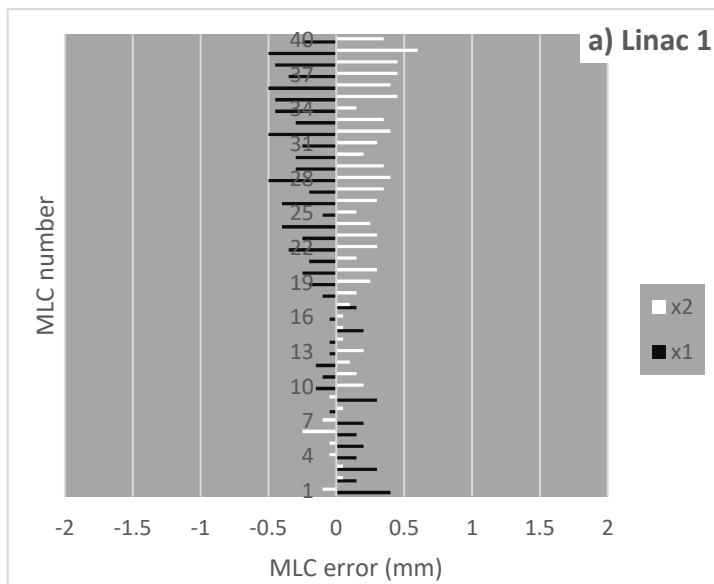
Figure 71 (a-e): MLC errors at set position +2 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. Overall most MLC errors are within 0.5 mm at this position and within 0.3 mm for Linac 5.
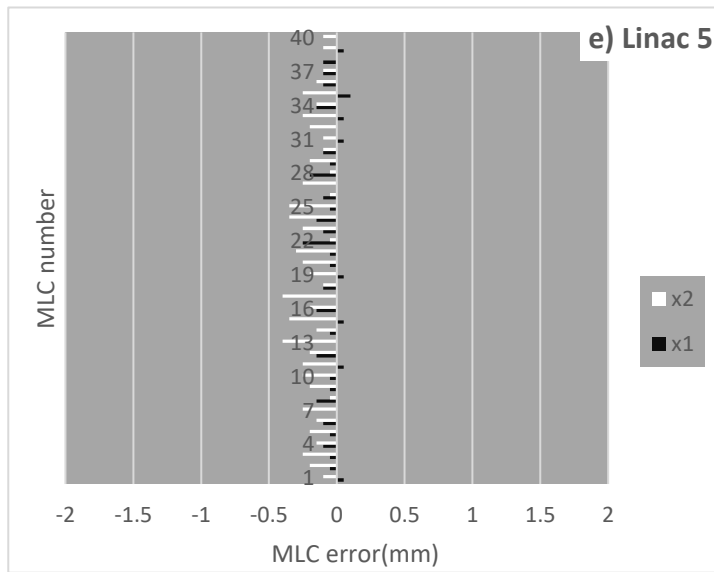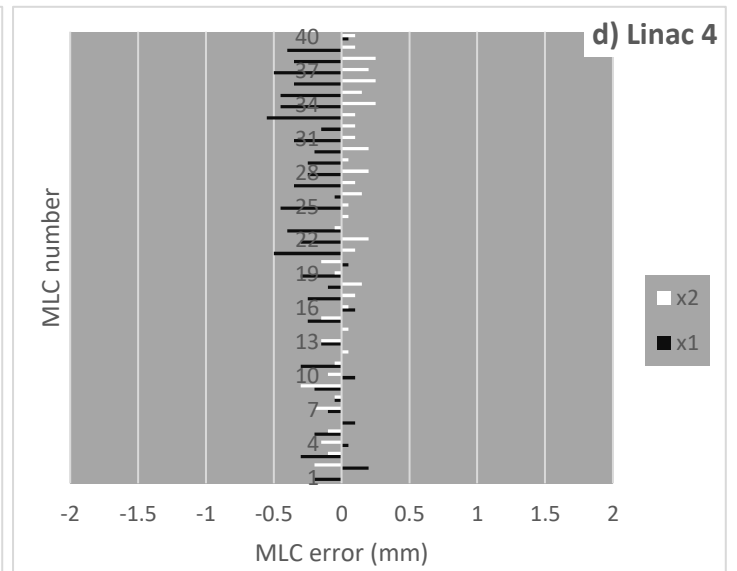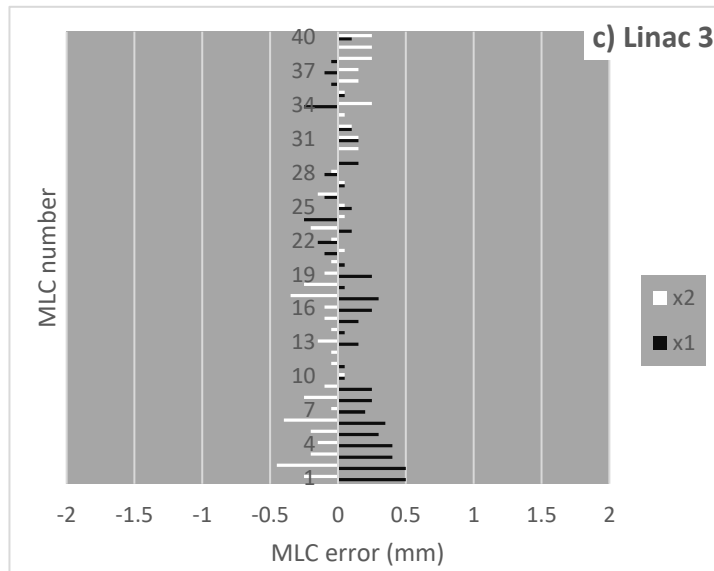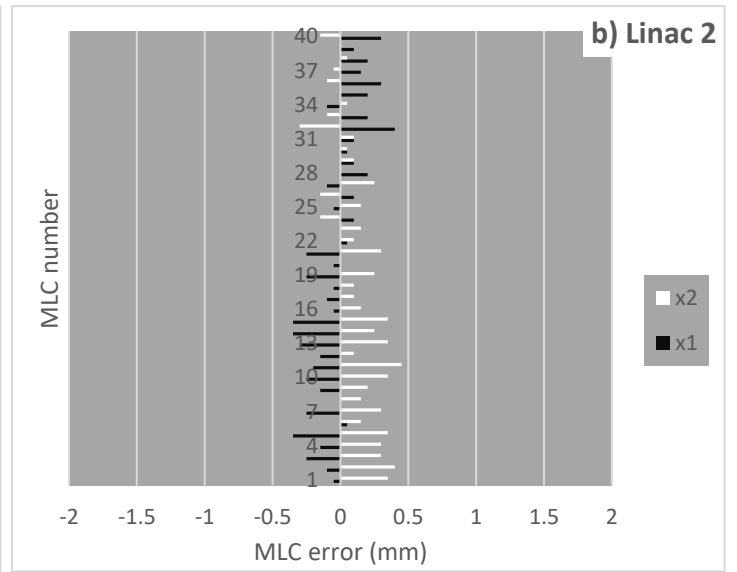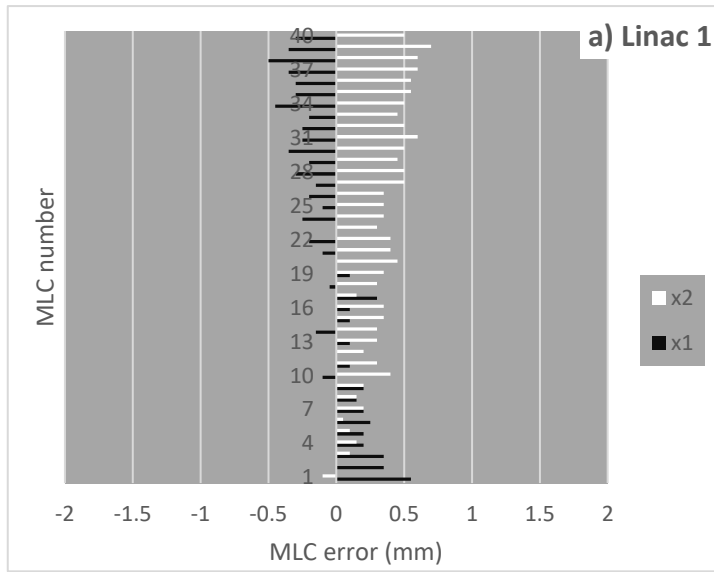
Figure 72 (a-e): MLC errors at set position +4 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. From a) Linac 1, a large deviation in MLC errors from the x1 bank is present as the difference between errors from MLC numbers 1 and 38 is +0.5 mm and -0.5 mm respectively.
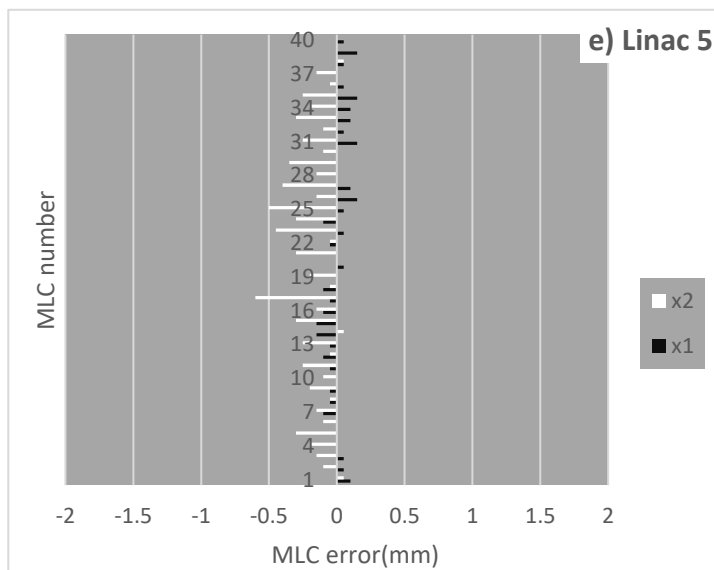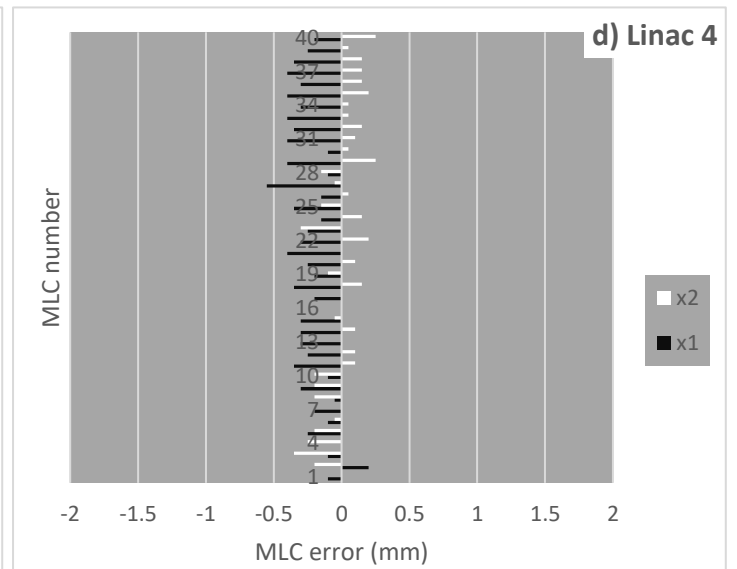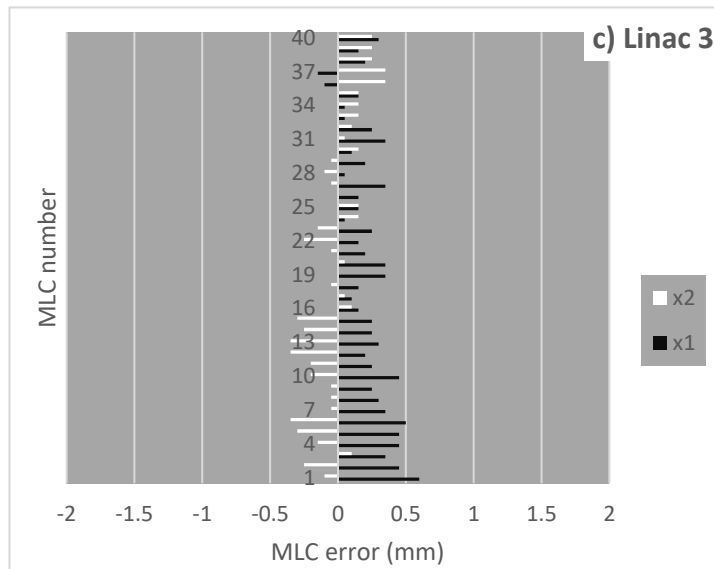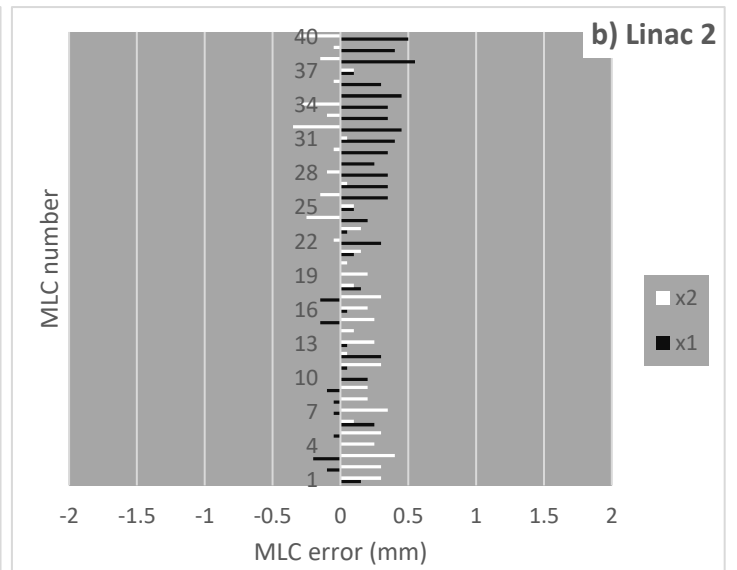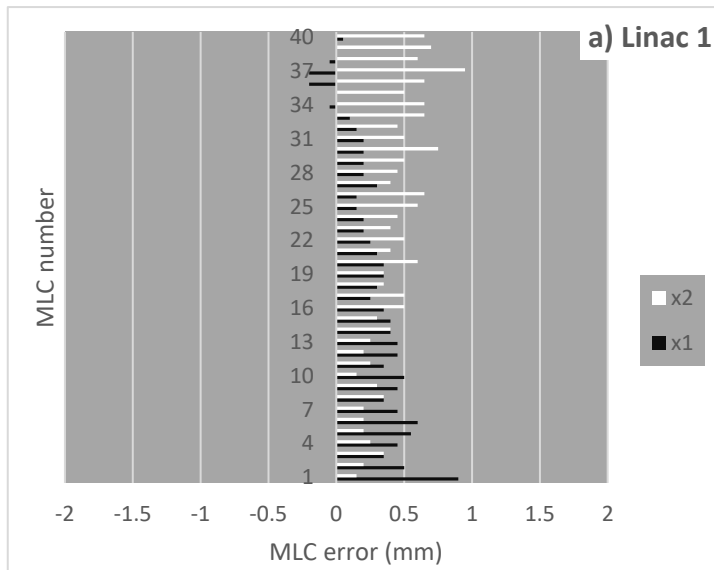
Figure 73 (a-e): MLC errors at set position +6 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5.
