

# Forecasting the Pharmacokinetics With Limited Early Frames in Dynamic Brain PET Imaging Using Neural Ordinary Differential Equation

Jimin Hong<sup>1</sup>, Student Member, IEEE, Matthias Brendel, Kjell Erlandsson<sup>2</sup>, Hasan Sari<sup>3</sup>, Jiaying Lu<sup>4</sup>, Christoph Clement<sup>5</sup>, Ngoc Vinh Bui<sup>6</sup>, Maria Meindl, Sibylle Ziegler<sup>7</sup>, Senior Member, IEEE, Henryk Barthel, Osama Sabri<sup>8</sup>, Hongyoon Choi<sup>9</sup>, Raphael Sznitman, Axel Rominger<sup>10</sup>, and Kuangyu Shi<sup>11</sup>, Member, IEEE

**Abstract**—In dynamic brain positron emission tomography (PET) studies, acquiring a time series of images, typically lasting more than an hour, is necessary to derive pharmacokinetic parameters. Analytically, these parameters are estimated by establishing kinetic models such as compartment models that consist of sets of ordinary differential equations (ODEs), and by fitting the sparse time-activity curve (TAC) of the tracer. Yet, these models are simplified approximations of highly complex underlying processes, and sufficient samples of TAC are required throughout the entire acquisition, which is not only impractical but also hindered by patient involuntary motion and intrinsic noise. Therefore, recovering samples in missing timeframes are often required, which, in practice, is achieved by interpolation or extrapolation. Here, we introduce a novel deep-learning-based method that utilizes neural ODE (N-ODE) to predict TAC in the extended timeframes by mimicking the analytical method in a data-driven manner. By training N-ODE to solve and fit sets of ODE such that the solution replicates the observed TAC, the

N-ODE converges to the functional shapes that best describe the underlying pharmacokinetic processes. We customized N-ODE to predict the full-dynamic images (12 frames, 60 min), hence pharmacokinetic parameters, given limited early-frame images (7–9 frames, 20–30 min). For proof of concept, the proposed N-ODE was applied to simulated and clinical <sup>18</sup>F-PI-2620 brain PET. We demonstrated that the proposed N-ODE delivered promising performance, indicated by bias, variance, and mean absolute error as well as pharmacokinetic parameters, such as rate constants, standardized uptake value ratio (SUVr), and binding potential (BP<sub>ND</sub>).

**Index Terms**—Artificial intelligence, dynamic positron emission tomography (PET), neuroimaging, positron emission tomography PET.

Manuscript received 4 October 2022; revised 21 January 2023; accepted 27 February 2023. Date of publication 7 March 2023; date of current version 3 July 2023. This work was supported in part by the Swiss National Science Foundation (SNF) under Grant 188350. (Corresponding author: Jimin Hong.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Local Ethics Committee under Application Nos. 17-569 and 19-022.

Jimin Hong, Christoph Clement, Axel Rominger, and Kuangyu Shi are with the Department of Nuclear Medicine, University of Bern, 3010 Bern, Switzerland (e-mail: jimin.hong@students.unibe.ch; christoph.clement@students.unibe.ch; axel.rominger@insel.ch; kuangyu.shi@dbmr.unibe.ch).

Matthias Brendel, Ngoc Vinh Bui, Maria Meindl, and Sibylle Ziegler are with the Department of Nuclear Medicine, Ludwig-Maximilians-University of Munich, 80539 Munich, Germany (e-mail: matthias.brendel@med.uni-muenchen.de; Vinh.Bui@outlook.de; Maria.Meindl@med.uni-muenchen.de; Sibylle.Ziegler@med.uni-muenchen.de).

Kjell Erlandsson is with the Institute of Nuclear Medicine, University College London, WC1E 6BT London, U.K. (e-mail: k.erlandsson@ucl.ac.uk).

Hasan Sari is with the Advanced Clinical Imaging Technology, Siemens Healthcare AG, 1015 Lausanne, Switzerland (e-mail: hasan.sari@siemens-healthineers.com).

Jiaying Lu is with the Huashan Hospital, Fudan University, Shanghai 200437, China (e-mail: jiaying.lu@students.unibe.ch).

Henryk Barthel and Osama Sabri are with the Department of Nuclear Medicine, University Hospital Leipzig, 04103 Leipzig, Germany (e-mail: henryk.barthel@medizin.uni-leipzig.de; osama.sabri@medizin.uni-leipzig.de).

Hongyoon Choi is with the Department of Nuclear Medicine, Seoul National University Hospital, Seoul 03080 Republic of Korea (e-mail: chy1000@snu.ac.kr).

Raphael Sznitman is with the ARTORG Center for Biomedical Engineering Research, University of Bern, 3012 Bern, Switzerland (e-mail: raphael.sznitman@unibe.ch).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TRPMS.2023.3253261>.

Digital Object Identifier 10.1109/TRPMS.2023.3253261

## I. INTRODUCTION

**P**OSITRON emission tomography (PET) is a molecular imaging technology that uses a radiotracer to visualize the metabolic and physiological processes in the body in vivo. For brain studies, the recent advances in PET tracers made it possible to assess several histopathological targets of neurodegenerative disorders, such as glucose metabolism, amyloid, tau, and neuroinflammation [1], [2], [3], [4], [5], [6]. The common application of PET in the clinics yields a simple snapshot of regional tracer concentration using a single timeframe post-injection, which is referred to as static PET. On the other hand, dynamic PET provides comprehensive information about the pharmacokinetics of the tracer, such as tracer delivery ( $K_1$ ), volume of distribution ( $V_T$ ), or binding potential (BP<sub>ND</sub>) by delineating both the spatial and temporal profiles of the tracer uptake [7], [8], [9]. Analytically, such pharmacokinetic parameters are estimated through kinetic models that translate hypotheses regarding the physiological system of dynamic processes in the tissue of interest, which is decomposed into several compartments, in a mathematical formulation [10], [11], [12], [13]. These models inevitably accompany a set of ordinary differential equations (ODEs), and the pharmacokinetic parameters are determined by solving the ODEs and fitting the time-activity curve (TAC) of each voxel.

Yet, analytical models are simplified approximations of highly complex underlying physiological and biological

processes and are not perfectly consistent with the measured data due to intrinsic PET noise, especially at the voxel level [14]. Additionally, sufficient numbers of samples across the entire protocol are required to perform the modeling, which is hindered by patient movement [15]. As such, in routine practice, the reduced acquisition time PET has been proposed, where kinetic modeling is performed with missing TAC samples either being ignored or being recovered by various interpolation or extrapolation methods [16], [17], [18], [19], [20], [21]. Scott et al. [16] combined arterial spinning labeling (ASL)-derived cerebral blood flow (CBF) data with  $^{18}\text{F}$ -florbetapir amyloid PET kinetic modeling and reduced the acquisition time from 60 to 30 min. In their work, the missing samples in reference region TAC were extrapolated. Heeman et al. [17] and Kolinger et al. [18] proposed the dual-time window protocols for  $^{18}\text{F}$ -flutemetamol,  $^{18}\text{F}$ -florbetaben, and  $^{18}\text{F}$ -MK-6240 which consist of two separate PET scans with a resting period in between. The omitted data points in the resting phase were interpolated either linearly or using compartment models. Viswanath et al. [19] shortened 65-min dynamic FDG protocol in long axial field-of-view (LAFOV) PET scans to 15–20-min dual-time window protocol where 10–15 min early frame and late 5 min were used. Wu et al. [20] demonstrated that a much-reduced scanning time can achieve the net influx rate,  $K_i$ , imaging using the nonlinear estimation method. Liu et al. [21] validated that an acquisition time reduction to 45 min was possible in a single-time window protocol to draw kinetic parameters of  $^{18}\text{F}$ -FDG as done by full dynamic protocol. In works from Varsha, Wu, and Liu, neither interpolation nor extrapolation was performed to construct absent samples in TAC.

Here, we introduce a novel deep-learning method based on neural ODE (N-ODE) [22], which recovers images in the missing timeframe by mimicking the analytical modeling in a data-driven manner. That is, by training N-ODE to solve sets of ODE such that the solution replicates the observed TAC, the N-ODE parameters converge to the functional shapes that best describe the underlying pharmacokinetic processes. To the best of our knowledge, N-ODE has been explored in various domains to leverage the existing well-established physical or chemical models that are explained by the ODE system in a machine-learning manner [23], [24], [25], yet, not in PET studies. In this article, we built an N-ODE-based model that learns the spatiotemporal features of the dynamic PET images and the underlying pharmacokinetic processes by embedding convolutional long short-term memory (LSTM), namely, seq2seq [26], [27] and variational autoencoder (VAE) [28] architect into N-ODE network. We customized the model to predict the total frames (12 frames, 60 min), hence the pharmacokinetic parameters, given a few early frames as input (7–9 frames, 20–30 min). For proof of concept, we applied this model to the simulated as well as the clinical dynamic  $^{18}\text{F}$ -PI-2620 brain PET. With limited early frames alone, we demonstrated that the proposed method is capable of recovering the images across full dynamic protocol, indicated by bias, variance, and mean absolute error (MAE), as well as the pharmacokinetic parameters, such as rate constants ( $K_1$ ,  $k_2$ ,  $k_3$ ,  $k_4$ ),  $\text{BP}_{\text{ND}}$ , and the standardized uptake value ratio (SUVr).

