



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria

Unterschrift Betreuer

DIPLOMARBEIT

Development of an Optimization Method for Implementation in the Computation of Kinetic Modeling in Nuclear Medicine

ausgeführt am Institut für
Angewandte Physik (E 134)
der Technischen Universität Wien

in Zusammenarbeit mit der

Universitätsklinik für Radiologie und Nuklearmedizin und dem
Zentrum für Medizinische Physik und Biomedizinische Technik
der Medizinischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. DI Dr. Martin Gröschl

(Technische Universität Wien)

Mag. Dr. Barbara Katharina Geist

Ao.Univ.Prof. Mag. Dr. Wolfgang Birkfellner

(Medizinische Universität Wien)

durch

Lukas Prager, BSc BSc

Viehtriftgasse 3/35

A-1210 Wien

Wien, am 7. Mai 2019

Ort, Datum

Unterschrift Student

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachtern und Betreuern beurteilten Arbeit überein.

Lukas Prager, BSc BSc

Wien, am 7. Mai 2019

Danksagung

An dieser Stelle möchte ich jenen Personen danken, welche mich während meines Studiums unterstützt und mir bei der Durchführung und Fertigstellung dieser Arbeit geholfen haben.

Zunächst möchte ich mich bei Herrn Ao.Univ.Prof. DI Dr. Martin Gröschl für seine wissenschaftliche Betreuung dieser Arbeit bedanken.

Ferner gilt mein außerordentlicher Dank Frau Mag. Dr. Barbara Katharina Geist, welche mich während meiner Arbeit an der Medizinischen Universität Wien unterstützt hat und immer für meine Anliegen da war, selbst wenn die Zeit drängte. Aufgrund der zahlreichen Gespräche auf intellektueller und persönlicher Ebene habe ich die Zeit, welche ich dort verbracht habe, sehr genossen. Außerdem möchte ich mich für die Begleitung während des Schreibens dieser Abhandlung und den zahlreichen Tipps für das Verfassen einer wissenschaftlichen Arbeit bedanken.

Besonders danken möchte ich auch meinem Lebensgefährten, Bernhard Schneider, welcher jederzeit für mich da war und mir geholfen hat, mein Ziel nicht aus den Augen zu verlieren. Ohne ihn hätte ich mein Studium nicht in dieser Schnelligkeit abgeschlossen.

Der wohl größte Dank aber ergeht an meine Eltern, Helga und Engelbert Prager, welche mich während meines ganzen Studiums unterstützt haben. Ohne sie wäre es mir nicht möglich gewesen, meine Interessen im technischen Bereich auf diese Weise zu verfolgen und meine Ausbildung in diesem Sinne abzuschließen. Insbesondere in schwierigen Zeiten hatten sie stets ein offenes Ohr und waren jederzeit für mich da, wofür ich ihnen immer dankbar sein werde.

Kurzfassung

Hintergrund: Auf Basis der Kinetischen Modellierung wurde ein Algorithmus entwickelt, welcher die Quantifizierung physiologischer Vorgänge im Körper erleichtert. Dafür werden mit Positron-Emissions-Tomographie (PET) aufgenommene Zeitaktivitätskurven (TAC) eingelesen und auf verschiedene mathematischen Charakteristika untersucht. Hierdurch bekommt man eine erste Einschätzung der Zeitkonstanten, welche für einen weiteren Kurvenfit verwendet werden können. *Methoden:* Um Korrelationen zwischen den Zeitkonstanten und den mathematischen Eigenschaften der TACs zu finden, wurden mit dem Programm EXCEL etwa 1200 TACs simuliert. Ein PYTHON Skript diente dabei der mathematischen Analyse der TACs. Um verschiedene Modelltypen miteinander vergleichen zu können, wurde ein t -Test durchgeführt. *Ergebnisse:* Signifikante Korrelationen konnten nur für die ersten beiden Zeitkonstanten gefunden werden ($r_{k_1} > 0.954, r_{k_2} > 0.799$). Die restlichen Zeitkonstanten wurden daher über Konfidenzintervalle zum Niveau 95% geschätzt. *Schlussfolgerungen:* Durch eine erste Abschätzung der Zeitkonstanten wird einerseits die Zuverlässigkeit von Programmen zum Fitten von Kurven erhöht, andererseits können dadurch auch neue Tracer in der Nuklearmedizin einfacher studiert und ihr Verhalten im Körper analysiert werden.

Abstract

Background: An algorithm was programmed to facilitate the quantification of physiological processes in kinetic modeling. In order to do so, it reads in a measured time activity curve (TAC) acquired via positron emission tomography (PET), analyzes it for various mathematical characteristics and outputs values for the rate constants, which can be used for further curve fitting tools. *Methods:* Around 1200 TACs with different rate constants and model types have been simulated via EXCEL. These TACs were used to find correlations between their rate constants and mathematical properties, which were calculated by a PYTHON script. Plus, a t -test was performed to see if there is a connection between different compartment model types. The obtained information can then be used for the analysis of real measured TACs, which yields good initial values for the rate constants. *Results:* Useful correlations were found for the first two rate constants only ($r_{k_1} > 0.954, r_{k_2} > 0.799$). The assessment of the remaining rate constants was done by providing two-sided 95% confidence intervals. *Conclusion:* By assessing the rate constants, both the reliability of curve fitting programs can be increased and new tracers in nuclear medicine can be studied more precisely.

Contents

Eidesstattliche Erklärung	i
Danksagung	ii
Kurzfassung – Abstract	iii
1 Introduction	1
1.1 Nuclear Medicine	1
1.1.1 Imaging via Positron Emission Tomography	2
1.1.2 Tracers used in Nuclear Medicine	4
1.1.3 Quantification	6
1.2 Kinetic Modeling	9
1.2.1 Compartmental Modeling	11
1.2.2 Mathematical Description	12
2 Motivation	17
2.1 Analytical Solution	17
2.2 Mathematical Examination for the Rate Constants	18
2.3 Numerical Approach	27
3 Methods	29
3.1 Simulation of Time Activity Curves	29
3.1.1 Generation of Input Functions	29
3.1.2 Stepwise Solution of the Differential Equations	32
3.2 Generation of Information regarding Curve Shape	34
3.2.1 Relative Slope and Intercept before a certain Time Point	34
3.2.2 Relative Slope and Intercept after a certain Time Point	36
3.2.3 Relative Integral	37
3.2.4 Curvature	37
3.2.5 1C/2C/3C and Trapped/Open	39
3.2.6 TAC Shape	39
3.2.7 Rate Constants k_i and underlying Input Function	44
3.3 Evaluation of obtained Data	44

3.3.1	Correlation Coefficient	45
3.3.2	Student's <i>t</i> -Test	46
3.4	Programming the Algorithm	47
3.4.1	General Structure of the Algorithm	48
3.4.2	Determination of the Input Function and the TAC shape	50
3.4.3	Determination of Feasible Intervals for the Rate Constants	51
3.4.4	Calculating Values for k_1 and k_2	52
3.5	Testing the Algorithm	53
4	Results	55
4.1	Simulation of Time Activity Curves	55
4.2	Generation of Information regarding Curve Shape	59
4.3	Evaluation of obtained Data	60
4.4	Programming the Algorithm	62
4.5	Testing the Algorithm	68
5	Discussion	71
6	Conclusion	74
	List of Abbreviations	75
	Nomenclature	76
	References	78
	List of Figures	83
	List of Tables	84

1 Introduction

Especially in the last decades the demand for imaging modalities in medicine has increased enormously, as they allow to monitor the human anatomy, visualize organ functions and facilitate early diagnoses of diseases such as cancer. Furthermore, these imaging techniques are important to assess the effectiveness of treatments performed. In nuclear medicine, Positron-Emission-Tomography (PET) is a common imaging technique since it allows to study the processes going on within the body and also to take a look at the human metabolism. For this purpose, a so called tracer has to be administered to the patient.

The next step in medical imaging is the quantification of the data acquired by the imaging device. In the case of monitoring via PET, the quantification is often performed by using kinetic modeling. To set up such a kinetic model, the concentration of the incorporated tracer has to be recorded over time, leading to a so called Time Activity Curve (TAC) representing the temporal behavior of the tracer. One subspecialty in kinetic modeling is compartmental modeling. By using compartment models to quantify the obtained data, one can assess the processes going on within the body. This master thesis concentrates on compartmental modeling only.

Apart from studying the tracer distribution and concentration, compartmental modeling is also important when it comes to exploring new tracers, which have to be assessed and analyzed. Thus, compartmental modeling is important for both quantifying human processes and doing research in nuclear medicine.

1.1 Nuclear Medicine

In medical imaging, there are different methods in order to monitor the human anatomy, its biological function or both of them simultaneously. Nowadays high-resolution 3D images can be obtained. Among other imaging techniques there are Computed Tomography (CT), Magnetic Resonance Imaging (MRI), PET, Single Photon Emission Computed Tomography (SPECT) and imaging via ultrasound [1]. With each of them having its particular field of application, the clinician has to decide which imaging modality to use. For further reading of the modalities' working procedure see [2].

A particular field in medical imaging is nuclear medicine. Here the imaging process is performed by using radioactive materials, so called tracers, which are incorporated by the

patient. All these tracers are subject to a certain decay process (for instance α decay, β^+ decay, β^- decay), which finally results in the emission of γ -photons. By detecting these photons, the location of the tracer and hence its track can be determined, which yields important information about the processes going on within the human body. Nuclear medicine thereby allows to gather information which cannot be drawn by conventional medical imaging [3].

With all the fields of application, nuclear medicine is indispensable when it comes to planning an appropriate treatment (for example for patients suffering from cancer), making disease prognoses and assessing the effectiveness of the treatment. Also, visualizing tumor progression may help the clinician to decide whether to stop, alter or continue the therapy [4, 5].

1.1.1 Imaging via Positron Emission Tomography

In oncology PET is a common imaging method since it allows to visualize the body's metabolism which is an essential indicator for evaluating tumor growth. A big advantage when it comes to imaging via PET is the large number of available tracers. For a PET scan especially low atomic-number positron emitting nuclides such as ^{11}C , ^{15}O and ^{18}F are suitable, which can easily be incorporated by the patient in order to visualize the behavior of various body organs or tumors.

Usually a PET in combination with a CT (PET/CT) is used for three reasons: first, in addition to the metabolic behavior gained from a PET scan, the anatomy recorded by a CT can be visualized in parallel. This provides a better understanding and orientation for the operator. The second important reason is the attenuation correction of a PET scan which has to be carried out, since the attenuation of γ -photons is a main factor influencing and worsening the image quality. This can easily be performed using a CT [6]. Another advantage when using PET/CT is both its higher sensitivity and increased specificity compared to imaging via PET or CT solely [4]. In figure 1.1 the images recorded via a CT and a PET separately and in parallel are shown.

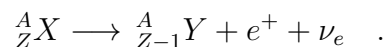
For imaging via PET, a tracer has to be administered to the patient. This tracer, also called radiopharmaceutical, is a bonding of a stable molecule and a radioactive nuclide. The stable molecule is responsible for undergoing normal metabolism, since often the chemical characteristics of these radiopharmaceuticals are the same as for their nonradioactive forms and the body thus cannot distinguish between these [7]. The radioactive nuclide is important for subsequent imaging. In order to avoid concentration saturation of the radiopharmaceutical and thereby prevent an alteration of chemical processes going on within the body, the

amount of incorporated radiopharmaceutical has to be kept low, which justifies the term "tracer" [8].

With the use of such radiopharmaceuticals it is possible to trace various physiological processes [7]. Ideally, the tracer represents one particular physiological parameter such as blood flow, tissue metabolism or molecular binding. Nevertheless, in reality there are other factors influencing the tracer's behavior, leading to a decreased relation between the tracer and the process of interest. For a proper tracer, however, all external factors having an impact on the tracer are negligible [8]. So a very important step in nuclear medicine is to decide which tracer to use.

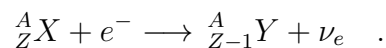
After a tracer is injected, it accumulates in certain body parts due to its physical, chemical and biological characteristics. As the decay of the tracer in the end causes an emission of γ -photons, the location of the tracer can be determined by means of a PET scan or other imaging modalities. Depending on the desired application, different tracers are used for imaging (see section 1.1.2).

There are many decay reactions a tracer can be subject to. One important decay mode especially when using a PET is the β^+ decay according to



By that, a positron (e^+) is emitted by the radionuclide which subsequently annihilates with an electron from the surrounding tissue. This annihilation leads to the emission of two anti-parallel γ -photons which can be recorded via a PET scan. The γ -photons hereby have an energy of 511 keV each, which corresponds to the rest masses of the initial particles (electron and positron) according to the mass-energy-equivalence $E = (m_{e^+} + m_{e^-})c^2 = 1022 \text{ keV}$ [9, 10].

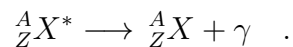
Another mode of decay is the electron capture (EC), where the following reaction takes place:



The captured electron often originates from the innermost shell, since it shows the largest probability for undergoing this process. This absorbed electron leads to a transformation of a proton to a neutron, causing an emission of a neutrino. Although it is difficult to detect the neutrino, also characteristic X-rays are emitted when the electron's void is filled by an

outer-shell electron. The energy of the γ -photons hereby is linked to the difference in binding energy of the two electron states. This radiation can be detected and used for further image formation. Since this reaction, as well as the β^+ decay, is preferred by nuclides having too much protons, the electron capture is considered as competing decay mode to the β^+ decay [11].

In nuclear medicine the decay by isomeric transition (IT) is important as well. A nuclear isomer is referred to a metastable form of a nuclide due to excited nuclei [12]. An often used metastable nuclide in nuclear medicine is ^{99m}Tc , where the "m" in the superscript indicates the metastable state of the atom. When this nuclear isomer falls back into its ground state, energy in form of γ -photons is emitted according to



The next step after detecting the photons emitted by a tracer administered to a patient is the image formation. In order to obtain a useful image of the patient, some kind of reconstruction of the recorded data has to be performed. Often used methods for data reconstruction are the filtered back-projection or an iterative reconstruction method based on a maximum likelihood estimation [6]. It is important to keep in mind that in the process of data reconstruction there are many factors the image has to be corrected for, such as detector normalization, photon scattering and attenuation as well as the detector's dead time [13]. As already mentioned, the use of a CT in parallel to a PET allows to correct for photon attenuation [8].

The spatial resolution of a PET is mainly limited by both the detector design and the lifetime of the ejected positron: with an average lifetime of about 10^{-10} s the positron has a free mean path length of a few millimeters resulting in a decreased spatial resolution of the PET scan [14].

Nevertheless, imaging via PET provides high-quality 3D images of radioactivity, from which the metabolic processes going on within the patient can be deduced [8].

1.1.2 Tracers used in Nuclear Medicine

As already mentioned above, there are many tracers used for imaging in nuclear medicine. Since they strongly differ by their behavior within the body, the decision upon a certain tracer depends on the intended application.

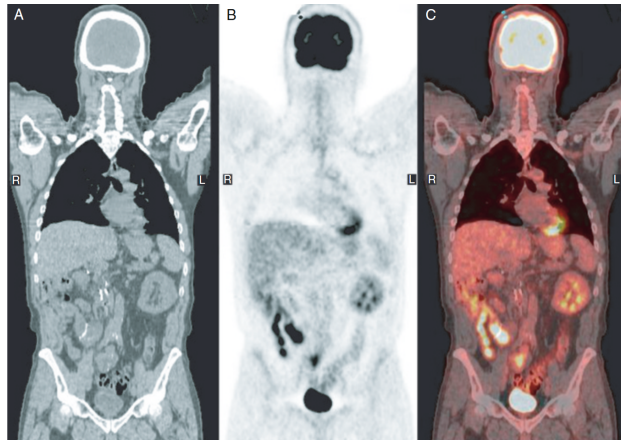


Figure 1.1: Imaging via PET/CT: the left plot shows the image acquired by a CT, the middle plot the corresponding PET scan and the right plot represents the fused image of a PET/CT combination [15]

The fundamental working principle for all tracers is their bonding to a radioactive nuclide, leading to the emission of γ -photons. By detecting these photons the tracer position can be determined. The tracers used in PET indirectly emit two anti-parallel photon by a previous β^+ decay (^{11}C , ^{15}O , ^{18}F , etc.). There are many positron emitting nuclides used in nuclear medicine, each of them having different half-lives. In contrast when performing SPECT, tracers undergoing isomeric transition (such as $^{99\text{m}}\text{Tc}$) and by that emitting a photon are suitable. In table 1.1 the most common nuclides with their respective half-lives, mode of decays and daughter products are listed. Hereafter, some tracers often used in nuclear medicine will be mentioned.

For radionuclides subject to β^+ decay the two anti-parallel photons are detected via a detector ring which is the main element in a PET scanner. By detecting these two photons, information about their origin can be obtained. A PET scanner is often operated in 3D mode, although 2D monitoring is possible as well.

