



Technical note

A Monte Carlo detector response model for the IRIS PET preclinical scanner

Alessandro Pilleri^{a,*,1}, Niccolò Camarlinghi^{a,b,1}, Alberto Del Guerra^{a,b}, Giancarlo Sportelli^{a,b}, Nicola Belcari^{a,b}^a University of Pisa, Department of Physics, Largo Bruno Pontecorvo 3, Pisa 56127, Italy^b INFN Sezione Pisa, Largo Bruno Pontecorvo 3, Pisa 56127, Italy

ARTICLE INFO

Keywords:

PET image reconstruction
 Monte Carlo simulation
 System model
 Detector response

ABSTRACT

PET preclinical studies require high spatial resolution due to the limited size of the animal under investigation. To achieve this target, iterative image reconstruction algorithms are commonly preferred over the analytical methods because they offer the possibility of accurately modeling the whole imaging process. In this work, we propose an accurate factorized system matrix for the INVISCAN IRIS preclinical PET scanner to be used with an iterative algorithm. The model includes two components: the geometric component and the detector response of the system. The main innovative aspect of the work is the creation of the detector matrix using a Monte Carlo simulation, with a particular focus on the optimization of the simulation process to reduce the calculation time. The new system model is compared with the current IRIS model to evaluate the image quality, following the NEMA Standards NU 4-2008. The comparison showed an enhancement of the image quality, in terms of uniformity and recovery coefficients. This work confirms that the inclusion of the detector response into the system model leads to improved reconstruction results.

1. Introduction

Preclinical Positron Emission Tomography (PET) [1] requires an excellent spatial resolution to investigate the small internal structures of animal models. Iterative algorithms are used for this purpose due to their superior performance in terms of spatial resolution and image quality if compared to analytical methods. The strength of the iterative algorithms lies in its capability of accurately modeling the physics involved in the whole imaging process [2]. Iterative algorithms depend on the system model, which relates the activity distribution to the data acquired by the scanner. The system model is usually represented as a matrix S in which each element s_{ji} represents the probability that two photons emitted by the annihilation of a positron in the voxel i are recorded by the *Line of Response* (LOR) j . There are three main methodologies used to obtain the whole system model: experimental, Monte Carlo and analytical [3]. The first approach is the most accurate if obtained under ideal experimental conditions [4]. However, it is also the most challenging due to the complexity of the experimental setup and to the time needed to perform the full acquisition [5]. An alternative method to directly measure the system matrix is to reproduce the experimental process with a Monte Carlo simulation [6]. The most important limitation of this approach is the time needed to perform the

whole simulation, which may be impractical using standard workstations. The last approach, that is the analytical implementation, allows for a much faster calculation. However, this class of methods usually disregards important physical aspects of the PET acquisition chain, such as the detector response and it does not account for statistical noise. Instead of computing the whole system matrix with only one of the aforementioned methods, it is possible to express the model as a multiplication of independent matrices. Each one accounting for one or a set of effects and can be calculated and stored separately [7].

The accuracy of the system model influences the image reconstruction process and the spatial resolution of the system. The main factors affecting the spatial resolution are described in [8]. Concerning the detector response, the deviation from the ideal behaviour is mostly due to the photon penetration across crystals and inter-crystal scatter (ICS) effects. These effects are particularly significant in small-animal PET due to the reduced size of the scanner ring. The inclusion of an accurate modeling of the detector response in the system matrix can lead to a reduction of the uncertainty in the LOR identification and therefore to an enhancement of the spatial resolution of the imaging system, as shown by other authors [9–12].

This study investigates the computation and validation of a factorized system model for the preclinical IRIS PET scanner [13] that

* Corresponding author.

E-mail address: alessandro.pilleri@df.unipi.it (A. Pilleri).¹ These authors contributed equally to this work.

includes the detector response. It represents an extension of a previous work [11] and it demonstrates that the method can be generalized to a complex polygonal scanner geometry. The previous geometry consisted of two opposite parallel detector modules in which each module is made of a matrix of scintillation crystals. This work extends the results to a multiring geometry and to non-parallel opposite detector modules. The IRIS system matrix is pre-computed and fully stored in memory. We compared the results obtained using the developed model to those obtained using the previously computed models. The first one is based on the Siddon's ray tracing algorithm [14] and it uses multiple rays to compute the volume of intersection between the LORs and the voxels of the field of view (FOV). Using this approach, the ICS and penetration effects are not accounted for. The model is indicated as *Geometrical* or simply *G*. The second approach uses a multi-ray tracing method [9] including an analytical component of the penetration effects, but it does not consider the ICS effect.

2. Material and methods

2.1. Experimental setup

The scanner used in this work is the IRIS PET/CT small-animal scanner distributed by Inviscan s.a.s (France). The PET component of the scanner consists of 16 *modular detectors* (modules) arranged in two octagonal rings divided by a gap of 6.84 mm. Each module, which can acquire coincidences with the six opposing modules, is composed of a matrix of 27×26 LYSO:Ce crystals of $1.6 \times 1.6 \times 12$ mm³ dimensions, with a crystal pitch of ~ 1.69 mm. Each matrix is directly coupled to a 64-anodes photomultiplier tubes (PMT) (H8500C, Hamamatsu Photonics K.K., Hamamatsu, Japan). The field of view is 95×80 mm² wide (axial and transaxial dimension) [15].

