A method to optimize patient specific HU-SP calibration curve from proton range measurements.

Jaroslav Albert\textsuperscript{1}, Rudi Labarbe\textsuperscript{2}, Edmond Sterpin\textsuperscript{3,4,5}

\textsuperscript{1}Université Catholique de Louvain, ELEN Place du Levant 2/L5.04.04 1348 Louvain-la-Neuve, Louvain-la-Neuve, Belgium
E-mail: jaroslav.albert@uclouvain.be

\textsuperscript{2}IBA Particle Therapy, Chemin du Cyclotron 3, B-1348 Louvain-la-Neuve, Belgium
E-mail: rudi.labarbe@iba-group.com

\textsuperscript{3}Université Catholique de Louvain, Molecular Imaging Radiotherapy and Oncology, Brussels, Belgium
\textsuperscript{4}Université Catholique de Louvain, Institut de Recherche Expérimentale et Clinique, Brussels, Belgium
\textsuperscript{5}KU Leuven, Department of Oncology, Laboratory of Experimental Radiotherapy, Leuven, Belgium
E-mail: edmond.sterpin@uclouvain.be

Abstract. Proton radiotherapy promises accurate dose delivery to a tumor and minimal dose deposition to all other tissues. However, in practice the planned dose distribution may not conform to the actual one due to noisy data and different types of errors. One such error comes in a form of potentially inaccurate conversion of the Hounsfield units (HU) to stopping powers (SP) of protons. We propose a method of improving the CC based on a planning CT and proton range measurements acquired during treatment. The range data were simulated using a virtual CC and a planning
The range data were given two types of noise: range shift due to patient setup errors; and range noise due to measurement imprecision, including a misalignment of the range measuring device. The method consists of two parts. The first part involves a Taylor expansion of the water equivalent path length (WEPL) map in terms of the range shift caused by the difference between the planning and the virtual CC. The range shift is then solved for explicitly, leading to a polynomial function of the difference between the two CCs. The second part consists in minimizing a score function relating the range due to the virtual CC and the range due to the optimized CC. Tested on ten different CCs, our results show that, with range data collected over a few fractions (less than 10), the optimized CC leads to an overall reduction of the range difference. More precisely, on average, the uncertainty of the CC was reduced from 2.67% to 1.62%, while the average reduction of the WEPL bias was reduced from 2.14% to 0.74%. The advantage of our method over others is 1) its speed, and 2) the fact that the range data it necessitates are acquired during the treatment itself, and as such it does not burden the patient with additional dose.

*Keywords:* Proton radiotherapy, calibration curve, stopping power, proton range
1. Introduction

Proton radiotherapy has the potential of delivering precise radiation dose to tumors while sparing healthy tissue more than conventional radiotherapy, especially in the distal area behind the tumor. Hindering this potential is protons’ sensitivity to errors, which may include changes in the anatomy due to inter/intra-fraction motion [1, 2], patient set up errors [3], but also imperfect translation of patient image data into dosimetric quantities. The latter includes: noise in the Hounsfield Units (HUs) of the CT scan [4, 5], the conversion of the HUs to stopping power (SP) [6] and artifacts in the CT. Although these errors may be small, their collective contribution can lead to a significant over or under-range inside a patient.

The HUs can be converted into SPs through various methodologies, either via single [7, 8, 9, 10, 11] or dual/multi energy [12, 13, 14] CT scanners. Both of these methods rely on a stoichiometric approximation [7, 10, 15] relating the HUs to the relative stopping power (RSP), which is defined as the stopping power (SP) of a medium (e. g. tissue) divided by the SP of water. However, uncertainties in the RSP remain due to inaccurate tissue segmentation (for instance, because of the noise) and inherent uncertainties linked to the mean excitation energies (I value) needed to compute stopping powers [16].

A common way to relate the RSP to HUs is via a calibration curve (CC). In recent years, a patient-specific approach to estimating the CC has been in development, one in which range data from an individual patient are collected via proton radiography (PR). In this approach, a patient is irradiated with protons of sufficient energy to traverse the patient and their Bragg profiles are detected by either a Multi-layer ionization chamber (MLIC), an integrating detector or a proton-by-proton counting device. From these profiles, it is possible to optimize the stoichiometric CC in order to be more suitable for the patient [17]. Alternative to the PR is a another invention for measuring the actual
(clinical) range inside a patient [18, 19, 20, 21, 22, 23]. Known as Prompt gamma camera (PGC), this device detects gamma radiation coming from nuclear reactions between the protons and the atomic nuclei whereby allowing one to reconstruct the Bragg profile from which the range can be determined. There exist other devices/methods, such as PET scan and ioniacoustic measurements [24, 25], that are able to measure, among other quantities, the proton range. In this paper, we do not focus on any specific device; we merely take as granted that range data can be collected during a treatment session.