The proposed method has the potential to solve the challenges largely in dynamic PET research, such as reducing the protocol acquisition studies or rapid dual-tracer studies [29], [30].

## II. MATERIALS AND METHODS

The mathematical concept of N-ODE in relation to the compartment model is explained in Section II-A, and the study flow is revisited in Section II-B. The simulation and the clinical data for the preparation of this article are described in Sections II-C and II-D. The indices to assess the model performance are summarized in Section II-E.

### A. Compartment Model and N-ODE

For simplicity, we consider N-ODE with regard to a simple compartment model. For each compartment, the change of tracer concentration is described in terms of a linear, first-order, constant-coefficient ODE as

$$\frac{d\mathbf{C}_i(t)}{dt} = f_i(\mathbf{C}_0(t), \mathbf{C}_1(t), \mathbf{C}_2(t), \dots, t, K_1, k_2, k_3, k_4, \dots) \quad (1)$$

where  $\mathbf{C}_i(t)$  is the concentration in compartment  $i$  in time  $t$ , and  $K_1$ ,  $k_2$ ,  $k_3$ , and  $k_4$  are rate constants.

For instance, the two-tissue compartment model using four rate constants (2T4CM) is written as

$$\frac{d\mathbf{C}_1(t)}{dt} = K_1\mathbf{C}_0(t) - (k_2 + k_3)\mathbf{C}_1(t) + k_4\mathbf{C}_2(t) \quad (2)$$

$$\frac{d\mathbf{C}_2(t)}{dt} = k_3\mathbf{C}_1(t) - k_4\mathbf{C}_2(t). \quad (3)$$

The typical interpretation is that  $\mathbf{C}_0(t)$  is the tracer concentration of the input function derived from arterial plasma.  $\mathbf{C}_1(t)$  represents free and nonspecifically bound tracer concentration, and  $\mathbf{C}_2(t)$  is specifically bound tracer concentration in tissue.

Analytic solutions for 2T4CM are derived [31] by using Laplace transformation, as

$$\mathbf{C}_1(t) = \frac{K_1}{\alpha_2 - \alpha_1} [(k_4 - \alpha_1)e^{-\alpha_1 t} + (\alpha_2 - k_4)e^{-\alpha_2 t}] \otimes \mathbf{C}_0(t) \quad (4)$$

$$\mathbf{C}_2(t) = \frac{K_1 k_3}{\alpha_2 - \alpha_1} [e^{-\alpha_1 t} - e^{-\alpha_2 t}] \otimes \mathbf{C}_0(t) \quad (5)$$

and the total concentration of tracer in the region of interest (ROI) can be expressed as

$$\mathbf{C}_T(t) = \mathbf{C}_1(t) + \mathbf{C}_2(t) \quad (6)$$

where

$$\alpha_1 = \frac{k_2 + k_3 + k_4 - \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}}{2}$$

$$\alpha_2 = \frac{k_2 + k_3 + k_4 + \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}}{2}.$$

N-ODE is a deep-learning method different from the conventional compartment models, as it uses trainable parameters to fit the data. That is, the key idea of N-ODE is built on parameterizing a continuous dynamical system that is in the format of ODE with an initial value, using a neural network (NN)

$$\frac{d\mathbf{h}(t)}{dt} = f(\mathbf{h}(t), t, \theta), \mathbf{h}(t_0) = \mathbf{h}_0. \quad (7)$$

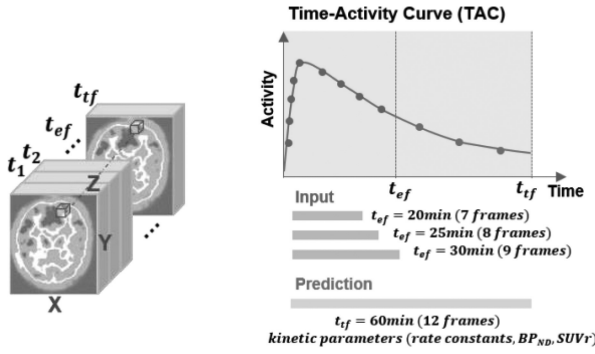


Fig. 1. Scheme of the study design. For simplicity, the 4-D (3-D image, 1-D time) spatiotemporal component of the study design (left) is drawn in 1-D voxel level (right). The proposed N-ODE-based model consumes the 3-D images across the time length of early frames ( $t_{ef}$ ) as an input. The model predicts the 3-D images across total frames ( $t_{tf}$ ), hence the pharmacokinetic parameters.

Here,  $f(\cdot)$  is parameterized by  $\Theta$ , which are the weights of the NN, instead of the rate constants described in (1)–(3).  $\mathbf{h}(t)$  can be defined as the function governing state evolution over time  $t$ , such as  $C_T(t)$  expressed in (6). The output state  $\mathbf{h}(t_i)$  in time  $t_i$  can be mapped using (7) and the ODE solver. N-ODE is trained by means of the loss function, evaluating the distance between the output  $\mathbf{h}(t_i)$  and the desired measurement  $\mathbf{h}(t_i)_{\text{measured}}$

$$\begin{aligned} L(\mathbf{h}(t_i)) &= L\left(\mathbf{h}(t_0) + \int_{t_0}^{t_i} f(\mathbf{h}(t), t, \theta) dt\right) \\ &= L(\text{ODEsolve}(\mathbf{h}(t_0), f, t_0, t_i, \theta)) \end{aligned} \quad (8)$$

where  $\text{ODEsolve}$  is given by the established ODE solvers, such as Runge–Kutta methods. For backward propagation, the adjoint sensitivity method is applied to the ODE solver to compute the gradients of loss with respect to parameters  $\Theta$  for memory efficiency [22].

### B. Study Design

Fig. 1 illustrates the study design. In this work, we set the timeframe for both simulation and the clinical brain PET data as  $[5 \times 120(\text{s}); 4 \times 300(\text{s}); 3 \times 600(\text{s})]$ . Our baseline was to predict the full dynamic 4-D images (12 frames, 60 min), given the initial input (9 frames, 30 min) using the proposed N-ODE model. We then analyzed the model with fewer early input frames (8 frames, 25 min; 7 frames, 20 min).

Fig. 2(a) outlines the proposed N-ODE-based method. The multilayered LSTM [26], namely, seq2seq architect [27], was built as a recognition network for VAE to map the early-frame images,  $x_{t_0}, x_{t_1}, \dots, x_{t_{ef}}$ , into a local initial state,  $z_{t_0}$ , through the hidden state,  $h_{t_0}, h_{t_1}, \dots, h_{t_{ef}}$ . Seq2seq consumed the early-frame images sequentially backwards in time, and the VAE encoder outputted  $q_{\theta}(z_{t_0}|x_{t_0}, \dots, x_{t_{ef}})$ . A global latent dynamics are shared across all time frames, and N-ODE produces a set of latent features for each frame,  $z_{t_1}, z_{t_2}, \dots, z_{t_{tf}}$ , by solving ODE using the Runge–Kutta of order 5 method, given the initial value of  $z_{t_0}$

$$z_{t_1}, z_{t_2}, \dots, z_{t_{tf}} = \text{ODEsolve}(z_{t_0}, f, t_0, t_1, \dots, t_{tf}, \theta_f). \quad (9)$$

The latent feature derived by the ODE solver is then translated back to image space by the VAE generator,  $p_{\theta}(x_{t_0}, \dots, x_{t_{tf}}|z_{t_0}, \dots, z_{t_{tf}})$ .

The classic VAE loss is expressed as

$$\begin{aligned} L(\varnothing, \theta) &= -E_{z \sim q_{\varnothing}(z|x)}(\log p_{\theta}(x|z)) \\ &\quad + KL(q_{\varnothing}(z|x)||p_{\theta}(z)) \end{aligned} \quad (10)$$

where the probabilistic generator can be expressed as  $p_{\theta}(x|z)$ , given the data sample,  $x$ , projected in the latent space,  $z$ . The posterior distribution  $p_{\theta}(z|x)$  is obtained by using the prior distribution  $p(z)$  and the probabilistic generator  $p_{\theta}(x|z)$ , such that  $p_{\theta}(z|x) \sim p(z)p_{\theta}(x|z)$ . The encoder learns an approximation  $q_{\varnothing}(z|x)$  to the posterior distribution  $p_{\theta}(z|x)$ , where  $\varnothing$  denotes the parameters of the encoder, and  $\theta$  stands for those of the generator. The first term in (10) defines the reconstruction loss, while the second term acts as a regularization term, in the form of the Kullback–Leibler (KL) divergence between the latent distribution learned and the prior distribution. In practice, the generator input is resampled by the encoded latent features  $z$

$$z_{\text{resampled}} = z_{\text{encoder}} + z_{sd} \times \varepsilon \quad (11)$$

where  $\varepsilon$  represents a random noise sample.

In this work, instead of  $q_{\varnothing}(z|x)$ , the model seeks  $q_{\varnothing}(z_{t_0}|x_{t_0}, \dots, x_{t_{ef}})$ . The model then finds  $z_{t_1}, z_{t_2}, \dots, z_{t_{tf}}$  as expressed in (9), hence,  $p_{\theta}(x|z)$  was replaced with  $p_{\theta}(x_{t_0}, \dots, x_{t_{tf}}|\text{ODEsolve}(z_{t_0}, f, t_0, t_1, \dots, t_{tf}, \theta_f))$ . The model was trained end-to-end and negative Gaussian log-likelihood loss was used as the reconstruction loss term together with KL divergence as described in (10).