Images are reconstructed with a multi-core LOR-based 3D *Maximum Likelihood Estimation Maximization* (ML-EM) [16] software. The reconstructed image set consists of a matrix of $101 \times 101 \times 120$ voxels of $0.855 \times 0.855 \times 0.855$ mm³ each.

2.2. System model

To model the detector response, the concept of *Line of Flight* (LOF) must be introduced. The LOF is the line in which the annihilation event took place and it is defined by the two γ -rays' flight path (we assume photon collinearity). The LOR instead is defined as the line connecting the two crystal pixels in which the photons are effectively detected. In an ideal detector, the LOR and the LOF are equivalent. However, in real situations, the LOF may be misplaced due to penetration and ICS effects.

A common decomposition of the system model is:

$$S = Z \cdot A \cdot D \cdot G \cdot R \quad (1)$$

- *Z* is a diagonal matrix that contains the normalization factors for each LOR;
- *A* is a diagonal matrix with the attenuation correction for each LOR;
- *D* $\in \mathbb{R}^{M \times M}$ is the detector matrix. The element d_{jk} represents the probability that the photons that travel along the LOF *k* are detected in the LOR *j*;
- *G* $\in \mathbb{R}^{M \times N}$ is the geometric matrix. The element g_{ki} represents the probability that back-to-back pair of photons emitted by the annihilation of a positron in the voxel *i* reach the surface of the crystals which identify the LOF *k*;
- *R* models the positron range, which could be calculated using Monte Carlo simulation [17].

where *N* is the number of voxels and *M* is the number of LOFs/LORs. We proposed a factorized model where only the analytical geometrical component and the detector component have been included. The

normalization factors are incorporated directly into the ML-EM algorithm, via a component-based approach. The inclusion of the attenuation correction and the positron range components have not been considered at this stage, as in this work we were interested only in the investigation of the detector response.

2.3. Detector matrix

We performed the study of the penetration and the ICS effects with a Monte Carlo simulation based on the GEANT4 toolkit [18]. The detector matrix is created adapting the method described in [11] to the IRIS geometry. The total simulation time is reduced by exploiting some of the rotational, translational and reflection symmetries of the scanner. No inter-module penetration/scatter effects have been included. A single LOF connecting two crystals of two opposed detectors is split into two half-lines representing the two photons flight path. A simulation is performed for each of the two photons. In the first one, a large number of photons are created at the surface of the starting crystal and sent toward the destination crystal. The second simulation consists of the same procedure but with the beam sent backward from the endpoint toward the starting crystal. As the surface of the crystal has a finite size, several lines actually connect the two crystals and the LOR concept needs to be generalized in a *Tube of Response* (TOR). To better reproduce this behaviour, the start and the end points of the line defining the photon flight path are chosen at a random position on the pixel surface instead of at the crystal centre.

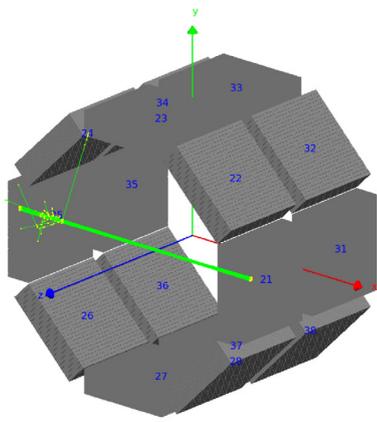
Fig. 1 shows the events distribution for three sample LOFs. The figures on the left show the IRIS scanner geometry within the GEANT4 environment, along with some tracks of the simulated photons. The boxed figures at the right show the distribution of the collected events as a 27×26 grey-scale image, which represents the detector module. Each image element represents the probability of detection of a photon in that specific crystal. Fig. 1 shows that the penetration and ICS effects can be responsible for a high LOR misplacement.

2.4. Simulation tuning

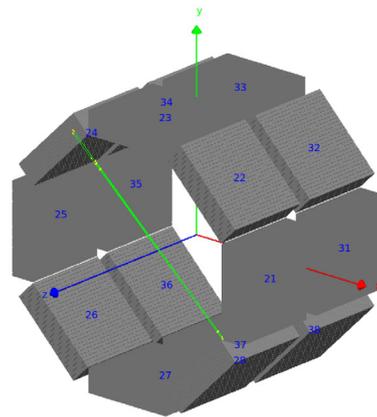
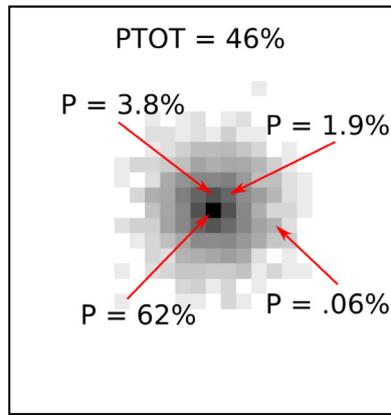
The most important GEANT4 parameters which regulate the simulation time are the number of photons generated for each TOR (N_c) and the variables that determine the accuracy of the particle tracking. N_c is adaptively selected using *inverse binomial sampling* [19]. This method allows one to relate the number of events (N_c) to be collected to the relative precision (ϵ) of the estimated probability (*p*); in this work, we used the approximate relation:

$$N_c = \epsilon^{-2}(1 - p) \leq \epsilon^{-2}$$

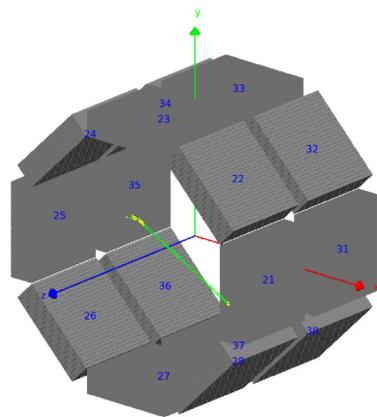
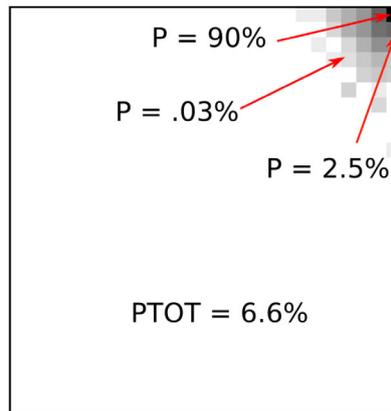
The tracking of the electrons within the scintillators can quickly become the bottleneck of the simulation. To improve this aspect it is possible to act on the physics of the simulation through the *Lowest Electron Energy* (LEE) parameter. The LEE is the energy beyond which the electrons that arise inside the crystals are no longer tracked by the simulator and are locally absorbed. To preserve the simulation accuracy, the LEE parameter should be set to an energy value associated with a low probability for the electron to cross the interface between two neighbouring crystals, in other words, to a value for which the distance travelled by the electron can be considered smaller than the crystal pitch. We chose the LEE value after the study of the electron range in the scintillator. Using the NIST database [20] it is possible to extract the *continuous slowing down approximation range* (CSDA range), the average path length travelled by a charged particle as it slows down to rest assuming that the rate of energy loss at each point along the track is equal to the total stopping power [21]. Instead of the CSDA range, a more representative quantity is the average electron *depth of penetration*, which is the average maximum distance reached by the particle from its starting point. We performed a GEANT4 simulation to study both these quantities and we used the NIST data to validate the



(a) Perpendicular incidence: most of the photons are recorded into the target crystal. In this case, the spread of the distribution is due to the inter-crystal scatter effect.



(b) Border incidence: the effect of crystal penetration brings to a lower total detection probability: most of the recorded events are revealed by the target crystal.



(c) Wide angle incidence: the penetration effect becomes significant and the total probability is shared between a higher number of crystals; the scatter effect spreads further the spatial distribution.

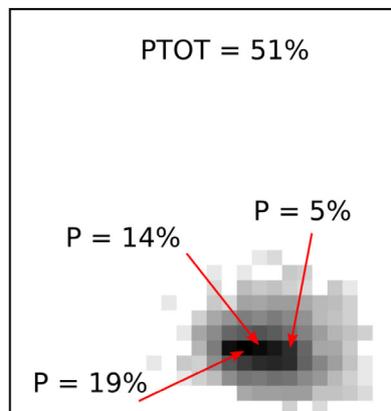


Fig. 1. Examples of simulated photon beams. The right sub-figures show the distribution of the collected events as a 27×26 grey-scale image: each pixel represents the probability of detection of a photon in that specific pixel. The value $PTOT$ represents the total probability of detection of the photon, that is the sum of the detection probabilities for every crystal of the module. The overall probability is lower than 100% due to the possibility that the photon completely penetrates the detector module without being detected. The P values, which are relative to $PTOT$, represent instead the detection probabilities of the sample pixels pointed by the arrows.

simulation results. Fig. 2 reports the results of the simulation: it shows that the depth of penetration is approximately fourfold lower than the CSDA range.

On the basis of the simulation, we set the LEE to 100 keV, which corresponds to an electron penetration depth of $8 \mu\text{m}$. This distance can be considered negligible with respect to the IRIS crystal pitch ($\sim 1.6 \text{ mm}$). Changing this value in the GEANT4 led to a reduction of the simulation time by a factor of ~ 3 .

Using this configuration, the simulation time for each module is ~ 4.5 days on a Xeon E5620 @ 2.40 GHz workstation.

2.5. System matrix construction

The computation of the system matrix is challenging, considering that during the multiplication of D and G the three matrices should coexist in memory. To overcome the memory limit, the $M \times N$ system matrix is computed with an *intermediate decomposition* technique. The $M \times M$ detector matrix is subdivided into $x = M/L$ sub-matrices of $M \times L$ elements each (M has to be an integral multiple of L) and the $M \times N$ geometrical matrix is subdivided into L matrices of $L \times N$ elements:

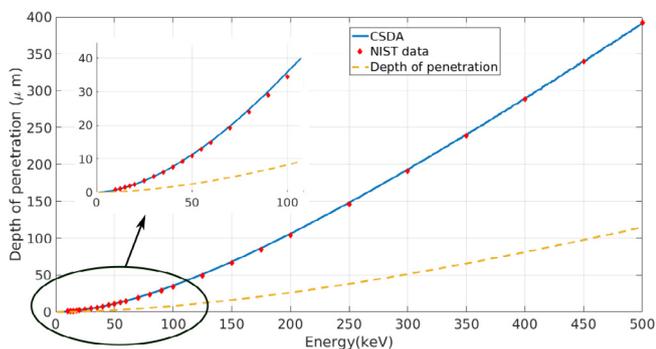


Fig. 2. Comparison of the CSDA range and average depth of penetration in LYSO for electrons up to 500 keV: data are obtained with a GEANT4 Monte Carlo simulation. The diamonds represent the data extracted from the NIST database. The NIST information is used to validate the simulated data. The depth of penetration is about four times smaller than the CSDA range.

$$D = \{D_1; D_2; \dots; D_x\} \tag{2}$$

$$G = \{G_1, G_2, \dots, G_x\} \tag{3}$$

The system matrix is obtained as:

$$S = \sum_i \cdot G_i \tag{4}$$

The intermediate products $D_i \cdot G_i$ are computed and stored on disk one-by-one and the final S matrix is assembled after all the products have been computed.

Since the S matrix is too large to be loaded into RAM, we applied an adaptive threshold to reduce its final size: the elements of the detector matrix with low probability values are neglected before the S matrix computation. The elements of each row of the detector matrix are sorted by descending order and they are saved until their sum reaches a

fixed threshold (T). The number of saved elements for each row is therefore variable. T is a fraction of the total detection probability, that is, the sum of the elements along the row. In this work, we set $T = 60\%$ which allows us to create a model which fits our memory requirements.

2.6. Computational aspects

The size of the final model depends strongly on the number of LORs and the number of voxels of the FOV. The most important challenge is the compression of the model to a suitable size, as the matrix has to be entirely loaded into the computer RAM to perform the reconstruction. A high level of compression is achieved by exploiting the symmetries of the system, using the method described in [22]. Since the detector matrix presents fewer symmetries with respect to the geometric matrix, after the multiplication it is not possible to reach a compression ratio as high as the one that is possible to obtain for the geometric matrix. The workstation used in this work has 64 GB of RAM and the most accurate final model has a size, after the compression, of 50 GB. With an increase of the accuracy of the computed detector matrix, that is increasing the threshold T , or in the number of voxels of the FOV, the size of the final model outreach the total RAM of the system and it is not possible to perform the reconstruction. The product between D and G was implemented in C++ using the EIGEN Library sparse matrices manipulation module [23].

3. Results

The image quality was evaluated following the NEMA NU 4-2008 standard recommendations [24]. The cylindrical chamber and the rods of the phantom were filled with a 3.7-MBq solution of $^{18}F - FDG$. The acquisition was performed for 20 min using the IRIS scanner at the *Institute of Clinical Physiology* of CNR in Pisa. The image reconstruction was performed using the ML-EM algorithm. No data correction was

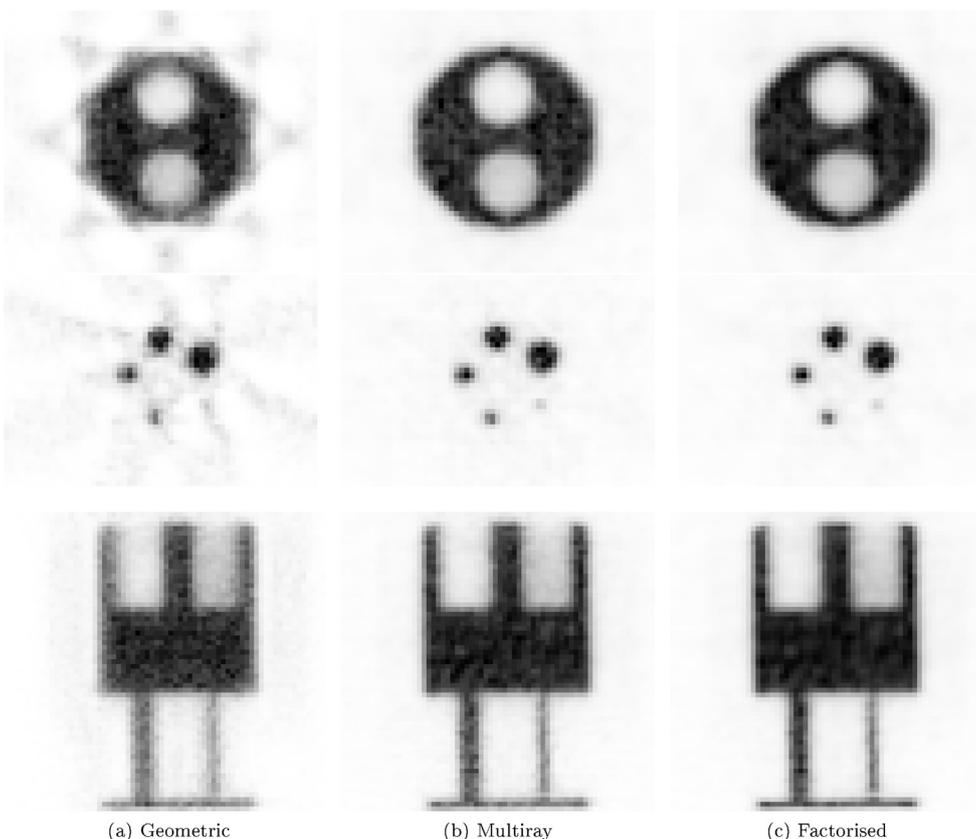


Fig. 3. Examples of reconstructed images of the phantom after 70 ML-EM iterations. Left Geometric, center Multiray, right Factorised.