What we are proposing is a method of optimizing the patient-specific CC that incorporates range measurements collected during a proton radiotherapy session. The method involves two main steps. The first step is dependent on the planning CT only, and can be carried out before the range measurements. Consequently, the second step, which leads to the improvement of the CC and is performed as soon as the range data are collected, can be performed in a matter of seconds. In order to alleviate the problem of multiple minima, engendered by noise/imperfect data, we place constraints on the way in which the CC can be adjusted. First, we partition the CC into sections corresponding to the highest frequency HU. Secondly, we restrict the deviations from the stoichiometric CC to be within 5% and require that the optimized CC increase monotonically.

The novelty of this method is two fold. The first is its speed. While a more accurate approach would involve a Monte Carlo simulation of the pencil beams (PBs), the time frame for such an approach would not be practical, as the input of the simulation is not one or two parameters, but an entire function, i.e. the CC. The second is its adaptability to any range measuring device/method. The source of these two novelties is what lies at the core of our method: an analytical expression relating the planning CT and the range measurements to the CC.
2. Materials and Methods

Our goal is to optimize the calibration curve (CC) using only the planning CT and range measurement data. The CC used at the planning stage will be referred to as $\text{CC}_{\text{pl}}$, while the optimized CC will be labeled as $\text{CC}_{\text{opt}}$. For testing purposes, we also define a virtual CC, $\text{CC}_v$, which represents an ideal CC, one that, if used instead of $\text{CC}_{\text{pl}}$, would produce a range map that is as close to the real range map as possible. In the context of this definition, the goal here is to modify $\text{CC}_{\text{pl}}$ so as to make $\text{CC}_{\text{opt}}$ as close to $\text{CC}_v$ as possible.

We begin by computing the water-equivalent path length (WEPL) map from $\text{CC}_{\text{pl}}$. The WEPL is defined as the distance a pencil beam would travel if the medium of interest, e.g., tissue, was replaced by water. For a proton moving along the z-axis, the WEPL map $W(x, y, z)$ can be computed using the continuous slowing down approximation (CSDA) scheme \[26\], which yields the formula

$$W(x, y, z) = \int_0^z \text{CC}(H(x, y, z')) dz',$$  \hspace{1cm} (1)

where $x, y$ are the coordinates in the plane perpendicular to the proton’s path and $\text{CC}(H(x, y, z))$ is the calibration curve as a function of $H$. Note that $\text{CC}$ does not depend on the beam’s energy, for the energy range applicable to proton therapy \[10\]. Once we have the WEPL map, we can compute the range $R(x, y)$ of a pencil beam of energy $E$ passing through any given pixel $(x, y)$ by solving for $z$ in Eq. (1). If $\text{CC} = \text{CC}_v$, $R(x, y)$ will correspond to the virtual range. If, however, $\text{CC} = \text{CC}_{\text{pl}}$, and $\text{CC}_{\text{pl}} \neq \text{CC}_v$, there will be a discrepancy, $\Delta R(x, y)$, between the virtual range $R_v(x, y)$ and the one computed using $\text{CC}_{\text{pl}}$, $R_{\text{pl}}(x, y)$: $\Delta R(x, y) \equiv R_v(x, y) - R_{\text{pl}}(x, y)$.

We proceed by noting that whatever $\text{CC}_v$ may be, it will not be different from $\text{CC}_{\text{pl}}$ by more than a few percent (max $\sim 5\%$). This means that the range computed using $\text{CC}_{\text{pl}}$, $R_{\text{pl}}$, and the one arising from $\text{CC}_v$, $R_v$, will also differ by only a few percent; that
This allows us to expand the WEPL map computed from $CC^v$, which will be denoted as $W^v$, in $\Delta R$:

$$W^v(x, y, R^v) = W^v(x, y, R_{pl}^v) + \frac{d}{dz} W^v(x, y, z) \bigg|_{z=R_{pl}^v} \Delta R + \frac{1}{2} \frac{d^2}{dz^2} W^v(x, y, z) \bigg|_{z=R_{pl}^v} \Delta R^2 + \ldots. \quad (2)$$

Let us examine the magnitudes of the second and third term. The former is a first derivative of $W^v$ with respect to $z$, which is merely $CC^v$ (see Eq. (1)), multiplied by $\Delta R$. The CC is of order 1, so this term will be of order $\Delta R$. The latter is a derivative of $CC^v$, multiplied by $(\Delta R)^2$. On average, the derivative of a CC is approximately the maximum value of the CC divided by the total range of HU: $CC(3000)/(3000 - (-1000)) \sim 10^{-3}$. Hence, the third term is of order $10^{-3}\Delta R^2$, which is smaller than the second term by a factor of $\sim 10^{-3}\Delta R$. Even for $\Delta R$ as large as 10 mm, this term can be safely neglected.