One of the most common tracers used for PET is ^{18}F FDG (Fludeoxyglucose), which in its chemical structure resembles a glucose. As malignant cells are characterized by an increased metabolism and therefore a higher demand for glucose as well, the ^{18}F FDG primarily accumulates in these cells. Once located within cells, ^{18}F FDG becomes phosphorylated and due to its now altered structure it is not able to undergo further metabolism. This state is referred to as metabolic trapped [16]. Due to its radioactivity, its localization and tracking is possible, plus the extend of accumulation can be determined as well [17]. After undergoing β^+ decay,

the daughter product ^{18}O is formed, by which further metabolism is now possible again. ^{18}F has a half-life of about 110 min, which makes ^{18}F FDG perfectly suitable for medical imaging.

Besides the increased glucose uptake of cancerous cells, their demand for choline, which is an important component of cell membranes, increases as well. This property is taken advantage of by using an analogue of choline, namely the radio-labeled ^{11}C -choline. With ^{11}C -choline being integrated into the membrane, cancerous cells can be identified. ^{11}C -choline for example is used for diagnosing brain tumors, prostate cancer and lung cancer [18]. The radionuclide ^{11}C is a positron emitting nuclide with a half-life of about 20 min. Apart from ^{18}F FDG and ^{11}C -choline there are many other tracers which are used for PET.

In contrast, for radionuclides emitting photons directly, gamma cameras are used. These cameras are equipped with a collimator ensuring to only detect photons coming from a certain direction. Hereby it is possible to perform planar scintigraphy yielding 2D images. When recording 3D images this imaging modality is referred to as SPECT.

An important tracer hereby is $^{99\text{m}}\text{Tc}$ -MAG3 (mercaptoacetyltriglycine), which is wielded for renal tubular scintigraphy. By using $^{99\text{m}}\text{Tc}$ -MAG3 clear images can be drawn even if the kidney function is decreased drastically [19]. Due to its low radiation exposure to the patient as well as its low costs, $^{99\text{m}}\text{Tc}$ is a favored radionuclide in nuclear medicine. Thus, apart from $^{99\text{m}}\text{Tc}$ -MAG3 there are many other radiopharmaceuticals labeled with $^{99\text{m}}\text{Tc}$, which is a metastable nuclide undergoing isomeric transition [7].

Another radiopharmaceutical suitable for SPECT will be ioflupane (^{123}I). Ioflupane is used for dementia and Alzheimer's disease as well as for diagnosis of movement disorders such as Parkinson's disease. Ioflupane binds on certain nerve cell endings which are responsible for dopamine transport. In the case of a loss of these nerve cells, the amount of bound ioflupane will be decreased, which can be seen in the SPECT image. Ioflupane comprises of the radionuclide ^{123}I , which is subject to electron capture leading to an emission of a γ -photon [20].

All these mentioned radiopharmaceuticals are merely a small selection of the numerous tracers available and used in nuclear medicine.

1.1.3 Quantification

After visualizing a certain tracer by means of a PET scan or SPECT, the interpretation and quantification of the obtained images is necessary in order to assess a physiological function

Radionuclide	Half-life	Decay mode	Decay Product
^{11}C	20.4 min	β^+	^{11}B
^{13}N	9.97 min	β^+	^{13}C
^{15}O	2.03 min	β^+	^{15}N
^{18}F	1.83 hr	β^+	^{18}O
^{67}Ga	3.26 d	EC, γ	^{67}Zn
^{82}Rb	1.27 min	β^+	^{82}Kr
$^{99\text{m}}\text{Tc}$	6.01 hr	IT, γ	^{99}Tc
^{123}I	13.2 h	EC, γ	^{123}Te

Table 1.1: Most common radionuclides in nuclear medicine with their half-lives, decay modes and decay products. EC stands for electron capture, IT for isomeric transition [1, 17].

of interest. Usually these particular functions are considered within a certain volume of interest (VOI) [4]. When imaging via PET, in each recorded plane a two-dimensional region of interest (ROI) is drawn. These ROIs subsequently are put together to form a three-dimensional volume of interest (VOI). Since the clinician’s interpretation is very subjective, the aim is to standardize such assessments by using generalized quantitative methods. Plus, it turned out that the assessment of processes going on within the body (for instance tumor response) is more precisely by using such quantitative methods [21].

Common methods for the quantification of images acquired via PET are the standardized uptake value (SUV) and the fractional uptake rate (FUR), which are both called semiquantitative methods [22]. The SUV can easily be inferred from PET images and represents the tracer uptake within a VOI with respect to the total dose administered to the patient [4]. It is defined as

$$\text{SUV} = \frac{C_{\text{PET}} \cdot w}{C_{\text{inc}}} ,$$

where C_{PET} is the radioactivity within a VOI derived from the obtained images given in units of MBq/g. w is the patient’s body weight in grams and C_{inc} is the incorporated activity in units of MBq. Since the body weight is taken into account, the SUV is an index without unit [5]. For a proper calculation of SUV, both the measured radioactivity C_{PET} and the incorporated dose A_{inc} have to be corrected for radioactive decay.

A big advantage of SUV compared to other quantitative methods is its simple and fast calculation. Plus, the SUV provides a good assessment of therapy effectiveness and also allows to make reliable prognoses even for more complex problems. The SUV thereby is

a popular method for image quantification [22]. Up to now there are different approaches on how to define SUV: the first one is called maximum SUV and uses the maximum tracer concentration within a VOI. Another possibility for defining the SUV is by calculating the mean concentration over a VOI. These two approaches for calculating the SUV are the most common definitions found in literature [4].

For quantification using SUV, the measurement of the tracer concentration C_{PET} should take place late after tracer injection. Especially when it comes to ^{18}F FDG, this is important since the amount of phosphorylated ^{18}F FDG within cells increases with time. By that, a delayed image acquisition by PET is favored since the signal to noise ratio increases and thereby also the SUV's reliability will be enhanced [5].

Another method of quantification to be coarsely introduced is the fractional uptake rate FUR. As well as the SUV, the FUR is a fast and simple approach for quantifying images recorded via PET. It is given as

$$\text{FUR} = \frac{C_{\text{PET}}(t)}{\int_0^t C_{\text{inc}}(s) ds} \quad .$$

In contrast to the SUV, for the calculation of the FUR however the temporal behavior of the input function $C_{\text{inc}}(t)$ is needed. The input function defines the tracer concentration over time measured in the blood or plasma supplying the tissue of interest. In order to measure the FUR, the tracer concentration C_{PET} has to be computed at a certain time point t only, while the input function $C_{\text{inc}}(t)$ has to be recorded over a time period. In the case of tracing with ^{18}F FDG, the plasma curve is used as input function [22].

Since the FUR is given in units of 1/min, it can be compared with the rate constants appearing in compartmental modeling (see section 1.2.1). While in compartmental modeling the aim is to estimate the individual rate constants, the FUR simply quantifies the overall process monitored via PET. Due to the FUR's simplicity, however, it does not provide as much information on the process going on within the human body as there is in compartmental modeling [23]. It should be mentioned that, regardless of constants such as the clearance rate, SUV and FUR can be considered proportional.

The problem when using SUV or FUR is that they only provide information about the overall process recorded within a certain VOI. However, for more complex structures these semiquantitative methods reach their limits. Another possibility for quantification are ki-

netic models, which are described in the next section. Hereby the dynamic processes going on can be examined, plus different locations and chemical states of the tracer can be taken into account as well.

1.2 Kinetic Modeling

In contrast to the aforementioned semiquantitative methods, kinetic models allow to investigate the dynamic behavior of a tracer within a certain VOI. Even though a pixel-by-pixel analysis is possible as well, performing kinetic modeling on a VOI reduces noise and hence increases the model's accuracy [22]. With the sequential images acquired via PET, the tracer kinetics can be quantified. Therefore a model describing the tracer behavior is needed, by which a certain physiological parameter of interest can be deduced [8, 24]. With this method of parameter estimation, kinetic modeling can be seen as the last step in PET data quantification before applying statistical analysis.

Especially in the last few years, kinetic modeling has become a more and more important tool when it comes to the quantification of PET data. Reason for this development is its increasing accuracy while keeping low or even reducing the data needed for the model. However, due to the complex structure of kinetic models, they are limited to medical institutions having good computation tools on one hand and advanced imaging devices on the other hand [22].

As already mentioned, a model is needed before one can investigate tracer kinetics. The difficulty hereby is to find a model, which is simple enough to allow relatively fast calculations while describing the tracer kinetics properly. For setting up such a model, in the beginning a complete model gets defined on basis of knowledge regarding the expected tracer behavior within the VOI. Since this complete model is far too complex for parameter investigation and thus unsuitable for general application, this model has to be simplified. The next step is to test the simplified model for its reliability and if necessary to get modified. After the model is adapted well enough, it now can be used for routine patient examination. A general difficulty in proposing a useful model is to find a good compromise between its accuracy and simplicity. While a very complex model tends to be impractical for general use, a simpler model is less accurate [8].

After defining a proper model, the tracer behavior can be analyzed and thus the physiological parameter of interest can be estimated. The tracer kinetics recorded by PET thereby

can relate to chemical transformations within the tissue (for instance the phosphorylation of ^{18}F FDG), degradation or recombination back into the blood vessel. All these processes can be described by individual functions over time. If the tracer is injected as a bolus, the functions describing the behavior are usually given as exponentials [22].

After injecting a tracer, it gets delivered to the recorded VOI by arterial blood or plasma. The temporal behavior of this supplier is referred to as input function. As the monitored tracer behavior strongly depends on the available tracer amount, it is necessary to know the underlying input function when it comes to kinetic models. For determining the input function, different approaches are available. A straightforward method hereby would be the measurement of arterial blood samples, which however is invasive and challenging to perform when imaging small animals [25].

Another possibility is the use of population-based input functions, by which the difficulties encountering in the measurement of the input function are reduced. Hereby a standard input function is utilized, which comprises of population-averaged blood data. For normalization and calibration this method still requires at least one blood sample [22, 26]. This technique turned out to be sufficiently accurate for many applications. Nevertheless, since the normalization factor depends on many parameters such as the patient's weight, the injected dose and the tissue recorded, a problem arising when it comes to population-based input functions is to find the right calibration factor [27].

Another method for assessing the input function is its estimation from the performed PET scan. Hereby the image derived input functions can be extracted directly from the acquired data which in contrast to population-based input functions is completely non-invasive. For this an artery in the PET images is used. For image derived input functions, however, the data has to be corrected for spillover effects and more analysis has to be performed [22].

A general problem when imaging via PET is the partial volume effect (PVE). Especially when monitoring an object smaller than the pixel size of the PET detectors, the object's intensity gets distributed over the whole pixel resulting in a decreased maximum intensity. For instance, for a tumor smaller than the pixel size, the PVE leads to an underestimation of its severity. But also for larger objects, the PVE results in a blurring of the image. Therefore the PVE has to be corrected for, which can be done by using correction factors derived from other imaging modalities such as MRI or CT [28].

In kinetic modeling there are different methods to assess the functions describing the tracer

behavior within a VOI. The most common method is compartmental modeling. It can be used to analyze complex processes including different chemical states and locations of the tracer. It also provides information about tracer uptake and clearance within a tissue [8]. This master thesis will concentrate on compartmental modeling only. Apart from compartmental modeling there are other approaches in kinetic modeling such as the Patlak method and the spectral analysis, which will not be discussed [22].

1.2.1 Compartmental Modeling

The most common implementation of kinetic models is the utilization of compartments. Each compartment thereby represents a different tissue region, body organ or even a certain chemical state of the tracer. By describing the tracer kinetics between all these compartments, the tracer behavior within the body as well as biochemical transformations can be investigated [8].

As an example, the behavior of the tracer ^{18}F FDG within the body can be described nicely by compartmental modeling: since the body cannot distinguish between ^{18}F FDG and glucose, both of them show the same transport across the cell membrane. From there they can either go back to the blood or undergo phosphorylation. In the case of the latter process, the phosphorylated ^{18}F FDG, in contrast to glucose, cannot enter the next step in glycolysis. So it partly accumulates in the phosphorylated state and partly goes back to the blood by getting dephosphorylated again. Especially in tumors the dephosphorylation takes place very slowly leading to an increased tracer accumulation within cancerous cells [29]. Each chemical state and location can be described by an individual compartment. The obtained data from the PET scan thereby represents the tracer concentration in both the interstitial and phosphorylated state plus a small fraction of the concentration within the blood as well. In compartmental modeling the overall concentration in the tissue compartments is called Time Activity Curve (TAC). The TAC represents the temporal behavior of the tracer within that tissue. By measuring the input function and the tracer concentration within these states, the tracer kinetics among the individual compartments can be estimated [22]. In the following the input function will be referred to as $B(t)$.

In the case of tissue modeling, compartment models are designed to describe the tracer kinetics between different compartments. The concentration within the individual compartments, however, is supposed to be spatially constant. This requirement of the tracer to be homogeneously distributed over each compartment is called the "well-mixed assumption".

Hence, no concentration gradients are taken into account. Another important assumption in compartmental modeling is that the tracers do not reach any saturation level. To ensure this requirement, the tracer has to be injected in negligible doses. As it is difficult to develop a proper kinetic model for heterogeneous tissues including concentration gradients, for this sort of problems other methods exist such as the spectral analysis or the multiple-time graphical analysis [4, 24].

Mathematically speaking, the tracer concentration within each compartment can be described by a function over time. After both setting up a proper model comprising of different compartments and defining the possible tracer exchange among them, the tracer concentration within each compartment can be described by a first-order differential equation (see next section 1.2.2) [24]. It is to mention that in the following master thesis, when speaking of kinetic modeling, the approach by compartmental modeling is meant only.

1.2.2 Mathematical Description

Usually, when dealing with compartmental modeling, the blood is presented as an individual compartment. It is responsible for the delivery of the tracer to the individual compartments. In literature, however, when speaking of a two-compartment model for instance, the model comprises of two tissue compartments plus the blood compartment. So the number of compartments follows the tissue compartments only.

The aim in compartmental modeling is to assess the tracer kinetics between all compartments. For this, so called rate constants k have to be introduced. These rate constants describe the fractional rate of change in compartmental tracer concentration and are usually given in units of 1/min. In figure 1.2, 1.3 and 1.4 a one-, a two- and a three-compartment model is shown with all possible rate constants k_i . The blood compartment is designated as $B(t)$, while the other compartments are labeled as $C_i(t)$. For a one-compartment model, there are only two rate constants describing the in- and outflow of the single compartment. A two-compartment model comprises of six rate constants in total and for a three-compartment model there are twelve possible rate constants. However, in compartmental modeling usually the number of rate constants taken into account is small (less than or equal to six) [8].

A further assumption in compartmental modeling is that the rate constants do not change over time. Hence, the processes going on have to be in a steady state, which allows the mathematical description by linear differential equations. Even if this assumption cannot be entirely met, the events are thought of as as being steady if the changes over time are slow

with respect to the data acquired over time.

As the rate constants define the processes going on, the goal in compartmental modeling is to assess these rate constants by looking at the measured tracer concentrations acquired by a PET. This would lead to an understanding of the overall tracer kinetics. However, a proper estimation of the rate constants can be very difficult [8].

To deduce the set of differential equations describing a compartment model, the total inflow and outflow of each compartment has to be characterized. The tracer flow from one compartment C_A to another compartment C_B is governed by the product of the corresponding rate constant $k_{A \rightarrow B}$ and the tracer concentration within the source compartment C_A . So the flow from C_A to C_B can be written as $C_A k_{A \rightarrow B}$. The net flow into a compartment is given by the sum of all inflows minus all outflows regarding this compartment [8]. This leads to a set of differential equations describing the concentration exchange between all the compartments [24]. In equation (1), (2) and (3) the set of differential equations for a one-compartment, a two-compartment and a three-compartment model are presented.

What has to be kept in mind is that tracer concentrations in blood are usually presented in kBq/ml, while the concentrations within the individual compartments are given in kBq/g. Hence, the rate constants describing the supply from the blood (or plasma) are given in units of ml/(g·min), while the rate constants among the individual compartments have units of 1/min. To underline this fact, in literature sometimes the constants describing the exchange between blood and tissue are capitalized [8].

As it is performed in most studies regarding compartmental modeling, this master thesis as well takes into account only three compartments at most. For more compartments the computation would get too complex, leading to an increase in statistical uncertainties and hence a lowered accuracy of the model. Nevertheless, theoretically one can extend the model to n compartments leading to n ordinary differential equations with each one describing the net flow into a different compartment [24]. By adding reasonable constraints to certain rate constants, the calculation of the rate constants will be facilitated leading to a better understanding of the compartment model [8].

To get familiar with compartmental modeling, a look at the simplest model, namely a one-compartment model, will be taken. A one-compartment model can be described by the following differential equation:

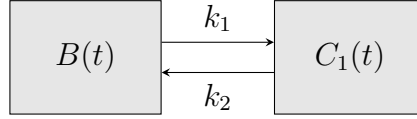


Figure 1.2: Kinetic model for a one-compartment model defined by two rate constants.

$$\dot{C}_1(t) = k_1 B(t) - k_2 C_1(t) \quad . \quad (1)$$

As one can see, the change in concentration for the compartment $C_1(t)$ is governed by the inflow given as $k_1 B(t)$ minus the outflow $k_2 C_1(t)$. Figure 1.2 shows a sketch of a one-compartment model.