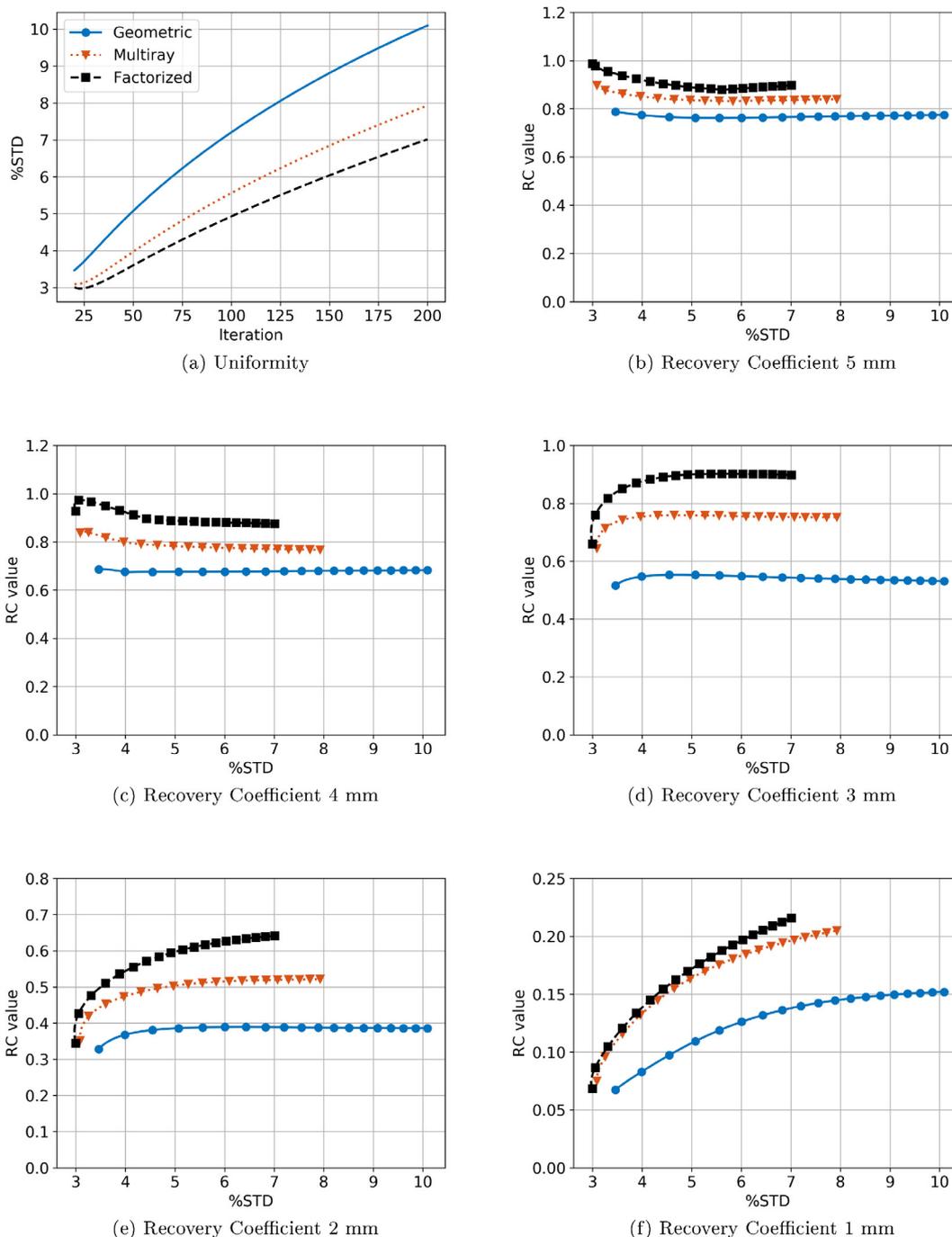


Fig. 4. (a): Uniformity measurement of the image quality phantom shown as the percentage standard deviation as a function of the iteration number. The factorized model presents less noise with respect to the others model. (b)–(f): Recovery coefficients of the five rods as a function of the percentage standard deviation of the uniform region. On the y-axis is shown the intensity of the pixels relative to the value of the uniform region. For the same uniformity values, the best results are obtained by the Factorized model. Markers are plotted every ten iterations to provide additional information about the trend of the RC as a function of the iteration number.

applied except for the detector normalization. The ML-EM process was stopped after 100 iterations. Fig. 3 shows examples of reconstructed images of the phantom.

3.1. Uniformity and noise

Fig. 4a shows the noise of the different models in terms of the percentage standard deviation of the central region of the phantom as a function of the iteration number. Using the ML-EM reconstruction method, the noise of the image typically increases with every iteration,

whilst the algorithm approaches the likelihood solution. The analysis shows that, for every iteration, the Geometric model has always a poorer uniformity than the other models. The best results are obtained using the Factorized model.