Solving for $\Delta R$, we obtain

$$\Delta R = \frac{W^v(x, y, R^v) - W^v(x, y, z)}{dW^v(x, y, z)/dz} \bigg|_{z=R_{pl}^v}. \quad (3)$$

Since $CC^v$ will differ from $CC_{pl}$ by only a few percent, so will $W^v$ from $W_{pl}$. Thus, we can express $W^v$ as

$$W^v(x, y, z) = W_{pl}(x, y, z) + \phi(x, y, z), \quad (4)$$

where $\phi(x, y, z) = W^v(x, y, z) - W_{pl}(x, y, z) \ll W^v(x, y, z)$, and then expand the right hand side of Eq. (3) in powers of $d\phi/dz = \phi'$:

$$\Delta R = \frac{W^v(x, y, R^v) - W^v(x, y, z)}{dW^v(x, y, z)/dz} \bigg|_{z=R_{pl}^v} = \frac{W^v(x, y, R^v) - W_{pl}(x, y, R_{pl}^v) - \phi(x, y, R_{pl}^v)}{CC_{pl} + \phi'(x, y, R_{pl}^v)} = \frac{-\phi(x, y, R_{pl}^v)}{CC_{pl}(x, y, R_{pl}^v)} + \frac{\phi(x, y, R_{pl}^v)\phi'(x, y, R_{pl}^v)}{(CC_{pl}(x, y, R_{pl}^v))^2} + \ldots, \quad (5)$$

where we set $dW_{pl}/dz = CC_{pl}$. The function $\phi$ is a correction to $W_{pl}$, while $\phi'$ is a correction to $CC_{pl}$. The term $W^v(x, y, R^v) - W_{pl}(x, y, R_{pl}^v)$ vanishes thanks to the
relation \( W^u(x, y, R^u) = W^{pl}(x, y, R^{pl}) \), or, more explicitly:

\[
\int_0^{R^u} CC^u(x, y, z')dz' = \int_0^{R^p} CC^{pl}(x, y, z')dz'
\]  

(6)

One way to construct \( \phi' \) is in terms of some bases functions \( \xi_j(x, y, z) \) such that

\[
\phi'(x, y, z) = \sum_{j=0}^{M} u_j \xi_j(x, y, z),
\]

(7)

with \( u_1, u_2, ..., u_M \) being some parameters. This allows Eq. (5) to be written in the form

\[
\Delta R = - \sum_{j=0}^{M} u_j \alpha_j(x, y) + \sum_{j=0}^{M} \sum_{k=0}^{M} u_j u_k \beta_{jk}(x, y) + ..., \]

(8)

where

\[
\alpha_j(x, y) = \int_0^{R^{pl}} \frac{\xi_j(x, y, z')}{CC^{pl}(x, y, R^{pl})}dz'
\]

\[
\beta_{jk}(x, y) = \int_0^{R^{pl}} \frac{\xi_j(x, y, z') \xi_k(x, y, R^{pl})}{(CC^{pl}(x, y, R^{pl}))^2}dz'
\]

(9)

Note that the parameters \( u_1, ..., u_M \) in Eq. (8) are outside of the brackets, which means that the brackets can be computed in advance, before optimizing the parameter set \( u_1, ..., u_M \).

Eq. (5) assumes that all the protons of a pencil beam pass through the pixel \((x, y)\). In reality this is far from true; the beam has a lateral Gaussian spread with standard deviation of approximately 3 mm, which means that protons as far away as 7 mm from \((x, y)\) contribute to the shape of the Bragg curve. In two recent studies [28, 29], it has been demonstrated that a simple convolution over a Gaussian with a \( \sigma \) equal to the width of the pencil beam provides an excellent approximation of the dose profile, and hence the average range. One of the upshots of these studies is that multiple Coulomb scattering can be neglected. Thus, in order to take into account the width of the pencil beams, we must perform a convolution of both sides of Eq. (8), \((\Delta R * G_\sigma)(x,y)\) where \( G_\sigma \) is a Gaussian centered at \((x,y)\) with standard deviation \( \sigma \), and the operator * means

\[
(f * G)(x, y) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} f(x', y')e^{-[(x'-x)^2+(y'-y)^2]/2\sigma^2}dx'dy'.
\]

(10)
Thus, we obtain
\[(\Delta R \ast G_\sigma)(x, y) = - \sum_{j=0}^{M} u_j (\alpha_j \ast G_\sigma)(x, y) + \sum_{j=0}^{M} \sum_{k=0}^{M} u_j u_k (\beta_{jk} \ast G_\sigma)(x, y),\]  
(11)
where \((\Delta R \ast G_\sigma) = (R^v \ast G_\sigma) - (R^{pl} \ast G_\sigma)\). The quantity \((R^v \ast G_\sigma)\) represents the measured range of a pencil beam centered at \((x, y)\) and will be denoted as \(R^m\). In order to compute \((R^{pl} \ast G_\sigma)\), we must solve Eq. (1) with \(CC^{pl}\) for all pixels \((x, y)\), interpolate the solution in the x-y plane, and then perform the convolution according to Eq. (10).