Using the Euler method for the corresponding homogeneous differential equation leads to an exponential function $Ae^{-k_2 t}$, where A is a constant. For solving the inhomogeneous differential equation (1), the method of variation of constants can be applied, providing the general solution following

$$C_1(t) = k_1 B(t) \otimes e^{-k_2 t} \quad .$$

The operator \otimes thereby designates the convolution [30] of two functions defined as

$$(f \otimes g)(t) := \int_{-\infty}^{\infty} f(s)g(t-s) ds \quad .$$

This example simply serves to get a first impression about compartmental modeling. In section 2.2 the calculation for a two-compartment model including the intermediate mathematical steps will be presented more thoroughly. Another elegant method for solving the inhomogeneous problem could be the Laplace transform [31]. As one can see, $C_1(t)$ is given by a decreasing exponential function convolved with the input function $B(t)$. Since the tissue concentration $C_1(t)$ can be measured via PET and the blood concentration $B(t)$ can be deduced from blood samples or directly from the PET scan, the rate constants k_1 and k_2 can be estimated via deconvolution or by other means. For models considering more compartments the solution gets more complex as well.

In case of two compartments, the model can be described by the following set of differential equations:

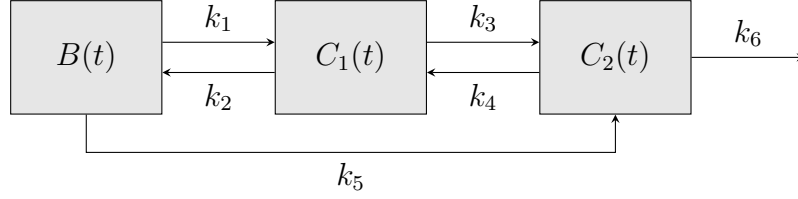


Figure 1.3: Kinetic model for a two-compartment model defined by six rate constants.

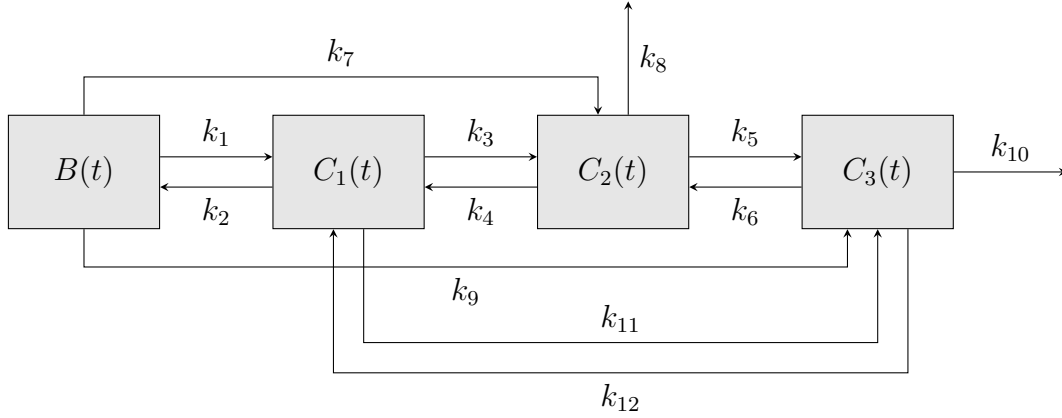


Figure 1.4: Kinetic model for a three-compartment model comprising of twelve rate constants in total.

$$\begin{aligned}
 \dot{C}_1(t) &= k_1 B(t) + k_4 C_2(t) - (k_2 + k_3) C_1(t) \\
 \dot{C}_2(t) &= k_5 B(t) + k_3 C_1(t) - (k_4 + k_6) C_2(t) \quad .
 \end{aligned}
 \tag{2}$$

Again, for both the first and second compartment, the total change in tracer concentration is given by the inflow resulting from other compartments minus the outflow of the compartment itself. The corresponding model is sketched in figure 1.3. As one can see, the number of rate constants increases to six constants for a two-compartment model. Hence, solving a two-compartment model is much more difficult than solving a one-compartment model.

The last model considered in this master thesis will be the three-compartment model, which comprises of twelve rate constants in total (see figure 1.4). This model can be described by using three linear differential equations:

$$\begin{aligned}
 \dot{C}_1(t) &= k_1 B(t) + k_4 C_2(t) + k_{12} C_3(t) - (k_2 + k_3 + k_{11}) C_1(t) \\
 \dot{C}_2(t) &= k_7 B(t) + k_3 C_1(t) + k_6 C_3(t) - (k_4 + k_5 + k_8) C_2(t) \\
 \dot{C}_3(t) &= k_9 B(t) + k_{11} C_1(t) + k_5 C_2(t) - (k_6 + k_{10} + k_{12}) C_3(t) \quad .
 \end{aligned}
 \tag{3}$$

However, it has to be mentioned that for this type of model not all rate constants are taken into account since this would explode the model's complexity. Instead, only a few rate constants are considered while the other ones are assumed to be zero. As already mentioned, the number of rate constants used for a model usually is less than or equal to six.

A special scenario in compartmental modeling is when the tracer is trapped within a specific compartment. For example, this can be the case when the tracer is irreversibly bound in a certain tissue [8]. Since in this case there is no escape from that compartment, the rate constants defining its outflow have to be zero. Such models are referred to as trapped model, otherwise a model is called open model. For a one-compartment model a trapped model is present if $k_2 = 0$, resulting in a differential equation according to

$$\dot{C}_1(t) = k_1 B(t)$$

with the simple solution

$$C_1(t) = k_1 \int_0^t B(s) ds \quad .$$

A trapped two-compartment model is present if either $k_2 = k_3 = 0$ (tracer is trapped in the first compartment) or $k_4 = k_6 = 0$ (tracer is trapped in the second compartment). For a three-compartment model at least one compartment has to be characterized by no outflow in order to be designated as trapped model.

Since some rate constants are set zero in case of a trapped model, further calculation and estimation of the remaining rate constants will be facilitated. In contrast, for an open model with no information about the tracer kinetics, the estimation of all rate constants is more complicated.

To solve such compartment models, different approaches exist such as the equilibrium method and the reference region method (see [8]). This master thesis, however, deals with a numerical approach in order to estimate the rate constants. All the models taken into account in this master thesis will be the trapped and open models for one, two and three compartments.

2 Motivation

As aforementioned, there are different approaches in solving kinetic models. The first try performed in this master thesis was done by solving the set of differential equations analytically (see section 2.1). Since the solution of the differential equation becomes very complex already for a simplified two-compartment model, this result motivated the approach of solving the differential equations by numerical means.

The next task was to find constraints for the rate constants k_i , which in consequence would facilitate further numerical calculations. As the TACs comprise of some characteristics regarding their shape (see figure 3.6), one can infer conditions which have to be fulfilled in order to guarantee a reasonably shaped TAC. However, as it turned out in section 2.2, no useful information could be drawn from this mathematical description. Therefore the set of differential equations had to be solved by numerical means with no available conditions for the rate constants at all.

To estimate the rate constants numerically, a curve fitting algorithm is used (see section 2.3). A general drawback of all these algorithms is the dependence of their reliability on the given initial values. Especially for a two- and a three-compartment model there are many rate constants to be estimated. Thus, it is important to feed the optimizing algorithm with proper initial values to enhance the outcome of the curve fit and hence to obtain a good estimation of the rate constants. The main goal in this master thesis was to program an algorithm which finds good initial values for the rate constants (see chapter 3).

2.1 Analytical Solution

This section deals with the analytical solution for a set of differential equations and shows how complex the analytical solutions become with increasing number of compartments. For simplification, a mono-exponential input function following $B(t) = e^{-bt}$ is assumed. For a one-compartment model comprising of two rate constants k_1 and k_2 , the analytical solution for equation (1) is given as

$$C_1(t) = \frac{k_1}{k_2 - b}(e^{-bt} - e^{-k_2t}) \quad .$$

Although this solution is simple, for a two-compartment it worsens dramatically. In case of a trapped two compartment model with $k_4 = k_5 = k_6 = 0$, the analytical solution for the

individual compartments is given as

$$C_1(t) = \frac{k_1}{k_2 + k_3 - b} (e^{-bt} - e^{-(k_2+k_3)t})$$

$$C_2(t) = \frac{k_1 k_3}{k_2 + k_3 - b} \left(-\frac{e^{-bt}}{b} + \frac{e^{-(k_2+k_3)t}}{k_2 + k_3} \right) + \frac{k_1 k_3}{b(k_2 + k_3)} .$$

The TAC concentration $C(t)$, which can be seen as the sum of the two individual compartment concentrations thereby can be written as

$$C(t) = C_1(t) + C_2(t)$$

$$= \frac{k_1}{k_2 + k_3 - b} \left(\left(1 - \frac{k_3}{b}\right) e^{-bt} - \left(1 - \frac{k_3}{k_2 + k_3}\right) e^{-(k_2+k_3)t} \right) + \frac{k_1 k_3}{b(k_2 + k_3)} .$$

With taking into account one more rate constant, namely k_4 , the complexity of the analytical solution explodes. With the computer algebra system MAXIMA and the abbreviations $\xi = k_2 + k_3 + k_4$ and $\zeta = \sqrt{\mu^2 - 4k_2 k_4}$ one obtains

$$C_1(t) = \frac{1}{k_2 k_4 + b(b - \mu)} \left[e^{-\frac{\mu}{2}t} \left(\frac{k_1[-\mu k_4 + b(k_4 - k_2 - k_3)]}{\zeta} \sinh\left(\frac{\zeta}{2}t\right) \right. \right.$$

$$\left. \left. - k_1(k_4 - b) \cosh\left(\frac{\zeta}{2}t\right) \right) + k_1(k_4 - b)e^{-bt} \right]$$

$$C_2(t) = \frac{1}{k_2 k_4 + b(b - \mu)} \left[e^{-\frac{\mu}{2}t} \left(\frac{k_1 k_3 (2b - \mu)}{\zeta} \sinh\left(\frac{\zeta}{2}t\right) - k_1 k_3 \cosh\left(\frac{\zeta}{2}t\right) \right) + k_1 k_3 e^{-bt} \right] .$$

Again, the measured TAC would be the sum of these two compartments. This example shows that by taking into account one more rate constant, the complexity of the analytical solution increases drastically. So for more complex compartment models and a more sophisticated input function, the solution could hardly be calculated by analytical means. Therefore, this master thesis concentrates on the solution via numerical techniques only.

2.2 Mathematical Examination for the Rate Constants

In the previous chapter the set of differential equations describing a kinetic model have already been presented (see equations (1), (2) and (3)). The more rate constants k_i considered in the model, the more complex the mathematical equations. Nevertheless, one can solve these differential equations non-explicitly by means of mathematical methods. Below, this

is done for a two-compartment model only. One would hope the solution to reveal some information about the rate constants, since some conditions have to be fulfilled as it will be explained hereafter.

For a two-compartment model the set of differential equations is given as

$$\dot{C}_1(t) = k_4 C_2(t) - (k_2 + k_3) C_1(t) + k_1 B(t) \quad (4)$$

$$\dot{C}_2(t) = k_3 C_1(t) - (k_4 + k_6) C_2(t) + k_5 B(t) \quad (5)$$

As one can see, equation (4) for $\dot{C}_1(t)$ also comprises of the term $C_2(t)$ and vice versa for equation (5). Therefore the first step is to uncouple these equations. This, for example, can be done by differentiating equation (4) and subsequently inserting equation (5) for the term $\dot{C}_2(t)$:

$$\begin{aligned} \ddot{C}_1(t) &= k_4 \dot{C}_2(t) - (k_2 + k_3) \dot{C}_1(t) + k_1 \dot{B}(t) \\ &= k_4 [k_3 C_1(t) - (k_4 + k_6) C_2(t) + k_5 B(t)] - (k_2 + k_3) \dot{C}_1(t) + k_1 \dot{B}(t) \\ &= k_3 k_4 C_1(t) + (k_4 + k_6) [-(k_2 + k_3) C_1(t) - \dot{C}_1(t) + k_1 B(t)] \\ &\quad - (k_2 + k_3) \dot{C}_1(t) + k_4 k_5 B(t) + k_1 \dot{B}(t) \quad . \end{aligned}$$

By that one obtains a second-order differential equation with an inhomogeneity, hereafter referred to as forcing function $F_1(t)$. For a better view, this equation can be rearranged:

$$\ddot{C}_1(t) + \underbrace{(k_2 + k_3 + k_4 + k_6)}_{=: \gamma} \dot{C}_1(t) + \underbrace{(k_2 k_4 + k_2 k_6 + k_3 k_6)}_{=: \omega_0^2} C_1(t) = \underbrace{(k_1 k_4 + k_1 k_6 + k_4 k_5) B(t) + k_1 \dot{B}(t)}_{=: F_1(t)} \quad .$$

The same can be done for the second differential equation (5), leading to a similar equation with a different forcing function $F_2(t)$ only. With the abbreviations

$$\begin{aligned} \gamma &= k_2 + k_3 + k_4 + k_6 \\ \omega_0^2 &= k_2 k_4 + k_2 k_6 + k_3 k_6 \\ F_1(t) &= (k_1 k_4 + k_1 k_6 + k_4 k_5) B(t) + k_1 \dot{B}(t) \\ F_2(t) &= (k_1 k_3 + k_2 k_5 + k_3 k_5) B(t) + k_5 \dot{B}(t) \end{aligned}$$

one obtains two uncoupled second-order differential equations with different forcing functions $F_1(t)$ and $F_2(t)$:

$$\ddot{C}_1(t) + \gamma\dot{C}_1(t) + \omega_0^2 C_1(t) = F_1(t) \quad (6)$$

$$\ddot{C}_2(t) + \gamma\dot{C}_2(t) + \omega_0^2 C_2(t) = F_2(t) \quad (7)$$

The corresponding homogeneous differential equation $\ddot{C}_i(t) + \gamma\dot{C}_i(t) + \omega_0^2 C_i(t) = 0$ with $i = 1, 2$ is the same for both equations (6) and (7). To solve these equations, the Euler approach $C_i(t) = Ae^{\lambda t}$ is utilized, leading to

$$\begin{aligned} \lambda^2 Ae^{\lambda t} + \lambda\gamma Ae^{\lambda t} + \omega_0^2 Ae^{\lambda t} &= 0 \\ \lambda^2 + \lambda\gamma + \omega_0^2 &= 0 \\ \lambda_{1,2} &= -\frac{\gamma}{2} \pm \frac{\xi}{2} \quad , \end{aligned}$$

where ξ is given by $\xi = \sqrt{\gamma^2 - 4\omega_0^2}$. The two possibilities for λ yield two individual functions

$$\begin{aligned} y_1(t) &= e^{\lambda_1 t} = e^{(-\frac{\gamma}{2} + \frac{\xi}{2})t} \\ y_2(t) &= e^{\lambda_2 t} = e^{(-\frac{\gamma}{2} - \frac{\xi}{2})t} \end{aligned}$$

solving the homogeneous differential equations. Since all the considered equations describing the kinetic model are linear, the general solutions for the homogeneous differential equations are given by a linear combination of those using the constants [32]:

$$\begin{aligned} C_1^h(t) &= A_1 y_1(t) + B_1 y_2(t) = A_1 e^{\lambda_1 t} + B_1 e^{\lambda_2 t} \\ C_2^h(t) &= A_2 y_1(t) + B_2 y_2(t) = A_2 e^{\lambda_1 t} + B_2 e^{\lambda_2 t} \quad . \end{aligned}$$

Hereby A_1, A_2, B_1 and B_2 are constants and the h in the superscript indicates that these are the solutions for the homogeneous differential equations. The inhomogeneous equations (6) and (7) with their forcing functions $F_1(t)$ and $F_2(t)$, respectively, can be solved by means of the method of variation of constants, considering the constants A_1, A_2, B_1 and B_2 as functions $A_1(t), A_2(t), B_1(t), B_2(t)$ over time.