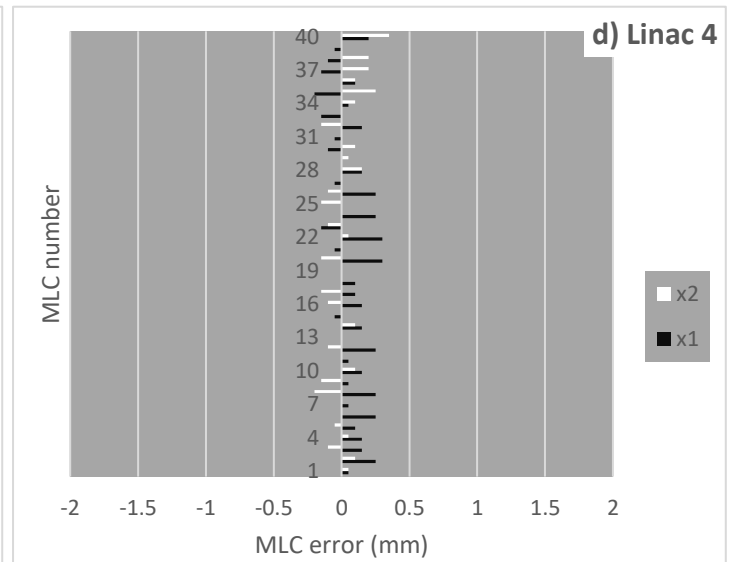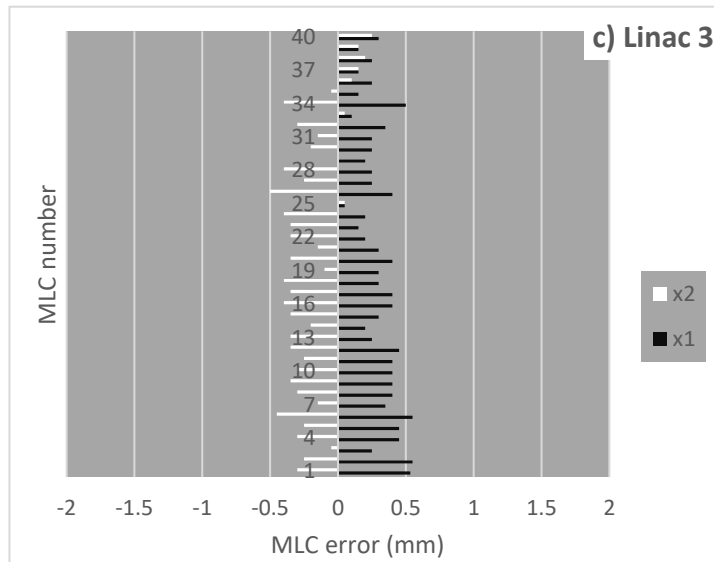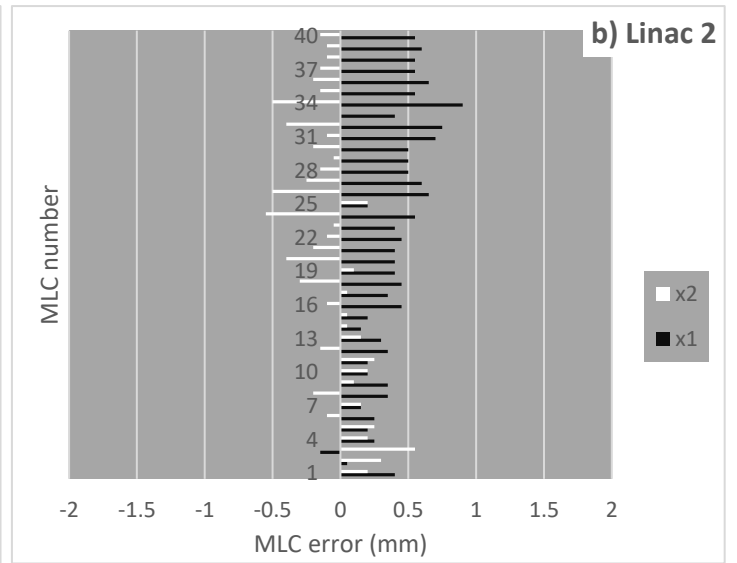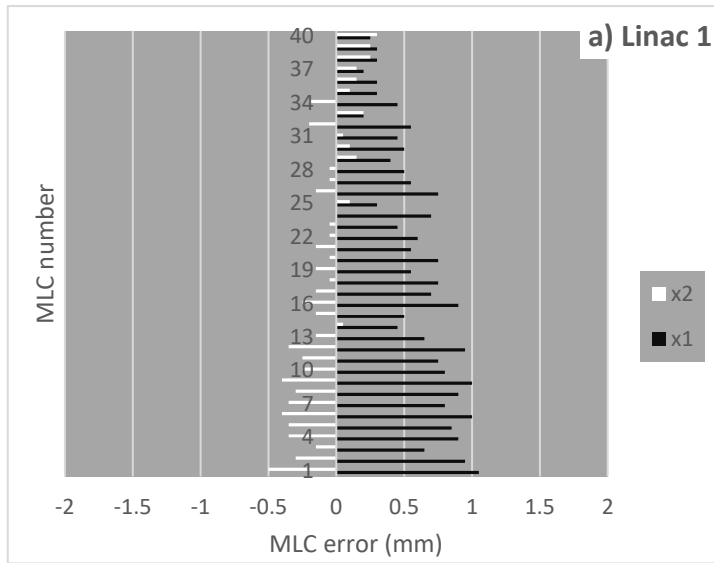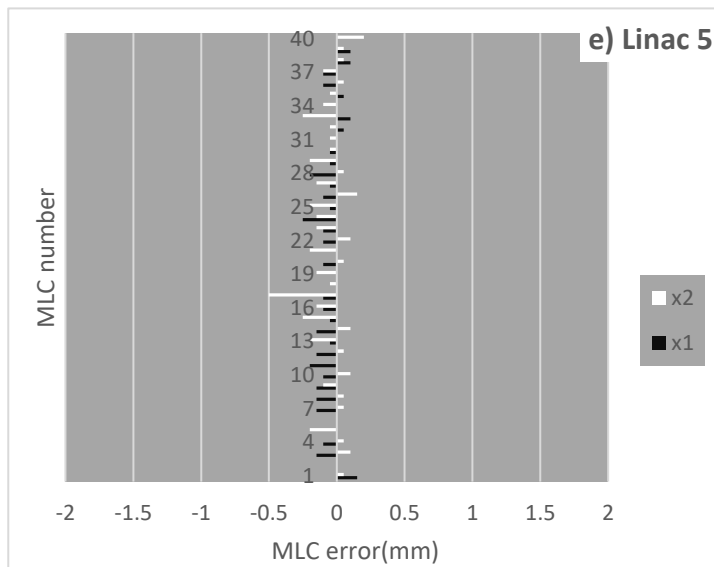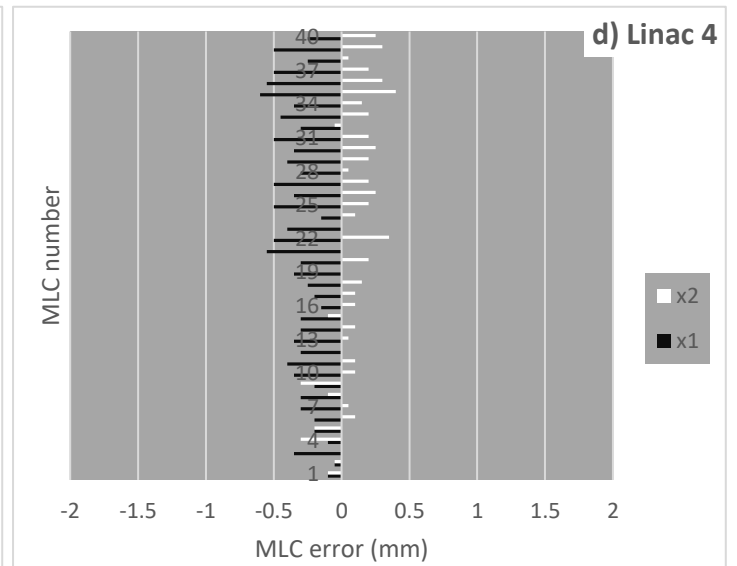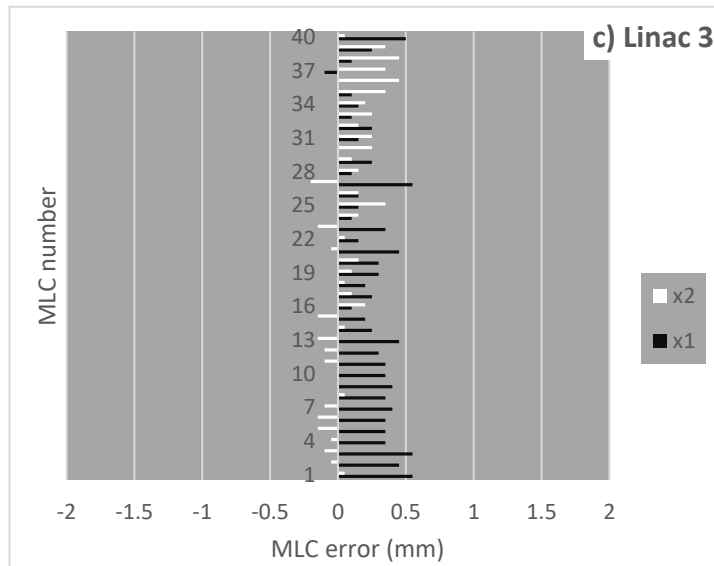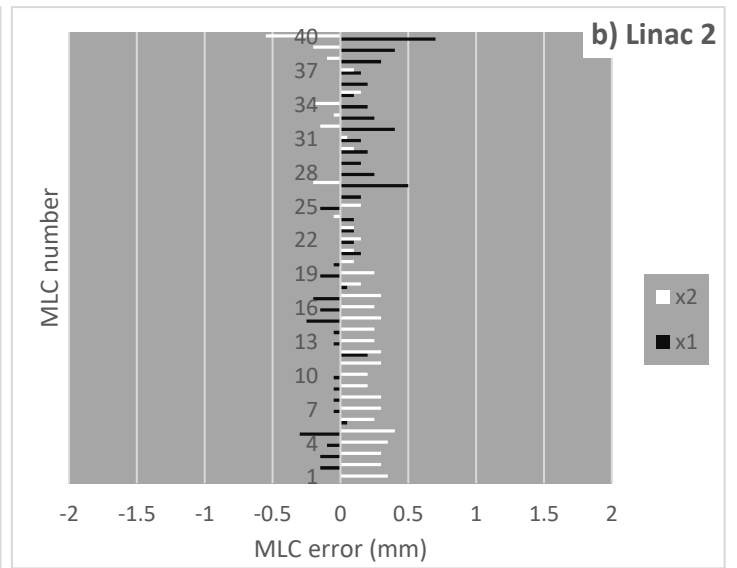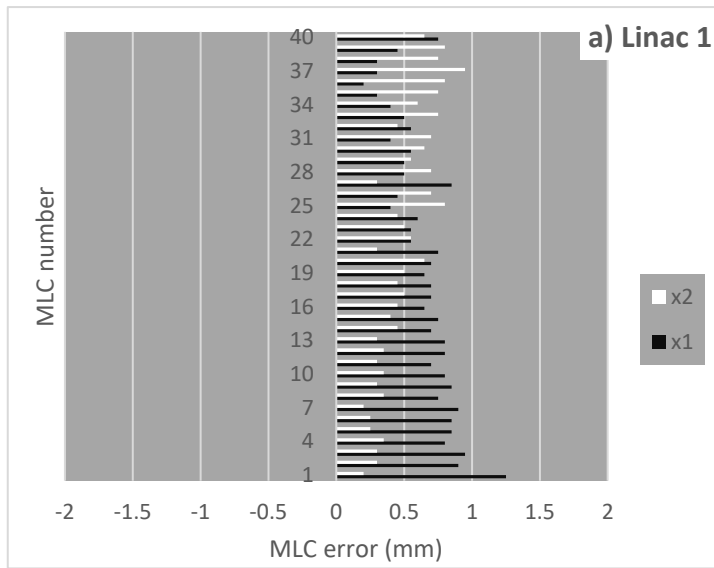
Figure 74 (a-e): MLC errors at set position +8 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5.
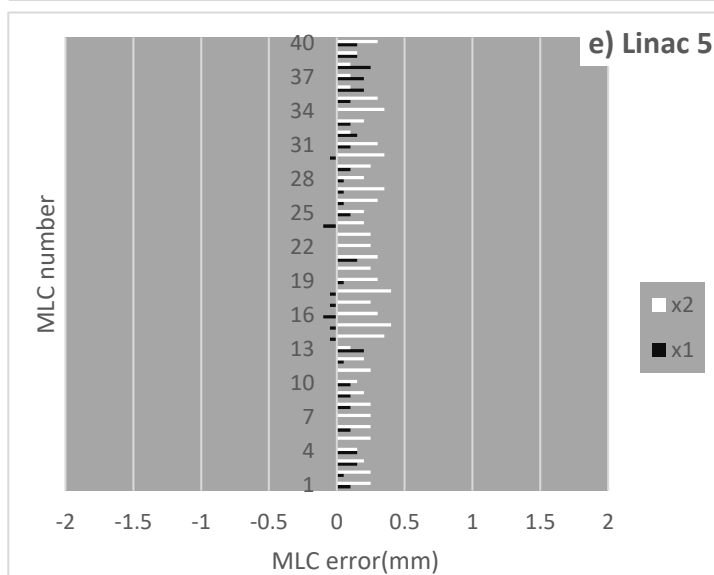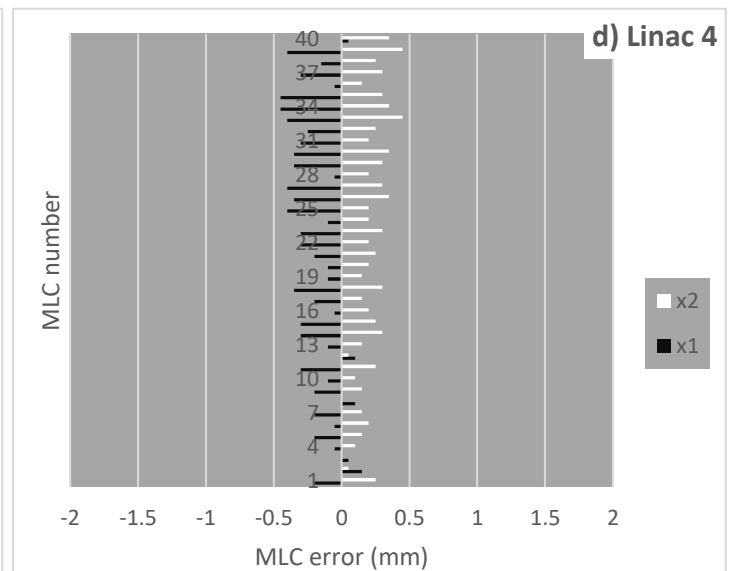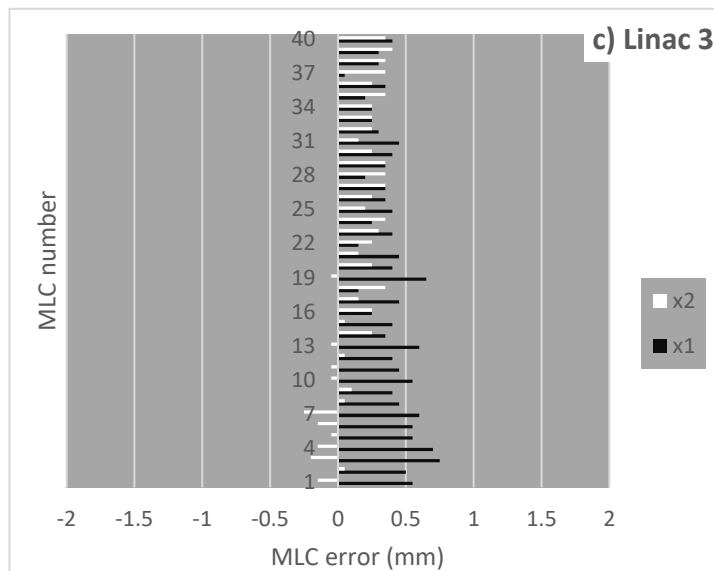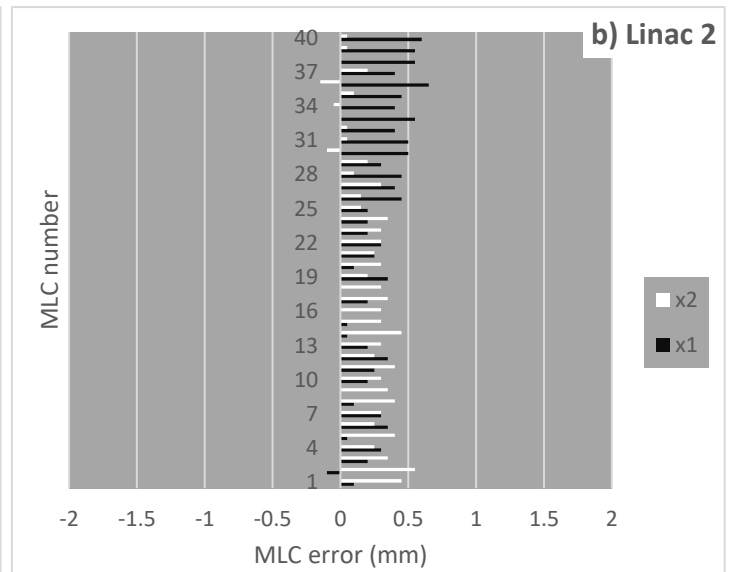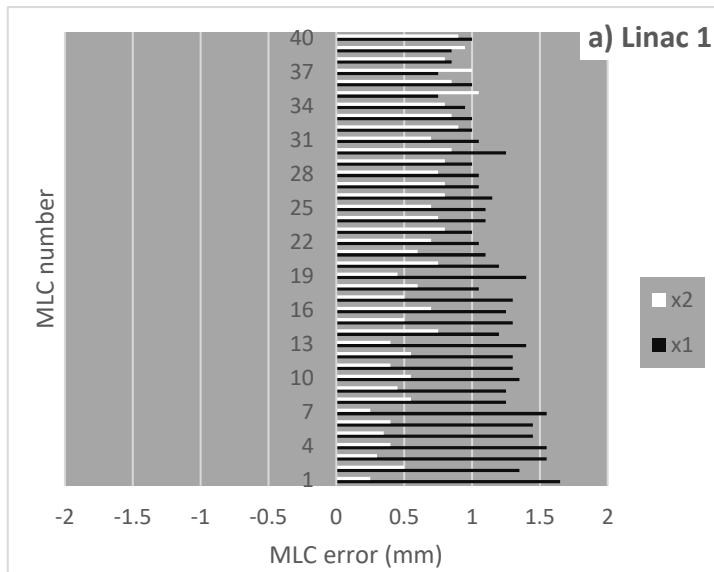
Figure 75 (a-e): MLC errors at set position +10 cm for (a) Linac1, (b) Linac2, (c) Linac3, (d) Linac4 and (e) Linac5. The largest MLC errors are from a) Linac 1 as both MLC banks are over travelling therefore a combined error up to 2mm at this position. On the contrary, from d) Linac 4, errors from opposing MLC's can potentially combine as smaller MLC errors (< 0.2 mm).