2.1. Application to proton therapy: head and neck

Let us now apply the techniques developed in the previous section to a patient. Figures 1 a) and b) show the CT and the dose from one beam (out of three) of a head and neck patient. The CT and the treatment plan were obtained from Cliniques universitaires Saint-Luc, MIRO. Since we do not know in advance the dependence of our range measuring device on the intensity of the PBs, we must test our methods only on PBs of the same intensity (or weight). In order to maximize the number of data points, we chose a group of PBs with weights between 0.015 and 0.025 shown as the black columns in Figure 1 c). Also shown are PBs from the chosen group for layers 8, 10 and 16.

Next, we must construct a \(CC^v\) that will serve as an ideal CC. We do this by introducing deviations to \(CC^{pl}\), which was obtained from Cliniques universitaires Saint-Luc, MIRO. To ensure that the deviations are smooth and correlated, we first add noise to the derivative of \(CC^{pl}\) sampled from a simple birth-death process, and then integrate the noisy \(dCC^{pl}/dHU\) up to \(HU\). Figure 2 a) shows ten CCs thus generated. As shown in Figure 2 b), by far the most frequent HUs in a volume of tissue surrounding the PTV are those near \(-1000\) and near zero, e. g. \(-120\) to 200. Since we are not interested in correcting the CC for air (\(HU \sim -1000\)), we will focus only on the region \([-120, 200]\).

Let us now chose the structure of \(\phi'\). The simplest choice might be a piece-wise
Figure 1. a) and b) A CT of a patient and the dose simulated by Monte Carlo. c) A histogram showing the distribution of PB weights. The black columns indicate the weights that were chosen for optimization. The CT on the right shows the PBs with the weight range of $0.015 - 0.025$ for three energy layers: blue dots = 113.72 MeV; red dots = 107.77 MeV; and green dots = 92.02 MeV.
linear function:
\[
\phi'(x, y, z) = \sum_{j=0}^{M-1} (m_j H(x, y, z) + b_j) \Omega_j(H(x, y, z)),
\]  
(12)

where
\[
m_j = \frac{u_{j+1} - u_j}{h_{j+1} - h_j},
\]  
(13)
is the slope of the line in the section \(h_j \leq H < h_{j+1}\) and
\[
b_j = \frac{h_{j+1}u_j - h_ju_{j+1}}{h_{j+1} - h_j}.
\]  
(14)
is the y-intercept. The set of coefficients \(h_1, ..., h_M\) represent HUs on the interval \([-120, 200]\). The function \(\Omega_j\) is defined as
\[
\Omega_j(x) = \begin{cases} 
1 & \text{if } h_{j+1} > x \geq h_j \\
0 & \text{otherwise}.
\end{cases}
\]  
(15)
Eq. (12) can be rearranged as follows:
\[
\phi'(H) = \sum_{j=0}^{M} u_j \left[ \frac{h_j+1 - H}{h_{j+1} - h_j} \Omega_j(H)(1-\delta_{jM}) + \frac{H - h_j-1}{h_{j+1} - h_j} \Omega_{j-1}(H)(1-\delta_{j0}) \right].
\]
(16)

Note that the expression in the square brackets is just \(\xi_j\) in Eq. (7). Finally, to optimize the parameter set \(u_i\), we can define a score function
\[
D^2 = \frac{1}{N} \sum_{i=1}^{N} \left[ R_{vi}^m - R_{vi}^{pl} + \sum_{j=0}^{M} u_j \alpha_j(\mathbf{p}_i, R_{vi}^{pl}) - \sum_{j=0}^{M} \sum_{l=0}^{M} u_j u_l \beta_{jl}(\mathbf{p}_i, R_{vi}^{pl}) \right]^2,
\]
(17)
where \(R_{vi}^{pl} = (R^{pl} \ast G_{\sigma})(\mathbf{p}_i)\), \(\mathbf{p}_i\) are the \(x, y\) coordinates of PB \(i\), and \(N\) is the number of PBs.

2.1.1. Optimization and constraints
There are many ways to chose both the \(M\) and the specific values for the \(h_i\)s. The only guidance we have in doing so is the trade off between the level of detail in \(CC^{opt}\) – which comes from choosing a large \(M\) – and the robustness of our method against errors and noise – which requires that \(M\) be small. With this in mind, we chose \(M = 5\): \(h_i = \{-120, 16.69, 50, 74, 200\}\), where \(h_1\) and \(h_5\) mark the boundary of the \(HU\) distribution (see Fig. 2 b)), \(h_2\) and \(h_4\) are the values at which the \(HU\) distribution is 50% of its maximum, which occurs at \(h_3\). In order to avoid unrealistic solutions, we need to impose constraints on the way \(CC^{opt}\) can behave. Based on experimental evidence, \(CC^v\) should not differ from \(CC^{pl}\) by more than 5% \[30\] for any given HU. Hence our first constraint. The second constraint is on the derivative of the CC. Some studies suggest that the slope is always positive \[10, 17, 31\]; however, other studies report that the CC can in fact have the opposite trend locally \[32\]. In addition, the CC can also be multi-valued. With the apparent complexity of a CC in mind, we first apply the constraint that the slope \([CC(h_{i+1})-CC(h_i)]/(h_{i+1} - h_i) = 0\), and then relax it by allowing the slope to be \(> -0.0005n\) for \(n = 1, ..., 10\). We take as the final CC the average of the eleven CCs yielded by this procedure.
2.1.2. Set-up errors and noise  The quality of the range measurements is subject to various errors and sources of noise. Chief among them are 1) the set-up error, which can be as large as ±3mm [3]; 2) the set-up of the range measuring device, which we assume to have uncertainty of 2 mm; and 3) the noise in the range data measured by the device, which, e. g. for a Prompt gamma can be of the order of 2 mm [23]. To stress our method, we chose it to be 3 mm.