By that, one obtains a system of equations according to

$$\begin{pmatrix} y_1(t) & y_2(t) \\ \dot{y}_1(t) & \dot{y}_2(t) \end{pmatrix} \cdot \begin{pmatrix} \dot{A}_i(t) \\ \dot{B}_i(t) \end{pmatrix} = \begin{pmatrix} 0 \\ F_i(t) \end{pmatrix} ,$$

where the index i again represents the compartment considered [33]. The solution of this system of equations is given by

$$\begin{aligned} \dot{A}_i(t) &= \frac{1}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} F_i(t) \\ \dot{B}_i(t) &= \frac{1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} F_i(t) \end{aligned}$$

and thus

$$\begin{aligned} A_i(t) &= \frac{1}{\xi} \int_0^t e^{-\lambda_1 s} F_i(s) ds \\ B_i(t) &= -\frac{1}{\xi} \int_0^t e^{-\lambda_2 s} F_i(s) ds \quad . \end{aligned}$$

One can write the general solution for equations (6) and (7) as following:

$$\begin{aligned} C_1(t) &= \frac{1}{\xi} \int_0^t e^{\lambda_1(t-s)} F_1(s) ds - \frac{1}{\xi} \int_0^t e^{\lambda_2(t-s)} F_1(s) ds \\ &= \frac{1}{\xi} \int_0^t e^{-\frac{\gamma}{2}(t-s)} \left[e^{\frac{\xi}{2}(t-s)} - e^{-\frac{\xi}{2}(t-s)} \right] F_1(s) ds \\ &= \frac{2}{\xi} \int_0^t e^{-\frac{\gamma}{2}(t-s)} \sinh \left(\frac{\xi}{2}(t-s) \right) F_1(s) ds \end{aligned} \quad (8)$$

$$C_2(t) = \frac{2}{\xi} \int_0^t e^{-\frac{\gamma}{2}(t-s)} \sinh \left(\frac{\xi}{2}(t-s) \right) F_2(s) ds \quad . \quad (9)$$

By using the convolution operator \otimes , the solutions for $C_1(t)$ and $C_2(t)$ (equation (8) and (9)) can be rewritten as

$$C_i(t) = \frac{2}{\xi} \left(e^{-\frac{\gamma}{2}t} \sinh \left(\frac{\xi}{2}t \right) \theta(t) \right) \otimes F_i(t)$$

with $\theta(t)$ being the Heaviside function defined as

$$\theta(t) = \begin{cases} 0 & , \text{ for } t < 0 \\ 1 & , \text{ for } t \geq 0 \end{cases} .$$

Since the upper integration limit in equation (8) and (9) is t , the Heaviside function $\theta(t)$ is important in order to consider this integration limit. The lower limit is set zero, as for the forcing function holds $F_i(t) = 0$ for $t < 0$. With the response function

$$X(t) = \frac{2}{\xi} e^{-\frac{\gamma}{2}t} \sinh\left(\frac{\xi}{2}t\right) \theta(t) \quad (10)$$

the solutions for $C_1(t)$ and $C_2(t)$ can also be written as

$$\begin{aligned} C_1(t) &= (X \otimes F_1)(t) \\ C_2(t) &= (X \otimes F_2)(t) \end{aligned} .$$

With having these solutions for the first and second compartment, one can try to infer some informations about the rate constants, as there are some important mathematical characteristics which have to be fulfilled. More precisely, there are two main conditions which have to be met by the individual compartments:

- By taking a look at real TACs (see figure 3.6), one can see that all these TACs are rather smooth, except for statistical noise leading to a jagged curve. Hence, the argument in the sine hyperbolic term in equation (10) needs to be real. Otherwise, in the case of a complex ξ , the sine hyperbolic would transform to a sine causing a sinusoidal oscillation within the TAC which is not reasonable.

Since the argument ξ in the sine hyperbolic is defined by $\xi = \sqrt{\gamma^2 - 4\omega_0^2}$, its radicand needs to be positive in order to guarantee a real ξ . By inserting the definitions of γ and ω_0^2 , one obtains the following condition:

$$\begin{aligned} \gamma^2 - 4\omega_0^2 &\geq 0 \\ (k_2 + k_3 + k_4 + k_6)^2 - 4(k_2k_4 + k_2k_6 + k_3k_6) &\geq 0 \\ k_2^2 + k_3^2 + k_4^2 + k_6^2 + 2(k_2k_3 - k_2k_4 - k_2k_6 + k_3k_4 - k_3k_6 + k_4k_6) &\geq 0 \end{aligned} . \quad (11)$$

This inequality can now be rearranged for the rate constants k_2, k_3, k_4 and k_6 :

- Rearranging for k_2 : the inequality can be rewritten as

$$k_2^2 + 2k_2(k_3 - k_4 - k_6) + k_3^2 + k_4^2 + k_6^2 + 2(k_3k_4 - k_3k_6 + k_4k_6) \geq 0 \quad , \quad (12)$$

which is only fulfilled for k_2 being not located within the interval

$$k_2 \notin [k_4 + k_6 - k_3 - \sqrt{-4k_3k_4}, k_4 + k_6 - k_3 + \sqrt{-4k_3k_4}] \quad .$$

As all rate constants are supposed to be positive, the radicand $-4k_3k_4$ is negative, leading to two complex interval boundaries for k_2 . Graphically spoken the quadratic function in equation (12) is positive for all values of k_2 . So there is no real forbidden interval for k_2 , which would result in a negative value for ξ and by that cause a sinusoidal oscillation within the response function $X(t)$. Hence, ξ is positive regardless of the rate constant k_2 (see left picture in figure 2.1).

- Rearranging for k_3 : after rewriting inequality (11), one gets

$$k_3^2 + 2k_3(k_2 + k_4 - k_6) + k_2^2 + k_4^2 + k_6^2 + 2(-k_2k_4 - k_2k_6 + k_4k_6) \geq 0 \quad .$$

By solving this quadratic inequality for k_3 , one obtains the forbidden interval for k_3 , where ξ would be negative and hence $X(t)$ would comprise of a sinusoidal oscillation:

$$k_3 \notin [k_6 - k_2 - k_4 - 2\sqrt{k_4(k_2 - k_6)}, k_6 - k_2 - k_4 + 2\sqrt{k_4(k_2 - k_6)}] \quad .$$

For $k_2 < k_6$, the radicand again would be negative leading to complex interval boundaries. In this case, ξ would be positive for all choices of k_3 and therefore no forbidden interval would exist similar to the previous outcome for k_2 .

In the case of $k_2 \geq k_6$, the radicand now would be positive causing a real forbidden interval for k_3 where ξ would become negative. The lower interval boundary

$$\underbrace{k_6 - k_2 - k_4 - 2\sqrt{k_4(k_2 - k_6)}}_{\leq 0}$$

is negative, as all individual terms are negative. One would hope the upper interval boundary to be positive to obtain a feasible forbidden interval for k_3 . Otherwise, the forbidden interval is completely located at negative values of k_3

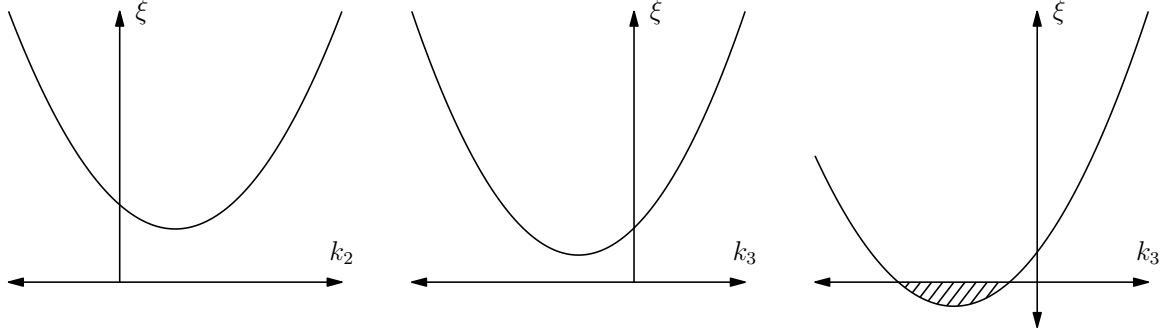


Figure 2.1: Dependence of ξ on the rate constants k_2 (left plot) and k_3 (middle and right plot). As one can see, ξ is positive for all values of k_2 . For the dependence on k_3 it is important to distinguish between the cases $k_2 < k_6$ (middle plot) and $k_2 \geq k_6$ (right plot). While for $k_2 < k_6$ the radicand ξ is positive for all values of k_3 , in the case of $k_2 \geq k_6$ there is a forbidden interval for k_3 highlighted in the right plot. However, since this interval is located in the negative range, no useful information can be drawn.

only, providing no useful restriction for k_3 . So it remains to check if the upper boundary is positive:

$$\begin{aligned}
 k_6 - k_2 - k_4 + 2\sqrt{k_4(k_2 - k_6)} &\stackrel{?}{>} 0 \\
 (k_2 + k_4 - k_6)^2 &< 4k_4(k_2 - k_6) \\
 k_2^2 + k_4^2 + k_6^2 + 2k_2k_4 - 2k_2k_6 - 2k_4k_6 &< 4k_2k_4 - 4k_4k_6 \\
 k_2^2 + k_4^2 + k_6^2 + 2(-k_2k_4 - k_2k_6 + k_4k_6) &< 0 \\
 (k_2 - k_4 - k_6)^2 &< 0 \quad .
 \end{aligned}$$

Obviously this inequality is not fulfilled, since the left-hand term is positive due to the square. Hence, in the case of $k_2 \geq k_6$, a forbidden interval for k_3 exists, however, since this interval is located in the negative range (both boundaries are negative), this interval only rules out negative values for k_3 and therefore no further information about k_3 can be drawn. In figure 2.1 the dependence of ξ on k_3 is shown for the case $k_2 < k_6$ (middle plot) as well as for $k_2 \geq k_6$ (right plot). In the right plot the forbidden interval for k_3 leading to a negative ξ is highlighted.

- Rearranging for k_4 and k_6 : since k_4 plays the same role as k_3 and k_6 the same as k_2 (see figure 1.3), rearranging for k_4 and k_6 would provide the same results as it does for k_2 and k_3 .

In conclusion, the condition for $\xi = \sqrt{\gamma^2 - 4\omega_0^2}$ to be real and thus $\gamma^2 - 4\omega_0^2$ to be positive does not yield any useful information for the individual rate constants, since ξ is real for all positive values of k_2, k_3, k_4 and k_6 .

- The other important factor which has to be considered in the response function (10) is the interaction between the exponential and the sine hyperbolic term. While the exponential term $e^{-\frac{\gamma}{2}t}$ causes a decrease, the sine hyperbolic $\sinh(\frac{\xi}{2}t)$ is responsible for an increase of $X(t)$. By rewriting the sine hyperbolic term and omitting the factor $\frac{2}{\xi}$ and the Heaviside function $\theta(t)$, which do not contribute to the overall shape of $X(t)$, one obtains the simplified response function \bar{X} :

$$\begin{aligned}\bar{X}(t) &= e^{-\frac{\gamma}{2}t} \sinh\left(\frac{\xi}{2}t\right) \\ &= \frac{1}{2} e^{-\frac{\gamma}{2}t} (e^{\frac{\xi}{2}t} - e^{-\frac{\xi}{2}t}) \\ &= \frac{1}{2} (e^{-\frac{1}{2}(\gamma-\xi)t} - e^{-\frac{1}{2}(\gamma+\xi)t})\end{aligned}\tag{13}$$

Since the response function must not explode, both exponential terms have to define a decrease of $\bar{X}(t)$ and by that also of $X(t)$. As the power in the second term $e^{-\frac{1}{2}(\gamma+\xi)t}$ is negative due to the positive γ and ξ , this term causes a decrease anyway. What one has to concentrate on is the first term $e^{-\frac{1}{2}(\gamma-\xi)t}$. To guarantee a negative power for this first term, γ must surpass ξ leading to the following condition:

$$\begin{aligned}\gamma &> \xi \\ \gamma &> \sqrt{\gamma^2 - 4\omega_0^2} \\ \gamma^2 &> \gamma^2 - 4\omega_0^2 \\ 0 < \omega_0^2 &= k_2k_4 + k_2k_6 + k_3k_6\end{aligned}$$

The last inequality is fulfilled as all the rate constants are supposed to be positive, hence the sum of the products is positive as well. Thus, both terms in equation (13) define a decrease of $\bar{X}(t)$ and $X(t)$, respectively.

The only case where the power in the first term transforms to zero is for a trapped model with $k_4 = k_6 = 0$ (see section 1.2.2). Inserting these rate constants into ω_0^2 , one

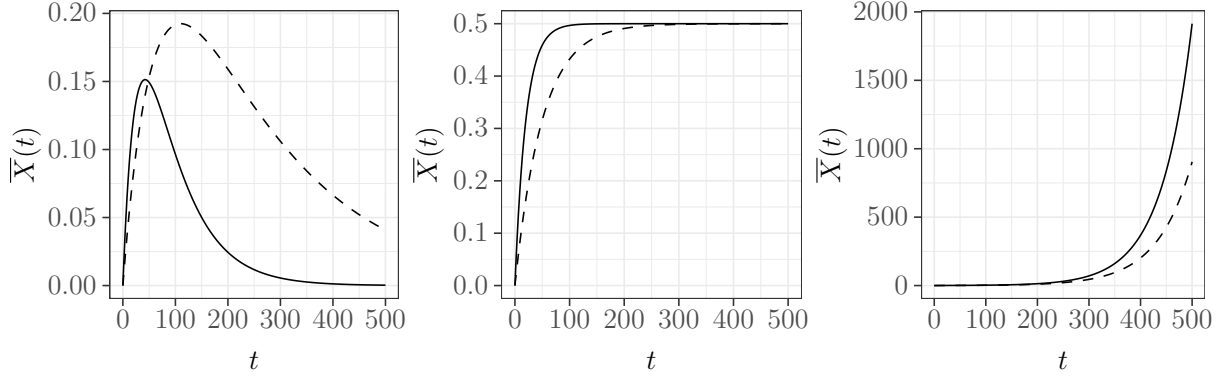


Figure 2.2: Shown are the different scenarios for the simplified response function $\bar{X}(t)$ (for different values of ξ and γ leading to the solid and the dashed line). In the left plot both exponential terms contribute to a decrease of $\bar{X}(t)$. The middle plot represents a trapped model, where the response function converges to a certain level and the right plot shows the forbidden case for $\bar{X}(t)$. Since all rate constants are positive, this will not be the case.

can easily see that $\omega_0^2 = 0$ and by that $\gamma = \xi$. In this case $\bar{X}(t)$ can be written as

$$\bar{X}(t) = \frac{1}{2}(1 - e^{-\frac{1}{2}(\gamma+\xi)t})$$

To analyze the temporal behavior of $\bar{X}(t)$, one let t go to infinity according to

$$\lim_{t \rightarrow \infty} \bar{X}(t) = \lim_{t \rightarrow \infty} \frac{1}{2}(1 - e^{-\frac{1}{2}(\gamma+\xi)t}) = \frac{1}{2} \quad .$$

For a trapped model with $k_4 = k_6 = 0$ the response function $X(t) = \frac{2}{\xi}\bar{X}(t)\theta(t)$ converges to the constant value $\lim_{t \rightarrow \infty} X(t) = \frac{1}{\xi}$. In contrast, for an open model both terms in equation (13) cause a decrease of $X(t)$. To sum up, the condition that $X(t)$ must not explode yields no information about the rate constants either. In figure 2.2 all three cases for the simplified response function $\bar{X}(t)$ are shown.

Above the attempt was made to solve a two-compartment model analytically. Although the response function has to fulfill some conditions, no useful information about the rate constants k_i can be drawn. This is due to the fact that the initial assumption of the rate constants to be non-negative already ensures that all other conditions are fulfilled already. So, apart from the rate constants being $k_i \geq 0$, all the conditions providing a reasonable shape of the response function do not yield any further information about the rate constants.

As the attempt of solving a kinetic model analytically is a dead end, another approach has

to be performed. This leads to the numerical solution of that problem, which is examined in this master thesis.

2.3 Numerical Approach

This master thesis follows up with a numerical estimation of the rate constants. There already are numerous optimization tools which adapt the rate constants k_i in such a way that the simulated TAC generated with the estimated rate constants (see section 3.1) matches the measured TAC as well as possible. For this, the optimization software takes initial values for all rate constants and gradually modifies them in order to increase the accordance with the measured TAC.

An often used optimization method for curve fitting is the Nelder-Mead simplex method, also called downhill simplex method [34]. With $\text{TAC}_{\text{real}}(t)$ being the real measured TAC and $\text{TAC}_{\text{sim}}(t_i)$ being the simulated TAC defined on a discrete time line and depending on the rate constants k_i , the objective function f to be minimized usually is defined as the sum of squared errors [35]:

$$f(k_1, k_2, \dots) = \sum_{t=0}^{t_{\text{end}}} (\text{TAC}_{\text{real}}(t) - \text{TAC}_{\text{sim}}(t, k_1, k_2, \dots))^2 \quad .$$

The Nelder-Mead simplex method gradually adapts the rate constants k_i in order to minimize f . This yields the rate constants providing the best local fit. A big advantage compared to other optimization tools is that the Nelder-Mead simplex method does not require any derivatives for its calculation. Plus, this method is rather robust against initial values. An accurate description of the setup of the Nelder-Mead simplex method is given in [36, 37].

While $\text{TAC}_{\text{real}}(t)$ is known, the discrete function $\text{TAC}_{\text{sim}}(t, k_1, k_2, \dots)$ has to be simulated. As the derivatives of the individual compartments are given explicitly (see equations (1), (2) and (3) in section 1.2.2), the simulation of the TAC can be performed by the Euler method, which can be seen as a first-order Runge-Kutta method [38]. This method provides a stepwise solution of the differential equations leading to the approximated $\text{TAC}_{\text{sim}}(t, k_1, k_2, \dots)$: for a two-compartment model the set of equations (2) can be written as

$$\begin{pmatrix} \dot{C}_1(t) \\ \dot{C}_2(t) \end{pmatrix} = \begin{pmatrix} f_1(t, C_1(t), C_2(t), B(t)) \\ f_2(t, C_1(t), C_2(t), B(t)) \end{pmatrix} \quad , \quad (14)$$

where f_1 and f_2 represent the respective right-hand sides in (2) [39]. Since a derivative can be approximated by

$$\dot{C}_i(t) = \frac{1}{h} (C_i(t+h) - C_i(t))$$

with step size h , equation (14) can be rewritten as

$$\begin{pmatrix} C_1(t+h) \\ C_2(t+h) \end{pmatrix} = \begin{pmatrix} C_1(t) + hf_1(t, C_1(t), C_2(t), B(t)) \\ C_2(t) + hf_2(t, C_1(t), C_2(t), B(t)) \end{pmatrix} . \quad (15)$$

With the initial condition $C_1(0) = C_2(0) = 0$ and the input function $B(t)$ being known, the concentrations in the individual compartments can be approximated [32]. For a three-compartment model the set of equations (3) can be rewritten analogously. A crucial factor hereby is the step size h : the smaller the step size, the better the approximation but the larger the computation time [40]. In section 3.1.2 the simulation of the TAC is performed with a step size $h = 1$ s.