3.2. Recovery coefficients

Fig. 4a–f show the values of the recovery coefficients (RCs) for every rod as a function of the percentage standard deviation of the uniform region (%STD). All the models show the same trend. As discussed by

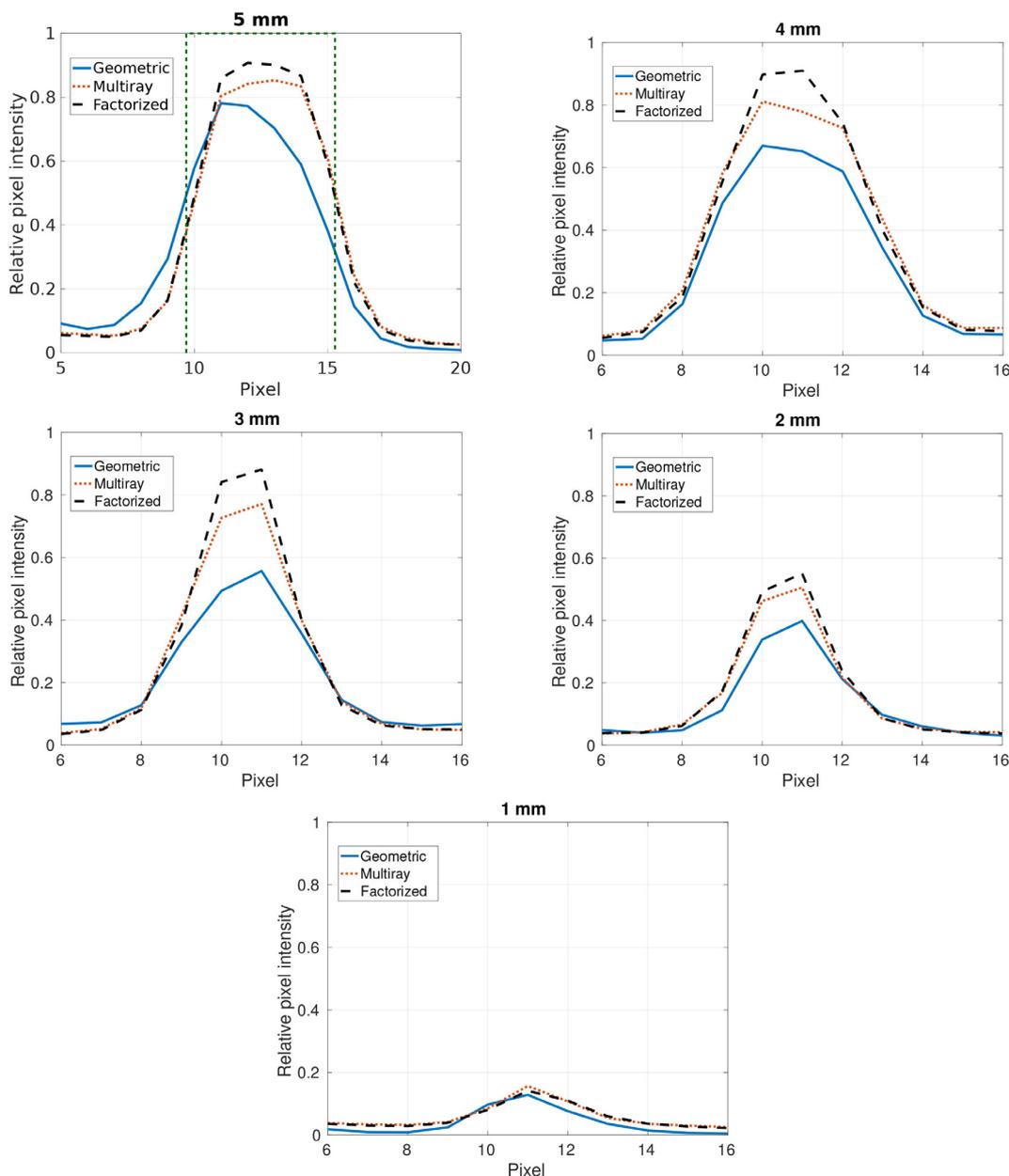


Fig. 5. Line profiles along the transverse plane, after 70 MLEM iterations. On the y-axis is shown the intensity of the pixels relative to the value of the uniform region. The Factorized model provides a slightly better rod shape (in terms of symmetry) and a slightly higher intensity values than the others two models.

[25], the convergence of the reconstruction process depends on the accuracy of the model and it can be slower including resolution recovery models. Therefore, it is worth investigating the trend of the recovery coefficient as a function of the %STD instead of as a function of the number of iterations. At the same %STD the Factorized model yields higher values of the RCs with respect to the other models. The slower convergence of the Factorized model than of the Multiray model is apparent in Fig. 4e and f.

Fig. 5 shows some examples of line profiles of the NEMA phantom rods, obtained by averaging the ten central slices of each rod (iteration 70). The intensity value is relative to the average intensity value of the uniform central region of the phantom. The Factorized model provides slightly better rod shape (in terms of symmetry) and slightly higher intensity values than the others two models.