Figures 76 to 79 summarizes the median MLC errors ranging between -0.4 to 1.2 mm and -0.5 to 1.3 mm for the x1 and x2 banks respectively. These figures contain multiple boxplots for each Linac, alphabetically labelled, summarizing MLC errors from each bank respectively. The median MLC error is represented with an "x" marker and horizontal markers are used for Q1 and Q3, representing the 25th and 75th percentile values respectively. The upper and lower whisker was defined in section 5.2.2. Values above or below the respective whiskers are plotted as outliers.

From graph 76(a), Linac2 and Linac5 have approximately the same average MLC error of 0.50 and 0.53 mm respectively. However, the 25th and 75th percentile values are 0.34 mm and 0.61 mm for Linac2 and 0.50 mm and 0.60 mm for Linac5 respectively. This means at the given stop position of -10 cm, Linac2 has a larger variation in MLC errors of same bank/adjacent MLCs compared to Linac5, previously shown in figure 65 (b) and (e). Also, lower and upper whiskers are -0.10 and 0.95 mm for Linac2 and 0.35 and 0.75 mm for Linac2 respectively. On the same graph, 25th and 75th percentile values range between 0.50 mm and 0.95 mm for Linac1, previously shown in figure 65 (a), which is the biggest range for all the Linacs. The larger range in MLC errors can be seen at all the stop positions of this Linac, for both MLC banks. Linac5 has got the smallest range between the 25th and 75th percentile values, mostly within 0.1 mm at all the stop positions for both MLC banks, except for the x2 bank at the CAX stop position (previously shown in figure 70(e)). It can be said that Linac5 has got the best grouped MLCs or the least varying MLCs in terms of stop position.
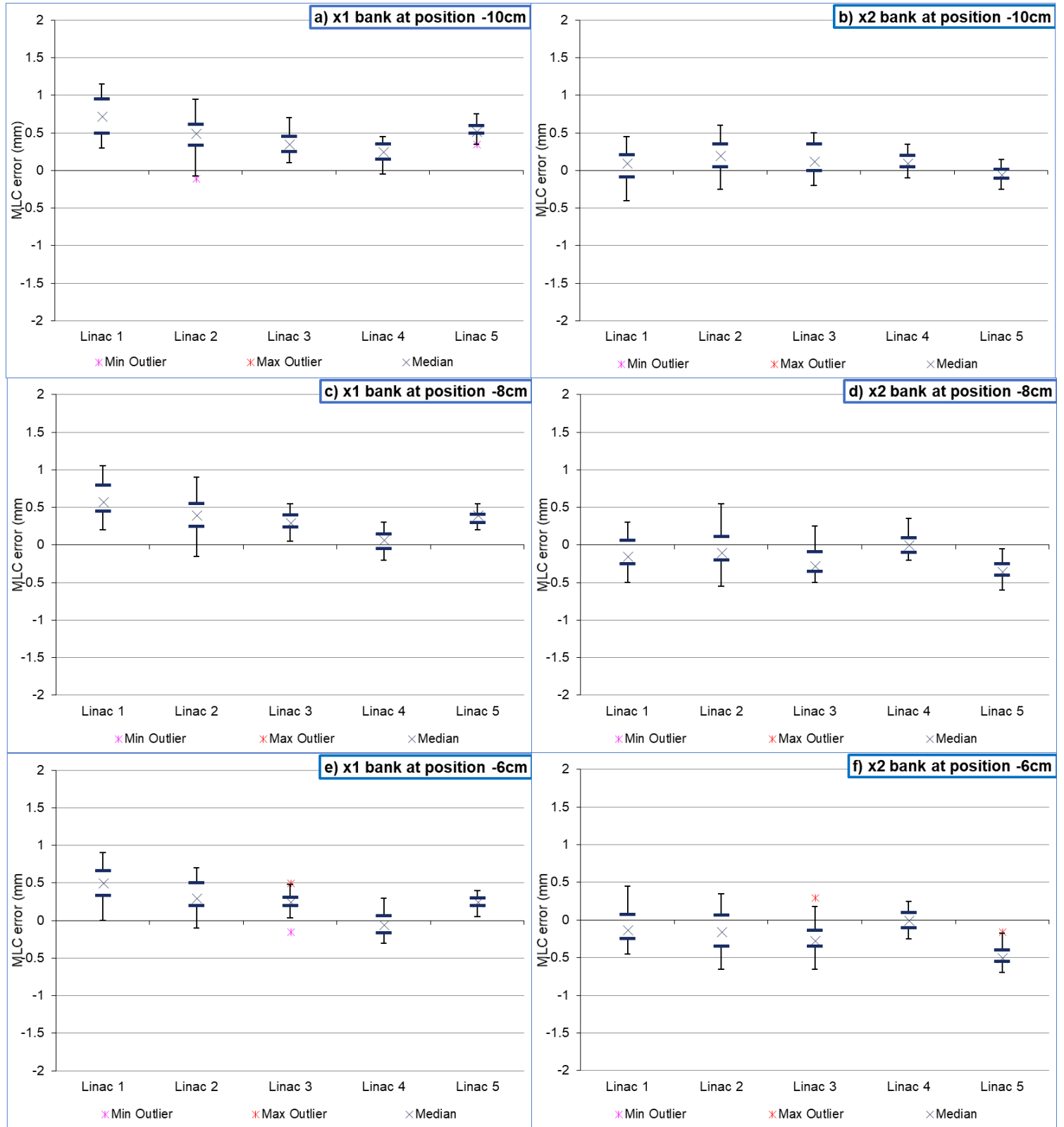
Figure 76: Box plot summaries of MLC errors per Linac for both MLC banks at position -10 cm (a and b), -8 cm (c and d) and -6 cm (e and f). Minimum and maximum outliers are only shown for values outside the 25th and 75th percentile range.
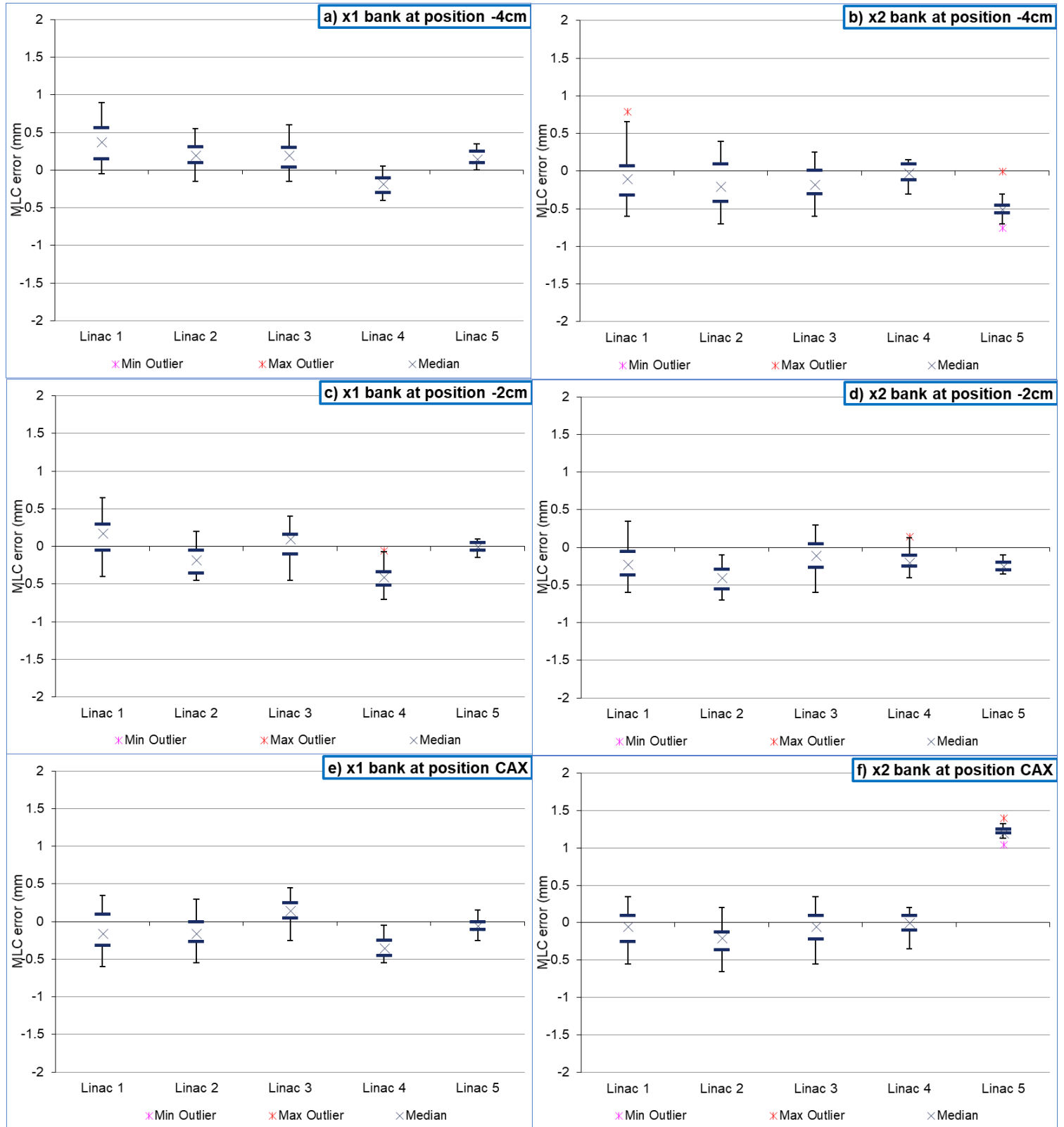
Figure 77: Box plot summaries of MLC errors per Linac for both MLC banks at position -4 cm (a and b), -2 cm (c and d) and CAX (e and f). Minimum and maximum outliers are only shown for values outside the 25th and 75th percentile range.
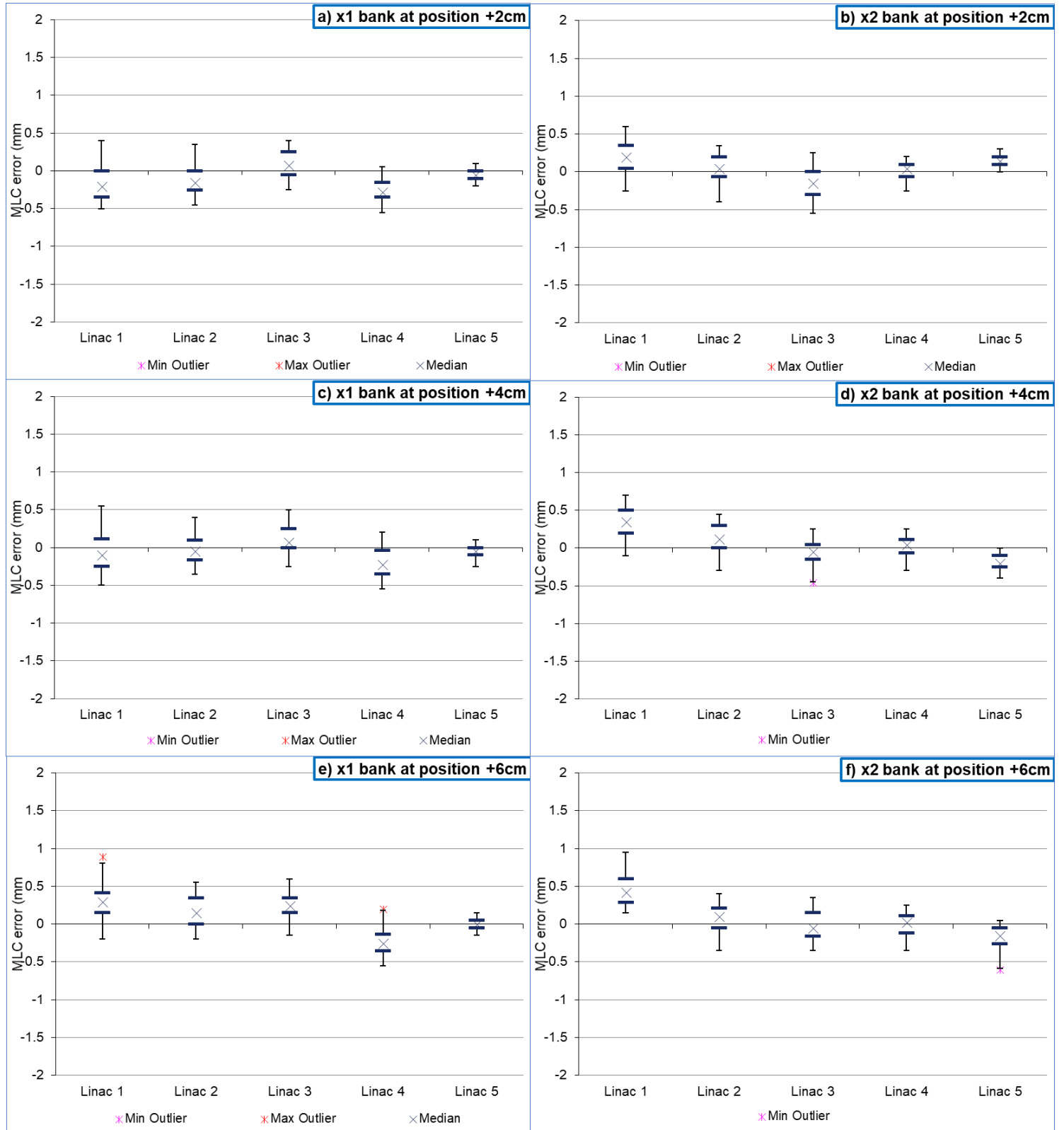
Figure 78: Box plot of MLC errors per Linac for both MLC banks at position +2 cm (a and b), +4 cm (c and d) and +6 cm (e and f). Minimum and maximum outliers are only shown for values outside the 25th and 75th percentile range.
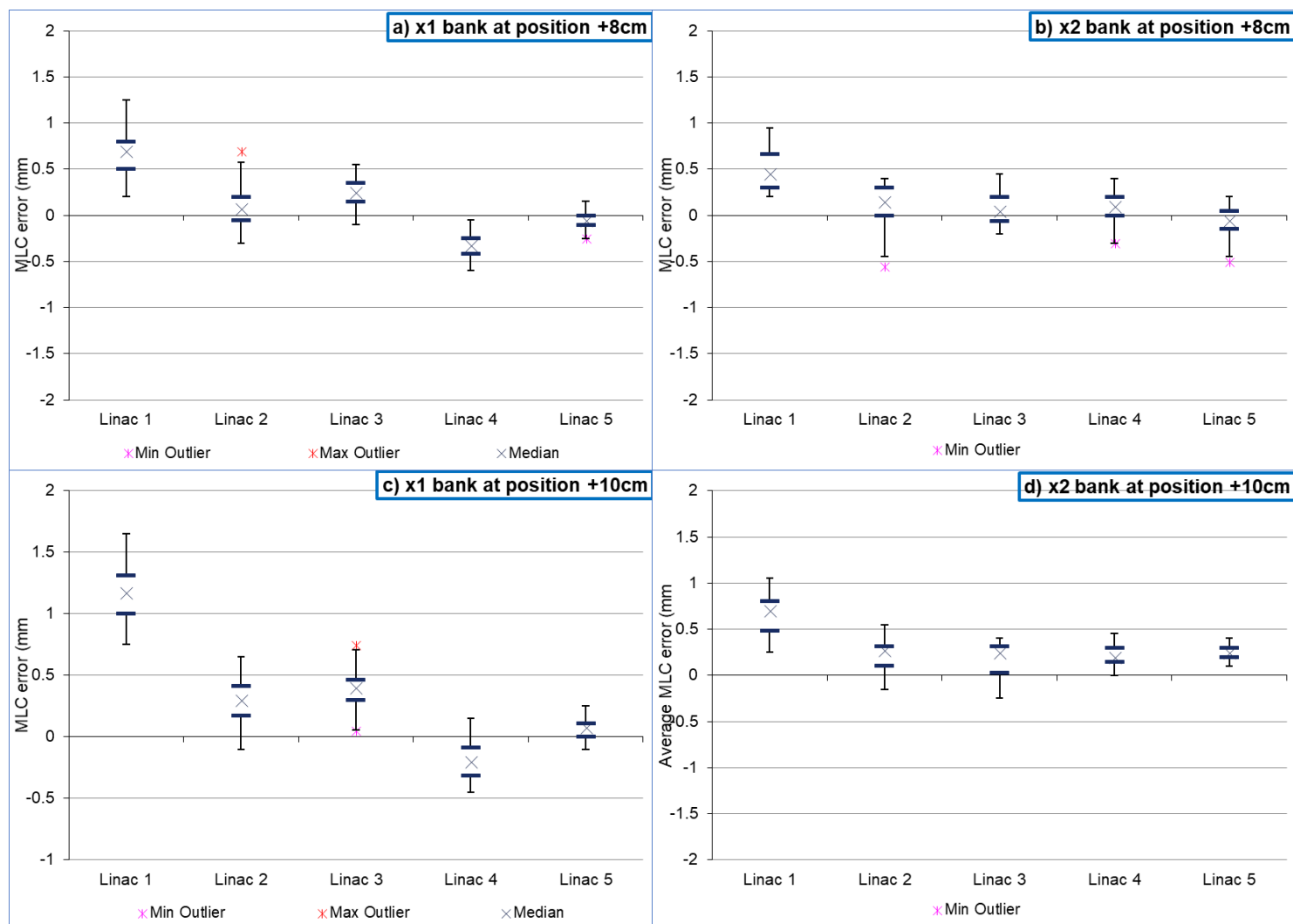
Figure 79: Box plot summaries of MLC errors per Linac for both MLC banks at position +8 cm (a and b) and +10 cm (c and d). Minimum and maximum outliers are only shown for values outside the 25th and 75th percentile range.

Generally, a non-linearity in MLC positions can be seen. Even though a more accurate MLC positioning is demonstrated for most of the Linacs around CAX ± 5 cm, average MLC positions are mostly within 0.5 mm from the intended position. Figures 80 and 81 summate the average MLC errors from all the evaluated stop positions.   Average MLC errors for the entire x1 bank from all 5 Linacs ranged between -0.20 and 0.4 mm. For this bank, the results from Linac1 should be highlighted, because 5 out of the 11 positions evaluated has average errors greater than 0.5 mm.   With an average deviation of the x1 and x2 banks at position +10 cm of 1.18 mm and 0.65 mm respectively, the positional error can potentially combine to 1.83 mm.

For bank x2 (figure 81), average MLC errors vary between ± 0.2 mm. Results of the x2 bank of Linac5 at stop position CAX should be highlighted, with an average MLC error of 1.2 mm, previously shown in figure 77(f). Also, this overtravel was clearly shown in figure 64(e) as a white line.
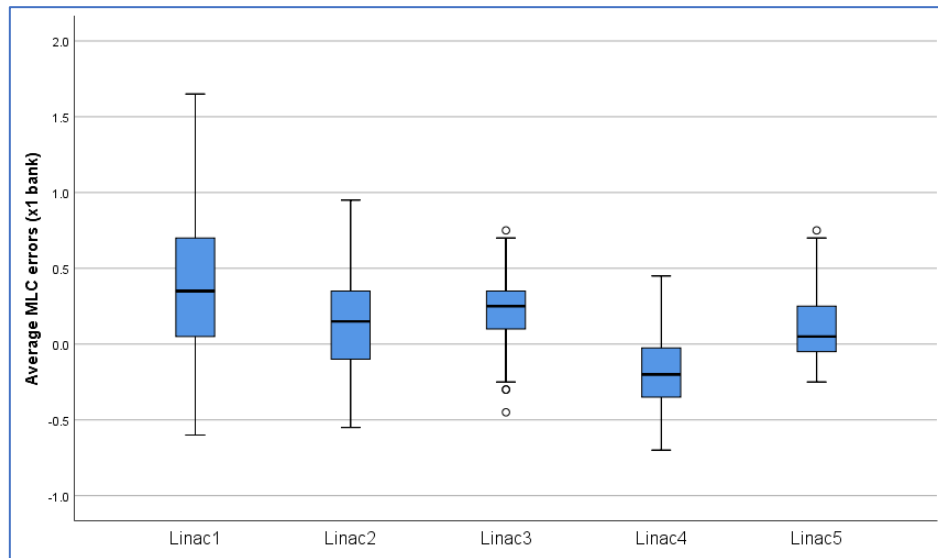
Figure 80: Average MLC errors of the x1 bank from 5 Linacs over a range of stop positions.
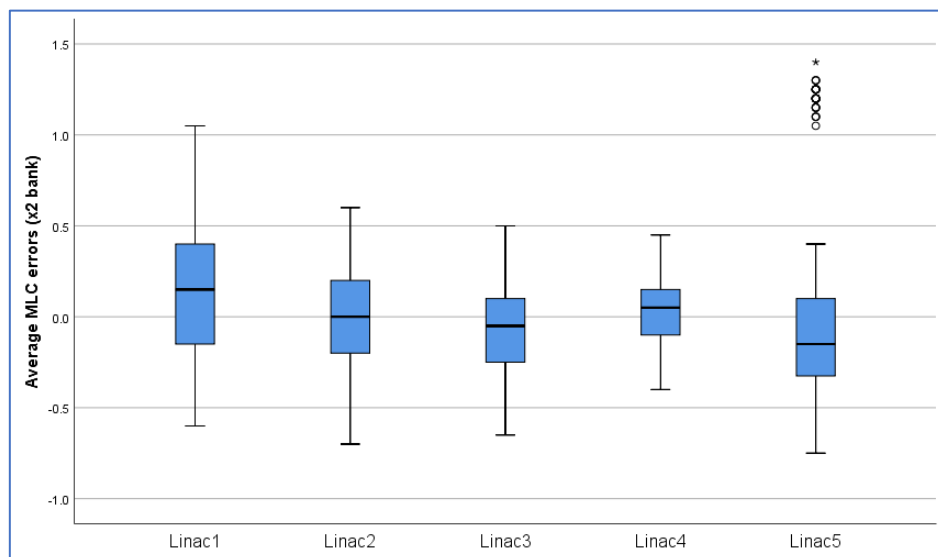
Figure 81: Average MLC errors of the x2 bank from 5 Linacs over a range of stop positions.