For a better approximation of the derivative, other methods exist such as the trapezoidal rule, which belongs to the class of second-order Runge-Kutta methods or the Adams method, which takes into account not only one time point t for calculating the value at $t+h$, but rather the last s time points to increase the accuracy of the numerical calculation [38, 41].

Although the simulation of a TAC as well as the optimization tool for minimizing the objective function f can be easily computed, a fundamental problem hereby is the strong dependence of the outcome for the rate constants on their initial values. Hence, the main goal is to find a reasonable estimation of the rate constants in order to feed the optimization software with good initial values. In this master thesis the attempt was made to find such a good estimation by numerical means.

By taking a look at equation (14) and (15), respectively, one can see that each compartment depends on the input function $B(t)$. Hence, the shape of the input function (described in section 3.1.1) is supposed to have a large impact on the shape of the simulated TAC.

With all the mathematical background being specified, chapter 3 deals with the step-by-step approach performed in this master thesis.

3 Methods

This chapter concentrates on the approach of programming an algorithm to obtain good initial values for the rate constants k_i . The whole procedure is divided into different subtasks with each one being described in the following sections. First, it was necessary to simulate both an input function and a TAC for several different rate constants, which was done by the Euler method explained in the previous chapter. These functions have then been used for gathering relations between the TACs and their underlying rate constants. In the next step these relations underwent several statistical analyses resulting in specific correlations between the rate constants k_i and the corresponding shape of the TAC. These connections have been used to estimate the rate constants by examining the TAC only. Finally, an algorithm was programmed which takes into account all the obtained information and by that provides good initial values for the rate constants k_i , which can be used for further optimization programs (see section 2.3).

In parallel, the results of all subtasks performed are mentioned and analyzed in the corresponding sections in chapter 4.

3.1 Simulation of Time Activity Curves

The first aim was the generation of TACs, which in consequence can be used for further examination. However, before simulating the compartment concentrations via the spreadsheet program EXCEL, it was important to generate different input functions $B(t)$, which represent suppliers for the various compartments in the kinetic model described in section 1.2.1. These input functions have then been used for the calculation of the individual compartment concentrations.

3.1.1 Generation of Input Functions

Since the collective behavior of the tracer concentration in the compartments strongly depends on the shape of the input function, first it was necessary to generate different input functions which then can be used for the following simulations of TACs. By that, one can abstract from the input function and investigate the compartment behavior solely. It turned out that for the problem considered in this master thesis there are three main types of input functions:

The first type shows a quick rise, then reaches a maximum followed by a rapid decrease of concentration as it can be observed for example with the tracer $^{68}\text{Ga-PSMA}$, which is used for prostate imaging in nuclear medicine [42]. This input function hereafter will be called pointed input function $B_p(t)$ and is characterized by a sharp peak. To generate an input function of this shape, the following mathematical formula is used:

$$B_p(t) = \underbrace{\theta'(t-10)\theta'(25-t) \cdot 10(t-10)}_{=:A(t)} + \theta'(t-25) \cdot (143e^{-0.05(t-25)} + 7) \quad , \quad (16)$$

where $\theta'(t)$ is the half-maximum convention of the Heaviside function defined by

$$\theta'(t) = \begin{cases} 0 & , \text{ for } t < 0 \\ \frac{1}{2} & , \text{ for } t = 0 \\ 1 & , \text{ for } t > 0 \end{cases} .$$

The term $A(t)$ in equation (16) can be thought of as the beginning term of the input function. $A(t)$ is the same for all input functions considered in this master thesis and defines a constant value of $B(t) = 0$ for $0 \leq t \leq 10$ followed by a linear rise from $B(10) = 0$ up to $B(25) = 150$ in the time interval $10 < t \leq 25$. The constant phase can be thought of as starting the concentration measurement before injecting the tracer. In this master thesis this delay was assumed to be 10s, hence $B(t) = 0$ for $0 \leq t \leq 10$. At $t = 10$ the injection takes place, thus from there on the concentration of the input function starts to rise.

For $t > 25$ the decreasing term is different for each type of input function. In case of a pointed input function $B_p(t)$ the quick exponential decrease is given by the term $143e^{-0.05(t-25)}$ in equation (16). With increasing time t the concentration converges to the constant level of $\lim_{t \rightarrow \infty} B_p(t) = 7$. In conclusion, the function $B_p(t)$ consists of a constant phase in the beginning with value $B(t) = 0$, then increases linearly until it reaches a maximum and from then on exponentially decreases rapidly again to a certain level. In the left graphic in figure 3.1 one can see the shape of the pointed input function.

The second type of input function, referred to as blunt input function $B_b(t)$, is similar to the pointed one. It is present for example when imaging via ^{18}FDG , a main tracer in oncology. The only difference to the pointed input function is the slower decrease after the peak. It is

given by

$$B_b(t) = A(t) + \theta'(t - 25) \cdot (143e^{-0.01(t-25)} + 7) \quad , \quad (17)$$

where $A(t)$ is the beginning term mentioned above. By comparison of equation (16) and (17), one can see that the only difference is the power in the exponential term. While for the pointed input function $B_p(t)$ the power is $-0.05(t - 25)$, for the blunt input function $B_b(t)$ it is given as $-0.01(t - 25)$. This smaller decay constant leads to a slower decrease after the peak, which can be seen in the middle plot in figure 3.1.

While the first two types of input functions comprise of a mono-exponential decreasing term, the third type is defined by a tri-exponential term. With these additional terms the shape of the input function can be changed tremendously. By varying the decay constants in each term a buckled input function can be generated. The third input function used for further simulations of TACs is defined by

$$B_k(t) = A(t) + \theta'(t - 25) \cdot \left[143 \left(\frac{1}{2} e^{-0.1(t-25)} + \frac{1}{4} e^{-0.01(t-25)} + \frac{1}{4} e^{-0.001(t-25)} \right) + 7 \right] \quad (18)$$

and will be called kinky input function $B_k(t)$. This type can be found in measurements with the tracer $^{99m}\text{Tc-MAG3}$, which is typically used for kidney examinations in nuclear medicine. While the constant and ascending phase is the same as for the pointed and blunt input function, the decreasing term differs drastically. In $B_k(t)$ there are three exponential terms leading to a fall down, each of them weighted with a factor. The first term $e^{-0.1(t-25)}$ causes a quick drop of concentration, while the second and in particular the third term, $e^{-0.01(t-25)}$ and $e^{-0.001(t-25)}$, are responsible for a slower decrease. After reaching the maximum, the concentration therefore drops down rapidly due to the first decreasing term. However, after a short time the input function shows a kink from where on a slower decrease can be pointed out. In the right plot in figure 3.1 the shape for a kinky input function $B_k(t)$ with its pronounced kink can be seen.

The main difference between the first two types and the third one is that for the kinky input function the concentration decreases slowly enough to not reach a constant level over a long time period. Due to that fact, the input function $B_k(t)$ continuously supplies the individual compartments with a relatively high tracer concentration, whereas for the pointed and the blunt input function the supply for large times is rather negligible.

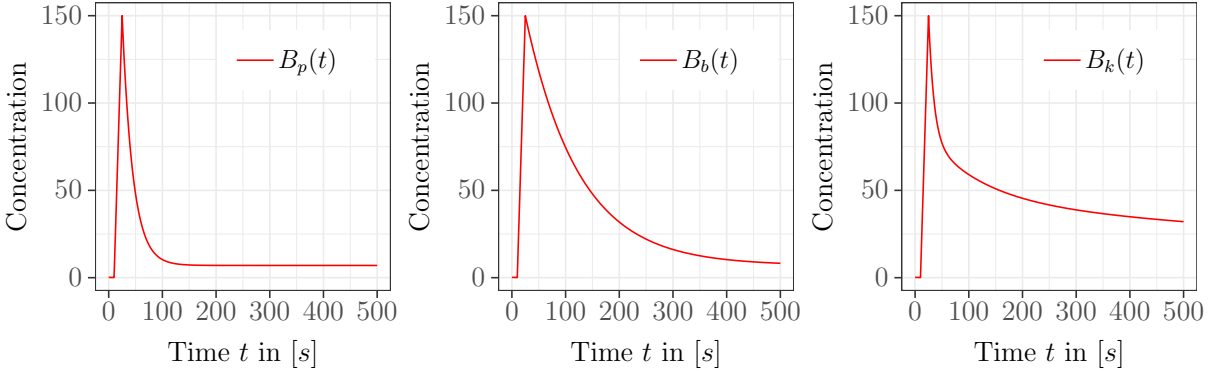


Figure 3.1: Shape of the different input functions $B(t)$. On the left-hand side the pointed input function with its sharp peak and the rapid fall down is shown. The middle figure represents the blunt input function comprising of a broader peak and a slower decrease afterwards. Both the sharp and the blunt input function reach a constant level after a certain time. On the right-hand side the kinky input function is shown. As one can see, after the peak it comprises of a rapid concentration decrease followed by a significant slower decrease leading to a kink in its shape.

3.1.2 Stepwise Solution of the Differential Equations

For the calculation of the individual compartment concentrations over time a stepwise solution of the set of differential equations (1), (2) and (3) described in section 1.2.2 has been established. For this, the Euler method with step of size $h = 1$ s was used. A smaller step size would increase the computation time while not yielding any further information, since dynamic measurements are usually recorded every second. In contrast, a larger step size could omit useful information. Plus, if h is chosen too large, the simulated TAC could comprise of oscillations which are not reasonable when looking at real TACs.

For a one-compartment model (see figure 1.2) the Euler approach with step $h = 1$ s has the following form:

$$C_1(t) = C_1(t - 1) + k_1 B(t - 1) - k_2 C_1(t - 1) \quad . \quad (19)$$

The compartment concentrations within a two-compartment model (see figure 1.3) can be calculated by

$$\begin{aligned} C_1(t) &= C_1(t - 1) + k_1 B(t - 1) + k_4 C_2(t - 1) - (k_2 + k_3) C_1(t - 1) \\ C_2(t) &= C_2(t - 1) + k_5 B(t - 1) + k_3 C_1(t - 1) - (k_4 + k_6) C_2(t - 1) \quad . \end{aligned} \quad (20)$$

For a three-compartment model (see figure 1.4) the set of equations above is expanded by a third compartment, hence also the number of rate constants k_i increases. The stepwise approach for the three-compartment model with step $h = 1$ s is given by:

$$\begin{aligned}
C_1(t) &= C_1(t-1) + k_1 B(t-1) + k_4 C_2(t-1) + k_{12} C_3(t-1) - (k_2 + k_3 + k_{11}) C_1(t-1) \\
C_2(t) &= C_2(t-1) + k_7 B(t-1) + k_3 C_1(t-1) + k_6 C_3(t-1) - (k_4 + k_5 + k_8) C_2(t-1) \\
C_3(t) &= C_3(t-1) + k_9 B(t-1) + k_{11} C_1(t-1) + k_5 C_2(t-1) - (k_6 + k_{10} + k_{12}) C_3(t-1) \quad .
\end{aligned}
\tag{21}$$

While the one- and two-compartment models comprise of two and six rate constants, respectively, the expansion to a three-compartment model increases the number of possible rate constants to up to twelve. However, as already mentioned, in compartmental modeling usually the number of rate constants is kept less than or equal to six. Hence, for a suitable model only some of the rate constants are of interest while the other ones are assumed to be zero.

For the discrete calculation of the compartment concentrations from the input function $B(t)$, the initial values for the compartments were set zero, which meets the biological requirement that there is no tracer concentration in the tissue before injection ($C_i(t_0) = 0$). It is to mention that the rate constants k_i in the models above are given in units of 1/sec while in medicine they are presented in units of 1/min. So it has to be kept in mind that one has to convert the results properly. The simulation was performed via EXCEL and subsequently exported to a csv-file, hereafter referred to as "TAC-simulation.csv".

The TAC representing the cumulated compartment concentration $C(t)$ thereby is given as the sum over all compartments $C_i(t)$ multiplied with a weighting factor $w \in [0, 1]$ plus the concentration of the input function $B(t)$ multiplied with $(1 - w)$:

$$C(t) = w \cdot \sum_{i=1}^n C_i(t) + (1 - w)B(t) \quad .$$

For this master thesis a weighting factor of $w = 0.95$ was used. n thereby stands for the number of compartments. With the last term $(1 - w)B(t)$ an impact of the input function on the acquired PET data of the tracer concentration is considered as well.

With the set of equations (19), (20) and (21), around 1200 different TACs have been generated by varying the rate constants k_i and using different number of compartments. In figure

4.1 simulated TACs for the two- and the three-compartment model are shown for the three types of input functions, while the k_i remain unchanged. In figure 4.2 the dependence of the TAC shape on a small change in the rate constants k_i is pointed out.

3.2 Generation of Information regarding Curve Shape

The next aim was to examine the generated TACs for various mathematical properties. For this purpose, a PYTHON script has been written which reads in the TACs from the "TAC-simulation.csv" file mentioned above, calculates values for specific mathematical properties and exports the obtained information into a new csv-file, called "information.csv". In the next step of writing the algorithm there was an attempt to find possible correlations between the rate constants k_i and the obtained information about mathematical characteristics.

Below the approach to calculate specific mathematical properties via a PYTHON script is explained. These procedures apply for the one- and two-compartment as well as for the three-compartment model. Before examining various mathematical characteristics, it was necessary to read in external parameters which then can be used for further calculations. The advantage of using external parameters is that they simply can be changed in an external file. Otherwise, the operator would have to change them in every PYTHON script separately when necessary. Each external parameters used will be explained in the corresponding subsection.

3.2.1 Relative Slope and Intercept before a certain Time Point

An interesting characteristic of the TAC is its slope in the beginning phase. It is supposed to represent to which extend the TAC is linked to the input function. This concentration exchange is governed by the rate constants, thus the slope of the TAC can be seen as an important indicator. One would think of high constants k_1 and k_2 in the case of a relatively high slope value leading to a rapid exchange between the input function and the first compartment (see the models described in figure 1.2, 1.3 and 1.4). In contrast, low values for k_1 and k_2 are expected to cause a slow concentration exchange inducing a small slope.

For the calculation of the slope of the TACs, a linear regression via the Least-Squares method of the TAC data has been performed. From the obtained best fit line the slope and intercept could be read off. For this purpose not the whole dataset but only the data within a certain

time window has been used. This window was given as the time interval $[t_1, t_2]$. Since the aim was to calculate the relative slope and intercept in the beginning phase, one had to decide which time points to use for t_1 and t_2 .

The time t_2 defining the end of the time window was set as the *peak time* t_{peak} with t_{peak} being the time where the input function $B(t)$ reaches its maximum:

$$t_{\text{peak}} = \underset{t}{\operatorname{argmax}}\{B(t)\} \quad .$$

The other time point t_1 was defined as the start of the measurement, which is determined by the input function $B(t)$. The *measurement start time* t_{start} is given as the time where the input function exceeds a certain level defined by an external parameter called *threshold level* c_{th} . This is inevitable, since in real measurements the input function comprises of noise even before tracer injection. Hence, to disregard that noise before starting the measurement, a starting time is necessary. t_{start} is the minimum time where $B(t) \geq c_{\text{th}}B(t_{\text{peak}})$ holds, thus

$$t_{\text{start}} = \min\{t : B(t) \geq c_{\text{th}}B(t_{\text{peak}})\} \quad .$$

In other words, c_{th} defines a certain percentage value with respect to the maximum of the input function $B(t_{\text{peak}})$ which has to be exceeded in order to be designated as a measurement start. By this, the time window began at $t_1 = t_{\text{start}}$. In equations (16), (17) and (18) it is obvious that the *measurement start time* was given as $t_{\text{start}} = 11$ s.

With the two time points $t_1 = t_{\text{start}}$ and $t_2 = t_{\text{peak}}$ the slope and intercept of a linear best fit ($y = mx + b$) within the time window $[t_{\text{start}}, t_{\text{peak}}]$ for both the input function and the TAC could be computed. By performing such a linear best fit, one obtained the slopes m_{TAC}, m_B and the intercepts b_{TAC}, b_B for the TAC and the input function $B(t)$, respectively. Since the slope m_{TAC} strongly depends on the input function and hence on its slope m_B as well, the relative slope $m_{\text{rel}} = \frac{m_{\text{TAC}}}{m_B}$ has been calculated. This is also done for the relative intercept $b_{\text{rel}} = \frac{b_{\text{TAC}}}{b_B}$. In the left graphic in figure 3.2 the time window $[t_{\text{start}}, t_{\text{peak}}]$ and the linear best fit lines for the input function and the TAC are presented as dotted lines. Since the input function $B(t)$ is already linear in the considered time window, it matches with its linear best fit, thus it is not visible in figure 3.2.