The recognition seq2seq network was built with 2 convolution layers with 32 channels of hidden states. We parameterized the dynamics function  $f(\cdot)$  with six convolutional layers for each encoder and generator. The original image size was  $(91, 109, 91, 12)$ , each of which accounts for the dimension across  $x, y, z$ , and the time,  $t$ , respectively. However, for computational efficiency, the patches are extracted randomly that are half of the original dimensions in the  $x$ -,  $y$ -, and  $z$ -axis for training as depicted in Fig. 2(b). For the test set, the regular grid patch was obtained, and the outputs of each patch are returned to the original image size, averaging the values in overlapping areas. All images are normalized to maximum value before feeding into the model. We trained on NVIDIA GeForce RTX 2080 Ti graphic card.

### C. Simulated Dynamic Brain PET

The input function was first simulated, and the dynamic brain PET was created by using 2T4CM described in (4)–(6).

Feng’s model was used to generate the input function [32]

$$C_0(t) = (A_1 t - A_2 - A_3)e^{\lambda_1 t} + A_2 e^{\lambda_2 t} + A_3 e^{\lambda_3 t} \quad (12)$$

where  $A_1, A_2$ , and  $A_3$  are amplitude coefficients, and  $\lambda_1, \lambda_2$ , and  $\lambda_3$  are eigenvalues of the system. For simulation, we did not consider the metabolite correction [8]. The dynamic brain PET image was generated using Hammers atlas to draw ROIs [33]. For each ROI, the concentration of the tracer,  $C_S(t)$ , is simulated as

$$C_S(t) = V_B C_0(t) + (1 - V_B) C_T(t) \quad (13)$$

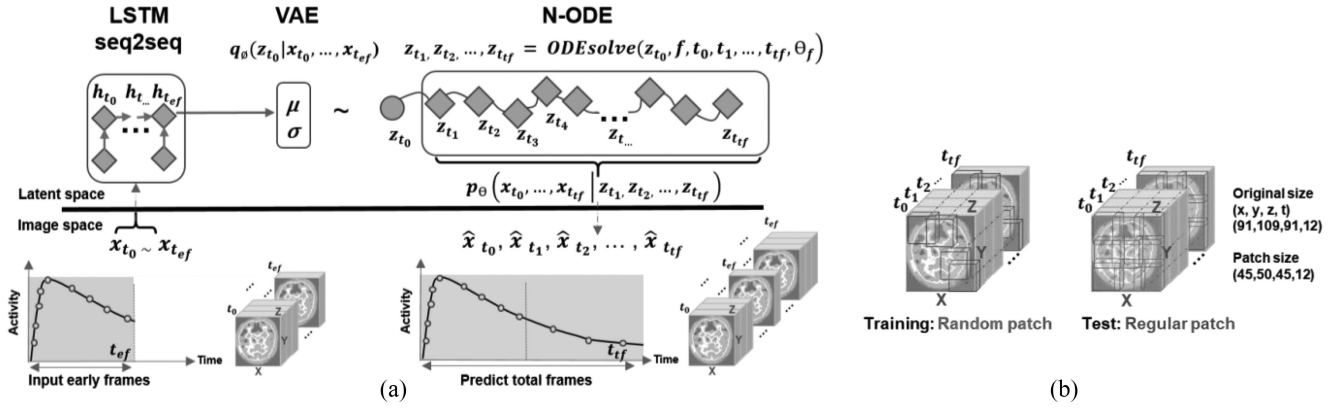


Fig. 2. Study design. (a) Customized N-ODE, using two-layer LSTM, seq2seq architect, and VAE. (b) Illustration of random patch extraction and regular grid extraction for each training and test.

TABLE I  
PREDEFINED MINIMUM AND MAXIMUM VALUE OF PARAMETERS FOR  
INPUT FUNCTION AND 2T4CM SIMULATION

Model	Parameter	Minimum, Maximum
Feng's Model	$A_1$	37, 370 (kBq/ml/min)
	$A_2$	3.7, 185 (kBq/ml)
	$A_3$	3.7, 185 (kBq/ml)
	$\lambda_1$	-5, -0.1 ( $\text{min}^{-1}$ )
	$\lambda_2$	-1, -0.1 ( $\text{min}^{-1}$ )
Two tissue – four parameter compartment model (2T4CM)	$\lambda_3$	-0.1, -0.01 ( $\text{min}^{-1}$ )
	$K_1$	0.1, 0.3 (mL/min/g)
	$k_2$	0.01, 0.3 ( $\text{min}^{-1}$ )
	$k_3$	0.01, 0.1 ( $\text{min}^{-1}$ )
	$k_4$	0.01, 0.05 ( $\text{min}^{-1}$ )

where  $V_B$  is the vascular volume fraction inside the ROIs, which is fixed as 0.05 for this work. The left and right lobes of each ROI are fixed to have the same concentration.

The parameters required for the input function and the rate constants for 2TCM were randomly sampled from the uniform distribution with the predefined minimum and maximum values as shown in Table I.

The noise was assumed to be proportional to the number of counts per time frame. The simulated 4-D brain images were reconstructed using an in-house ordered-subsets expectation-maximization (OSEM) algorithm [30]. 600 and 100 samples were created for training and test set, respectively.

#### D. Clinical Dynamic $^{18}\text{F}$ -PI-2620 PET

$^{18}\text{F}$ -PI-2620 PET scans in combination with computed tomography (CT) was performed in a full dynamic setting (0–60 min post-injection). The injected dose was  $18966 \pm 13.19$  MBq and applied as a bolus injection. All data were acquired in Munich on a Siemens Biograph True point 64 PET/CT (Siemens, Erlangen, Germany) or a Siemens mCT (Siemens, Erlangen, Germany). The dynamic brain PET data were acquired in 3-D list mode over 60 min, and reconstructed into a  $336 \times 336 \times 109$  matrix (voxel size:  $1.02 \times 1.02 \times 2.03$  mm<sup>3</sup>) using the built-in OSEM algorithm with 4 iterations, 21 subsets, and a 5-mm Gaussian on the Siemens Biograph and with 5 iterations and 24 subsets on the

Siemens mCT. A low-dose CT served for attenuation correction. All images were spatially normalized to the Montreal Neurological Institute (MNI) space using an in-house tracer-specific PET image template (PMOD V4.1). Fifty eight sets were collected in total, where 56 sets and 17 sets are used as training and test, respectively. Amongst 56 training dataset, 22 were from mCT and 34 were from Biograph. Out of 17 test data set, 6 were from mCT and 11 were from Biograph.

#### E. Evaluation of the Performance

We assessed the performance of the proposed method by calculating the following properties for each ROI and timeframe:

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - \hat{x}_i|}{x_i} \quad (14)$$

$$\text{Variance} = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}_N}{x_i} \right)^2 \quad (15)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad (16)$$

where  $N$  is the total number of voxels in one ROI,  $x_i$  and  $\hat{x}_i$  are the ground truth and the prediction of the  $i$ th voxel, respectively, and  $\bar{x}_N$  represents the average predicted value of voxels in the ROI. While MAE describes the average magnitude of error produced by the proposed model, bias characterizes how far the model's predictions are off from the ground truth. As bias is normalized to the prediction itself, it represents the error taking the magnitude of TAC into consideration.

In addition, the pharmacokinetic parameters from both the ground truth and the predicted images are estimated and compared. For simulation datasets, the rate constants  $K_1$ ,  $k_2$ ,  $k_3$  and  $k_4$  were evaluated for each ROIs. On the other hand, voxel-wise indirect parameters, such as SUVr and binding potential ( $\text{BP}_{\text{ND}}$ ), are assessed for  $^{18}\text{F}$ -PI-2620 datasets, as an arterial input function was not available. The average image of late 30 min-3 frames was normalized by the mean uptake in cerebellum gray matter to derive SUVr [34].  $\text{BP}_{\text{ND}}$  images were estimated using a noninvasive Logan graphical plot