Table 6 is showing correlation coefficients (from the Paired sample T test) of different Linac combinations, comparing MLC errors at the 11 stop positions for both MLC banks of each Linac to another individually. A value of 0.05 and above indicates that zero is included in the confidence interval, therefore the null hypothesis is not rejected, meaning that a significant difference between the two Linacs could not be indicated with a 95% probability. However, in such cases it does not mean that the MLCs of the two Linacs are the same at that stop position. The best comparison of MLC behaviour was seen between Linac 2 and 3, with 9 out of 22 stop positions having a p value higher than 0.05. However, the difference between these Linacs were still significant for 13 out of 22 (59%) stop positions. Overall it can be said that the difference in MLC errors observed between Linacs at various stop positions were statistically significant. In other words, all 5 Linacs had different MLC behaviour in terms of accuracy.

p significance (95% confidence interval)

	#	Linac1 - Linac2	Linac1 - Linac3	Linac1 - Linac4	Linac1 - Linac5	Linac2 - Linac3	Linac2 - Linac4	Linac2 - Linac5	Linac3 - Linac4	Linac3 - Linac5	Linac4 - Linac5
x1	1	0.000	0.000	0.000	0.000	0.006	0.000	0.069	0.000	0.000	0.000
x1	2	0.002	0.000	0.000	0.000	0.024	0.000	0.242	0.000	0.023	0.000
x1	3	0.023	0.000	0.000	0.000	0.053	0.000	0.004	0.000	0.989	0.000
x1	4	0.005	0.000	0.000	0.000	0.624	0.000	0.010	0.000	0.326	0.000
x1	5	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.176	0.000
x1	6	0.919	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000
x1	7	0.732	0.000	0.022	0.052	0.000	0.005	0.009	0.000	0.000	0.000
x1	8	0.736	0.000	0.001	1.000	0.008	0.001	0.542	0.000	0.000	0.000
x1	9	0.072	0.072	0.000	0.000	0.136	0.000	0.000	0.000	0.000	0.000
x1	10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
x1	11	0.000	0.000	0.000	0.000	0.012	0.000	0.000	0.000	0.000	0.000
x2	12	0.015	0.000	0.067	0.029	0.350	0.113	0.000	0.446	0.000	0.000
x2	13	0.484	0.000	0.001	0.000	0.002	0.110	0.000	0.000	0.001	0.000
x2	14	0.280	0.000	0.003	0.000	0.106	0.002	0.000	0.000	0.000	0.000
x2	15	0.270	0.022	0.427	0.000	0.697	0.022	0.000	0.002	0.000	0.000
x2	16	0.001	0.008	0.461	0.133	0.000	0.000	0.000	0.158	0.001	0.000
x2	17	0.019	0.341	0.063	0.000	0.002	0.576	0.000	0.111	0.000	0.000
x2	18	0.006	0.000	0.000	0.147	0.001	0.036	0.001	0.000	0.000	0.000
x2	19	0.000	0.000	0.000	0.000	0.001	0.156	0.000	0.004	0.000	0.000
x2	20	0.000	0.000	0.000	0.000	0.080	0.515	0.000	0.462	0.000	0.000
x2	21	0.000	0.000	0.000	0.000	0.460	0.000	0.000	0.849	0.000	0.000
x2	22	0.000	0.000	0.000	0.000	0.130	0.893	0.575	0.005	0.008	0.284
	p<0.05	15	20	18	18	13	16	18	17	19	21
	p>0.05	7	2	4	4	9	6	4	5	3	1

Table 6: p correlation coefficients of different Linac combinations which include all 11 stop positions for both MLC banks. A value below 0.05 indicates a significant difference between Linacs and has an incidence of 75% in this table. Insignificant differences (values above 0.05) are highlighted in yellow.

Large deviations within a single MLC bank of ± 1 mm, also for a Siemens Linac with 41 MLCs of 1 cm width each has been reported [90]. Generally, results obtained are similar to that measured by Tacke *et al.* (2008), who reported MLC errors of ±0.6 mm for the same MLC design, but also reported a maximum value of 0.8 mm for a single profile [61]. Another study also reported MLC errors of ±0.5 mm for a double focused Siemens Linac, also scanning in 1 mm resolution using a liquid ionization chamber linear array (PTW LA48) [91]. The authors highlighted the importance of an accurate setup procedure and showed a reproducibility of ±0.1 mm. It should be highlighted that the MLC carriage position backlash, that may cause an error (hysteresis effect) moving outward compared to moving inward [65], was not considered in the current study.

Tacke *et al.* (2008) also showed a difference in the triangular tongue and groove (T&G) effect due to the alternating pattern of upper and lower leaves. As previously shown in Figure 4, the MLC's are tilted towards a point 2.6 mm from the focal spot to avoid a straight open-air gap or an easy pass for the radiation beam between MLCs. This pattern of MLC construction in relation to an incident radiation beam are shown in Figure 82. This means the amount of interleaf leakage, or the degree of the T&G effect, differs for any two neighbouring MLCs; a larger under dosage dip occur at positions where the upper part of an upper leaf overlaps with the lower edge of a lower leaf resulting in a wider overlap region. Therefore, a non-uniform triangular T&G effect was reported.

Further, Kluter *et al.* (2011) studied the leakage behaviour of 3 structurally identical Siemens 160 MLC's and reported different interleaf leakage profiles [92]. Results were attributed to asymmetric focal spots: a difference in the position of the intensity from the focal spot and the rotational axis of the collimator. They also reported the effect of beam steering on 6MV interleaf leakage values, and showed the higher the steering value, the more the beam is shifted which leads to a different effective tilt of the MLCs. In other words, as the focal spot moves the MLCs become directly aligned with larger or smaller parts of the radiation source and the leakage varies. Interleaf leakage values between 1.1% to 3.3% were measured for different steering settings. However, these values referred to individual peaks in the leakages profile whereas average values between Linacs varied between 0.4 and 0.6%.

Therefore, interleaf leakage may be unique for each Siemens Artiste[TM] Linac due to different positions of the focal spot and the amount of beam steering. Even though this effect and the potential impact it may have on MLC positional accuracy were not investigated in the current study, it could contribute to the large deviation in MLC errors reported.
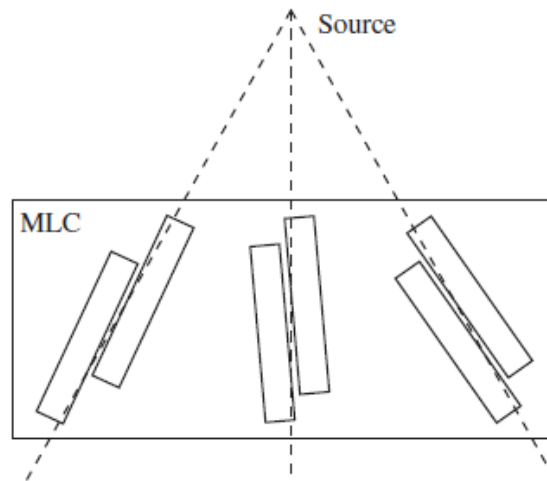
Figure 82: Varying triangular tongue and groove effect [61]

5.3.1.2)   MLC radiological calibration

Results of the variation in MLC positional accuracies, maybe due to different MLC tilts and subsequently different triangular T&G effects, can potentially be compensated for by applying software corrections (discussed in section 3.1). Radiological calibration curves for each of the 5 Linacs are shown in figures 83 and 84 respectively.

Graphs 83 and 84 show the offset in millimetre applied at various MLC stop positions (CAX ± 20 cm) for each Linac individually. Generally, a change in trend can be seen around CAX ± 5 cm, most drastically for the x2 bank, and then a gradual increase with distance form CAX. Therefore, it appears such drastic offset values are required to obtain leaf end positional accuracy or MLC errors within 0.5 mm. However, a complete change in trend of x2 offset values around CAX can be seen for Linac 5 (figure 84), which explains the abnormal overtravel of this bank at this position as were shown in figures 64(e), 70(e) and 77(f). Also, the MLCs titled toward a virtual focus point, other than the real focal spot, could clarify why both leave banks do not have the same radiological calibration.

Interestingly Linac 2 and 4 have the same MLC calibration curves. However, a difference in MLC errors were obtained (up to 0.4 mm for the x1 bank and 0.3 mm for the x2 bank as previously shown in section 5.3.1.1). Probable causes were discussed above and in the previous section and could also include a combination of measurement accuracy and reproducibility.
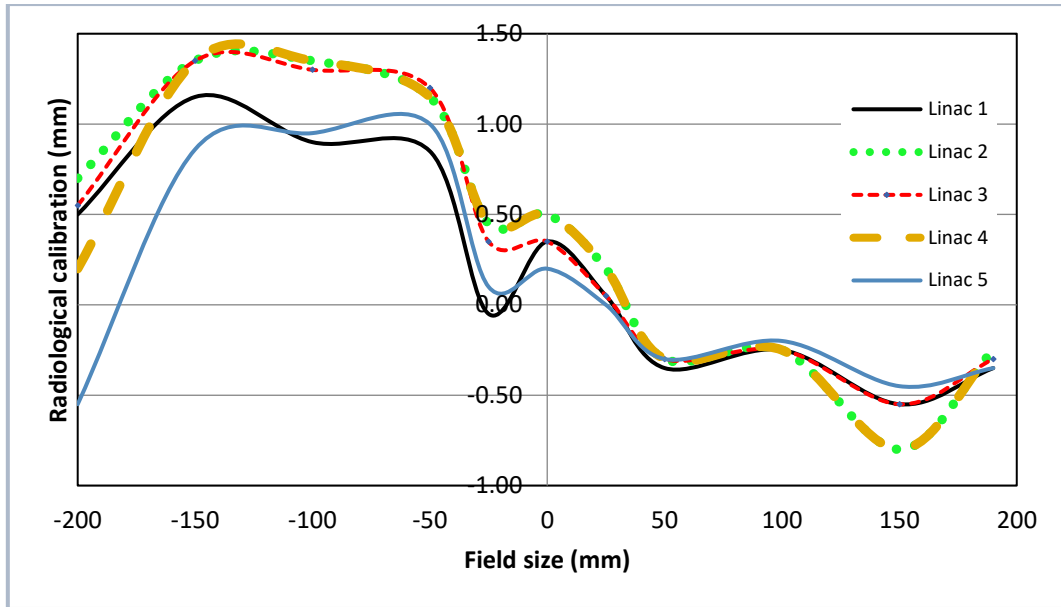
Figure83: Current MLC offset values of the x1 bank of each Linac respectively. The y-axis shows the amount of MLC offset in mm per distance from CAX (x-axis). A similar trend between the 5 Linacs can be seen.
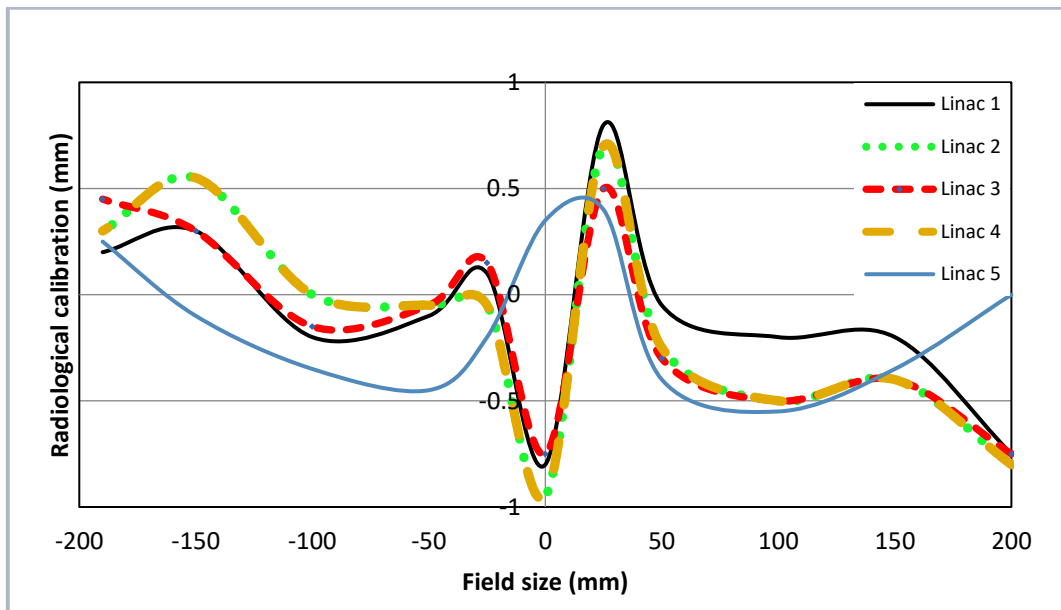
Figure84: Current MLC offset values of the x2 bank of each Linac respectively. The y-axis shows the amount of MLC offset in mm per distance from CAX (x-axis). Linac5 has different negative offset values, with a major difference around CAX. Linac 1 also has different offset values for positive field sizes.

Based on the average MLC error results from Linacs 2 to 5 for the x1 bank (Linac 1 excluded) and from Linacs 2 to 4 for the x2 bank (Linacs 1 and 5 excluded), new proposed radiological calibration curves up to CAX ± 10 cm are shown in figures 85 and 86. The curves follow the same trend as before. Therefore, the average measured MLC error at a particular bank position are corrected for by applying a software offset of the same magnitude in the opposing direction. To summarize from these graphs: for the x1 bank less overtravel is proposed at positions -10 cm (0.5 mm) and +10 cm (0.2 mm); for bank x2 more overtravel is proposed from positions -0.6 cm to CAX (between 0.2 mm and 0.4 mm) and less overtravel or more under travel from +6 cm to +10 cm (between 0.1 mm and 0.2 mm).

In his thesis Tacke (2009) mentioned "after modification of the calibration tools MLC positional accuracy improved from ±0.6 mm to ±0.3 mm", but unfortunately did not include details on how this was achieved [93]. It is believed a single radiological calibration curve (representative of all 5 Linacs) could potentially be constructed to at least improve general non-linearity in MLC positions across the radiation field to well within ±0.5 mm as well as removing larger deviating results as was obtained for Linacs 1 and 5. However, even though a single radiological calibration curve will most probably rectify the large overtravel of x2 for Linac 5 at position CAX, it may not be sufficient to rectify the above average overtravel of both banks for Linac 1 at positions ± 10 cm.

Even though MLC offset values are fully customizable, the manufacturer advised these should only be changed by an installation engineer (with vendor equipment) and unfortunately an agreement could not be reached between the vendor and customer due to service contract issues. Therefore, the proposed MLC offset curves could not be applied and tested in the current study.
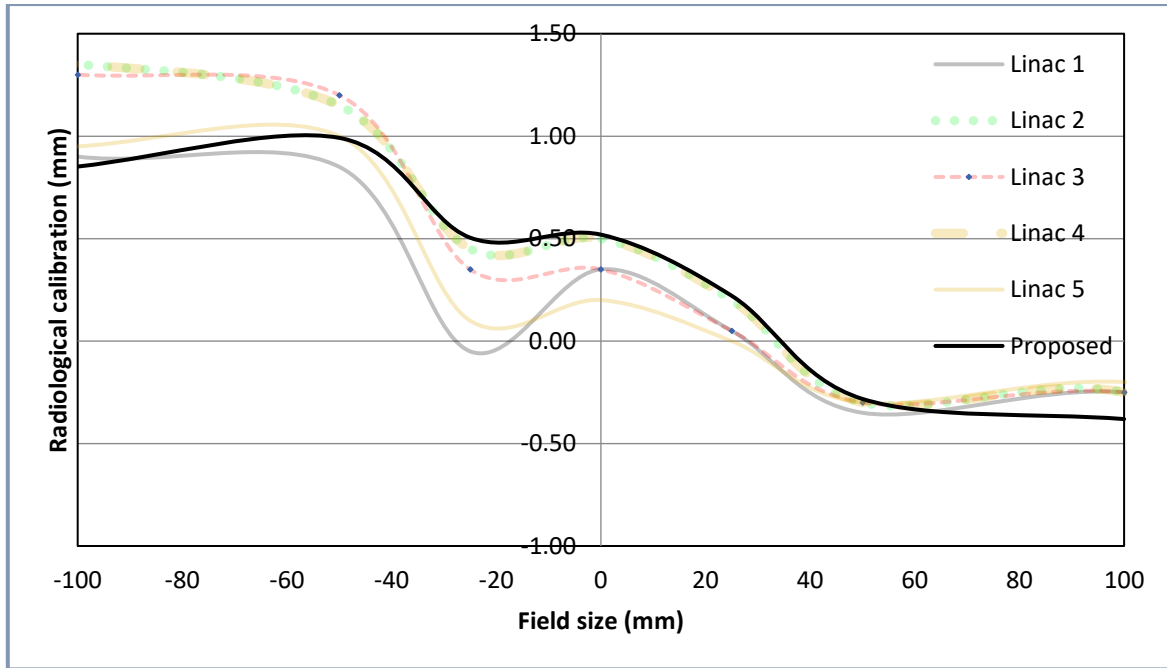
Figure85: Proposed MLC offset values for the x1 bank. The black line indicates the average of Linacs 2 to 5. Less overtravel of approximately 0.5 and 0.2 mm is proposed at positions -100 and +100 mm respectively.
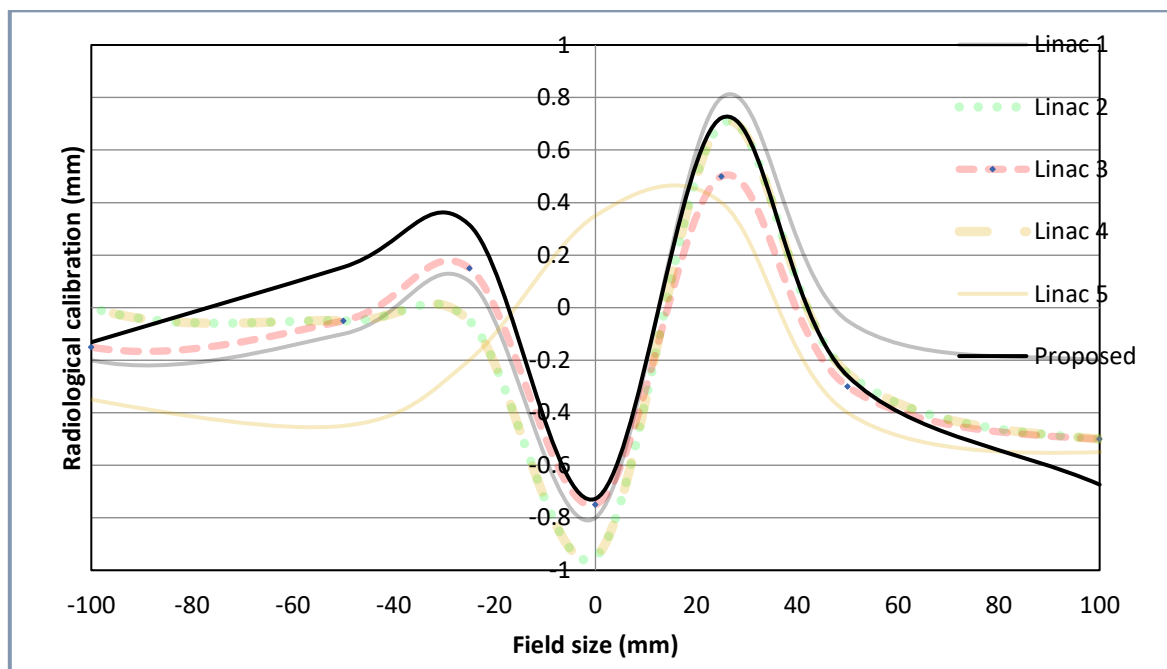
Figure86: Proposed MLC offset values for the x2 bank. The black line indicates the average of Linacs 2 to 4. More overtravel is proposed from positions -60 mm to CAX (between 0.2 mm and 0.4 mm) and less overtravel or more under travel from +60 mm to +100 mm (between 0.1mm and 0.2mm).