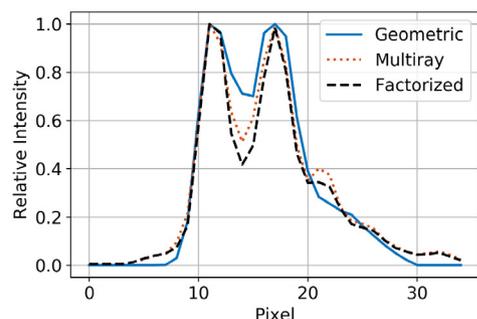
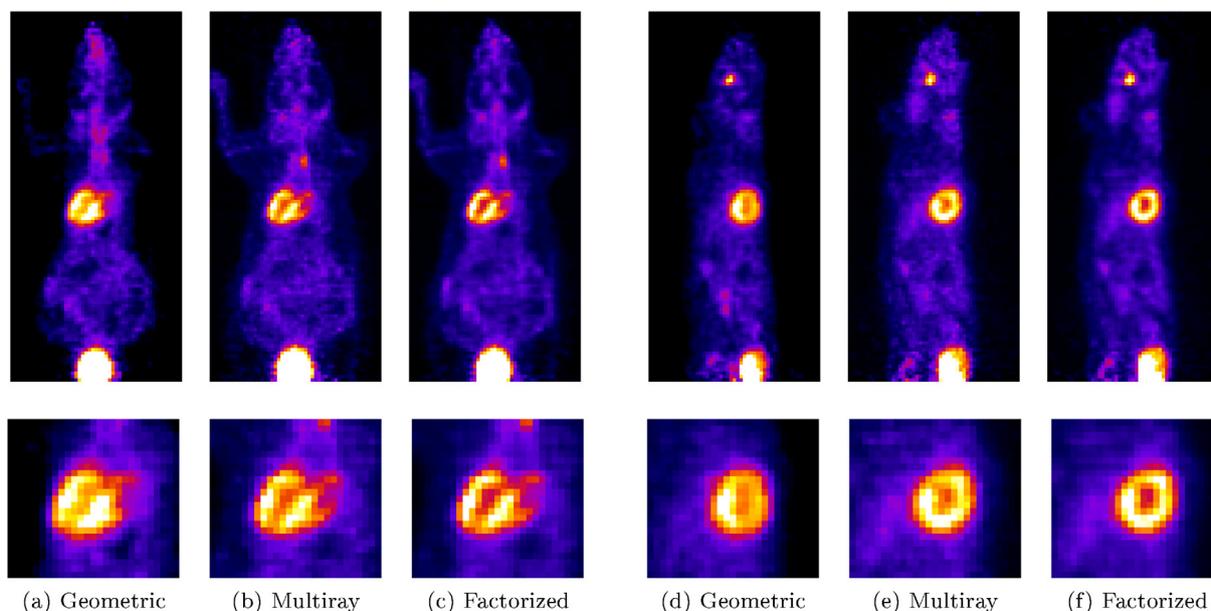
3.3. Preclinical study

Fig. 6 shows an example of reconstruction for a ^{18}F – FDG cardiac

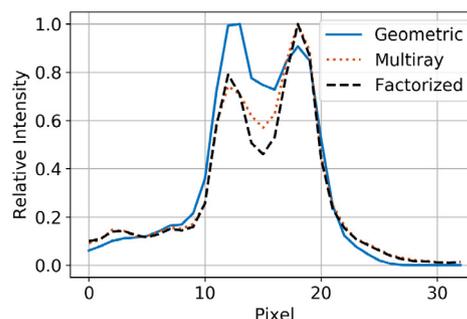
mouse study: top row shows a full body view of the subject and center row shows a magnification of the cardiac region. Fig. 6g shows a sample line profile that intercepts horizontally the cardiac walls shown in Fig. 6a, b and c. Fig. 6h shows a sample line profile that intercepts on the diagonal the cardiac walls shown in Fig. 6d, e and f. The y-value represents the pixel intensity relative to the maximum value. The results show that the Factorized model is able to provide a better delineation of the cardiac walls as the difference between the relative maximum and minimum intensity values (between the two cardiac walls) is higher in the Factorized model with respect to the other models.

4. Conclusions

We proposed a factorized system model for the IRIS PET preclinical scanner. It comprises two components: a geometrical component and a detector response component. The geometric component was analytically computed in previous work with a Siddon based method. The



(g) Horizontal line profile through the cardiac walls.



(h) Horizontal line profile through the cardiac walls.

Fig. 6. Comparison of the reconstructed images by the Geometric, Multiray and Factorised models for a preclinical study, ML-EM 100 iterations. Top row shows a full body view of subject. Center row shows a magnification of the cardiac region relative to the top row figures. Fig. (g) shows a sample line profile that intercepts horizontally the cardiac walls shown in figures (a), (b) and (c). Fig. (h) shows a sample line profile that intercepts on the diagonal the cardiac walls shown in figures (d), (e) and (f). The y-value represents the pixel intensity relative to the maximum value. The cardiac walls of the mouse are better resolved when reconstructing with the factorized model. (Courtesy of Piero Salvadori and Daniele Panetta, Institute of Clinical Physiology of CNR Pisa).

detector model is calculated in this work using a Monte Carlo simulation which allows modeling of the penetration and inter-crystal scatter effects inside the crystal matrices of the detector modules.