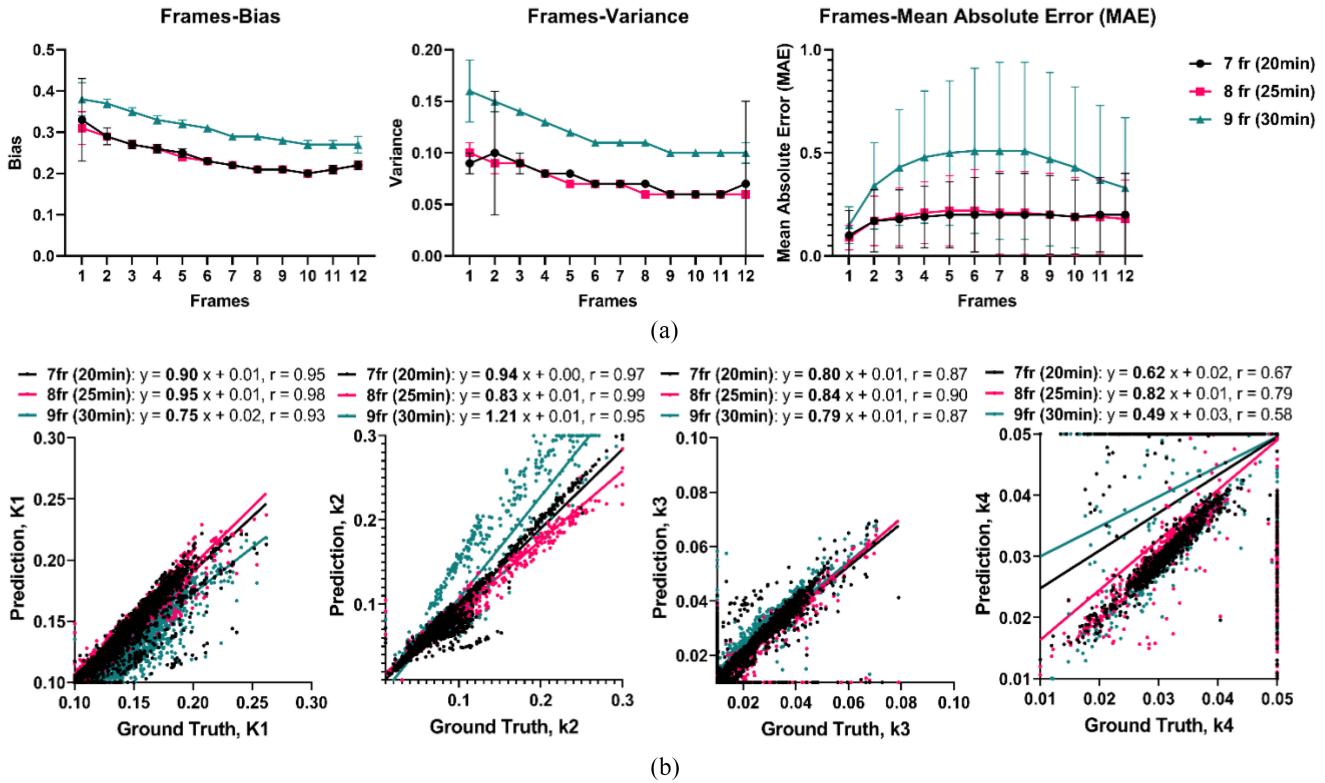


Fig. 3. Model performance. (a) Bias, variance, and the MAE of the proposed model. The model was evaluated using 7 frames (20 min), 8 frames (25 min), and 9 frames (30 min) as an input. (b) Linear regression of rate constants ( $K_1$ ,  $k_2$ ,  $k_3$ , and  $k_4$ ) estimated by the ground-truth dynamic image and by the proposed model, with different numbers of input frames.

( $r^* = 20$  min,  $k_2' = 0.22$  min $^{-1}$ ) and cerebellar gray matter as the reference region [34].

### III. RESULTS

#### A. Simulated Data Result

Fig. 3(a) illustrates the average bias, variance, and MAE of the outputs resulting from the proposed N-ODE-based model, in each of the 12 frames. Lower bias and variance were observed in the later frames than in the early frames regardless of the number of early input frames. MAE was higher in the middle frames when the model was given more input frames (9 frames, 30 min). However, with fewer input frames (7 frames, 20 min; 8 frames 25 min), MAE was similar across all frames except for the first 1 or 2 frames. The average bias, variance, and MAE indicated the better performance with fewer input frames (7 frames, 20 min) than with more input frames (9 frames, 30 min). Yet, the profiles of bias, variance, and MAE were almost identical when the model was given with 7 input frames (20 min) and 8 input frames (25 min).

Fig. 3(b) depicts the linear fitting result between the rate constants estimated from the ground truth and the predicted dynamic image. The model performed worse in predicting  $k_4$ , compared to predicting  $K_1$ ,  $k_2$ , and  $k_3$ . In general, eight input frames resulted in the best prediction of the rate constants, except for  $k_2$  where seven input frames achieved best.

Fig. 4 illustrates the examples of the test data sets that showed the best and the worst MAE. For all three settings with different numbers of input frames, the same data sets

were identified as the best and worst results. Fig. 4 depicts the examples when seven frames (20 min) were given as an input. The worst prediction that was made in this case presented high bias near edges, however, preserved the overall contrast between the ROIs across all timeframes. For both best and worst examples, the horizontal and vertical strikes were observed, which resulted from patch-based learning.

#### B. Clinical $^{18}\text{F}$ -PI-2620 Data Result

Fig. 5 illustrates the voxel-level parametric images of SUVr and BP<sub>ND</sub> in two examples of  $^{18}\text{F}$ -PI-2620 test datasets with different numbers of input frames. For a visual assessment at the image level, the region with the high uptake was well preserved in both test sets, regardless of the number of input frames given to the model. Table II shows the average bias, variance, and MAE resulting from the proposed model on the  $^{18}\text{F}$ -PI-2620 data for each of the 12 frames. The average bias, variance, and MAE improved with more input frames, although the differences were small. Fig. 6 is the heat map representing the average bias, variance, and MAE in ROIs drawn from the Hammers template across all 12 frames. On average, the bias, variance, and MAE decreased with more input frames, yet, their general aspects were unchanged. For instance, the central structures, such as caudate nucleus (CaudateNucl) and corpus callosum (Corp\_callosum) as well as ventricle such as frontal horn (FrontalHorn) presented higher bias than other ROIs, regardless of the number of input frames. Overall, the central structures, including putamen, thalamus, and substantia

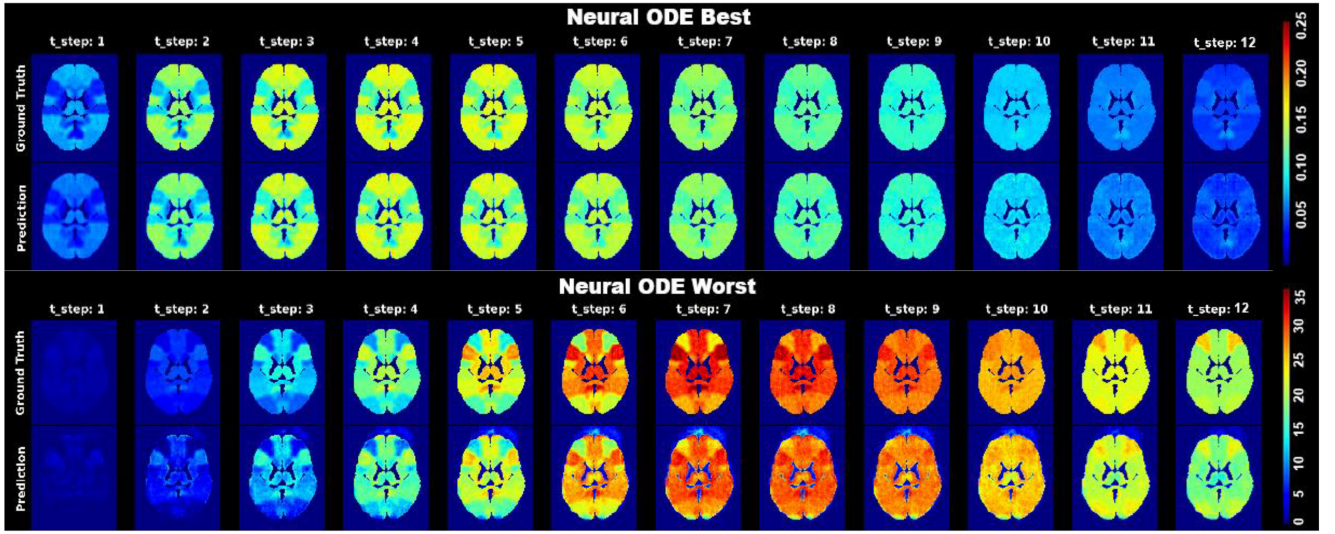


Fig. 4. Best and worst test examples of the proposed N-ODE-based model with 7 early frames (20 min) as an input. The first row of each example indicates the ground-truth image and the second row depicts the predicted images by the N-ODE-based model in axial view.