5.3.2)   Test beam doses

Table 7 shows γ analysis results for measured vs calculated test beam doses (previously discussed in sections 3.2.2 and 4.4.2). All simple fields (20 × 20 cm^2 and 10 × 10 cm^2) pass above 95.0% on a 2%/2mm criterion for all 5 machines with an average pass rate of 97.2 ± 0.87%. Lower pass rates for MLC characterization fields (3ABUT and 7SegA) were seen, with a lowest pass rate of 82.7% for the 7SegA field on Linac 4 and an average pass rate between Linacs for this field of 89.8 ± 4.75%. A CAX profile screenshot between the measured and planned 3ABUT and 7SegA fields of each Linac are shown in figures 87 and 88 respectively.

		Field Name				
Linac	*γ criterion*	20×20	10×10	3AUBT	7SegA	HIMRT
Linac 1	3%/3mm	99.3	98.8	96.8	95.6	97.7
	2%/2mm	97.7	95.6	93	86.3	92
Linac 2	3%/3mm	99.7	99.1	98.8	94.4	99.5
	2%/2mm	98.7	97	96.1	87.7	95.8
Linac 3	3%/3mm	99.8	99.4	96.5	90.1	99.4
	2%/2mm	98.2	96.8	91.7	83.5	96.2
Linac 4	3%/3mm	99.3	99.9	98.4	92.1	98.5
	2%/2mm	97.1	96.7	94.6	82.7	96.9
Linac 5	3%/3mm	99.3	99.6	99	96.1	100
	2%/2mm	96.6	97.2	94.1	88.6	97.9

Table 7: Gamma percentage pass rates from the test beam package for each Linac respectively. Very good pass rates for the open fields (10 × 10 and 20 × 20 cm^2) as well as the IMRT field on all the Linacs. However, MLC positional accuracy test fields (3ABUT and 7SeA) have lower γ pass rates, especially on a 2%/2mm evaluation criterion, again showing a non-linearity in MLC positional accuracy for all the Linacs.

As expected, these results clearly show the different MLC behaviour of each machine as was discussed in section 5.3.1. However, the results of this section are for comparison purposes only and not to adjust MLC parameters of the beam model as was discussed in section 3.2.2. Model parameters were historically obtained from

Linac 1, although it could not be optimized using this method as it was unclear how to change the MLC parameters to gain agreement, given the non-linearity in MLC positional accuracy.

Comparing γ pass rates show that 75% of fields (15 out of 20) measured on Linacs other than the reference Linac1 had higher pass rates. Lower than reference Linac pass rates were mostly similar (within 1.5%), except for the 7SegA field of Linacs 3 and 4 with a 2.8% and 3.6% lower pass rate respectively. Even though satisfactory results for the IMRT field on all machines were obtained, Linacs 2 to 5 have higher pass rates for this field (average of 4.7% on 2%/2mm criterion) compared to Linac 1. Figure 89 shows profile screenshots of TPS calculated and measured points of the IMRT field for each Linac individually.
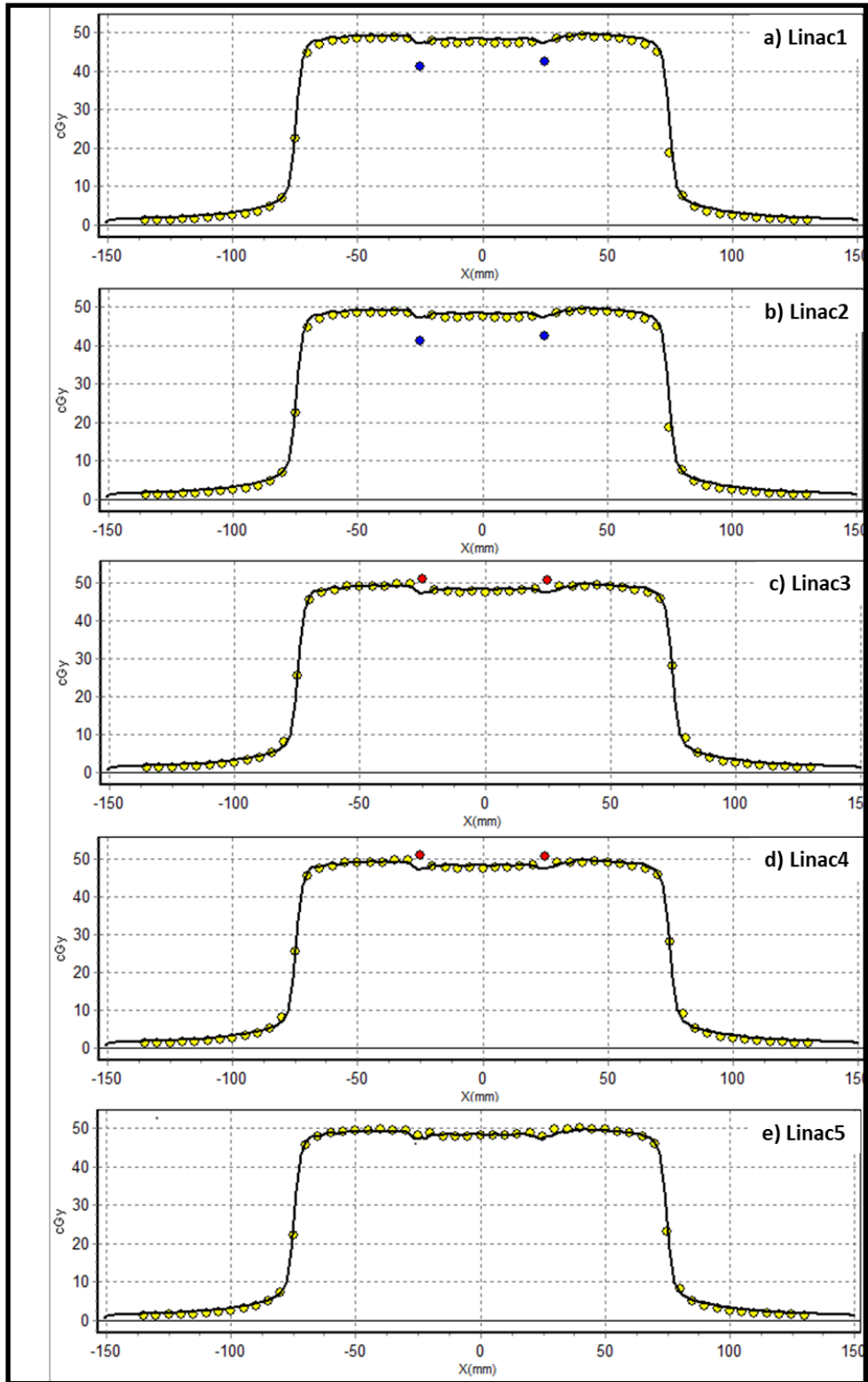
Figure 87: Comparison between TPS calculated and measured cross plane profiles through the central axis crosshair of the 3ABUT test beam for each Linac respectively: a) Linac1, b) Linac2, c) Linac3, d) Linac4 and e) Linac5. The black profile is TPS calculated and yellow dots indicate the measured profile. Blue and red dots indicate points failing the γ criterion: blue dots show measured doses lower than TPS calculated and red dots show measured doses higher than TPS calculated.
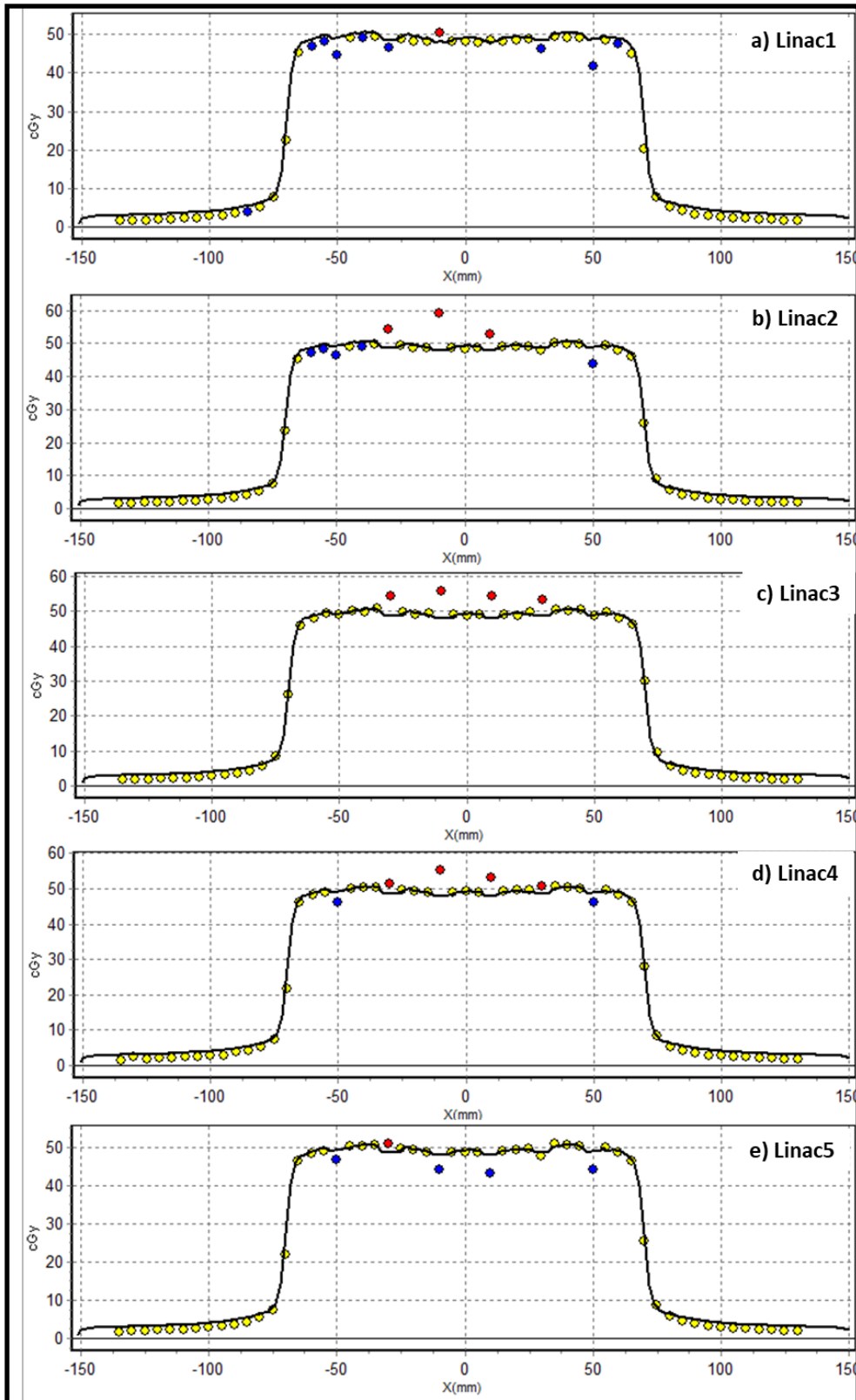
Figure 88: Comparison between TPS calculated and measured cross plane profiles through the central axis crosshair of the 7SegA test beam for each Linac respectively: a) Linac1, b) Linac2, c) Linac3, d) Linac4 and e) Linac5. The black profile is TPS calculated and yellow (including red and blue) dots indicate the measured profile. Non-linearity in MLC positional accuracy are clearly highlighted by failing γ points that are different for each Linac.
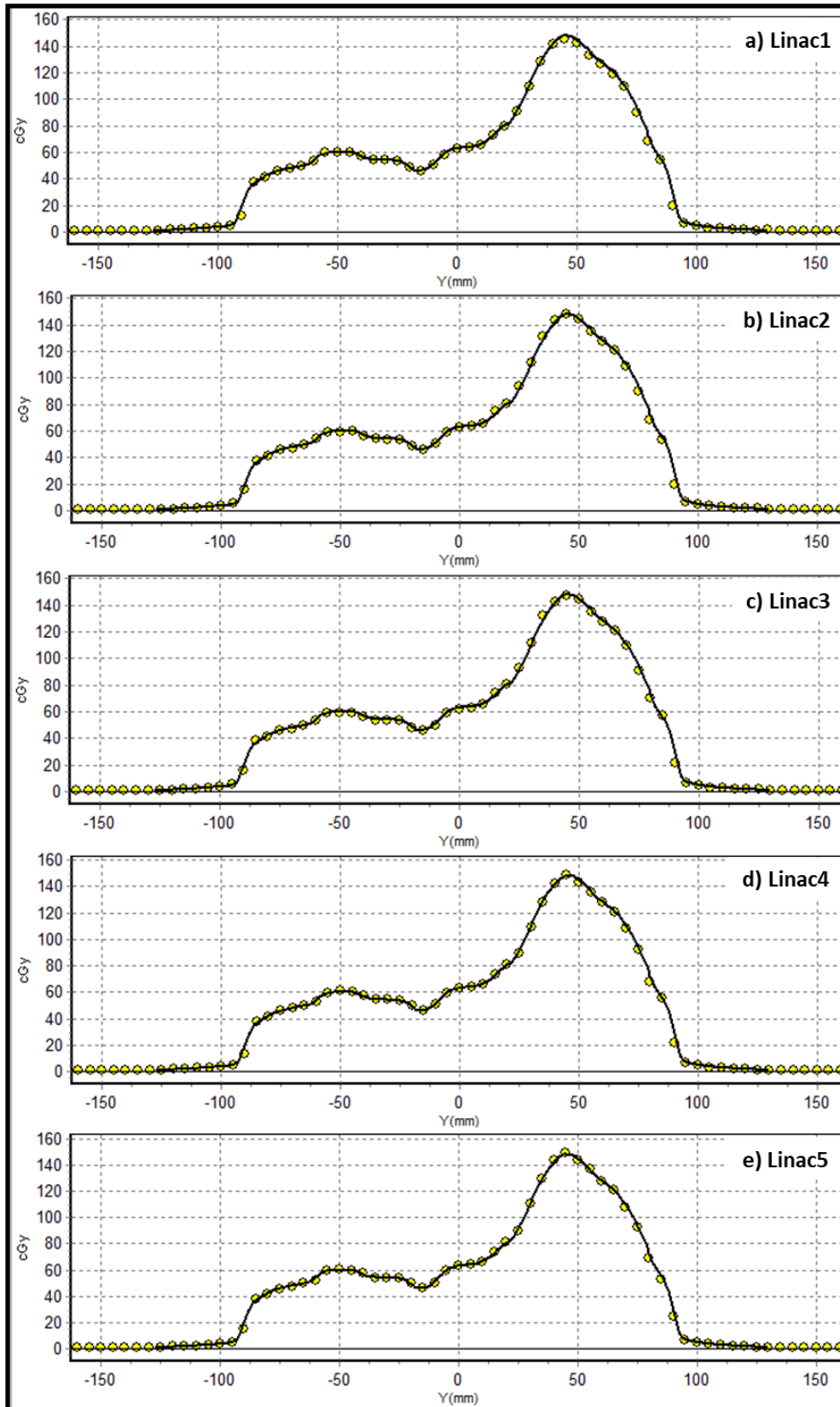
Figure 89: Comparison between TPS calculated and measured cross plane profiles through the central axis crosshair of the HNIMRT test beam for each Linac respectively: a) Linac1, b) Linac2, c) Linac3, d) Linac4 and e) Linac5. The black profile is TPS calculated and yellow dots indicate the measured profile. No failing γ points at the profile position.

5.3.3)   Clinical IMRT fields

Gamma pass rates obtained from each of the 5 IMRT fields measured on all 5 Linacs compared to the calculated doses from the reference TPS model are shown in tables 8 (field 1) to 12 (field 5). These include results from all 4 γ evaluation methods: "Mapcheck single" (a single Mapcheck measurement), "Mapcheck double" (double exposure Mapcheck measurements), "OmniPro single" (a single Mapcheck measurement analysed in OmniPro I'mRt) and "OmniPro 1mm" (10 Mapcheck measurements analysed in OmniPro I'mRt). Also included in Table 8 to 12 are CAX point dose percentage differences that range between -1.46% and 1.73%. This is satisfactory given the uncertainties in point dose measurements, such as the Linac output calibration with a Farmer chamber (previously mentioned as 1.5%) and daily Linac output variation (approximately 1%) as well as the statistical uncertainty in the MC dose calculation (0.5%).