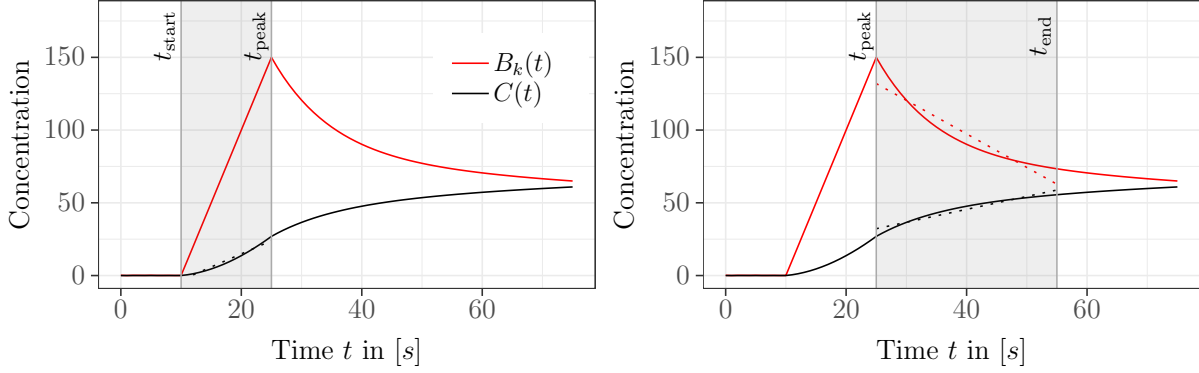


Figure 3.2: Relative slope and intercept before (left) and after (right) a certain time point. In the left graphic the time window is given as $[t_{\text{start}}, t_{\text{peak}}]$, while in the right one it is defined by $[t_{\text{peak}}, t_{\text{end}}]$. The linear best fit functions are shown as dotted lines for the input function $B(t)$ (red) and the TAC (black).

3.2.2 Relative Slope and Intercept after a certain Time Point

For calculating the relative slope and intercept of a linear best fit function after a specific time point the same procedure as described above has been performed. The only difference was the time window considered. For that purpose, the time points t_1 and t_2 of the time window had to be shifted. While t_1 now is set $t_1 = t_{\text{peak}}$, the determination of t_2 requires another time point, referred to as *window end time* t_{end} . To define the *window end time* t_{end} , both t_{start} and t_{peak} as well as the *slope range* r_s are necessary with r_s being an external parameter. The *window end time* thereby is given as

$$t_{\text{end}} = t_{\text{start}} + r_s(t_{\text{peak}} - t_{\text{start}}) \quad .$$

As one can see, r_s determines the width of the integral with respect to the time difference of the *measurement start time* t_{start} and the *peak time* t_{peak} . Within the time window $[t_{\text{peak}}, t_{\text{end}}]$ again a linear best fit has been performed for the TAC and the input function $B(t)$ yielding the slopes m_{TAC}, m_B and the intercepts b_{TAC}, b_B of the linear regression lines $y = mx + b$. For further analysis described in section 3.3, the relative slope $\frac{m_{\text{TAC}}}{m_B}$ and the relative intercept $\frac{b_{\text{TAC}}}{b_B}$ were calculated similarly to the relative slope and intercept before a certain time point.

In the right graphic in figure 3.2 one can see the shifted time window compared to the time window for the calculation before a certain time point. The dotted lines correspond to the linear best fit functions for both the input function and the TAC.

3.2.3 Relative Integral

The next mathematical characteristic considered was the integral I_{TAC} of the TAC in a certain time region. However, since the area under the TAC also depends on the input function $B(t)$, the relative integral $I_{\text{rel}} = \frac{I_{\text{TAC}}}{I_B}$ had to be calculated. To set the time interval $[t_{\text{start}}, t_{\text{end}}]$ for calculating the integral, t_{end} has to be defined again. With the external parameter *integral range* r_I the *window end time* t_{end} is given as

$$t_{\text{end}} = t_{\text{start}} + r_I(t_{\text{peak}} - t_{\text{start}}) \quad .$$

For continuous TACs and input functions one would calculate the relative integral by

$$I_{\text{rel}} = \frac{\int_{t_{\text{start}}}^{t_{\text{end}}} \text{TAC}(t) dt}{\int_{t_{\text{start}}}^{t_{\text{end}}} B(t) dt} \quad .$$

However, since the TACs generated in section 3.1 are not continuous but discrete instead, the integrals above are replaced by sums. Therefore, the relative integral above transforms to

$$I_{\text{rel}} = \frac{\sum_{t=t_{\text{start}}}^{t_{\text{end}}} \text{TAC}(t)}{\sum_{t=t_{\text{start}}}^{t_{\text{end}}} B(t)} \quad .$$

One would think of a high correlation between the relative integral and the rate constants k_i , especially k_1 and k_2 , since they are mainly responsible for the general in- and outflow of the compartments.

3.2.4 Curvature

Despite the slope, the intercept and the integral of the TAC, the curvature is an interesting characteristic as well. It determines the bending of the curve and therefore represents the interaction between the input function and the TAC. For calculating the curvature, an

approximative approach similar to the centered difference quotient for a derivative was used:

$$\begin{aligned} \left. \frac{\partial^2 f(t)}{\partial t^2} \right|_{t=t_0} &= \frac{\partial}{\partial t} \left(\lim_{\Delta t \rightarrow 0} \frac{f(t_0 + \Delta t) - f(t_0 - \Delta t)}{2\Delta t} \right) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{2\Delta t} \left(\lim_{\Delta s \rightarrow 0} \frac{f(t_0 + \Delta t + \Delta s) - f(t_0 + \Delta t - \Delta s)}{2\Delta s} - \right. \\ &\quad \left. \frac{f(t_0 - \Delta t + \Delta s) - f(t_0 - \Delta t - \Delta s)}{2\Delta s} \right) . \end{aligned} \quad (22)$$

This equation can be approximated numerically by small Δt and Δs . For simplification $\Delta t = \Delta s$ is assumed. Equation (22) thereby transforms to

$$\left. \frac{\partial^2 f(t)}{\partial t^2} \right|_{t=t_0} \approx \frac{1}{4\Delta t} (f(t_0 + 2\Delta t) - 2f(t_0) + f(t_0 - 2\Delta t)) . \quad (23)$$

Since a real TAC comprises of noise caused by the measurement itself, it is difficult to calculate a reasonable curvature within a small time frame. To disregard such noise a wider time window for the calculation had been used. For the TACs the curvature C in the PYTHON script is defined as

$$C = \text{TAC}(t_1) + \text{TAC}(t_0) - 2\text{TAC}\left(\frac{t_0 + t_1}{2}\right) , \quad (24)$$

where the factor $\frac{1}{4\Delta t}$ in equation (23) is constant for fixed values of t_0 and t_1 and therefore can be ignored, as it only leads to a proportional modification. Since one was interested in the TAC curvature in the beginning only, t_0 was set $t_0 = t_{\text{start}}$ according to the *measurement start time* explained in section 3.1. The other time point t_1 was defined by the *peak time* t_{peak} , the measurement start time t_{start} and the *curvature range* r_C given as external parameter. Again the *window end time* is defined by

$$t_{\text{end}} = t_{\text{start}} + r_C(t_{\text{peak}} - t_{\text{start}}) .$$

Hence, the term $\frac{t_0+t_1}{2}$ in equation (24) is the mean value t_{mean} of t_{start} and t_{end} ensuring equidistant time points. By this, the approximate curvature was calculated according to equation (24) with $t_0 = t_{\text{start}}$, $t_1 = t_{\text{end}}$ and $\frac{t_0+t_1}{2} = t_{\text{mean}}$. In figure 3.3 one can see the time window defined by the marginal and the intermediate time points. The calculation of the curvature according to equation (24) can be thought of as the mean value of the slope

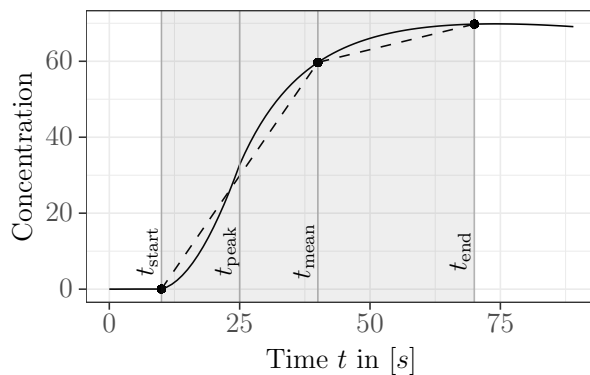


Figure 3.3: Curvature calculated for a certain time window $[t_{\text{start}}, t_{\text{end}}]$. The curvature within the time window is approximated by the difference of the slopes of two linear best fit functions defined by the corresponding time points. The linear regression functions are shown as dashed lines.

differences of the linear lines from t_{start} to t_{mean} and t_{mean} to t_{end} , which are shown as dashed lines.

3.2.5 1C/2C/3C and Trapped/Open

Another important characteristic for the TAC is its underlying model type. Therefore the PYTHON script read in the rate constants k_i of the individual TACs and determined the type of kinetic model. For the differentiation between a one-, a two- and a three-compartment model the script simply checked if all the additional rate constants appearing in the respective model (two- or three-compartment model) are equal to $k_i = 0$. Depending on that, the underlying model was known.

The rate constants k_i also allow to investigate if the model is open or trapped. In section 1.2.2 the difference between an open and a trapped model is described. A trapped model is given if it comprises of at least one compartment having an inflow but no outflow. This information about the underlying model type was important for further analysis described in section 3.3. Therefore the PYTHON script exported this information as well.

3.2.6 TAC Shape

Another characteristic was the general shape of the TAC. For this purpose, it has been investigated if the TAC comprises of a maximum in the beginning phase (hereafter referred to as max_1), if it shows a maximum at a later time (called max_2) and if it falls down at the

end of the measurement.

The main difference between a \max_1 and a \max_2 is their time point of appearance. While a \max_1 has to occur before a certain time point called *limit time* t_{limit} , which is defined by the external parameter *limit range* r_{limit} , a \max_2 needs to appear after t_{limit} . Therefore t_{limit} designates a time point distinguishing a \max_1 from a \max_2 . The *limit time* is defined by

$$t_{\text{limit}} = t_{\text{start}} + r_{\text{limit}}(t_{\text{peak}} - t_{\text{start}}) \quad (25)$$

with t_{start} being the *measurement start time* and t_{peak} being the *peak time*. Another important external parameter for the examination of these maximums is the *percentage step* p .

For a \max_1 first the maximum concentration within the time region $[t_{\text{start}}, t_{\text{limit}}]$ was determined. If the time point of the maximum lay in-between this time interval, the TAC was considered as having a maximum in the beginning phase. However, if the time point of the maximum matched the upper interval boundary t_{limit} , then in the next step the considered time interval got shortened by a factor p and the search for a maximum started again. The shortening of the time window was done by reducing the upper boundary t_{limit} while leaving t_{start} unchanged. In this case the procedure started all over again with the new time interval. By that one avoids a wrong consideration of a maximum which in fact was only the maximal value on the upper boundary. In figure 3.4 one can see the strategy of this procedure. For better understanding a pseudo code is shown below:

```

# reading in external parameters/already calculated values:
1:  $p :=$  percentage step
2:  $r_{\text{limit}} :=$  limit range
3:  $t_{\text{start}} :=$  measurement start time,  $t_{\text{peak}} :=$  peak time
# procedure start:
4:  $i := 0$ 
5: while  $i \geq 0$ :
6:    $t_{\text{limit}} := t_{\text{start}} + r_{\text{limit}}(t_{\text{peak}} - t_{\text{start}}) \cdot (1 - p \cdot i)$ 
7:   if  $t_{\text{limit}} \leq t_{\text{start}}$ : return "max1 does not exist"
8:    $t_{\text{max}} :=$  time of maximum concentration in interval  $[t_{\text{start}}, t_{\text{limit}}]$ 
9:   if  $t_{\text{max}} \neq t_{\text{limit}}$ : return "max1 exists"
10:  else:  $i = i + 1$ 

```

In the code above the lines 1-3 refer to the input of both the external parameters, namely p and r_{limit} as well as the previously calculated values for t_{start} and t_{peak} . The procedure to determine whether the TAC comprises of a max_1 or not starts at line 4. With each round in the while loop the *limit time* t_{limit} gets decreased due to the term $(1 - p \cdot i)$ in line 6. The while loop ends if either a maximum is found in-between the interval (line 9) or t_{limit} is lower than t_{start} due to the increasing number of i leading to an unfeasible interval (line 7). Thus, the output of the procedure can be "max₁ exists" or "max₁ does not exist".

Comparing the pseudo code and the plots in figure 3.4, one can see that the left plot refers to the first run ($i = 0$) of the procedure. Since the maximum is found at the upper interval boundary, t_{limit} gets reduced to t'_{limit} and the while loop starts again. In the next run ($i = 1$) again the maximum is located at the upper boundary leading to a further reduction of t'_{limit} to t''_{limit} . In the third run ($i = 2$) the new maximum is found in between the interval, forcing the while loop to stop and return "max₁ exists".

The existence of a maximum max_2 after a certain time point has been determined using a similar approach as for max_1 . For a max_2 , however, an additional external parameter of importance was the *concentration difference* ε . The time interval considered now was given by $[t_{\text{limit}}, t_{\text{end}}]$ with t_{end} being the *measurement end time* (in contrast to the previous sections, where t_{end} designated the *window end time* of a certain time interval).

Like for the max_1 the procedure for finding a max_2 starts with the search for the maximum concentration within the time interval $[t_{\text{limit}}, t_{\text{end}}]$. In contrast to the max_1 , for a max_2 three conditions have to be fulfilled: the time point of the maximum must not be located at the lower interval boundary t_{limit} , nor is it allowed to be at the upper boundary t_{end} and the concentration value of the maximum has to exceed the concentrations given at the boundaries $t = t_{\text{limit}}$ and $t = t_{\text{end}}$ by at least $\varepsilon \cdot B(t_{\text{peak}})$, where the *concentration difference* ε is defined on a percentage basis. The last condition guarantees the existence of a significant maximum rather than a maximum appearing due to noise only.

The PYTHON code for the max_2 is similar to the code for finding a max_1 with an additional line including the third condition, namely

- $\text{TAC}(t_{\text{max}}) - \text{TAC}(t_{\text{limit}}) > \varepsilon \cdot B(t_{\text{peak}})$ and
- $\text{TAC}(t_{\text{max}}) - \text{TAC}(t_{\text{end}}) > \varepsilon \cdot B(t_{\text{peak}})$.

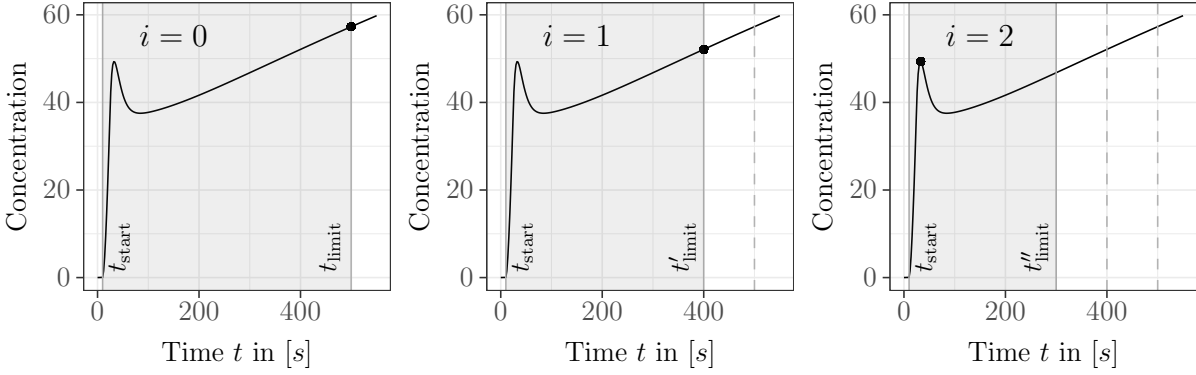


Figure 3.4: Examining the TAC for the existence of a maximum in the beginning phase (\max_1). Shown are the first three runs ($i = 0$ to $i = 2$) according to the procedure mentioned in the text. The calculated maximum in each run is marked by a dot. One can see that with increasing number i of runs the interval gets shortened until the real maximum is found in between the considered time interval (right plot).

Again, the considered time region $[t_{\text{limit}}, t_{\text{end}}]$ got shortened for each run: if the maximum was found at the lower boundary, t_{limit} was shifted towards t_{end} . In contrast, if the maximum was located at the upper boundary, t_{end} has been reduced towards t_{limit} . In both events the considered time window got shortened by the factor p and the procedure started again.