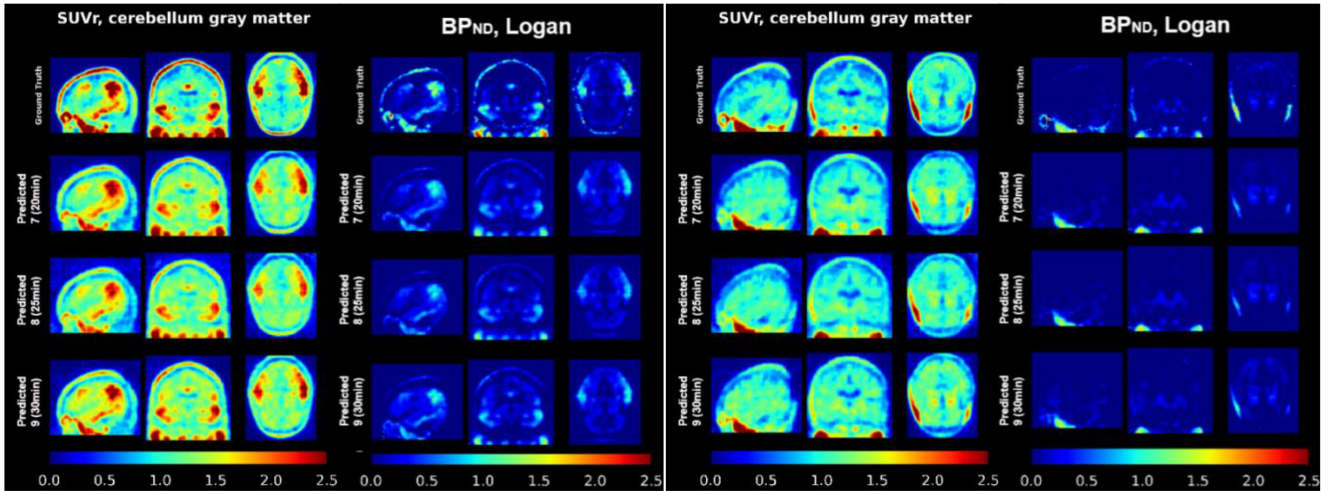


Fig. 5. Voxel-wise parametric image of SUVr and binding potential ( $BP_{ND}$ ) in two examples of  $^{18}F$ -PI-2620 test sets. Both SUVr and  $BP_{ND}$  images are estimated by the ground truth (first row), and the predicted images with different numbers of input frames [second-last row: 7 frames (20 min), 8 frames (25 min), and 9 frames (30 min)].

nigra ( $S_{nigra}$ ), showed relatively high MAE. However, low MAE was observed in the pallidum, corpus callosum, and frontal horn. The frontal lobe, including the middle frontal gyrus ( $FL_{mid\_fr\_G}$ ), inferior frontal gyrus ( $FL_{inf\_fr\_G}$ ), and superior frontal gyrus ( $FL_{sup\_fr\_G}$ ), reported relatively high variance. The variance was more dependent on ROI than on the timeframes. In other words, the ROIs which showed comparably higher variance presented higher variance throughout the entire timeframes. However, bias and MAE were dependent not only on the ROIs but also on the timeframes. Relatively high MAE was observed in early 1-5 frames, compared to later frames. Conversely, higher bias was observed in the later frames, in comparison to the early frames.

#### IV. DISCUSSION

In this article, we introduced N-ODE and proposed the N-ODE-based method that recovers the missing images in dynamic brain PET, allowing the pharmacokinetic parameters

such as rate constants or  $BP_{ND}$  to be estimated with a few numbers of early frames alone. The proposed N-ODE-based method, which is established to resemble the analytical methods such as compartment modeling, exploited the intrinsic pharmacokinetics of the tracer by solving and fitting sets of ODEs, however, in a deep-learning manner. We demonstrated the result of the proposed method with the simulated and the clinical  $^{18}F$ -PI-2620 dynamic brain PET images. We then assessed the performance with various numbers of early input frames that are given to the model as an initial value decreasing from 9 to 7, which is equivalent to 30–20 min. We believe that this work will contribute to the dynamic PET studies in general, not only in a direction to reducing the scanning time but also to rapid multitracer studies [29], [30].

##### A. Time Frames Used in the Proposed N-ODE-Based Model

In this proposed study, the duration of the early frames was extended in comparison to typical common settings. We

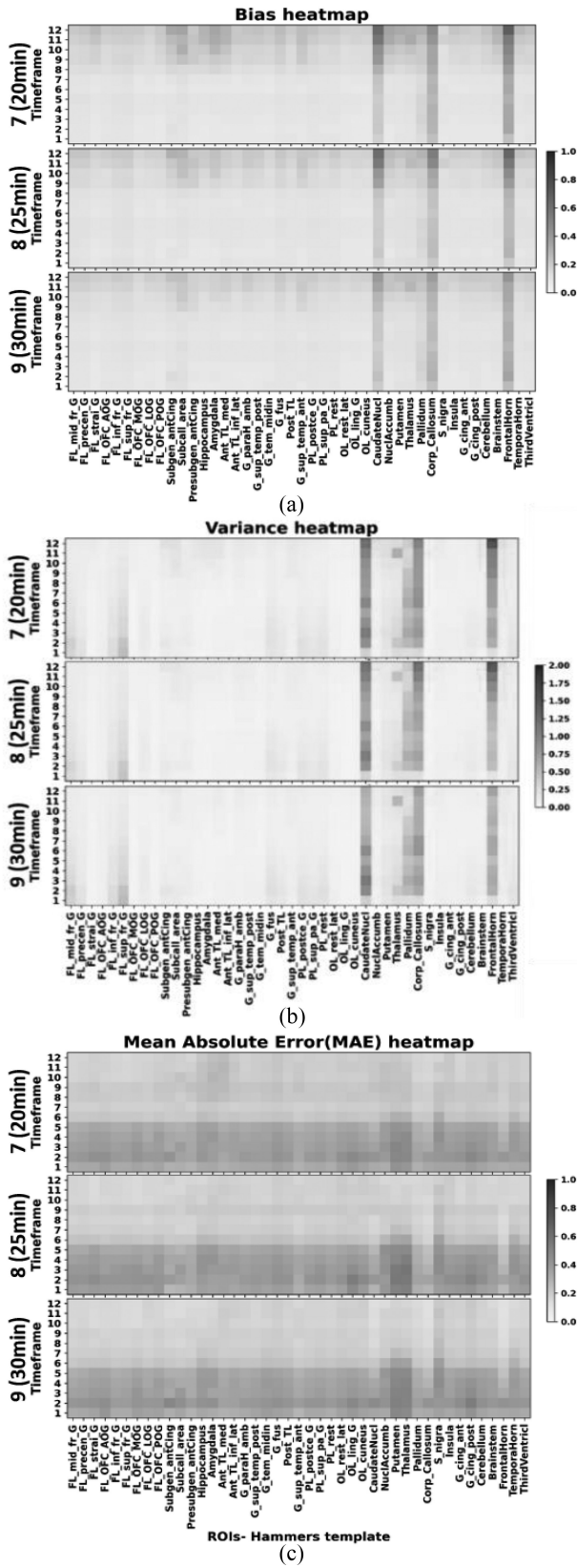


Fig. 6. Heat map of (a) bias, (b) variance, and (c) MAE. ROIs are drawn by the Hammers template [33].

observed that the initial value holds significant importance in solving ODEs using the N-ODE-based model. Therefore, to mitigate the impact of noisy initial values, the duration of the

early frames was increased. Additionally, implementing a post-processing or reconstruction algorithm that reduces noise may further enhance the performance of the proposed method.

*B. N-ODE-Based Model on Simulation and <sup>18</sup>F-PI-2620 Data*

The results of the simulation data indicate that the extended duration of early frames did not harm the prediction of  $K_1$  and  $k_2$ . However, the prediction of  $k_3$  and  $k_4$  was suboptimal (Fig. 3). As the proposed model predicts the entire 12 frames given 7, 8, and 9 frames, it may be simpler for the model to forecast the initial 7–9 frames than last 5–3 frames as they are seen by the model. Due to the fact that  $K_1$  and  $k_2$  are closely related to the perfusion, or early frames, these parameters may be easier for models to predict than  $k_3$  or  $k_4$ . Furthermore,  $k_4$  was estimated more poorly than  $k_3$ , possibly because of its relatively small value and error propagation.

The proposed method using N-ODE is demonstrated to be sensitive to the initial value as evidenced by the heatmaps of bias and variance from <sup>18</sup>F-PI-2620 data presented in Fig. 6. Poor prediction of the initial value in certain ROIs, such as central structures, resulted in a lack of improvement in predictions made for subsequent time frames. However, the MAE was observed to be relatively lower in some of the central structures. This can be attributed to the fact that MAE is calculated as the difference between the ground truth and prediction value, and these ROIs have lower values. On the other hand, these ROIs resulted in a high bias and variance, as both indices are normalized by the ground-truth value itself [(14) and (15)]. The ROIs which are small in size and value suffered more in prediction when the noise was presented. On the other hand, distinctive horizontal lines were found in bias and MAE heatmaps, indicating that both indices were affected also by the time frame. Specifically, <sup>18</sup>F-PI-2620 peaks before the fifth frame, which is equivalent to 10 min [34], resulting in higher MAE up until the fifth frame. However, the bias was low until the last few frames. As a result, we observed that when MAE loss was selected as the reconstruction loss, predictions were poorly made in later frames as the model sought to find an optimal solution by primarily focusing on the early frames.