In order to test the robustness of our method against patient set-up errors, measurement noise, and measurement setup errors, we computed the range from Eq. (1) using CC" to the convolution around the center of each PB with the center shifted by a vector \( s = (s_x, s_y) \), \( (R^v \ast G_\sigma)(x + s_x, y + s_y) \), and added a random variable \( \eta \), sampled from a Normal distribution \( \mathcal{N}(r, \eta) \), where \( \eta = 3 \) mm, and \( r \) is a random variable sampled from another Normal distribution \( \mathcal{N}(0, 2) \). Hence, \( \eta \) represents the random uncertainty of the measurement for each PB, and \( r \) is the random set-up error of the range-measuring device. The variables \( s_x \) and \( s_y \) are the set-up errors in the \( x \) and \( y \)-direction respectively; they were sampled from a Normal distribution \( \mathcal{N}(\mu, \sigma) \) with \( \sigma = 3 \) mm and \( \mu \) that was sampled from another Normal distribution \( \mathcal{N}(0, \sigma) \), representing a systematic error.

Because of the errors present in the setup and in the range measurement, optimizing the score function \( (17) \) for each fraction, using the range data from that fraction alone, may not lead to a reliable CC\textsuperscript{opt}. It is even conceivable that CC\textsuperscript{opt} may fit CC\textsuperscript{v} worse than CC\textsuperscript{pl} does. For this reason, it may be necessary to accumulate range data over several fractions, thus averaging out the noise and, to some degree, the set-up errors. Hence, the measured range, \( R^m_i \) in the score function must read

\[
R^m_i = \frac{1}{K} \sum_{k=1}^{K} R^m_i(k),
\]

where \( R^m_i(k) \) is the measured range of PB \( i \) for fraction \( k \), and \( K \) is the number of
fractions already delivered.

2.1.3. Method of optimization  The search for the optimum set \((u_1, ..., u_5)\) was performed on Mathematica using the “Nminimize” function with the “DifferentialEvolution” method. The aforementioned constraints were stated explicitly in the “Nminimize” function.

3. Results

In Fig. 3, we show the ten \(CC^v\)’s from Fig. 2a) (dashed), \(CC^{pl}\) (solid black) and the \(CC^{opt}\) for ten fractions (marked by “x”). The gray line connects the average values over ten fractions for each of the five HUs. The smaller frames in the bottom left corner show the x- and y-components of the set-up error in millimeters for ten fractions; the systematic set-up error is indicated by the black arrow. By visual inspection, we see that not all \(CC^{opt}\) are an improvement on the \(CC^{pl}\). For instance, in CC 4 the average \(CC^{opt}\) is too large in the domain \([0, 100]\), where it really counts. In CC 10 we see a similar behavior; however, this case is even worse: in the domain \([-120, 0]\) the average \(CC^v\) is below \(CC^{pl}\) when it should be above.

Let us go beyond mere visual inspection and look at the average relative deviation of \(CC^{opt}\) and \(CC^{pl}\) from \(CC^v\), defined by

\[
\Delta SPR = \frac{1}{320} \int_{-120}^{200} dy \left[ 1 - \frac{CC^{pl/opt}(y)}{CC^v(y)} \right].
\]  

Fig. 4a) shows this quantity for \(CC^{pl}\) (black dots) and for \(CC^v\) for ten fractions (marked by “x”); the center of the circle indicates the 5\(^{th}\) fraction, while the tip of the triangle overlaps with the 10\(^{th}\) fraction. This figure reveals that, for the 5\(^{th}\) fraction, \(CC^{opt}\) in CC 4 is indeed worse on average than \(CC^{pl}\). However, the converse is true for the 10\(^{th}\) fraction; \(CC^{opt}\) gives about 50% improvement over \(CC^{pl}\). By the measure defined in
Figure 3. $CC_v$ (dashed gray), $CC_{pl}$ (black solid) and $CC_{op}$ for the first ten fractions (crosses). The frames in the lower right corner show the systematic (arrow) and random (black dot) setup errors for the first ten fractions.
Eq. (19), $CC^{opt}$ in CC 10, although the best among the ten fractions, is still worse than $CC^{pl}$.