For examining the TAC for a fall off at the end, a linear regression has been performed. For this, the last i data points (determined by the external parameter called *fall off range* $r_{\text{fall off}}$) of the TAC were utilized. Hence, the considered time window was given by $[t_{\text{fall off start}}, t_{\text{end}}]$, where t_{end} is the *measurement end time* and $t_{\text{fall off start}}$ is defined as

$$t_{\text{fall off start}} = t_{\text{end}} - r_{\text{fall off}}(t_{\text{end}} - t_{\text{start}}) \quad .$$

If the slope of the linear regression line was negative, the TAC is designated to comprise of a fall off at the end.

By examining whether or not the TAC comprises of a \max_1 , a \max_2 or a fall off at the end, one can classify the TAC shape. In figure 3.5 all possible combinations of these characteristics are shown. As one can see, there are only five possible classes considered in this master thesis.

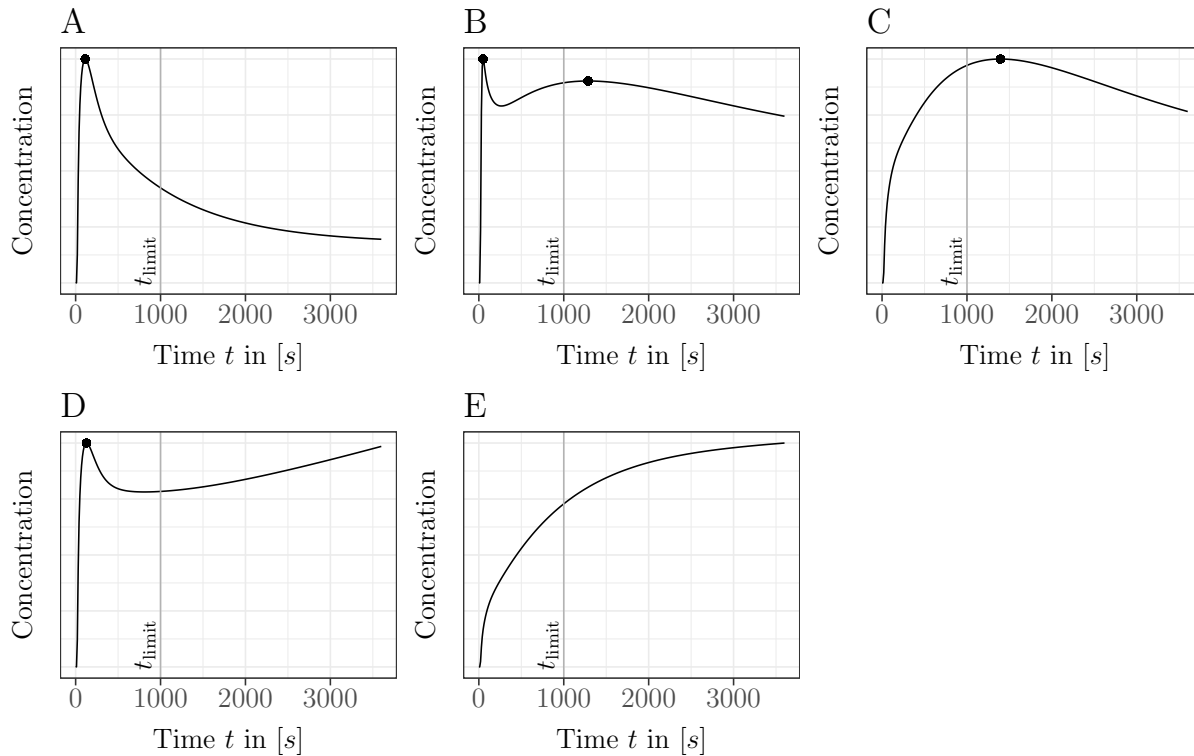


Figure 3.5: All five possible combinations of a \max_1 , a \max_2 and a fall off at the end. In figure A a \max_1 exists with a fall off at the end, figure B comprises of a \max_1 , a \max_2 and a fall off as well. Figure C shows both a \max_2 and a fall off. In figure D only a \max_1 is visible and in figure E no \max_1 , no \max_2 nor a fall off at the end is present. In all plots the *limit time* t_{limit} distinguishing a \max_1 from a \max_2 and the maximums, if present, are shown.

3.2.7 Rate Constants k_i and underlying Input Function

In addition to the underlying model type the rate constants k_i and the type of input function $B(t)$, by which the TACs had been generated, were important. Since in the next step described in the following section 3.3 the evaluation of all these mathematical properties and the rate constants took place, the rate constants had to be written in the exported csv-file "information.csv" as well. For that, the PYTHON script simply took the rate constants from the "TAC-simulations.csv" file (see section 3.1.2) and transferred them to the "information.csv" file.

The type of the input function (pointed, blunt or kinky) has been extracted as well. By this, further analysis could concentrate on individual input function types solely. For the determination of the input function type the exponential terms defining the input function were examined. All input functions described in section 3.1.1 comprise either of a mono-exponential or a tri-exponential decreasing term of the form ae^{-bt} or $ae^{-bt} + ce^{-dt} + fe^{-gt}$, respectively. The PYTHON script therefore checked, if the input function consists of a tri-exponential term, which corresponds to a kinky input function $B_k(t)$. Otherwise the power b in the first term was considered. Depending on b either a pointed input function $B_p(t)$ ($b = 0.05$) or a blunt input function $B_b(t)$ ($b = 0.01$) was present.

In summary, all the calculated mathematical information (relative slope and intercept before and after a certain time point, relative integral, curvature) as well as the information about the underlying model type (1C, 2C or 3C, open or trapped, TAC shape, rate constants k_i , input function type $B(t)$) have been written into an external "information.csv" file. In the next quest this csv-file has been used for further analysis described in the following section.

In section 4.2 an overview of the calculated values for all aforementioned mathematical properties is given. Table 4.1 gives a rough insight into the "information.csv" file generated by the PYTHON script.

3.3 Evaluation of obtained Data

The next goal was the analysis of the obtained data described in the previous section. The aim was to find correlations between the rate constants k_i and the specific mathematical properties such as the relative slope and intercept before and after a certain time point, the relative integral, the curvature and the TAC shape. Another method performed was the analysis via a Student's t -test to determine whether the obtained mathematical values differ

for different kinds of input functions $B(t)$ and model types.

Hereafter both the correlation coefficient and the Student's t -test with their application to the obtained data are closer examined.

3.3.1 Correlation Coefficient

The correlation coefficient has been used to determine if there is a relationship between the rate constants k_i of the underlying model type and the obtained mathematical data.

In general, the correlation coefficient r_{xy} for two datasets x_i and y_i , $i = 1, \dots, n$ is defined as

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad . \quad (26)$$

Hereby, \bar{x} and \bar{y} present the arithmetic mean of the set of values x_i and y_i , respectively. For an according to amount high correlation coefficient ($|r_{xy}|$ close to 1) a strong correlation between the two datasets is valid. In contrast, for $r_{xy} \approx 0$ no clear correlation can be pointed out leading to a negligible linear relation between the individual datasets.

Since one is interested in the correlation between the rate constants k_i and the various mathematical properties of the simulated TAC, each property had to set into relation with each rate constant. In section 4.3 the individual correlations are listed and discussed. The computation of all these correlation coefficients provided a good estimation whether a specific rate constant k_i has an identifiable influence on a particular mathematical property.

If there happened to be a good correlation between a certain rate constant k_i and a specific mathematical value v_i , there had to be a good linear fit according to $v_i = m_i k_i + b_i$. Therefore, one was able to estimate the rate constant by back calculation of the mathematical value according to

$$k_i = \frac{v_i - b_i}{m_i} \quad .$$

In table 4.2 all correlation coefficients r_{xy} are listed with their significance coded by color.

3.3.2 Student's t -Test

Besides the interest in significant correlations between the mathematical properties and the rate constants, the existence of possible impacts of the model type and input function on the TAC shape were of high importance as well. For this, all the information written in the "information.csv" file has been subdivided into different classes.

The dataset could be sorted by different characteristics, such as the type of input function $B(t)$ (pointed, blunt or kinky), the number of compartments $\#C$ (one-, two- or three-compartment model), the presence of an open or trapped model as well as the general shape of the TAC (\max_1 , \max_2 , fall off). The various classifications considered are listed in section 4.3.

After sorting the dataset, one was interested in whether there is a connection between two different classes or not. For example, is there a remarkable difference in the relative integral I_{rel} when having a pointed or a kinky input function? If so, then the input function is supposed to have a considerable influence on the relative integral. However, if there is no significant difference in the data between the two considered classes, one would expect the input functions to have no impact on the relative integral.

Mathematically speaking, the aim was to compare two samples $\{x_1, \dots, x_{n_x}\}$ and $\{y_1, \dots, y_{n_y}\}$ with n_x and n_y being the size of the respective samples. For the example mentioned above, x_i would be the data of the relative integral I_{rel} for a pointed input function and y_i that one for a kinky input function. The comparison was done by looking at the sample means \bar{x} and \bar{y} . According to the Central Limit theorem, the sample means can be thought of as being normally distributed due to the large sample sizes n_x and n_y , even though the data entries themselves are not [43]. Since both the sample means and variances had to be estimated by the sample itself, the datasets were no longer normally distributed but t -distributed instead. By using the t -test one could determine whether the two datasets are significantly different from each other. Since the datasets presumably were unpaired with different variances, the Welch's t -test had to be used instead of a Student's t -test [44].

The null hypothesis \mathcal{H}_0 and the alternative hypothesis \mathcal{H}_1 hereby were given as

$$\begin{aligned}\mathcal{H}_0 &: \bar{x} = \bar{y} \\ \mathcal{H}_1 &: \bar{x} \neq \bar{y} \quad ,\end{aligned}$$

therefore a double-ended t -test was necessary. By this, one tested the null hypothesis whether the two sample means can be considered equal against the alternative hypothesis. The t -statistic for the Welch's t -test is given by

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} ,$$

where s_x^2 and s_y^2 are the respective sample variances. The degrees of freedom k have to be approximated using the Welch-Satterthwaite equation (see [45, 46]).

With the test statistic T and the degrees of freedom k the p -value is given as

$$p = 2(1 - F_{t,k}(|T|)) \quad , \quad (27)$$

where $F_{t,k}$ describes the cumulative distribution function of a t -distribution with k degrees of freedom. A high p -value would state that the two datasets can be thought of as being equal, therefore the classification would have no significant impact on the data. In contrast, if a low p -value is present there seems to be a not negligible difference between the two datasets. In this case the classification is supposed to have an influence on the individual data.

In table 4.4 the results of the t -test with their p -values are shown. The entries are highlighted by color according to their significance.

3.4 Programming the Algorithm

The aim of this master thesis was to implement an algorithm which reads in an external TAC, analyzes it by different methods already mentioned above and outputs feasible values for the rate constants k_i . These values can then be used as initial values for an optimization tool such as the Nelder-Mead simplex method (see section 2.3). Since the reliability of these optimization programs strongly depends on the initial values, it is necessary to feed them with reasonable values appropriate for the specific model type.

In the previous section first the attempt was made to find reasonable correlations between the rate constants and the various mathematical properties. Secondly, the equality of these properties between any two classifications was tested via a Student's t -test (or more precisely a Welch's t -test). As one can see in section 4.3, the presence of reasonable correlations between the rate constants k_i and the various mathematical properties only applied for the first

two rate constants k_1 and k_2 . For the other rate constants there seemed to be no significant relation. The analysis via a t -test pointed out that there is no remarkable connection among the various mathematical properties between different classifications. Thus, the mathematical properties were assumed to not only depend on the rate constants k_i , but on the number of compartments $\#C$, the input function $B(t)$, the existence of an open or trapped model and the TAC shape as well. Therefore, the rate constants had to be estimated for each model type individually.

Since for k_1 and k_2 good results of the correlation coefficients have been obtained, these rate constants were analyzed separately from the other ones. This led to a subdivision of the algorithm's implementation into two subtasks, namely the determination of feasible intervals for all rate constants for one thing and secondly the estimation of the rate constants k_1 and k_2 . The first task was done by the determination of confidence intervals for each rate constant, since for calculating an exact value there were too many uncertainties in the statistical methods. The latter task has been performed by taking a look at the correlation coefficients and subsequently estimating k_1 and k_2 via back-calculation of the corresponding mathematical properties.

For a three-compartment model with twelve possible rate constants the algorithm is limited to the rate constants k_1, k_2, k_3, k_4, k_7 and k_8 , which comply with the rate constants k_1 to k_6 of a two-compartment model (see figure 1.4). All other rate constants in a three-compartment model were not considered since the analysis of the data via statistical methods turned out to be less informative than expected. In contrast to a three-compartment model, for the one-compartment model there is no limitation in the algorithm since the model only comprises of two rate constants anyway.

3.4.1 General Structure of the Algorithm

As mentioned in section 4.3, the outcome of the t -test showed that there are relevant differences between each classification. Hence, for the algorithm the data from the "information.csv" file was subdivided into different subclasses. Since this classification had to be done for the t -test anyway, the data used for calculating the correlation coefficients was classified as well to increase the algorithm's reliability.

In conclusion the data from the "information.csv" file had to be subdivided into the type of input function $B(t)$, the number of compartments $\#C$ and the existence of either an open or a trapped model. For the TAC shape the existence or absence of a \max_1 , a \max_2 and a fall

off at the end had been considered separately. For a one-compartment model the distinction between an open and a trapped model was not performed due to its simple structure. This led to the following classes considered in the algorithm for each type of input function $B(t)$:

- 1C
- 2C open
- 2C trapped
- 3C open
- 3C trapped

So the first goal of the algorithm is to analyze the input function and figure out if it rather is a pointed, a blunt or a kinky input function. By knowing that, the algorithm then analyzes the recorded TAC via the PYTHON script for various mathematical properties necessary for both the correlation coefficient and the TAC shape. With this information it determines feasible intervals for the rate constants on the one hand and on the other hand finds values for k_1 and k_2 via back-calculation of the already computed correlation coefficients. This is done for each compartment model listed above.

Since the type of input function is known by then (see section 3.4.2), the algorithm performs all the calculations on basis of the specific input function. Hence, the output of the algorithm has the following form:

Approximated Input Function Type: ____ (pointed, blunt, kinky)
 $a = _ , b = _ , c = _$
 1C:
 $k_1 = _ , k_2 = _$
 2C open:
 $k_1 = _ , k_2 = _ , k_3 : [_ , _] , k_4 : [_ , _] , k_5 : [_ , _] , k_6 : [_ , _]$
 2C trapped:
 $k_1 = _ , k_2 = _ , k_3 : [_ , _] , k_5 : [_ , _]$
 3C open:
 $k_1 = _ , k_2 = _ , k_3 : [_ , _] , k_4 : [_ , _] , k_7 : [_ , _] , k_8 : [_ , _]$
 3C trapped:
 $k_1 = _ , k_2 = _ , k_3 : [_ , _] , k_4 : [_ , _] , k_7 : [_ , _] , k_8 : [_ , _]$

The parameters a, b and c are the constants of the approximated input function described in the following subsection. By this, the input function can be determined. With respect to this input function the algorithm provides values and feasible intervals for the rate constants for all model types. Due to the lack of information about the underlying model type, the algorithm does not confine itself to determine the model type but provides the values and feasible intervals for all model types instead. Therefore, the operator has to decide which model type could fit best for the specific problem.

3.4.2 Determination of the Input Function and the TAC shape

First of all the algorithm needs to determine the type of the measured input function. Since neither of the equations (16), (17) and (18) will exactly describe the measured input function, it is approximated by a shifted mono-exponential function $B_{\text{app}}(t)$ following

$$B_{\text{app}}(t) = ae^{-bt} + c \quad . \quad (28)$$

This is done via the least squares method yielding values for the constants a, b and c . These values are displayed by the algorithm (see the boxed sketch above). To distinguish between a pointed, a blunt and a kinky input function, one needs to consider these parameters. As both the parameters a and c depend on the injected tracer concentrations, their absolute values themselves are not representative. Therefore the relative value $\frac{c}{a}$ is considered instead.

For a pointed and blunt input function (see equations (16) and (17)) with $a_p = 143, b_p = 0.05, c_p = 7$ and $a_b = 143, b_b = 0.01, c_b = 7$, this ratio is given as $\frac{c_p}{a_p} = \frac{c_b}{a_b} = 49.0 \cdot 10^{-3}$. For a kinky input function following equation (18) the best mono-exponential fit including a shift yields the constants $a_k = 58.13, b_k = 2.179 \cdot 10^{-3}$ and $c_k = 10.45$ leading to a ratio of $\frac{c_k}{a_k} = 179.7 \cdot 10^{-3}$. Thus, to discriminate between a kinky input function and the other two types, a threshold of $\frac{c}{a} = 100 \cdot 10^{-3}$ is used. The threshold for the power b differentiating between a pointed and a blunt input function is set $b = 3 \cdot 10^{-2}$.

So the measured input function is approximated by $B_{\text{app}} = ae^{-bt} + c$ providing the parameters a, b and c . For $\frac{c}{a} \geq 100 \cdot 10^{-3}$, the input function is thought of as being a kinky input function. In the case of $\frac{c}{a} < 100 \cdot 10^{-3}$ the input function could either be pointed or blunt. Therefore, in the next step the power b is considered. For $b \geq 3 \cdot 10^{-2}$ a pointed input function is assumed and for $b < 3 \cdot 10^{-2}$ the input function is supposed to be blunt. The type of the approximated input function (pointed, blunt or kinky) is displayed in the algorithm as well.