*C. Characteristics of Kinetics and the N-ODE-Based Model*

The number of early frames used as input to achieve the best result was different between the simulation and the clinical <sup>18</sup>F-PI-2620 data. The bias, variance, and MAE indicated that the proposed model performed better with fewer numbers of input frames (7 frames, 20 min; 8 frames, 30 min) for the simulated data as shown in Fig. 3(a), however, the reversed trend was found for the clinical data as summarized in Table II. We reflect that this is coming from the intrinsic dynamic profile of the tracer and the model architect. For <sup>18</sup>F-PI-2620, the TAC peaks at around 5 min [34] and drops rapidly, while, for the simulation data, TAC washes out slowly. As the convolutional LSTM model, seq2seq, is utilized as a recognition network for embedding the early frames backwards in time, the proposed model seemed to have a tradeoff between the

TABLE II  
BIAS, VARIANCE, AND MAE OF THE PROPOSED N-ODE-BASED MODEL ON THE 18F-PI-2620 TEST DATASETS

Frames Input Output	Bias			Variance			Mean Absolute Error		
	7(20min)	8(25min)	9(30min)	7(20min)	8(25min)	9(30min)	7(20min)	8(25min)	9(30min)
1	0.09 ±0.01	0.10 ±0.01	0.09 ±0.01	0.12 ±0.05	0.10 ±0.04	0.12 ±0.05	0.33 ±0.07	0.36 ±0.08	0.34 ±0.08
2	0.09 ±0.01	0.09 ±0.01	0.10 ±0.01	0.17 ±0.08	0.16 ±0.08	0.18 ±0.08	0.40 ±0.07	0.41 ±0.06	0.42 ±0.07
3	0.10 ±0.01	0.10 ±0.01	0.10 ±0.01	0.17 ±0.08	0.15 ±0.08	0.17 ±0.08	0.38 ±0.07	0.39 ±0.07	0.37 ±0.07
4	0.11 ±0.01	0.11 ±0.01	0.11 ±0.01	0.14 ±0.07	0.13 ±0.06	0.14 ±0.06	0.36 ±0.06	0.37 ±0.06	0.36 ±0.06
5	0.11 ±0.01	0.11 ±0.02	0.11 ±0.01	0.13 ±0.06	0.12 ±0.06	0.12 ±0.06	0.33 ±0.06	0.34 ±0.06	0.34 ±0.06
6	0.10 ±0.01	0.10 ±0.01	0.10 ±0.01	0.13 ±0.07	0.12 ±0.06	0.13 ±0.06	0.22 ±0.03	0.23 ±0.04	0.23 ±0.04
7	0.11 ±0.01	0.10 ±0.02	0.11 ±0.01	0.12 ±0.05	0.11 ±0.05	0.11 ±0.05	0.19 ±0.03	0.19 ±0.03	0.19 ±0.03
8	0.14 ±0.02	0.12 ±0.02	0.11 ±0.02	0.12 ±0.05	0.10 ±0.05	0.10 ±0.04	0.21 ±0.03	0.17 ±0.02	0.17 ±0.02
9	0.18 ±0.03	0.16 ±0.03	0.14 ±0.02	0.14 ±0.06	0.12 ±0.06	0.11 ±0.04	0.21 ±0.04	0.19 ±0.03	0.17 ±0.03
10	0.19 ±0.04	0.16 ±0.04	0.15 ±0.03	0.15 ±0.07	0.13 ±0.07	0.10 ±0.04	0.19 ±0.04	0.16 ±0.03	0.16 ±0.03
11	0.23 ±0.06	0.20 ±0.05	0.18 ±0.04	0.18 ±0.09	0.16 ±0.09	0.12 ±0.06	0.19 ±0.04	0.17 ±0.04	0.16 ±0.03
12	0.26 ±0.06	0.23 ±0.04	0.21 ±0.03	0.20 ±0.09	0.18 ±0.09	0.13 ±0.06	0.18 ±0.04	0.16 ±0.04	0.16 ±0.04
<b>Total</b>	<b>0.14 ±0.10</b>	<b>0.13 ±0.09</b>	<b>0.13 ±0.07</b>	<b>0.15 ±0.31</b>	<b>0.13 ±0.29</b>	<b>0.13 ±0.26</b>	<b>0.27 ±0.11</b>	<b>0.26 ±0.12</b>	<b>0.26 ±0.12</b>

early-frame seq2seq encoder and the complexity of the ODE system, which requires optimizing the numbers of early input frame images for the best result. The embedding of the early-frame images relies heavily on the pharmacokinetics defined by the hidden state in the final output of the seq2seq architect [27]. Hence, it is challenging for the seq2seq network to deal with the limited number of input frames. On the other hand, as the seq2seq runs backwards in time in the proposed model, the excessive number of early frames may result in a suboptimal initial point as described in (9) through hidden states. There was no consistent trend found in terms of the number of input frames when it comes to predicting pharmacokinetic parameters such as the rate constants, SUV<sub>r</sub>, and BP<sub>ND</sub>. Particularly, 7 frames (20 min) alone achieved robust performance in deriving a voxel-level image of SUV<sub>r</sub> and BP<sub>ND</sub> as shown in Fig. 5, and we believe this method has the potential to lead a much-reduced scanning time protocols in dynamic PET studies.

One common aspect of the proposed model on both data was that it learns the pharmacokinetics more efficiently when the average spatial variance of the initial frame is large, which is the case for tau or amyloid PET scans in the Alzheimer's disease (AD) [2], [3], [4]. As shown in Fig. 4 best example, the proposed method performed the best with the image that appeared to have a relatively larger contrast between different ROIs in early frames. For <sup>18</sup>F-PI-2620 datasets, it indicated that the proposed method worked better with the test set 0 that has regional high uptakes in the occipital and frontal lobe as well as the limbic area. However, we observed that the proposed model resulted in better prediction results when the temporal variance is small throughout the entire timeframes. As illustrated in Fig. 4, the overall ground-truth TAC of the best example test data increased consistently until the 5th frame (10 min) and then decrease with time, whereas the overall ground-truth TAC of the worst example peaked at the 8th frame (25 min) and washed out more slowly. Since the model imprints the early-frame information through the seq2seq architect, the dramatic changes in TAC within early frames seemed to be lost in the hidden state of

the seq2seq model, which resulted in poor estimation for the data that contain the extreme gradient of TAC in the early frames.

#### D. N-ODE-Based Model and the Auto-Regression-Only Model

As a tool for regression, it has been reported that the N-ODE-based method has a few advantages over the auto-regression-only model, such as LSTM-based models in learning dynamic data [22], [23], [26], [27]. One of the benefits is that the N-ODE-based method leverages less computational burden than LSTM models [22], [23]. Furthermore, the N-ODE-based models maintain the satisfying performance with irregular time points [23], which is often the case in dynamic PET cases. To our best knowledge, the investigation has not been made to evaluate both models on dynamic PET images, let alone each of the models. We evaluated the regression-only models using the same architect described in Fig. 2. Instead of utilizing the VAE-N-ODE architect, we exploited the seq2seq-only model to directly predict the dynamic <sup>18</sup>F-PI-2620 brain PET. The resulting voxel-wise SUV<sub>r</sub> of the predicted dynamic PET is illustrated in Fig. 7. The seq2seq-only model poorly estimated the pharmacokinetics given the limited number of input frames (7 frames, 20 min) in both test datasets. To that end, we believe that the N-ODE-based model is preferable, particularly when only very early perfusion images are available.

Additionally, we observed the different textures of the output images between the two models. The seq2seq-only model tends to smooth the output more than the N-ODE-based model. As a result, the prediction made from the seq2seq-only model exhibited a clearer outline of the ROIs such as fissures, in output. We speculate this occurred from the architect of models. In the seq2seq-only model, the predictions are made from the very first frames to the respective frames by regressing over convolutional layers. However, for the N-ODE-based model, the latent features that represent the predictions are first estimated using ODE solvers, and the final predictions are made



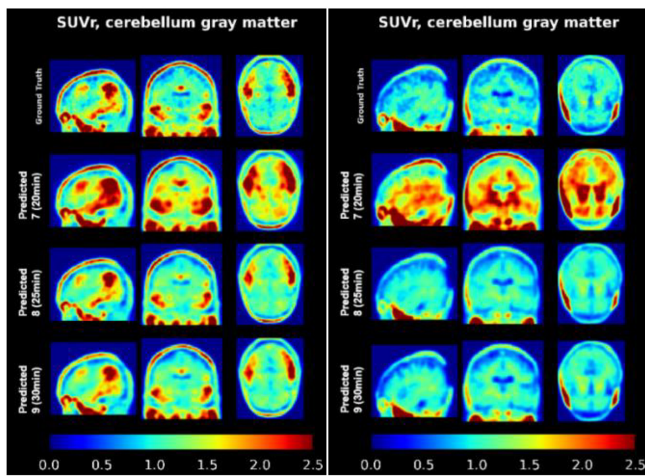


Fig. 7. SUVr using the regression-only model. Test set 0 and test set 1 are illustrated in left and right, respectively. The first row illustrates the ground truth, and the following row depicts the prediction made by 7 input frames (20 min), 8 input frames (25 min), and 9 input frames (30 min).

using convolutional layers. In other words, the lump of convolution layers that is used to regress the final predicted outputs is massive in size in the seq2seq-only model, compared to the N-ODE-based model. This contributes to the overall smoothness of the output images, which was an unexpected benefit for the seq2seq-only model. However, due to this aspect, the images with the input presenting less contrast in the first place, such as test set 1 showed a highly biased output, especially with the limited number of input frames (7 frames, 20 min).

### E. Limitations

Although we have conducted our study both on simulated and clinical dynamic data, one limitation of our work is that it is trained by the restricted amount of clinical brain PET data. We initially adopted patch-based learning to meet the computing resources, however, this approach served as data augmentation, resulting in agreeable results with confined training data. However, the model repetitively presented high bias, variance, and MAE in central structures that are relatively small in size, regardless of the numbers of the input frames. In this regard, we believe that the model has room for improvement when the patches that contain the central structures are visited more often during training. Another limitation which originated from patch-based learning is that the outputs inevitably display the horizontal and vertical strikes.