Evaluating 25 measured IMRT fields (5 fields on 5 Linacs) with a 3%/3mm criterion, an average γ pass rate of 99.4% ± 0.48 is obtained. These exceptional results (summarized in figure 95) could indicate accurate IMRT delivery on each of the 5 Linacs in the per-field mode or it could be that the use of the Mapcheck device and other analysis parameters provide too loose criteria to quantify the degree of deviation. Therefore, results from criterion 2%/2mm (graphically shown in Figures 90 to 94 and summarized in figure 96) may be more meaningful as some differences could be seen between the different γ evaluation methods. However, all fields passed the 2%/2mm criterion with at least 92%. Yet, from figures 90, 91 and 94 it is clear that lower γ pass rates were obtained with the "OmniPro single" γ evaluation method.

γ Criterion	Evaluation Method	Linac1	Linac2	Linac3	Linac4	Linac5
3%/3mm	Mapcheck single	99.2	100.0	99.7	99.2	100.0
	OnmiPro single	99.5	99.9	99.8	99.8	99.7
	Mapcheck double	98.7	99.7	99.7	99.2	99.6
	OmniPro 1mm	99.9	99.8	99.9	99.4	100.0
2%/2mm	Mapcheck single	96.5	98.5	98.0	97.4	97.9
	OnmiPro single	94.2	96.2	96.4	96.7	97.2
	Mapcheck double	95.4	98.1	96.6	97.5	97.2
	OmniPro 1mm	97.5	98.5	97.9	97.7	99.4
Measured Dose (CAX cGy)	TPS dose = 38.8	39.4	40.8	40.4	40.0	37.9

Table 8: Gamma pass rates obtained by measuring an identical IMRT field (field 1) on all 5 Linacs and comparing each measurement to the TPS calculated dose respectively. Four different evaluation methods were used using both evaluation criteria's (3%/3mm and 2%/2mm); CAX doses are also included.

γ Criterion	Evaluation Method	Linac1	Linac2	Linac3	Linac4	Linac5
3%/3mm	Mapcheck single	100.0	99.7	99.1	99.7	99.4
	OnmiPro single	98.6	99.2	99.0	98.3	98.1
	Mapcheck double	98.3	99.2	98.6	99.8	98.5
	OmniPro 1mm	99.5	99.2	99.0	99.7	99.5
2%/2mm	Mapcheck single	99.1	99.7	98.2	99.1	98.8
	OnmiPro single	92.6	93.5	92.9	94.2	92.8
	Mapcheck double	98.3	99.4	96.6	99.7	98.8
	OmniPro 1mm	96.3	97.8	96.5	98.0	97.7
Measured Dose (CAX cGy)	TPS dose = 65.8	64.1	64.9	67.7	66.3	62.4

Table 9: Gamma pass rates obtained by measuring an identical IMRT field (field 2) on all 5 Linacs and comparing each measurement to the TPS calculated dose respectively. Four different evaluation methods were used using both evaluation criteria's (3%/3mm and 2%/2mm); CAX doses are also included.

γ Criterion	Evaluation Method	Linac1	Linac2	Linac3	Linac4	Linac5
3%/3mm	Mapcheck single	99.7	100	100	99.5	100
	OnmiPro single	99.7	99.8	99.9	99.8	99.7
	Mapcheck double	99.6	99.5	99.7	99.7	99
	OmniPro 1mm	99.4	99.6	99.6	99.9	99.8
2%/2mm	Mapcheck single	99.5	99	98.5	99.2	98.4
	OnmiPro single	96.3	97.8	97.2	97.9	96.7
	Mapcheck double	98.2	97.5	98.3	98.3	97.8
	OmniPro 1mm	97.6	98	97.8	98.6	98.9
Measured Dose (CAX cGy)	TPS dose = 65.8	54.5	54.5	55.4	54.3	54.6

Table 10: Gamma pass rates obtained by measuring an identical IMRT field (field 3) on all 5 Linacs and comparing each measurement to the TPS calculated dose respectively. Four different evaluation methods were used using both evaluation criteria's (3%/3mm and 2%/2mm); CAX doses are also included.

γ Criterion	Evaluation Method	Linac1	Linac2	Linac3	Linac4	Linac5
3%/3mm	Mapcheck single	99.4	99.8	99.4	99.2	98.5
	OnmiPro single	99.6	99.6	99.8	99.6	99.6
	Mapcheck double	99.5	99.0	99.0	99.1	98.6
	OmniPro 1mm	99.5	99.5	99.8	99.9	99.9
2%/2mm	Mapcheck single	95.8	97.5	96.5	96.4	95.8
	OnmiPro single	94.7	97.0	96.3	96.8	96.3
	Mapcheck double	96.3	97.0	95.5	96.7	95.7
	OmniPro 1mm	97.1	98.5	97.2	98.8	98.9
Measured Dose (CAX cGy)	TPS dose = 4.3	3.9	4.0	4.0	3.9	4.0

Table 11: Gamma pass rates obtained by measuring an identical IMRT field (field 4) on all 5 Linacs and comparing each measurement to the TPS calculated dose respectively. Four different evaluation methods were used using both evaluation criteria's (3%/3mm and 2%/2mm); CAX doses are also included.

γ Criterion	Evaluation Method	Linac1	Linac2	Linac3	Linac4	Linac5
3%/3mm	Mapcheck single	99.5	99.0	99.7	98.2	99.2
	OnmiPro single	99.7	99.6	99.8	99.4	99.3
	Mapcheck double	99.6	99.0	98.7	98.3	99.2
	OmniPro 1mm	99.6	99.7	99.9	99.3	99.9
2%/2mm	Mapcheck single	98.7	96.9	97.9	96.7	97.7
	OnmiPro single	95.2	95.4	96.4	95.6	94.4
	Mapcheck double	98.8	97.0	95.9	96.3	97.1
	OmniPro 1mm	98.0	98.4	98.4	98.7	98.5
Measured Dose (CAX cGy)	*TPS dose = 67.2*	*65.5*	*67.7*	*66.9*	*66.5*	*69.0*

Table 12: Gamma pass rates obtained by measuring an identical IMRT field (field 5) on all 5 Linacs and comparing each measurement to the TPS calculated dose respectively. Four different evaluation methods were used using both evaluation criteria's (3%/3mm and 2%/2mm); CAX doses are also included.
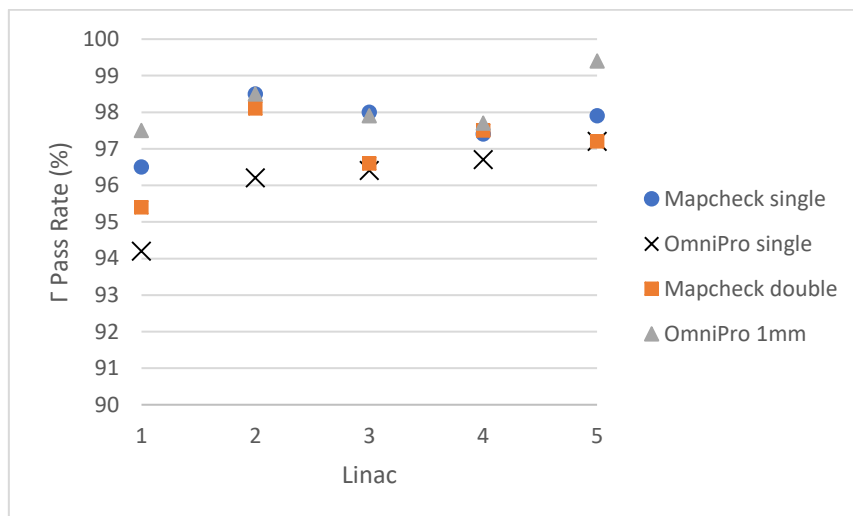
Figure 90: Per Linac γ pass rates (2%/2mm) for field 1 using 4 evaluation methods. For this field, on all 5 Linacs, the high-resolution measurement ("OnmiPro 1mm") shows a better γ pass rate and "OmniPro single" (a single Mapcheck measurement evaluated in OmniPro) a lower γ pass rate.
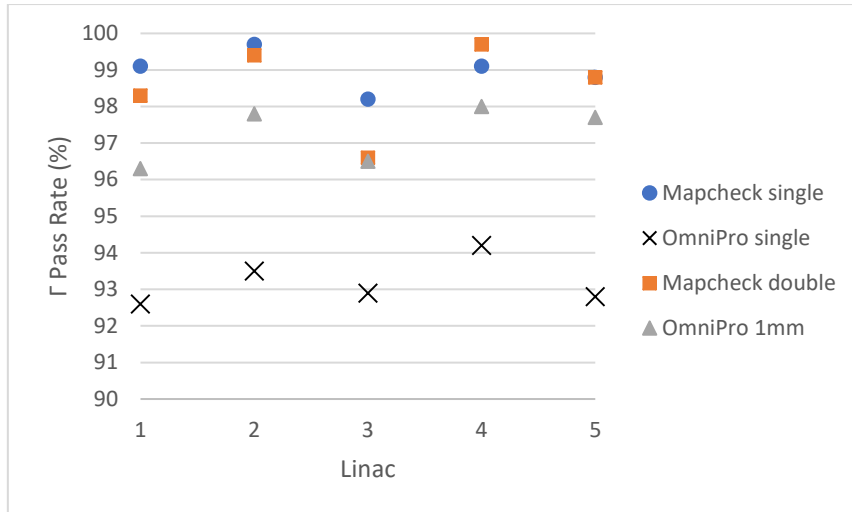
Figure 91: Per Linac γ pass rates (2%/2mm) for field 2 using 4 evaluation methods. For this field, on all 5 Linacs, the "OmniPro single" evaluation method produces lower γ pass rates; on average 5.8% compared to "Mapcheck single".
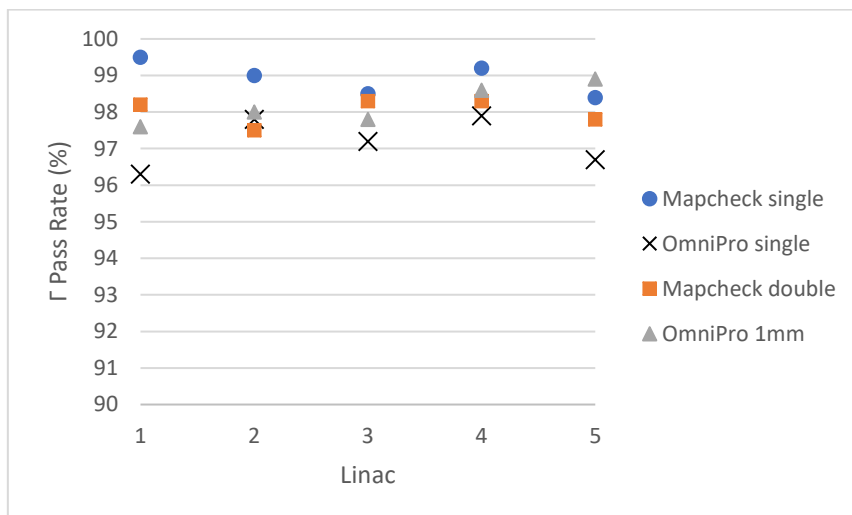
Figure 92: Per Linac γ pass rates (2%/2mm) for field 3 using 4 evaluation methods. For this field, on all 5 Linacs, "Mapcheck single" show higher γ pass rates and once again "OmniPro single" lower γ pass rates, yet the lowest pass rate is 96.3% (Linac1).
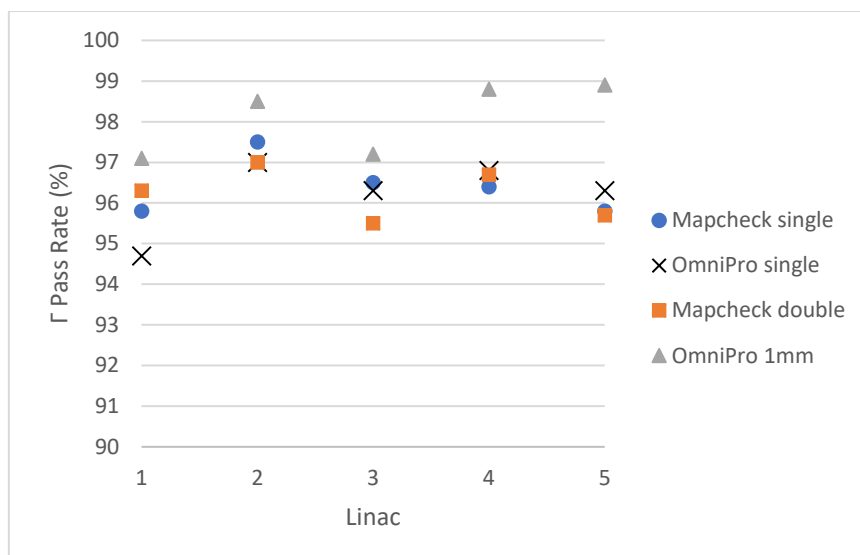
Figure 93: Per Linac γ pass rates (2%/2mm) for field 4 using 4 evaluation methods. For this field, on all 5 Linacs, "OmniPro 1mm" shows the highest γ pass rate.
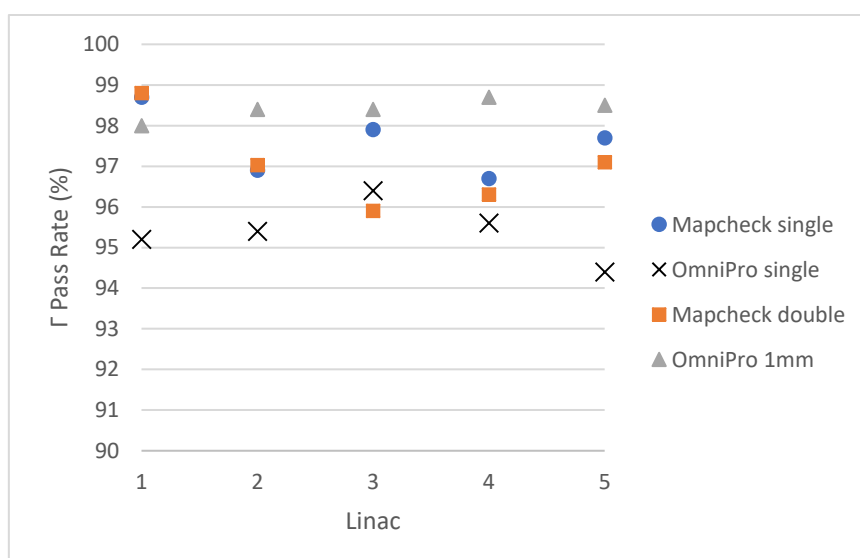
Figure 94: Per Linac γ pass rates (2%/2mm) for field 5 using 4 evaluation methods. For this field, on all 5 Linacs, "OmniPro single" shows a lower γ pass rate for most of the Linacs, except for Linac 3 with a "Mapcheck double" lowest pass rate of 95.9%.
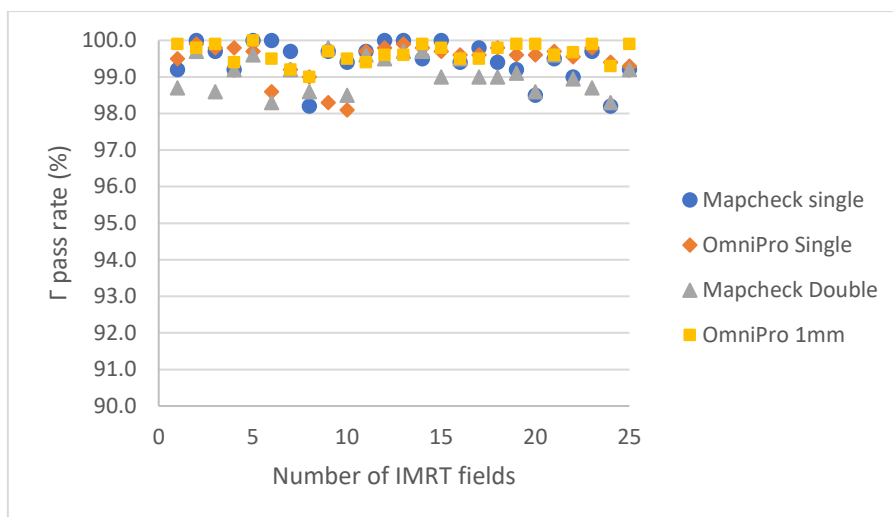
Figure 95: Summary of γ pass rates obtained for criterion 3%/3mm from all 25 IMRT fields measured (5 fields on each of the 5 Linacs). The average γ rate percentage from all 4 evaluated methods combined is 99.4 ± 0.48. The lowest pass rate is 98.1% on Linac 5 using the "OmniPro single" evaluation method.
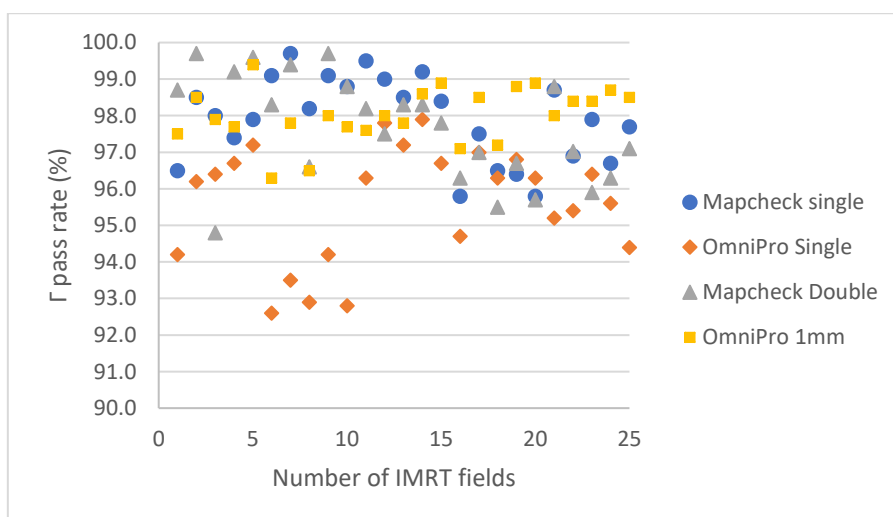
Figure 96: Summary of γ pass rates obtained for criterion 2%/2mm from all 25 IMRT fields measured (5 fields on each of the 5 Linacs). The average γ rate percentage from all 4 evaluated methods combined is 97.3 % ± 1.59. A trend of lower pass rates using "OmniPro single" can be seen, and for this evaluation method the average percentage γ pass rate is 95.6 ± 1.57.

Per linac results are summarized in Table 13 (a-d), showing a summary of the average γ pass rate, with standard deviations and confidence limits for each evaluation method (2%/2mm). As mentioned before, differences could be seen using this criterion and therefore the results from the 3%/3mm criterion are excluded.