The most challenging aspects of this part of the work are the development of a simulation able to provide useful information in a reasonable time, the algebraic computation of the system matrix and the constraint of its size to a value that allows its entire loading into the computer central memory to perform fast image reconstruction. To limit the simulation time, the Monte Carlo is optimized by the tuning of physical and statistical parameters. The computation of the system matrix is obtained with the use of sparse matrix algorithm and intermediate decomposition. The memory requirements of the model are reduced by limiting the number of included LORs with respect to each LOF or, in other words, by limiting the number of stored non-null elements of each row of the matrix. After a suitable compression, the model can be used by the reconstruction software and the reconstructed images can be compared to the ones obtained with the models currently used by the IRIS scanner. The comparison is performed in terms of image quality according to the NEMA NU 4-2008 standard, using real data acquired by the scanner at the *Institute of Clinical Physiology of CNR (Pisa)*.

The results show that the Factorised model has both lower noise and higher recovery coefficients with respect to the other models and it

provides better image quality for small-animal studies. This work also proves the feasibility of the inclusion of the detector matrix in the system model factorization in the case of a complex polygonal scanner geometry.

References

- [1] Del Guerra A, Belcari N, Giuseppina M. Positron emission tomography: Its 65 years. *La Rivista del Nuovo Cimento* 2016;39:155–223.
- [2] Reader AJ. The promise of new pet image reconstruction. *Phys Med* 2008;24(2):49–56.
- [3] Iriarte A, Marabini R, Matej S, Sorzano C, Lewitt R. System models for PET statistical iterative reconstruction: a review. *Comput Med Imaging Graph* 2016;48:30–48.
- [4] Li K, Safavi-Naeini M, Franklin DR, Han Z, Rosenfeld AB, Hutton B, et al. A new virtual ring-based system matrix generator for iterative image reconstruction in high resolution small volume PET systems. *Phys Med Biol* 2015;60(17):6949–73.
- [5] Panin V, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging* 2006;25(7):907–21.
- [6] Rafecas M, Mosler B, Dietz M, Pogl M, Stamatakis A, McElroy D, et al. Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data. *IEEE Trans Nucl Sci* 2004;51(5):2597–605.
- [7] Mumcuoglu E, Leahy R, Cherry S, Hoffman E. Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET. *IEEE Nucl Sci Symp Conf Rec* 1996;3:1569–73.
- [8] Moses WW. Fundamental limits of spatial resolution in pet. *Nucl Instrum Methods Phys Res, Sect A* 2011;648:S236–40.

- [9] Moehrs S, Defrise M, Belcari N, Del Guerra A, Bartoli A, Fabbri S, et al. Multi-ray-based system matrix generation for 3D PET reconstruction. *Phys Med Biol* 2008;53(23):6925–45.
- [10] Alessio A, Kinahan P, Lewellen T. Modeling and incorporation of system response functions in 3-D whole body PET. *IEEE Trans Med Imaging* 2006;25(7):828–37.
- [11] Cecchetti M, Moehrs S, Belcari N, Del Guerra A. Accurate and efficient modeling of the detector response in small animal multi-head PET systems. *Phys Med Biol* 2013;58(19):6713–31.
- [12] Angelis GI, Kotasidis FA, Matthews JC, Markiewicz PJ, Lionheart WR, Reader AJ. Full field spatially-variant image-based resolution modelling reconstruction for the HRRT. *Phys Med* 2015;31(2):137–45.
- [13] Inviscan s.a.s. Imaging System. Online: accessed May 2017. URL: <http://www.inviscan.fr>.
- [14] Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys* 1985;12(2):252–5.
- [15] Belcari N, Camarlinghi N, Ferretti S, Iozzo P, Panetta D, Salvadori PA, et al. NEMA NU-4 Performance Evaluation of the IRIS PET/CT Preclinical Scanner. *IEEE Trans Radiat Plasma Med Sci* 2017;1(4):301–9.
- [16] Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* 1982;1(2):113–22.
- [17] Bertolli O, Eleftheriou A, Cecchetti M, Camarlinghi N, Belcari N, Tsoumpas C. PET iterative reconstruction incorporating an efficient positron range correction method. *Phys Med* 2016;32(2):323–30.
- [18] Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. GEANT4 – a simulation toolkit. *Nucl Instrum Methods PhysRes, Sect A* 2003;506(3):250–303.
- [19] Mendo L, Hernando J. A simple sequential stopping rule for Monte Carlo Simulation. *IEEE Trans Commun* 2006;54(2):231–41.
- [20] Berger M. Stopping-power and range tables for electrons, protons, and helium ions: Physical reference data 1999.
- [21] NIST: Significance of Calculated Quantities. Online: accessed May 2017. URL: <http://physics.nist.gov/PhysRefData/Star/Text/appendix.html>.
- [22] Camarlinghi N, Sportelli G, Del Guerra A, Belcari N. An automatic algorithm to exploit the symmetries of the system response matrix in pet iterative reconstruction. *Phys Med Biol* 2018.
- [23] Eigen: a C++ template library for linear algebra. Online; version 3.3.2. URL: <http://eigen.tuxfamily.org>.
- [24] National Electrical Manufacturers Association. Performance measurements of small animal positron emission tomographs. NEMA Standards Publication, NU4-2008 2008:1-23.
- [25] Rahmim A, Qi J, Sossi V. Resolution modeling in pet imaging: theory, practice, benefits, and pitfalls. *Med Phys* 2013;40(61):064301 .