The measure that is of real interest is the range; more specifically the distribution of the over/under-range defined as $\sum_{x,y}|R^v(x,y) - R^{pl/opt}(x,y)|$. Fig. 4b) shows the average over/under-range defined as

$$\langle \Delta R^{pl/opt} \rangle = \frac{1}{L_p L_E} \sum_{m=1}^{L_E} \sum_{x,y} [R^v(x,y) - R^{pl/opt}(x,y)],$$

where $L_p$ is the total number of pixels in the sum and $L_E = 20$ the number of layers. The area covered by the sum was chosen so as to fit all PBs of the plan, plus ten additional millimeters surrounding the border in anticipation of patient set-up errors. In some anatomical regions, e.g. near the interface between the skin and the air, even a small difference between the $CC^v$ and $CC^{opt}$ can result in either $R^v$ or $R^{pl/opt}$ to traverse the tissue, giving a range equal to the limit of the CT grid. Computing $R^v(x,y) - R^{pl/opt}(x,y)$ for such pixels would not be a fair assessment of the effect coming from the difference between $CC^v$ and $CC^{pl/opt}$. For this reason they were omitted from the sum. Finally, Fig. 4c) shows the standard deviation of the over/under-range defined as

$$\langle \Delta \Delta R^{pl/opt} \rangle = \left[ \frac{1}{L_p L_E} \sum_{m=1}^{L_E} \sum_{x,y} [R^v(x,y) - R^{pl/opt}(x,y) - \langle \Delta R^{pl/opt} \rangle]^2 \right]^{1/2},$$

The results of Figs. 4 a), b) and c) are shown explicitly in Table 1.

4. Discussion

The method of optimizing a CC presented in this paper was tested against various sources of error and noise. Patient set-up errors, systematic and random, in the plane transverse to the beam, both with $\sigma = 3$ mm, were added. The simulated virtual range of each PB, representing the range as measured by a range-measuring device, was skewed by a a value sampled from a normal distribution with $\sigma = 3$ mm; this value was meant
Figure 4. a) RSP uncertainty as defined in Eq. (19). b) Range bias and c) range uncertainty as defined in Eqs. (20) and (21), respectively.

Table 1. This table shows explicitly the values from Figure 4 a) b) and c).

| CC | CC_{pl} error | CC_{opt} error | R_{pl} bias | R_{opt} bias | \sigma_{R_{pl}} | \sigma_{R_{opt}} |
|----|----------------|----------------|-------------|-------------|----------------|----------------|
| 1  | 4.68 %         | 2.57, 1.77 %   | 2.52        | 0.35, -0.23 | 1.24           | 1.01, 1.01     |
| 2  | 3.85 %         | 1.37, 1.38 %   | -3.24       | -0.73, -0.75| 1.03           | 0.71, 0.71     |
| 3  | 1.44 %         | 0.92, 1.02 %   | -0.87       | 0.32, -0.02 | 0.80           | 0.71, 0.71     |
| 4  | 1.73 %         | 2.90, 0.90 %   | -0.71       | 1.12, -0.35 | 0.82           | 1.09, 0.79     |
| 5  | 1.93 %         | 1.23, 1.89 %   | -1.63       | -1.18, -1.05| 0.82           | 0.70, 0.68     |
| 6  | 2.61 %         | 1.48, 1.03 %   | -2.45       | -1.62, -1.12| 0.96           | 0.67, 0.64     |
| 7  | 3.24 %         | 2.63, 2.72 %   | 3.53        | 0.40, 0.40  | 1.42           | 1.34, 1.37     |
| 8  | 3.32 %         | 2.30, 1.99 %   | -2.90       | -0.86, -1.81| 1.05           | 0.82, 0.82     |
| 9  | 1.88 %         | 1.03, 1.55 %   | -2.00       | -0.35, -0.21| 0.90           | 0.68, 0.69     |
| 10 | 2.02 %         | 3.57, 2.93 %   | -0.36       | 0.48, 1.19  | 1.03           | 1.11, 1.11     |
to simulate the uncertainty in the range measurement. Another source of error with $\sigma = 2$ mm was added to capture the error in the set-up of the measuring device.

Due to so much uncertainty, one cannot expect this method to be reliable for a given fraction but rather must be applied after several fractions, taking in as input range data that have been averaged and hence partly filtered. This filtering, however, only applies to random errors, not systematic ones. Nevertheless, according to Figures 4 a), b) and c), even in the presence of systematic errors the CC, the range bias and the range uncertainty tend to improve after ten fractions. Even after five fractions, in eight out of ten cases, the average deviation of $CC_{\text{opt}}$ from $CC_{v}$ was lower than the average deviation of $CC_{pl}$ from $CC_{v}$. Same was true for the range bias and the range uncertainty. Regarding the range bias in the last case, CC 10, the fifth fraction is actually better than the tenth fraction. In addition, for several fractions (marked by the crosses), the range bias is close to zero. This suggests that for this particular case of a $CC_{v}$, our method is less robust to noise and errors.