Confidence limits were calculated as follows for each evaluation method: 4.5% ("Mapcheck single"), 7.5% ("OmniPro single"), 5.0% ("Mapcheck double") and 3.1% ("OmniPro 1mm") respectively. This means that, for example for Linac 1 using the "Mapcheck single" method, the percentage of points passing the criterion should exceed 94.7% (100 – 5.3%) approximately 95% of the time (table 13(a)).

a) Mapcheck single	Linac				
	1	2	3	4	5
Mean	97.9	98.3	97.8	97.8	97.7
Standard deviation (σ)	1.7	1.1	0.8	1.3	1.2
Local confidence limit (%)	5.3	3.9	3.7	4.8	4.5

b) OmniPro single	Linac				
	1	2	3	4	5
Mean	94.6	96.0	95.8	96.2	95.5
Standard deviation (σ)	1.4	1.6	1.7	1.4	1.8
Local confidence limit (%)	8.1	7.3	7.5	6.5	8.1

c) Mapcheck double	Linac				
	1	2	3	4	5
Mean	98.1	98.1	96.2	98.0	97.8
Standard deviation (σ)	1.0	1.3	1.3	1.5	1.5
Local confidence limit (%)	3.9	4.5	6.4	4.9	5.2

d) OmniPro 1mm	Linac				
	1	2	3	4	5
Mean	97.3	98.2	97.6	98.4	98.7
Standard deviation (σ)	0.6	0.3	0.7	0.5	0.6
Local confidence limit (%)	4.0	2.4	3.9	2.6	2.6

Table 13 (a-d): Percentage of points passing γ criterion of 2%/2mm, averaged over the 5 test fields, with associated confidence limits for evaluation methods a) "Mapcheck single", b) "OmniPro single", c) "Mapcheck double" and d) "OmniPro 1mm". From all fields and methods evaluated, the average confidence limit of "OmniPro single" is the highest on 7.5% and "OmniPro 1mm" the lowest with 3.1%.

In principle one would expect results from Mapcheck and OmniPro single evaluation methods to be the same as the same matrix of measured data is analysed. Instead, from figure 96 and table 13, OmniPro was calculating a lower γ pass rate (average of 2.3% with maximum of 6.5%). The same effect was reported before, SNC Patient™ (measured with ArcCheck™) calculated on average a 2.7% higher γ pass rate than OmniPro I'mRT (measured with Gafchromic film) [94].

In both SNC Patient™ and OmniPro I'mRT software packages, global normalization was applied in other words percentage difference in dose was with respect to the maximum point in the region and not the local point. Nevertheless, the two software packages are calculating different γ results. Possible reasons for this include a difference in: a) number of γ evaluation points or b) software corrections applied:

a) From table 14 the number of γ evaluation points for OmniPro single is approximately 3x that of Mapcheck single, even though the number of measured points is the same. Known facts for both matrixes are: the matrix size is 3445 points; a 10% threshold was applied; removing this threshold increases the number of evaluation points by approximately 3x, depending on the size of the treatment field. It seems SNC Patient™ is removing some of the data points before γ evaluation. As stated by the manufacturer, if there is a planned point that is higher than the 10% threshold but there is no corresponding measured point at the same position, then that point is not included in the analysis and vice versa.

b) In SNC Patient™, a measurement uncertainty correction factor was applied (as mentioned in section 4.4.3). Recently this factor was studied in detail and it was reported that it can inflate γ pass rates by as much as 9-14% [95]. The authors recommended that it's use be clearly indicated, especially in published reports. Interestingly, when not applied, the average γ pass rate with 2%/2mm criterion for Mapcheck single decreases to 95.3% ± 2.3% with a confidence interval of 92.6%, which is highly comparable to the results from OmniPro single.

	Number of γ evaluation points				
Analysis software	Field 1	Field 2	Field 3	Field 4	Field 5
Mapcheck single	396	328	389	479	389
OmniPro single	1142	1081	1181	1329	1185
Mapcheck double	787	656	774	948	778
OmniPro 1mm	5753	5408	5895	6705	5974

Table 14: Per field number of γ evaluation points from each evaluation method. "OmniPro single" has approximately 3x more evaluation points compared to "Mapcheck single".

It has been noted that a difference in γ passing rates can occur between individual software packages and even using different versions of the same software. A call was done for manufacturers of these software packages to disclose software implementation details [96], addressed recently in AAPM TG-218 [97]. A vendor survey questionnaire on the implementation of IMRT QA analysis software from 8 different vendors were reported. From this, SNC[TM] is using a DTA search radius of 8 mm by default (not customizable) whereas the value is customizable in OmniPro. However, recalculating OmniPro results with an updated DTA search radius (from 3 mm to 8 mm) did not change the γ pass rates obtained in the current study.

Recommendations of the AAPM Task Group No. 218, on tolerance limits for IMRT measurement-based verification QA, are as follows: a tolerance limit on γ pass rates of ≥ 95% and a universal action limit on γ pass rates of ≥ 90%, using a 3%/2 mm criterion and a 10% dose threshold. Results from the current study show an average γ pass rate (from all the evaluation methods and Linacs combined) of 97.3%, ranging between 96.9% (Linac 3) and 97.7% (Linac2), on a 2%/2mm criterion. Therefore, based on these tolerance recommendations, IMRT dose calculations from the single TPS model was successfully verified on all 5 Linacs, even with a tighter γ criterion of 2%/2mm.

Evaluating the effect of increased measurement resolution from OmniPro I'mRT software only (from "OmniPro single" to "OmniPro 1mm"), an increase in mean γ pass rate and confidence interval of 2.4% and 4.4% respectively can be seen. Although this result seems noteworthy, it is statistically insignificant (p<0.0001).

Similar results have been reported: increasing film resolution by a factor 4, using the same software package, increased the γ pass rate by 4.3% for 2%/2mm [48]. Also, in evaluating 28 clinical IMRT fields, an average γ pass rate of 95.5±4% (2%/2mm criterion) have been reported for the same TPS but a different Linac and measurement device [82]. This result is very comparable to the "OmniPro single" evaluation method in the current study, with the lowest average pass rate of 95.6%.

Results obtained in this study agree with clinically acceptable published results discussed below, including the limitation of the IMRT QA method: planar measurements at nominal gantry angles using a γ criterion of 3%/3mm might not detect treatment delivery errors [98] [99] [100] [94] [101] [102].

In comparing IMRT results from 7 different sites, an overall average γ pass rate of 97.9% (using a 3%/3mm criterion) with a standard deviation of 2.5% and confidence limit of 7.0% was reported [24]. In that study 5 of the 7 sites also used the Mapcheck measurement device. A IROC Houston phantom study, a complete plan and treat audit, done at 855 different institutions showed that planar IMRT QA was a poor tool of predicting failure of the phantom results [98]. An institutional failure rate of only 0.6% was reported, whereas the IROC Houston phantom reported a 14% failure rate. Their study also showed the low sensitivity of the Mapcheck device in detecting IMRT QA failures and recommended an increased threshold on overall γ pass rate of 97% using a 3%/3mm criterion. Although not part of the current study design, an overall agreement within 3% was obtained irradiating the IROC Houston phantom on Linac3 and comparing the thermo luminescent detector measurements to doses calculated from the reference TPS model.

Another more stringent γ pass rate tolerance for the use of the Mapcheck device for per field QA at nominal angle, including a 95% confidence interval, was set at 96 ± 4.6% (3%/3mm) and 85% ± 8% (2%/2mm) [103]. These recommendations for increased tolerances are based on a sensitivity (the ability of the dosimeter to accurately label an unacceptable plan as failing) and a specificity (the ability to label an acceptable plan as passing) value calculated for the Mapcheck device. Therefore increasing γ pass rate tolerances, like a criterion of 2%/2mm, can be beneficial to distinguish

between clinically acceptable and unacceptable IMRT plans based on planar IMRT QA evaluations [99] [81].

From the study by Kruse (2010), clinically unacceptable plans generally had an average γ passing percentage of less than 90% using an ion chamber array with 2%/2mm criterion. In the current study a mean γ pass rate of 95.6% (worst case from 4 evaluation methods) were obtained using the same criterion. Their study further showed that effective patient specific IMRT QA should still include a single measurement of composite dose in the complete plan, not done in the current study.

Contradictory to results obtained in this study, for the same Linac and measurement device, but calculated on different TPSs, Yao and Farr (2015) reported lower average IMRT pass rates. With the Pinnacle and Eclipse TPSs an average γ pass rate for criterion 3%/3mm of 94.0% and 97.7% and for criterion 2%/2mm of 78.8% and 88.3% was obtained respectively [62]. However, in that study the Van Dyk dose comparison function (discussed in section 4.4.3) was not applied, which explains the lower γ pass rates achieved. They concluded to achieve higher dosimetry agreement or IMRT QA pass rates with Siemens® 160 MLC^{TM}, small fields should be avoided, segment widths be larger than 2cm and T&G extensions be less than 1 cm. However, the purpose of that study was to determine the optimal dosimetric leaf gap (DLG), a parameter related to the gap between light and radiation fields and more intrinsic to other vendor Linacs (e.g. Clinac® linear accelerator, TrueBeam® system) in combination with the TPSs mentioned. These TPSs model the MLC ends as square and applies the DLG to account for leakage through rounded leaf ends. Nevertheless, they reported the slightly tilted MLCs of the Siemens Artiste result in unbalanced transmission ratios along the tilt direction and this increase uncertainties in the dose delivered.

Therefore, it cannot be excluded that, to obtain high accuracy for small IMRT field treatments (smaller than the fields used in the current study), the TPS model may need to be individually adjusted for each Linac respectively. This should be further investigated. Small IMRT fields could be evaluated using the high measurement resolution method used in the current study (1mm spacing between measurement points). However, it could also be that an inaccuracy in the dose delivered for small

fields is from a physical Linac limitation (due to the unbalanced nature of the MLC's) that cannot be accurately accounted for or improved in the TPS model.

The ability of γ analysis in detecting specifically MLC errors have been questioned. A weak correlation between γ pass rates and MLC errors of 0.5 up to 2 mm have been reported for planar measurements [104] [105]. Similarly, from the current study, significant differences in MLC positional accuracy between Linacs were observed, but the difference in IMRT QA pass rates between Linacs are not significant. However, a decrease in average γ pass rate could be seen for Linac1, with the worst performing MLC's. Also, it could be that the lower pass rates of Linac5 is a consequence of the large overtravel of the x1 bank at CAX position. On the contrary, it has been reported that adjustments of MLC parameters in a TPS model, for example the MLC offset value (with an outcome similar to MLC positional errors), had relatively little effect on a clinical IMRT field [69]. However, the authors did note that the effect can vary, and possibly become more significant, as it is dependent on the plans average gap width; with small average gap width plans being more sensitive to MLC offset values.

Therefore, in future work special attention should be given to the shape and size of individual IMRT segments. A good start off point would be to investigate failing γ criterion points [81] and possibly relate these to the off-axis position of the MLC bank and to the location of individual, and adjacent, MLCs within these segments.

Other studies highlighted the importance of the location and overlap of these per-beam errors in terms of critical volumes (targets and organs at risk) and not about per-beam passing rates in a phantom [100] [106]. It was concluded that steps should be taken to analyse where the per-beam errors overlap in 3D space in relation to critical structures, something that can possibly be used in future to improve the QA methodology used in the current study. For H&N IMRT delivery on a Varian 120 MLC, a requirement for MLC tolerance of 0.3 mm was reported in order to meet a 2% dose deviation in the CTV [107]. Further, random errors of up to 2 mm was shown tolerable, although the incidence thereof was not reported. Nevertheless, the importance of per field γ analysis when commissioning a TPS model was mentioned as a useful tool to compare and optimize measured vs planned doses. In the current study it was a useful tool to further compare Linacs in terms of treatment delivery.

6. CONCLUSION

A method of using a 2D array detector for TPS model verification was evaluated. Results showed with increased resolution the sensitivity of the array are comparable to a microDiamond detector for small field profiles, with most γ values ≤ 1 using a 1%/1mm criterion. High resolution dose measurements are beneficial to compare beam-matched Linacs, specifically for accurate MLC field size measurements and subsequent quantification of MLC errors. A range of MLC errors from zero up to 1.65 mm was measured. A general non-linearity in MLC stop positions was shown and differences between MLCs from 5 beam-matched Linacs quantified. These differences were significant and new, unified MLC software correction values proposed. Gamma analysis were performed using different commercial software with decreased γ criterion (2%/2mm), thereby increasing confidence in IMRT QA results obtained. An average γ percentage pass rate of at least 95.6% with a confidence limit of 92.5% was obtained. These results showed that a single MC beam model can be used for IMRT treatment planning for multiple beam-matched linear accelerators. Even though differences could be seen in γ pass rates between Linacs and analysing software, and potential causes therefore discussed, increasing measurement resolution for IMRT dose verification up to 1mm and thereby removing measurement interpolation points, has proven unnecessary as it does not change γ pass rates significantly. Also, the dosimetric accomplishment of the method evaluated may be weakened by acquisition and processing workload. Even with increased measurement resolution, γ pass rates failed to detect seemingly significant MLC errors. Further investigation is required to possibly relate MLC errors to individual IMRT segments delivered and to estimate the clinical significance thereof.

REFERENCES

1. ICRP Publication 9. *Recommendations of the International Commission on Radiological Protection.*; 1965.

2. Van Dyk J, Rosenwald J, Fraass B, Cramb J, Ionescu-Farca F. *IAEA, International Atomic Energy Agency. Technical Report Series No. 430: Commissioning and Quality Assurance of Computerized Planning Systems for Radiation Treatment of Cancer.* Vienna, Austria; 2004.

3. Brahme A. Dosimetric precision requirements in radiation therapy. *Acta Oncol (Madr).* 1984;23(5):379-391. doi:10.3109/02841868409136037.

4. Mijnheer BJ, Battermann JJ, Wambersie A. What degree of accuracy is required and can be achieved in photon and neutron therapy? *Radiother Oncol.* 1987;8(3):237-252. doi:10.1016/S0167-8140(87)80247-5.

5. Mayles W, Lake R, McKenzie A, Macaulay E, Morgan H, Jordan T, et al. *Physics Aspects of Quality Control in Radiotherapy (Report No. 81).*; 1999. doi:10.1088/0031-9155/45/3/501.

6. Rock L, Battista JJ, Regional L, Centre C, Boyer AL, Kappas C, et al. *American Association of Physicists in Medicine Radiation Therapy Committee Task Group 65: Tissue Inhomogeneity Corrections for Megavoltage Photon Beams: AAPM Report No. 85.*; 2004.

7. Gagneur JD, Ezzell GA. An improvement in IMRT QA results and beam matching in linacs using statistical process control. *J Appl Clin Med Phys.* 2014;15(5):190-195.

8. Sjölin M, Edmund JM. Incorrect dosimetric leaf separation in IMRT and VMAT treatment planning: Clinical impact and correlation with pretreatment quality assurance. *Phys Medica.* 2016;32:918-925.

9. Podgorsak EB, Andreo P, Evans MDC, Hendry JH, Horton JL, Izewska J, et al. Review of Radiation Oncology Physics : A Handbook for Teachers and Students. In: *Chapter 6.* ; 2004:133-178.

10. Fraass B, Doppke K, Hunt M, Kutcher G, Starkschall G, Stern R, et al. American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: Quality assurance for clinical radiotherapy treatment planning. *Med Phys.* 1998;25(10):1773.

11. Bhangle JR, Narayanan VKS, Kumar NK, Vaitheeswaran R. Dosimetric analysis of beam-matching procedure of two similar linear accelerators. *J Med Phys*. 2011;36(3):176-180.

12. Hrbacek J, Depuydt T, Nulens A, Swinnen A, Van den Heuvel F. Quantitative evaluation of a beam-matching procedure using one-dimensional gamma analysis. *Med Phys*. 2007;34(7):2917.

13. Das IJ, Cheng C-W, Watts RJ, Ahnesjö A, Gibbons J, Li XA, et al. Accelerator beam data commissioning equipment and procedures: Report of the TG-106 of the Therapy Physics Committee of the AAPM. *Med Phys*. 2008;35(9):4186.

14. Spezi E, Lewis G. An overview of Monte Carlo Treatment Planning for Radiotherapy. *Radiat Prot Dosimetry*. 2008;131(1):123-129.

15. Fraas BA, Smathers J, Deye J. Summary and recommendations of a National Cancer Institute workshop on issues limiting the clinical use of Monte Carlo dose calculation algorithms for megavoltage external beam radiation therapy. *Med Phys*. 2003;30(12):3206-3216.

16. Reynaert N, van der Marck SCC, Schaart DRR, Van der Zee W, Van Vliet-Vroegindeweij C, Tomsej M, et al. Monte Carlo treatment planning for photon and electron beams. *Radiat Phys Chem*. 2007;76(4):643-686.

17. Fippel M, Haryanto F, Dohm O, Nüsslin F, Kriesen S, Nüsslin F, et al. A virtual photon energy fluence model for Monte Carlo dose calculation. *Med Phys*. 2003;30(3):301.

18. Chetty IJ, Curran B, Cygler JE, DeMarco JJ, Ezzell G, Faddegon B a., et al. Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning. *Med Phys*. 2007;34(12):4818.

19. Sikora MP. Virtual Source Modelling of Photon Beams for Monte Carlo Based Radiation Therapy Treatment Planning. 2011.

20. McCullough EC, Krueger AM. Performance Evaluation of Computerized Treatment Planning Systems for Radiotherapy: External Photon Beams. *Int J Radiat Oncol Biol Phys*. 1980;6(11):1599-1605.

21. Gifford KA, Followill DS, Liu HH, Starkschall G. Verification of the accuracy of a photon dose-calculation algorithm. *J Appl Clin Med Phys*. 2002;3(1):26-45.

22. Van Dyk J, Barnett R, Gygler J, Shragge P. Commissioning and Quality Assurance of Treatment Planning Comupters. *Int J Radiat Oncol Biol Phys*. 1993;261:261-273.

23. Jamema S, Upreti R, Sharma S, Deshpande D. Commissioning and comprehensive quality assurance of commercial 3D treatment planning system using IAEA Technical Report Series-430. *Australas Phys Eng Sci Med*. 2008;31(3):207-215.

24. Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys*. 2009;36(11):5359.

25. Sjöström D, Bjelkengren U, Ottosson W, Behrens CF. A beam-matching concept for medical linear accelerators. *Acta Oncol*. 2009;48(2):192-200. doi:10.1080/02841860802258794.

26. Marshall MG. Matching the 6MV photon beam characteristics of two dissimilar linear accelerators. *Med Phys*. 1993;20(6):1743.