### F. Clinical Implication and Future Direction

In clinical routine, the bottleneck for dynamic PET imaging is the long acquisition time, which is not only impractical but also hampered by patient movement. Furthermore, the intrinsic noise of PET images hinders the performance of the kinetic modeling, especially at the voxel level. The proposed method can address these two main limitations in clinical dynamic PET studies.

1) *Reduced Time Protocol*: The recent efforts were made to reduce the scanning time while allowing the parametric images to be drawn [17], [19], [20], [35], [36], [37], [38]. As

the proposed model predicts the entire frames of the dynamic PET given a few early frames alone, it provides the solution to reduce the scanning time while deriving the pharmacokinetics of interest in brain studies such as perfusion or binding perfusion. Compared to the previous studies, the proposed method utilized not only the temporal information of the dynamic PET images but also spatial information through a mass of convolutional layers and the ODE solving mechanism, which contributed to the promising performances. We demonstrated reducing from 60 to 20 min is possible in  $^{18}\text{F}$ -PI-2620 dynamic PET scans. With the recent development of the new scanner [39], the proposed method potentially can achieve a very-short dynamic scanning time [21]. Yet, all tracers display different spatiotemporal patterns and extensive studies on fine-tuning the proposed method for other tracers will be part of future studies.

The advantage of the proposed method in reducing the time protocol is that it does not require the AIF, which is a challenge for the analytical method. One straightforward way to predict the later frames in dynamic PET analytically is to use compartment modeling to fit the images only with early frames and use the derived rate constants to predict the late frames [21]. However, this requires the determining AIF and the accurate kinetic model, both of which add the uncertainty of the prediction. Furthermore, with only a few early frames, the model suffers from the sparse sample size for the fitting. In particular, predicting  $k_3$  or  $k_4$  is challenging, as both rate constants are associated with the later frames. Instead, the proposed method avoids this issue by learning from the data themselves.

Yet, the AIF is still required in reduced scanning time, when the aim is on deriving micro parameters such as rate constants. One way is to infer from predicted images using an image-derived input function (IDIF), which will be incorporated in the future study.

2) *Dual-Tracer Studies*: Contrary to the dual time window protocol, or coffee break protocol [17], [18], [19], the proposed method does not require the late frame images to derive the full dynamic or parametric images, which leaves the margin for other tracers to be administered after the cut-off frame set for seq2seq architect as shown in Fig. 2. Thus, another critical application of the proposed method is dual-tracer studies.

Dual-tracer brain PET is a promising technique that measures the distribution of two tracers in the brain by a single scan with tracer administrations staggered in time, characterizing dual aspects of function without repeated sessions of scanning separated by hours or days [29]. Integrating complementary information of both tracers may enhance the sensitivity and specificity for presymptomatic staging as well as the differential diagnosis of heterogeneous pathogenesis with similar clinical phenotypes [6]. However, recovering the separate scans of each tracer from the dual-tracer scan is challenging, as all tracer signals for the PET scanner are indistinguishable 511-KeV annihilation photon pairs, and no explicit information is available that indicates which tracer gave rise to the PET coincidence-pair measurement.

Similar to the single-tracer dynamic PET studies, the analytical explanation for deriving the pharmacokinetic parameter from dual-tracer studies is led by establishing a compartmental model. However, for dual-tracer studies, the parallel configuration is often used [29]. On the other hand, studies were held recently that use deep-learning models, which typically solve the blind source separation problem. Ruan and Liu [40] and Xu and Liu [41] proposed the deep-learning model which seeks to single out the contribution of each tracer from the mixture of signals using a deep NN. Particularly, Ruan et al. utilized stacked autoencoder (SAE) and Xu et al. employed a deep belief network (DBN).

Instead of focusing on the separation of the signal, the proposed method offers the novel approach to recover the signal from each individual tracer by aiming to predict the full dynamic images of the first administered tracer, given early-frame images. In this way, the dynamic images of the second tracer are readily estimated by subtracting the signal of the first administered tracer from the total measurement where both overlap. The future direction of the current work includes the feasibility study of recovering single tracers in dual-tracer signal using the proposed method.

## V. CONCLUSION

In this work, we established the deep-learning model using N-ODE that predicts the full dynamic frame images, hence the pharmacokinetic parameters, such as rate constants and  $BP_{ND}$ , given the limited number of input frame images. We demonstrated that the proposed method showed promising results with both simulated and clinical brain PET data, indicated by bias, variance, MAE, and pharmacokinetic parameters. In particular, we showed that  $SUV_r$  and  $BP_{ND}$  parametric images can still be derived with 20 min  $^{18}F$ -PI-2620, instead of 60 min. Additionally, we believe that this approach sheds new light on rapid dual-tracer studies.

## ACKNOWLEDGMENT

Matthias Brendel received speaker honoraria from GE healthcare, Roche, and Life Molecular Imaging (LMI) and is an advisor of LMI; Henryk Barthel received reader honoraria from LMI; and Osama Sabri received research support from LMI. The remaining authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this article.

## REFERENCES

- [1] S. Hellwig et al., “[ $^{18}F$ ]FDG-PET is superior to [ $^{123}I$ ]IBZM-SPECT for the differential diagnosis of parkinsonism,” *Neurology*, vol. 79, no. 13, pp. 1314–1322, Sep. 2012, doi: [10.1212/WNL.0b013e31826c1b0a](https://doi.org/10.1212/WNL.0b013e31826c1b0a).
- [2] B. Dubois et al., “Cognitive and neuroimaging features and brain  $\beta$ -amyloidosis in individuals at risk of Alzheimer’s disease (INSIGHT-preAD): A longitudinal observational study,” *Lancet Neurol.*, vol. 17, no. 4, pp. 335–346, Apr. 2018, doi: [10.1016/S1474-4422\(18\)30029-2](https://doi.org/10.1016/S1474-4422(18)30029-2).
- [3] H. Barthel et al., “Cerebral amyloid- $\beta$  PET with florbetaben ( $^{18}F$ ) in patients with Alzheimer’s disease and healthy controls: a multicentre phase 2 diagnostic study,” *Lancet Neurol.*, vol. 10, no. 5, pp. 424–435, May 2011, doi: [10.1016/S1474-4422\(11\)70077-1](https://doi.org/10.1016/S1474-4422(11)70077-1).
- [4] B. A. Gordon et al., “Tau PET in autosomal dominant Alzheimer’s disease: relationship with cognition, dementia and other biomarkers,” *Brain*, vol. 142, no. 4, pp. 1063–1076, Apr. 2019, doi: [10.1093/brain/awz2019](https://doi.org/10.1093/brain/awz2019).
- [5] M. Brendel et al., “Perfusion-phase [ $^{18}F$ ]THK5351 Tau-PET imaging as a surrogate marker for neurodegeneration,” *J. Alzheimer’s Disease Rep.*, vol. 1, no. 1, pp. 109–113, Jan. 2017, doi: [10.3233/ADR-170023](https://doi.org/10.3233/ADR-170023).
- [6] M. Grossman, “Integrated multimodal imaging in neurodegenerative disease,” *Lancet Neurol.*, vol. 14, no. 10, pp. 973–975, Oct. 2015, doi: [10.1016/S14744422\(15\)00182-9](https://doi.org/10.1016/S14744422(15)00182-9).
- [7] G. Wang, A. Rahmim, and R. N. Gunn, “PET parametric imaging: Past, present, and future,” *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 4, no. 6, pp. 663–675, Nov. 2020, doi: [10.1109/TRPMS.2020.3025086](https://doi.org/10.1109/TRPMS.2020.3025086).
- [8] A. Dimitrakopoulou-Strauss, L. Pan, and C. Sachpekidis, “Kinetic modeling and parametric imaging with dynamic PET for oncological applications: General considerations, current clinical applications, and future perspectives,” *Eur. J. Nucl. Med. Mol. Imag.*, vol. 48, no. 1, pp. 21–39, Jan. 2021, doi: [10.1007/s00259-020-04843-6](https://doi.org/10.1007/s00259-020-04843-6).
- [9] X. Zhang et al., “Total-body dynamic reconstruction and parametric imaging on the uEXPLORER,” *J. Nucl. Med.*, vol. 61, no. 2, pp. 285–291, Feb. 2020, doi: [10.2967/jnumed.119.230565](https://doi.org/10.2967/jnumed.119.230565).
- [10] K. Erlandsson, “Tracer kinetic modeling: Basics and concepts,” in *Basic Sciences of Nuclear Medicine*, M. M. Khalil, Ed. Cham, Switzerland: Springer Int., 2021, pp. 531–549.
- [11] K. Heurling, C. Buckley, K. Van Laere, R. Vandenberghe, and M. Lubberink, “Parametric imaging and quantitative analysis of the PET amyloid ligand [ $^{18}F$ ]flutemetamol,” *NeuroImage*, vol. 121, pp. 184–192, Nov. 2015, doi: [10.1016/j.neuroimage.2015.07.037](https://doi.org/10.1016/j.neuroimage.2015.07.037).
- [12] O. Barret et al., “Kinetic modeling of the tau PET tracer  $^{18}F$ -AV-1451 in human healthy volunteers and alzheimer disease subjects,” *J. Nucl. Med.*, vol. 58, no. 7, pp. 1124–1131, Jul. 2017, doi: [10.2967/jnumed.116.182881](https://doi.org/10.2967/jnumed.116.182881).
- [13] M.-K. Chen et al., “Comparison of [ $^{11}C$ ]UCB-J and [ $^{18}F$ ]FDG PET in Alzheimer’s disease: A tracer kinetic modeling study,” *J. Cerebr. Blood Flow Metab.*, vol. 41, no. 9, pp. 2395–2409, Sep. 2021, doi: [10.1177/0271678X21100431](https://doi.org/10.1177/0271678X21100431).
- [14] L. Pan, C. Cheng, U. Haberkorn, and A. Dimitrakopoulou-Strauss, “Machine learning-based kinetic modeling: A robust and reproducible solution for quantitative analysis of dynamic PET data,” *Phys. Med. Biol.*, vol. 62, no. 9, pp. 3566–3581, May 2017, doi: [10.1088/1361-6560/aa6244](https://doi.org/10.1088/1361-6560/aa6244).
- [15] F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, and Y. Ouchi, “4D deep image prior: Dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network,” *Phys. Med. Biol.*, vol. 66, no. 1, Jan. 2021, Art. no. 015006, doi: [10.1088/1361-6560/abcd1a](https://doi.org/10.1088/1361-6560/abcd1a).
- [16] C. J. Scott et al., “Reduced acquisition time PET pharmacokinetic modelling using simultaneous ASL–MRI: Proof of concept,” *J. Cerebr. Blood Flow Metab.*, vol. 39, no. 12, pp. 2419–2432, Dec. 2019, doi: [10.1177/0271678X18797343](https://doi.org/10.1177/0271678X18797343).
- [17] F. Heeman et al., “Parametric imaging of dual-time window [ $^{18}F$ ]flutemetamol and [ $^{18}F$ ]florbetaben studies,” *NeuroImage*, vol. 234, Jul. 2021, Art. no. 117953, doi: [10.1016/j.neuroimage.2021.117953](https://doi.org/10.1016/j.neuroimage.2021.117953).
- [18] G. D. Kolinger et al., “A dual-time-window protocol to reduce acquisition time of dynamic tau PET imaging using [ $^{18}F$ ]MK-6240,” *EJNMMI Res.*, vol. 11, no. 1, p. 49, Dec. 2021, doi: [10.1186/s13550-021-00790-x](https://doi.org/10.1186/s13550-021-00790-x).
- [19] V. Viswanath et al., “Abbreviated scan protocols to capture  $^{18}F$ -FDG kinetics for long axial FOV PET scanners,” *Eur. J. Nucl. Med. Mol. Imag.*, vol. 49, pp. 3215–3225, Mar. 2022, doi: [10.1007/s00259-022-05747-3](https://doi.org/10.1007/s00259-022-05747-3).
- [20] Y. Wu et al., “Whole-body parametric imaging of  $^{18}F$ -FDG PET using uEXPLORER with reduced scanning time,” *J. Nucl. Med.*, vol. 63, no. 4, pp. 622–628, Apr. 2022, doi: [10.2967/jnumed.120.261651](https://doi.org/10.2967/jnumed.120.261651).
- [21] G. Liu et al., “Short-time total-body dynamic PET imaging performance in quantifying the kinetic metrics of  $^{18}F$ -FDG in healthy volunteers,” *Eur. J. Nucl. Med. Mol. Imag.*, vol. 49, pp. 2493–2503, Aug. 2021, doi: [10.1007/s00259-021-05500-2](https://doi.org/10.1007/s00259-021-05500-2).
- [22] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, “Neural ordinary differential equations,” Dec. 2019, *arXiv:1806.07366*.
- [23] Y. Rubanova, R. T. Q. Chen, and D. Duvenaud, “Latent ODEs for irregularly-sampled time series,” Jul. 2019, *arXiv:1907.03907*.
- [24] Z. Lai, C. Mylonas, S. Nagarajaiah, and E. Chatzi, “Structural identification with physics-informed neural ordinary differential equations,” *J. Sound Vib.*, vol. 508, Sep. 2021, Art. no. 116196, doi: [10.1016/j.jsv.2021.116196](https://doi.org/10.1016/j.jsv.2021.116196).