The success of our method is on a par with other published methods of optimizing the CC. For example, it has been demonstrated [ref] that the use of dual energy CT (DECT) can reduce uncertainty in the stopping power from 1.59% to 0.61%, and offer a reduction in bias from -0.88% to -0.58% and -0.14%. On average, our method reduced the error in a CC from 2.67% to 1.62% and the WEPL bias from 2.14% to 0.74%. Other studies... plus a discussion of advantages/disadvantages, e.g. no additional dose to patient.

5. Conclusions

The method of optimizing the HU-SP calibration curve presented herein can provide a new way for clinicians to monitor and/or adapt the course of a treatment. The main
novelty of this method lies in its structure: the bulk of the computation is performed before the treatment so that the computation required upon the range data acquisition is very efficient (a few seconds). What makes this possible is the fact that our method relies on an analytical expression, rather than an algorithm, which takes as input the planning CT, a planning (usually a stoichiometric) CC, and the range data, and yields the optimized CC as output. Controlling the quality of this output are constraints placed upon the minimization procedure. We have shown that our method, when applied over several fractions, yields a CC that better serves the patient, i.e., a CC that results in the overall reduction of the over/under-range. Although more work is required to ascertain the true potential of this method, we submit that the present study sufficiently demonstrates its usefulness in proton radiotherapy.

Lastly, it is worth reiterating that our method is based on an analytical formula whose only inputs are a planning CT and range data. As such, it is in principle able to incorporate any range measuring device/method, e.g., proton radiography, Prompt gamma camera, PET scan, ionoacoustic range measurements. The quality of results yielded by our method will, of course, depend on the quality of the range data provided by the range measuring device. The second advantage of having an analytical formula as a basis is the speed with which the CC can be optimized. While direct methods, such as Monte Carlo, are preferable to others due to their accuracy, they are not practical in cases where optimization of many parameters is desired. This is where our method can be of great benefit.

6. Acknowledgments

This research project was made possible by ImagX-R research program co-funded IBA. JA would like to thank E Ejkova for her technical support.
References

[1] Langen KM and Jones DTL 2001 Organ motion and its management. *Int. J. Radiat. Oncol. Biol. Phys.* 50(1) 265278

[2] Young-Kee O, Jong-Geun B, Ok-Bae K and Jin-Hee K 2014 Assessment of setup uncertainties for various tumor sites when using daily CBCT for more than 2200 VMAT treatments. *J. Appl. Clin. Med. Phys.* 15(2)

[3] Létourneau D, Martinez AA, Lockman D, Yan D, Vargas C, Ivaldi G and Wong J 2005 Assessment of residual error for online cone-beam CT-guided treatment of prostate cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.* 62(4); 12391246

[4] Fleischmann D, Boas FE, Tye GA, Sheahan D and Molvin LZ 2011 Effect of low radiation dose on image noise and subjective image quality for analytic vs iterative image reconstruction in abdominal CT. *RSNA Chicago*

[5] Chvetsov AV and Paige SL 2010 The influence of CT image noise on proton range calculation in radiotherapy planning *Phys. Med. Biol.* 55(6)

[6] Paganetti H 2012 Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys. Med. Biol.* 57(11) R99R117

[7] Schaffner B and Pedroni E 1998 The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. *Phys. Med. Biol.* 43(6) 1579-1592.

[8] Jiang H, Seco J and Paganetti H 2007 Effects of Hounsfield number conversion on CT based proton Monte Carlo dose calculations. *Med. Phys.* 34(4) 14391449 doi: 10.1118/1.2715481

[9] Yang M, Virshup G, Clayton J, Zhu XR, Mohan R and Dong L, 2010 Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Phys. Med. Biol.* 55 134362

[10] Schneider U, Pedroni E and Lomax A 1996 The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys. Med. Biol.* 41(1) 111-124

[11] Moyers MF, Sardesai M, Sun S and Miller DW 2010 Ion stopping powers and CT numbers. *Med. Dosim.* 35(3) 179-194

[12] Hünemohr N, Krauss B, Tremmel C, Ackermann B, Jkel O and Greilich S. 2014 Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates. *Phys. Med. Biol.* 59(1) 83-96
[13] Saito M and Sagara S 2017 Simplified derivation of stopping power ratio in the human body from dual-energy CT data. *Med. Phys.* doi: 10.1002/mp.12386

[14] Grantham KK, Li H, Zhao T and Klein EE 2015 The impact of CT scan energy on range calculation in proton therapy planning. *J. Appl. Clin. Med. Phys.* 16(6)

[15] Bourque AE, Carrier JF, Bouchard H 2014 A stoichiometric calibration method for dual energy computed tomography. *Phys. Med. Biol.* 59(8) 2059-88