27. Venselaar J, Welleweerd H, Mijnheer B. Tolerances for the accuracy of photon beam dose calculations of treatment planning systems. *Radiother Oncol*. 2001;60(2):191-201.

28. Born E, Fogliata-Cozzi A, Ionescu F, Ionescu V, Tercier P-A. *SSRMP, Swiss Society of Radiobiology and Medical Physics. Quality Control of Treatment Planning Systems for Teletherapy: Recommendations No. 7*.; 1999.

29. Attalla EM, Abou-Elenein H, Ammar H, Eldesoky I. Dosimetric evaluation of a beam matching procedure. *Chinese-German J Clin Oncol*. 2014;13(2):89-93.

30. Li JG, Yan G, Liu C. Comparison of two commercial detector arrays for IMRT quality assurance. *J Appl Clin Med Phys*. 2009;10(2):62-74.

31. Alber M, Broggi S, De Wagter C, Esichwuzel I, Engström P, Fiorino C, et al. QUASIMODO, Quality Assurance of Intensity Modulated Radiation Oncology. ESTRO Booklet No. 9: Guidelines for the verification of IMRT. In: Belguim, Brussels; 2008.

32. Klein EE, Hanley J, Bayouth J, Yin F-F, Simon W, Dresser S, et al. American Association of Physicists in Medicine Radiation Therapy Committee Task Group 142 report: Quality assurance of medical accelerators. *Med Phys*. 2009;36(9):4197.

33. Chen Z, D'Errico F, Nath R. Principles and requirements of external beam dosimetry. *Radiat Meas*. 2007;41(SUPPL. 1):2-21.

34. Laub WU, Wong T. The volume effect of detectors in the dosimetry of small fields used in IMRT. *Med Phys*. 2003;30(3):341-347.

35. Yorke E, Alecu R, Ding L, Fontenla D, Kalend A, Kaurin D, et al. *American Association of Physicists in Medicine Radiation Therapy Committee Task Group 62: Diode in Vivo Dosimetry for Patients Receiving External Beam Radiation Therapy, AAPM Report No. 87*.; 2005.

36. Haryanto F, Fippel M, Laub W, Dohm O, Nüsslin F. Investigation of photon beam output factors for conformal radiation therapy - Monte Carlo simulations and measurements. *Phys Med Biol*. 2002;47(11):N133-N143.

37. Ade N, Nam TL. The influence of detector size relative to field size in small-field photon-beam dosimetry using synthetic diamond crystals as sensors. *Radiat Phys Chem*. 2015;113:6-13.

38. Yan G, Fox C, Liu C, Li JG. The extraction of true profiles for TPS commissioning and its impact on IMRT patient-specific QA. *Med Phys*. 2008;35(8):3661-3670.

39. Spezi E, Angelini A, Romani F, Ferri A. Characterization of a 2D ion chamber array for the verification of radiotherapy treatments. *Phys Med Biol*. 2005;(50):3361-3371.

40. Stelljes TS, Harmeyer A, Reuter J, Looe HK, Chofor N, Harder D, et al. Dosimetric characteristics of the novel 2D ionization chamber array OCTAVIUS Detector 1500. *Med Phys*. 2015;42(4):1528-1537.

41. Bushberg JT, Deibert JA, Leidholdt EM, Boone J. *The Essential Physics of Medical Imaging*. 3rd editio. Philadelphia: Lippincott Williams & Wilkens; 2012.

42. Poppe B, Djouguela A, Harder D, Kollhoff R. Spatial resolution of 2D ionization chamber arrays for IMRT dose verification : single-detector size and step width. *Phys Med Biol*. 2007;(52):2921-2935.

43. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys*. 1998;25(5):656-661.

44. Schreiner LJ, Holmes O, Salomons G. Analysis and evaluation of planned and delivered dose distributions: Practical concerns with γ- and χ- evaluations. *J Phys Conf Ser*. 2013;444(1). doi:10.1088/1742-6596/444/1/012016.

45. Gago-Arias A, Brualla-Gonzalez L, Gonzalez-Castano D, Gomez F, Sanchez Garcia M, Luna Vega V, et al. Evaluation of chamber response function influence on IMRT verification using 2D commercial detectors. *Phys Med Biol*. 2012;57:2005-2020.

46. Buonamici FB, Compagnucci A, Marrazzo L, Russo S, Bucciolini M. An intercomparison between film dosimetry and diode matrix for IMRT quality assurance. *Med Phys*. 2007;34(4):1372-1379.

47. Low DA, Moran JM, Dempsey JF, Dong L, Oldham M. Dosimetry tools and techniques for IMRT. *Med Phys*. 2011;38(3):1313-1338.

48. Huang JY, Pulliam KB, McKenzie EM, Followill DS, Kry SF. Effects of spatial resolution and noise on gamma analysis for IMRT QA. *J Appl Clin Med Phys*. 2015;15(4).

49. Herzen J, Todorovic M, Cremers F, Platz V, Albers D, Bartels A, et al. Dosimetric evaluation of a 2D pixel ionization for implementation in clinical routine. *Phys Med Biol*. 2007;52:1197-1208.

50. Poppe B, Blechschmidt A, Djouguela A, Kollhoff R, Rubach A, Willborn KC, et al. Two-dimensional ionization chamber arrays for IMRT plan verification. *Med Phys*. 2006;33(4):1005-1015.

51. Ayouth JE. Siemens Mulileaf Collimator Characterization and Quality Assurance Approaches for Intensity-Modulated Radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008;71(1 (supplement)):93-97.

52. Wang S, Li Z, Chao KSC, Chang J. Calibration of a detector array through beam profile reconstruction with error-locking. *J Appl Clin Med Phys*. 2014;15(6):13-29.

53. Wang S, Chao K, Chang J. Inaccurate Positioning Might Introduce Significant Calibration Errors to a Detector-Array QA Device for Flatten Filter Free Beams. *Austin J Radiat Oncol Cancer*. 2015;1(2):1-5.

54. Van Esch A, Huyskens DP, Basta K, Ghislain M, Delvaux R. The quest for sensible data analysis in clinical routine: Study case on the new Octavius1500 array and its associated phantoms. *8th Int Conf 3D Radiat Dosim (IC3DDose), J Phys*. 2015;Series 573.

55. Van Esch A, Basta K, Evrard M, Ghislain M, Sergent F, Huyskens DP. The Octavius1500 2D ion chamber array and its associated phantoms: Dosimetric characterization of a new prototype. *Med Phys*. 2014;41(9).

56. Sumida I, Yamaguchi H, Kizaki H, Koizumi M, Ogata T, Takahashi Y, et al. Quality assurance of MLC leaf position accuracy and relative dose effect at the MLC abutment region using an electronic portal imaging device. *J Radiat Res*. 2012;53:798-806.

57. Van Esch A, Bohsung J, Sorvari P, Tenhunen M, Paiusco M, Iori M, et al. Acceptance tests and quality control (QC) procedures for the clinical implementation of intensity modulated radiotherapy (IMRT) using inverse planning and the sliding window technique: Experience from five radiotherapy departments. *Radiother Oncol*. 2002;65:53-70.

58. Letourneau D, Gulam M, Yan D, Oldham M, Wong JW. Evaluation of a 2D diode array for IMRT quality assurance. *Radiother Oncol*. 2004;70:199-206.

59. Moreno RD, Venencia D, Garrigo E. A method to enhance spatial resolution of a 2D ion chamber array for quality control of MLC. *J Appl Clin Med Phys*. 2011;12(4):63-73.

60. Graves M, Thompson A, Martel M, McShan D, Fraass B. Calibration and quality assurance for rounded leaf-end MLC systems. *Med Phys*. 2001;28(11):2227-2233.

61. Tacke MB, Nill S, Häring P, Oelfke U. 6 MV dosimetric characterization of the 160 MLC ™ , the new Siemens multileaf collimator. *Med Phys*. 2008;35(5):1634-1642.

62. Yao W, Farr JB. Determining the optimal dosimetric leaf gap setting for rounded leaf-end multileaf collimator systems by simple test fields. *J Appl Clin Med Phys*. 2015;16(4):65-77.

63. Prah E, Kainz K, Peng C, Li XA. The Dosimetric and Delivery Advantages of a New 160-leaf MLC. *Technol Cancer Res Treat*. 2011;10(3):219-229.

64. Klüter S, Sroka-Perez G, Schubert K, Debus J. Impact of Beam Alignment on Leakage of the Siemens 160 MLC ™. In: *IFMBE Proceedings*. ; 2009:496-497.

65. Siemens Digital Linear Accelarator: Physics Primer. 2014:133-135.

66. IMPAC Medical Systems I. Monaco Technical Reference: Post Modeling Adjustment of MLC Parameters. 2012:Document ID: LRMMON0003.

67. Sikora M, Dohm O, Alber M. A virtual photon source model of an Elekta linear accelerator with integrated mini MLC for Monte Carlo based IMRT dose calculation. *Phys Med Biol*. 2007;52(15):4449-4463.

68. IMPAC Medical Systems I. Monaco Dose Calculation Technical Reference. 2013:Document ID: LRMMON0001.

69. Kinsella P, Shields L, Mccavana P, Mcclean B, Langan B. Determination of MLC model parameters for Monaco using commercial diode arrays. *J Appl Clin Med Phys.* 2016;17(4):37-47.

70. Sun Nuclear Corporation. MapCHECK 2 Reference Guide Raising the Bar in IMRT QA. 2017.

71. Blanck O, Masi L, Chan MKH, Adamczyk S, Albrecht C, Damme M, et al. High resolution ion chamber array delivery quality assurance for robotic radiosurgery: Commissioning and validation. *Phys Medica.* 2016;32:838-846.

72. Poppe B, Stelljes TS, Looe HK, Chofor N, Harder D, Willborn K. Performance parameters of a liquid filled ionization chamber array. *Med Phys.* 2013;40(8).

73. Ciancaglioni I, Marinelli M, Milani E, Prestopino G, Verona-Rinati C, Verona-rinati G, et al. Dosimetric characterization of a synthetic single crystal diamond detector in clinical radiation therapy small photon beams. *Med Phys.* 2013;39(7):4493-4501.

74. Laub WU, Crilly R. Clinical radiation therapy measurements with a new commercial synthetic single crystal diamond detector. *J Appl Clin Med Phys.* 2014;15(6):92-102.

75. Andreo P, Burns D, Hohlfed K, Huq M, Kanai T, Laitona F, et al. *IAEA, International Atomic Energy Agency. Technical Report Series No. 398: Absorbed Dose Determination in External Beam Radiotherapy: An International Code of Practice for Dosimetry Based on Standards of Absorbed Dose to Water.* Vienna, Austria; 2004.

76. Siemens Medical Solutions. 550 TxT Treatment Table: Accuracy and Strength for Advanced Therapy Applications. 2005.

77. Kantz S, Troeller Mcdermott A, Söhn M, Reinhardt S, Belka C, Parodi K, et al. Practical implications for the quality assurance of modulated radiation therapy techniques using point detector arrays. *J Appl Clin Med Phys.* 2017;18(6):20-31. doi:10.1002/acm2.12157.

78. Gordon JD, Krafft SP, Jang S, Smith-Raymond L, Stevie MY, Hamilton RJ. Confidence limit variation for a single IMRT system following the TG119 protocol. *Med Phys.* 2011;38(3):1641.

79. Lopez-Tarjuelo J, Garcia-Molla R, Juan-Senabre XJ, Quiros-Higueras JD, Santos-Serra A, de Marco-Blancas N, et al. Acceptance and commissioning of a treatment planning system based on Monte Carlo calculations. *Technol Cancer Res Treat*. April 2013.

80. Narayanasamy G, Saenz DL, Defoor D, Papanikolaou N, Stathakis S. Dosimetric validation of Monaco treatment planning system on an Elekta VersaHD linear accelerator. *J Appl Clin Med Phys*. 2017;18(6):123-129. doi:10.1002/acm2.12188.

81. Smilowitz JB, Das IJ, Feygelman V, Fraass BA, Kry F, Marshall IR, et al. AAPM Medical Physics Practice Guideline 5 . a .: Commissioning and QA of Treatment Planning Dose Calculations — Megavoltage Photon and Electron Beams. *J Appl Clin Med Phys*. 2015;16(5):14-34. doi:10.1120/jacmp.v16i5.5768.

82. Grofsmid D, Dirkx M, Marijnissen H, Woudstra E, Heijmen B. Dosimetric validation of a commercial Monte Carlo based IMRT planning system. *Med Phys*. 2010;37(2):540-549. doi:10.1118/1.3284359.

83. Court LE, Tishler R, Xiang H, Allen AM, Makrigiorgos M, Chin L. Experimental evaluation of the accuracy of skin dose calculation for a commercial treatment planning system. *J Appl Clin Med Phys*. 2008;9(1):2792.

84. Ishmael Parsai E, Shvydka D, Pearson D, Gopalakrishnan M, Feldmeier JJ. Surface and build-up region dose analysis for clinical radiotherapy photon beams. *Appl Radiat Isot*. 2008;66(10):1438-1442. doi:10.1016/j.apradiso.2008.02.089.

85. Hsu S-H, Moran JM, Chen Y, Kulasekere R, Roberson PL. Dose discrepancies in the buildup region and their impact on dose calculations for IMRT fields. *J Med Phys*. 2010;37(5). doi:10.1118/1.3377769.

86. Venselaar J, Welleweerd H. Application of a test package in an intercomparison of the photon dose calculation performance of treatment planning systems used in a clinical setting. *Radiother Oncol*. 2001;60:203-213.

87. Fippel M, Haryanto F, Dohm O, Nüsslin F, Kriesen S. A virtual photon energy fluence model for Monte Carlo dose calculation. *Med Phys*. 2003;30(3):301-311.

88. Kry SF, Bednarz B, Howell RM, Dauer L, Followill D, Klein E, et al. AAPM TG 158: Measurement and calculation of doses outside the treated volume from external-beam radiation therapy. *Med Phys*. 2017;44(10):e391-e429. doi:10.1002/mp.12462.

89. Simon TA, Simon WE, Kahler D, Li J, Liu C. Wide field array calibration dependence on the stability of measured dose distributions. *Med Phys.* 2010;37(7):3501-3509.

90. Antypas C, Floros I, Rouchota M, Armpilia C, Lyra M. MLC positional accuracy evaluation through the Picket Fence test on EBT2 films and a 3D volumetric phantom. *J Appl Clin Med Phys.* 2015;16(2):189-197.

91. Lopes MC, Chaves A, Capela M. A dosimetric calibration method for a double-focused multileaf collimator. *Med Phys.* 2007;34(9):3473-3474.

92. Kluter S, Sroka-Perez G, Schubert K, Debus J. Leakage of the Siemens 160 MLC multileaf collimator on a dual energy linear accelerator. *Phys Med Biol.* 2011;56:N29–N37.

93. Tacke M. Adaptation of High-Precision Radiotherapy to Moving Target Volumes in Real-Time Using Dynamic Multileaf Collimators. 2009.

94. Hussein M, Rowshanfarzad P, Ebert MA, Nisbet A, Clark CH. A comparison of the gamma index analysis in various commercial IMRT / VMAT QA systems. *Radiother Oncol.* 2013;109(3):370-376.

95. Bailey DW, Spaans JD, Kumaraswamy LK, Podgorsak MB. The MapCHECK Measurement Uncertainty function and its effect on planar dose pass rates. *Med Phys.* 2016;17(2):165-173.

96. Hussein M, Clementel E, Eaton DJ, Greer PB, Haworth A, Ishikura S, et al. A virtual dosimetry audit – Towards transferability of gamma index analysis between clinical trial QA groups. *Radiother Oncol.* 2017;125(3):398-404. doi:10.1016/j.radonc.2017.10.012.

97. Miften M, Olch A, Mihailidis D, Moran J, Pawlicki T, Molineu A, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA : Recommendations of AAPM Task Group No . 218. *Med Phys.* 2018. doi:https://doi.org/10.1002/mp.12810.

98. Kry SF, Molineu A, Kerns JR, Faught AM, Huang JY, Pulliam KB, et al. Institutional Patient-specific IMRT QA Does Not Predict Unacceptable Plan Delivery. *Int J Radiat Oncol Biol Phys.* 2014;90(5):1195-1201.

99. Kruse JJ. On the insensitivity of single field planar dosimetry to IMRT inaccuracies. *Med Phys.* 2010;37(6):2516-2525.

100. Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys.* 2011;38(2):1037-1044.

101. Pulliam KB, Followill D, Court L, Dong L, Gillin M, Prado K, et al. A six-year review of more than 13 , 000 patient-specific IMRT QA results from 13 different treatment sites. 2014;15(5):196-206.

102. Stojadinovic S, Ouyang L, Gu X, Pompoš A, Bao Q, Solberg TD. Breaking bad IMRT QA practice. 2015;16(3):154-165.

103. McKenzie EM, Balter P a, Stingo FC, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans. *Med Phys*. 2014;41(12):121702.

104. Wang J, Jin X, Peng J, Xie J, Chen J, Hu W. Are simple IMRT beams more robust against MLC error? Exploring the impact of MLC errors on planar quality assurance and plan quality for different complexity beams. *J Appl Clin Med Phys*. 2016;17(3):147-157. doi:10.1120/jacmp.v17i3.6022.

105. Shang Q, Godley A, Huang L, Qi P, Xia P. Sensitivity of array detector measurements in determining shifts of MLC leaf positions. *J Appl Clin Med Phys*. 2017;18(5):80-88. doi:10.1002/acm2.12148.

106. Song J, Kim Y-H, Jeong J-U, Yoon MS, Ahn S-J, Chung W-K, et al. Dosimetric evaluation of MapCHECK 2 and 3DVH in the IMRT delivery quality assurance process. *Med Dosim*. 2014;39:134-138.

107. Rangel A, Dunscombe P. Tolerances on MLC leaf position accuracy for IMRT delivery with a dynamic MLC. 2009;3304(36 (7)). doi:10.1118/1.3134244.

AKNOWLEDGEMENTS

This thesis was a very rewarding process of constant learning, both professionally and personally. I would like to thank each and everyone that showed interest in, supported or impacted my work.

To my supervisor, dr Willie Shaw, my gratitude for all the hard work and advice he has put into this study. The knowledge I have gained and role you have played in improving my skills are of great value.

Thank you to my co-supervisor as well, dr Feek du Plessis, for your input in this manuscript.

Thank you to Lourens Strauss for writing the code for "Mapcheck Combine" and for sharing your thoughts and ideas.

I would like to thank my lovely wife, Hiloise for your encouragement and support throughout, amid our two beautiful children, Louise and Roulph being born during this time. You are my motivation in everyday life and I love you all very much.

I am also thankful to my colleague, mentor and friend, Hennie Smit, for years of support and for promoting the essence of logical thinking in the field of medical physics.