- [25] W. Bonnaffé, B. C. Sheldon, and T. Coulson, “Neural ordinary differential equations for ecological and evolutionary time-series analysis,” *Methods Ecol. Evol.*, vol. 12, no. 7, pp. 1301–1315, 2021, doi: [10.1111/2041-210X.13606](https://doi.org/10.1111/2041-210X.13606).
- [26] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- [27] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” Dec. 2014, *arXiv:1409.3215*.
- [28] D. P. Kingma and M. Welling, “Auto-encoding variational Bayes,” May 2014, *arXiv:1312.6114*.
- [29] S. C. Huang, R. E. Carson, E. J. Hoffman, D. E. Kuhl, and M. E. Phelps, “An investigation of a double-tracer technique for positron computerized tomography,” *J. Nucl. Med.*, vol. 23, no. 9, pp. 816–822, 1982.
- [30] X. Cheng et al., “Direct parametric image reconstruction in reduced parameter space for rapid multi-tracer PET imaging,” *IEEE Trans. Med. Imag.*, vol. 34, no. 7, pp. 1498–1512, Jul. 2015, doi: [10.1109/TMI.2015.2403300](https://doi.org/10.1109/TMI.2015.2403300).
- [31] Y. T. Hong and T. D. Fryer, “Kinetic modelling using basis functions derived from two-tissue compartmental models with a plasma input function: General principle and application to [18F]fluorodeoxyglucose positron emission tomography,” *NeuroImage*, vol. 51, no. 1, pp. 164–172, May 2010, doi: [10.1016/j.neuroimage.2010.02.013](https://doi.org/10.1016/j.neuroimage.2010.02.013)
- [32] D. Feng, K.-P. Wong, C.-M. Wu, and W.-C. Siu, “A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: Theory and simulation study,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 1, no. 4, pp. 243–254, Dec. 1997, doi: [10.1109/4233.681168](https://doi.org/10.1109/4233.681168).
- [33] A. Hammers et al., “Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe,” *Human Brain Map.*, vol. 19, no. 4, pp. 224–247, Aug. 2003, doi: [10.1002/hbm.10123](https://doi.org/10.1002/hbm.10123).
- [34] A. Mueller et al., “Tau PET imaging with  $^{18}\text{F}$ -PI-2620 in patients with alzheimer disease and healthy controls: A first-in-humans study,” *J. Nucl. Med.*, vol. 61, no. 6, pp. 911–919, Jun. 2020, doi: [10.2967/jnumed.119.236224](https://doi.org/10.2967/jnumed.119.236224).
- [35] A. Sanaat, E. Mirsadeghi, B. Razeghi, N. Ginovart, and H. Zaidi, “Fast dynamic brain PET imaging using stochastic variational prediction for recurrent frame generation,” *Med. Phys.*, vol. 48, no. 9, pp. 5059–5071, Sep. 2021, doi: [10.1002/mp.15063](https://doi.org/10.1002/mp.15063).
- [36] T. Torizuka et al., “Short dynamic FDG-PET imaging protocol for patients with lung cancer,” *Eur. J. Nucl. Med.*, vol. 27, no. 10, pp. 1538–1542, Oct. 2000, doi: [10.1007/s002590000312](https://doi.org/10.1007/s002590000312).
- [37] T. Monden, N. Kudomi, Y. Sasakawa, Y. Yamamoto, N. Kawai, and Y. Nishiyama, “Shortening the duration of [18F]FDG PET brain examination for diagnosis of brain glioma,” *Mol. Imag. Biol.*, vol. 13, no. 4, pp. 754–758, Aug. 2011, doi: [10.1007/s11307-010-0384-z](https://doi.org/10.1007/s11307-010-0384-z).
- [38] Z. Wang et al., “Comparison between a dual-time-window protocol and other simplified protocols for dynamic total-body 18F-FDG PET imaging,” *EJNMMI Phys.*, vol. 9, no. 1, p. 63, Sep. 2022, doi: [10.1186/s40658-022-00492-w](https://doi.org/10.1186/s40658-022-00492-w).
- [39] I. Alberts et al., “Clinical performance of long axial field of view PET/CT: a head-to-head intra-individual comparison of the Biograph Vision Quadra with the Biograph Vision PET/CT,” *Eur. J. Nucl. Med. Mol. Imag.*, vol. 48, no. 8, pp. 2395–2404, Jul. 2021, doi: [10.1007/s00259-021-05282-7](https://doi.org/10.1007/s00259-021-05282-7).
- [40] D. Ruan and H. Liu, “Separation of a mixture of simultaneous dual-tracer PET signals: A data-driven approach,” *IEEE Trans. Nucl. Sci.*, vol. 64, no. 9, pp. 2588–2597, Sep. 2017, doi: [10.1109/TNS.2017.2736644](https://doi.org/10.1109/TNS.2017.2736644).
- [41] J. Xu and H. Liu, “Deep-learning-based separation of a mixture of dual-tracer single-acquisition PET signals with equal half-lives: A simulation study,” *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 6, pp. 649–659, Nov. 2019, doi: [10.1109/TRPMS.2019.2897120](https://doi.org/10.1109/TRPMS.2019.2897120).