[16] Doolan JP, Collins-Fekete C-A, Dias MF, Ruggieri TA, D’Souza D and Seco J 2016 Inter-comparison of relative stopping power estimation models for proton therapy. *Phys. Med. Biol.* 61(22)

[17] Doolan PJ, Testa M, Sharp G, Bentefour EH, Royle G and Lu H-M 2015 Patient-specific stopping power calibration for proton therapy planning based on single-detector proton radiography. *Phys. Med. Biol.* 60 19011917

[18] Xie Y, Bentefour H, Janssens G, Smiets J, Vander Stappen F, Hotoiu L, Yin L, Dolney D, Avery S, OGrady F, Prieels D, McDonough J, Solberg TD, Lustig R, Lin A and Teo B-KK 2017 Prompt gamma imaging for in vivo range verification of pencil beam scanning proton therapy. *Int. J. Radiat. Oncol. Biol. Phys.* DOI: 10.1016/j.ijrobp.2017.04.027

[19] Priegnitz M, Barczyk S, Nenoff L, Golnik C, Keitz I, Werner T, Mein S, Smiets J, Vander Stappen F, Janssens G, Hotoiu L, Fiedler F, Prieels D, Enghardt W, Pausch G and Richter C 2016 Towards clinical application: prompt gamma imaging of passively scattered proton fields with a knife-edge slit camera. *Phys. Med. Biol.* 61 7881

[20] Smiets J, Roellinghoff F, Janssens G, Perali I, Celani A, Fiorini C, Freud N, Testa E and Prieels D 2016 Experimental Comparison of Knife-Edge and Multi-Parallel Slit Collimators for Prompt Gamma Imaging of Proton Pencil Beams. *Front. Oncol.* 6 156 doi: 10.3389/fonc.2016.00156

[21] Xing Y, Macq B, Janssens G and Bondar L 2016 Prompt Gamma Imaging-Based Prediction of Bragg Peak Position for Realistic Treatment Error Scenarios. *Med. Phys.* 43(6) 3456

[22] Nenoff L, Priegnitz M, Trezza A, Smiets J, Janssens G, Vander Stappen F, Hotoiu L, Prieels D, Enghardt W, Pausch G and Richter C 2017 Sensitivity evaluation of prompt -ray based range verification with a slit camera. *Radiother. Oncol.* 123 S76S77 DOI: [http://dx.doi.org/10.1016/S0167-8140(17)30596-0](http://dx.doi.org/10.1016/S0167-8140(17)30596-0)

[23] Janssens G, Smiets J, Stappen FV, Prieels D, Clementel E, Hotoiu EL and Sterpin E 2016 Sensitivity study of prompt gamma imaging of scanned beam proton therapy in heterogeneous
anatomies. *Radiother. Oncol.* **118**(3) 562-7

[24] Kellnberger S, Assmann W, Lehrack S, Reinhardt S, Thirolf P, Queirs D, Sergiadis G, Dollinger G, Parodi K, Ntziachristos V 2016 Ionoacoustic tomography of the proton Bragg peak in combination with ultrasound and optoacoustic imaging *Scientific Reports* **6** 29305

[25] Lehrack S, Assmann W, Bertrand D, Henrotin S, Herault J, Heymans V, Stappen FV, Thirolf PG, Vidal M, Van de Walle J, Parodi K. 2017 Submillimeter ionoacoustic range determination for protons in water at a clinical synchrocyclotron. *Phys. Med. Biol.* **62**(17) L20-L30

[26] Newhauser WD and Zhang R 2015 The physics of proton therapy *Phys. Med. Biol.* **60** R155R209

[27] Bortfeld T 1997 An analytical approximation of the Bragg curve for therapeutic proton beams. *Med. Phys.* **24**(12) 2024-33

[28] Farace P, Righetto R, Defeit S, Meijers A, Vander Stappen F 2016 Technical Note: A direct ray-tracing method to compute integral depth dose in pencil beam proton radiography with a multilayer ionization chamber. *Med. Phys.* **43**(12) 6405

[29] Defeit S, Macq B, Righetto R, Vander Stappen F, Farace P 2017 Registration of pencil beam proton radiography data with X-ray CT. *Med. Phys.* **44**(10) 5393-5401

[30] Yang M, Zhu XR, Park PC, Titt U, Mohan R, Virshup G, Clayton JE and Dong L. 2012 Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration *Phys. Med. Biol.* **57** 4095115

[31] Ainsley CG and Yeager CM 2014 Practical considerations in the calibration of CT scanners for proton therapy. *J Appl Clin Med Phys* **15**(3) 4721

[32] Hudobivnik N, Schwarz F, Johnson T, Agolli L, Dedes G, Tessonnier T, Verhaegen F, Thieke C, Belka C, Sommer WH, Parodi K and Landry G 2016 Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images. *Med. Phys.* **43**(1) 495504 [http://doi.org/10.1118/1.4939106](http://doi.org/10.1118/1.4939106)