The TAC shape can easily be determined by means of the PYTHON script. Thereby it can be figured out whether or not the TAC comprises of a \max_1 , a \max_2 and a fall off at the end. Knowing that, the TAC can be categorized as one of the five considered TAC shapes (see figure 3.5).

By knowing both the input function best describing the measured input function and the TAC shape, the rate constants can now be estimated for each model type (1C, 2C open, 2C trapped, 3C open, 3C trapped).

3.4.3 Determination of Feasible Intervals for the Rate Constants

To obtain confidence intervals for the rate constants, besides the input function and the TAC shape the model type is considered as well. Hence, the data from the "information.csv" file had to be classified by all these categories. Subsequently a two-sided 95% confidence interval has been calculated for all combinations of input function, model type and TAC shape by means of a Student's t -statistic. The t -statistic was necessary due to the unknown variance which therefore had to be estimated by the sample itself leading to a Student's t -distribution. To determine the 95% confidence interval, the sample mean \bar{x} , the sample variance s and the sample size n as well as the level of significance $\alpha = 0.05$ were necessary. By calculating the confidence intervals for all model types (1C, 2C open, 2C trapped, 3C open, 3C trapped), all types of input functions (sharp, blunt, kinky) and the various TAC shapes (existence or absence of a \max_1 , a \max_2 and a fall off at the end), one obtained feasible ranges for the rate constants k_1 and k_2 in a one-compartment model, k_1 to k_6 in a two-compartment model and k_1, k_2, k_3, k_4, k_7 and k_8 in a three-compartment model. Since in the computation of confidence intervals it is possible for the lower boundary to be negative, in this case they were set 0 instead.

For example, for the first TAC shape (plot A in figure 3.5) comprising of a \max_1 and a fall off at the end but without any \max_2 , one obtains three confidence intervals for each rate constant and each input function according to

$$\begin{aligned} I_i^{(\max_1)} &= [a_i^{(\max_1)}, b_i^{(\max_1)}] \\ I_i^{(\text{no } \max_2)} &= [a_i^{(\text{no } \max_2)}, b_i^{(\text{no } \max_2)}] \\ I_i^{(\text{fall off})} &= [a_i^{(\text{fall off})}, b_i^{(\text{fall off})}] \quad . \end{aligned}$$

Hereby the subscript i stands for the rate constant k_i considered. To obtain one single interval of feasibility, the intersection of the three confidence intervals is taken. This was performed for each rate constant k_i , for each possible TAC shape (see figure 3.5), for all model types and for all input functions as well. All the feasible intervals are shown in tables 4.5, 4.6 and 4.7.

For the rate constants k_3 to k_6 in a two-compartment model, respectively k_3, k_4, k_7 and k_8 in a three-compartment model, the algorithm outputs the corresponding feasible intervals. In addition, for the first two rate constants a back-calculation using the correlation coefficients is performed.

3.4.4 Calculating Values for k_1 and k_2

Apart from determining feasible intervals, for the rate constants k_1 and k_2 explicit values can be calculated. Table 4.2 states that there are high correlations between the first two rate constants and specific mathematical properties. Hence, a look at the corresponding linear regression lines can be taken. Since for both k_1 and k_2 there is no solely but a few mathematical properties having a good correlation, these rate constants are determined with their best two correlating mathematical properties via back-calculation of the regression lines.

With the linear regression lines for the two mathematical properties showing the best correlation coefficient, one obtains two equations for both k_1 and k_2 according to

$$\begin{aligned} v_1^{(1)} &= m_1^{(1)}k_1^{(1)} + b_1^{(1)} & , & & v_1^{(2)} &= m_1^{(2)}k_1^{(2)} + b_1^{(2)} \\ v_2^{(1)} &= m_2^{(1)}k_2^{(1)} + b_2^{(1)} & , & & v_2^{(2)} &= m_2^{(2)}k_2^{(2)} + b_2^{(2)} \end{aligned} .$$

Hereby $m_1^{(i)}$ and $b_1^{(i)}$ are the slopes and intercepts of the two linear regression lines applying for the first rate constant k_1 . The superscript indicates whether it belongs to the first mathematical property $v_1^{(1)}$ or the second one given as $v_1^{(2)}$. The same holds for the second rate constant k_2 . These equations can be rearranged to define the rate constants $k_1^{(1)}, k_1^{(2)}$ and $k_2^{(1)}, k_2^{(2)}$. Since the equations yield two values for each constant k_i , the arithmetic mean is taken according to

$$k_i = \frac{k_i^{(1)} + k_i^{(2)}}{2} = \frac{1}{2} \left(\frac{v_i^{(1)} - b_i^{(1)}}{m_i^{(1)}} + \frac{v_i^{(2)} - b_i^{(2)}}{m_i^{(2)}} \right) . \quad (29)$$

The arithmetic mean provides a good estimation for k_1 and k_2 .

The slopes $m_i^{(j)}$ and intercepts $b_i^{(j)}$ depend on the TAC shape, the model type as well as on the input function. Since both the TAC shape and the input function can be determined by the PYTHON script (see section 3.4.2), the values for k_1 and k_2 have to be calculated for each model type (1C, 2C open, 2C trapped, 3C open, 3C trapped) separately. The task of the algorithm therefore is to analyze the TAC for the various mathematical properties $v_i^{(j)}$ and use these values together with the corresponding slopes and intercepts to estimate the rate constants according to equation (29).

By computing the rate constants via back-calculation of the mathematical values, it cannot be guaranteed that this method results in values for k_1 and k_2 which are located in their feasible intervals $I_1 = [a_1, b_1]$ and $I_2 = [a_2, b_2]$ calculated before. If the rate constants k_i lie below their lower boundaries a_i , they are simply set $k_i = a_i$. In the case of too large constants $k_i > b_i$, they are set $k_i = b_i$. By that, on the one hand the rate constants k_1 and k_2 are calculated by means of a linear regression line resulting in a good estimation. On the other hand it is ensured that they lie within their feasible intervals.

In figure 4.4 the linear regression lines for both k_1 and k_2 for an open two-compartment model supplied by a kinky input function are shown.

3.5 Testing the Algorithm

As the implementation of the algorithm was done on basis of perfect input functions on the one hand and ideal TACs without any measurement uncertainties on the other hand, after programming the algorithm it had to be tested for its reliability. Therefore the algorithm had to be applied on real measurement curves.

There are some remarkable differences between real and simulated input functions and TACs: for one thing, the simulated input functions described in section 3.1.1 comprise of a perfect mono- or tri-exponential term, respectively. They also show a constant beginning term, namely $B(t) = 0$. In contrast to these simulated input functions an input function appearing in real measurements comprises of not negligible fluctuations. These are carried forward to the individual compartments leading to a noise in the TAC as well. Moreover, one would expect the approximation via a mono- or tri-exponential function including an shift (see equation (28)) to be rather poor, since real input functions show a noteworthy difference from the simulated ones.

Another important factor is the TAC analyzed by the algorithm: while the generated TACs

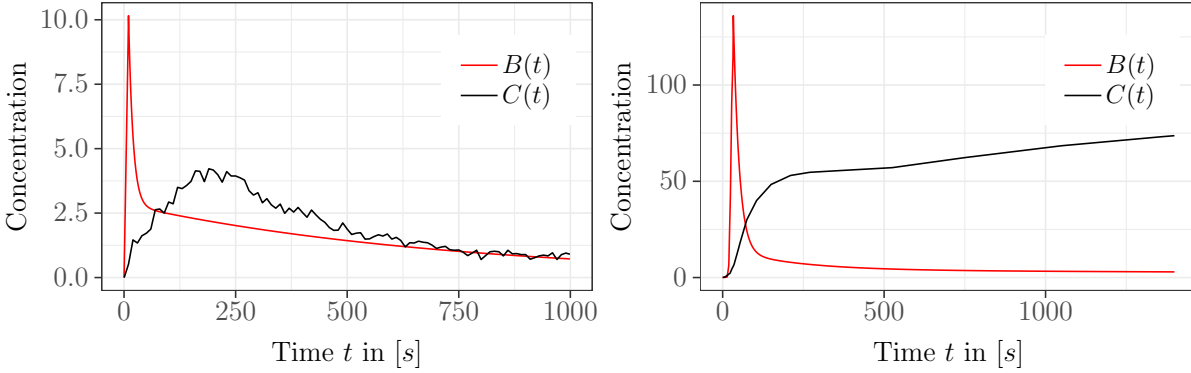


Figure 3.6: Real TACs and their input functions: the left plot shows a TAC of a right kidney with its input function coming from a heart. As tracer $^{99\text{m}}\text{Tc-MAG3}$ is used. The kink within the input function is clearly visible. The right graphic uses the tracer $^{18}\text{F-PSMA}$ for imaging. The input function originates from the iliac artery and the TAC represents the concentration within a prostate tumor.

described in section 3.1 are supplied by an input function defined on a discrete time line, real TACs are continuously fed by their input functions. In addition, due to measurement uncertainties real TACs comprise of noise which affects the information calculated by the algorithm and therefore possibly alters the outcome for the rate constants k_i tremendously.

In figure 3.6 two examples of real TACs with their input functions are shown. These two recordings are completely different regarding their input function, the tracer used as well as the observed VOI for measuring the TAC. As one can see in the left graphic, the TAC did not undergo a smoothing procedure and therefore comprises of large fluctuations. In contrast, the right plot shows a smoothed TAC. Furthermore the input function in the left plot comprises of a pronounced kink and hence is expected to be classified as kinky input function. The graphic on the right-hand side shows an input function, which one would refer to as pointed input function.

4 Results

This chapter follows up with the results of each subtask explained in the previous chapter. Since the specific overall progress of this master thesis was not fixed in the beginning, the outcome of each task determined the next approaches to be performed and therefore had an influence on the course of programming the algorithm.

4.1 Simulation of Time Activity Curves

For the simulation of TACs, first it was necessary to generate different input functions which reasonably cover all sorts of input functions considered in this master thesis. In reality, there are many more possibilities for injecting a tracer, for example a constant injection over time. Nevertheless, the input functions observed in nuclear medicine can be classified roughly into three types described in section 3.1.1.

To generate the TACs, a stepwise solution according to equation (19), (20) and (21) was performed. By varying the rate constants k_i around 400 TACs have been generated for each kind of input function. So in total approximately 1200 TACs were simulated which subsequently have been used for the examination of various mathematical properties.

In figure 4.1 one can see the impact of the input functions on the shape of the TACs. In addition to the TACs and the input functions the tracer concentrations in the individual compartments are shown as well. By this, a better understanding of the overall process is provided. For the pointed input function (upper plots), after the peak there is a rapid fall down which leads to a low tracer inflow into the various compartments for later time points t . In this case the TAC shows a peak almost simultaneously to the peak of the input function, whereas after the peak no significant events can be pointed out. The blunt input function (middle plots) comprises of a broader peak which in consequence leads to a broader peak in the TAC as well. The tracer inflow into the compartments hereby is more pronounced than for a pointed input function. In contrast to the pointed and the blunt input function which are quite similar regarding their shape, the kinky input function (lower plots) has a significant impact on the TAC even for later time points t . This is due to the fact that the kinky input function has a comparatively slow descending phase. Thereby the input function continuously supplies the individual compartments with a relatively high tracer amount, by which the shape of the TACs can be altered tremendously. All the properties described above can be seen in figure 4.1.

Figure 4.2 shows the dependence of the TAC shape on the choice of the rate constants k_i for a kinky input function in a two-compartment model. This is done by slightly changing the rate constants k_1 and k_6 while leaving the other constants k_2 to k_5 unchanged. While k_1 is responsible for the inflow into the first compartment, k_6 determines the outflow of the second compartment (see figure 1.3). Therefore a large k_1 in combination with a low k_6 leads to a high tracer accumulation in the compartments and hence a high maximum of the TAC whereas a low k_1 and a high k_6 induce a relatively small maximum. Figure 4.2 comprises of four individual plots with different k_1 and k_6 . Upper left plot: $k_1 = 1, k_6 = 0.1$, upper right: $k_1 = 0.3, k_6 = 0.1$, lower left: $k_1 = 1, k_6 = 0.3$ and lower right: $k_1 = 0.3, k_6 = 0.3$.

It can be pointed out that higher rate constants k_i lead to a TAC shape which roughly follows the shape of the input function $B(t)$. This is shown in figure 4.3. For high rate constants the concentration exchange between the compartments and the input function as well as among the compartments themselves is rather quick. Hence the TAC more or less resembles the shape of the input function. For low rate constant values, however, the exchange is considerably slower causing a temporal delay and by that leading to a completely different TAC shape.

As it is apparent in figures 4.1 and 4.2 both the input function and the rate constants k_i have a large effect on the shape of the TACs. However, since one is interested in the impact of the rate constants k_i only, it is necessary to abstract from the influence of the input functions. Therefore TACs have been generated for each kind of input function in a one-, a two- and a three-compartment model.

By taking a look at the TACs generated for a specific input function, it was possible to investigate the dependence of the TAC shape on the k_i solely. For this examination a total of around 1200 TACs has been simulated.

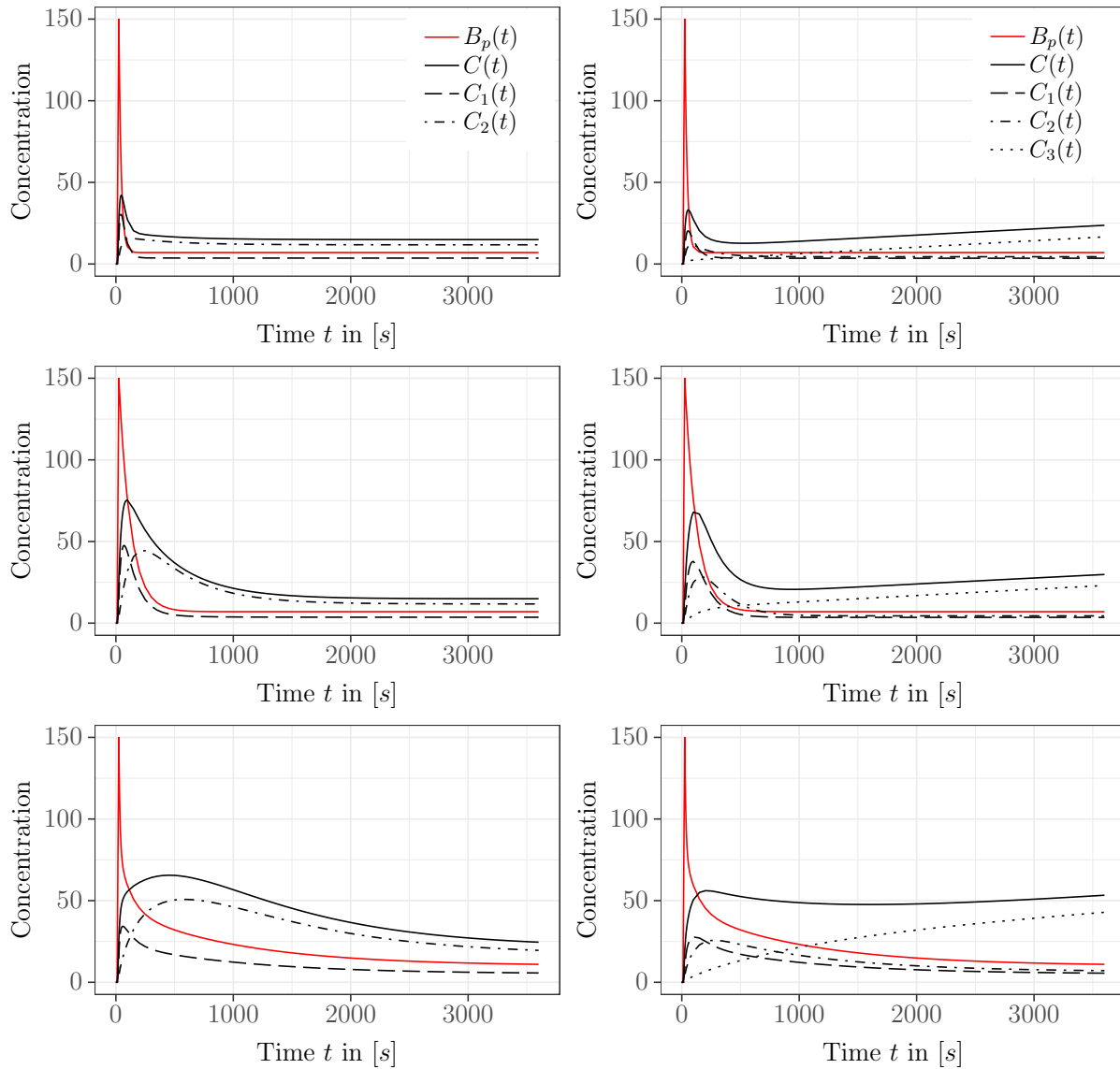


Figure 4.1: Generated TACs via EXCEL: On the left-hand side the TACs for a two-compartment model and on the right-hand side these for a three-compartment model are shown. From up to down the TACs are simulated for a pointed, a blunt and a kinky input function, respectively.

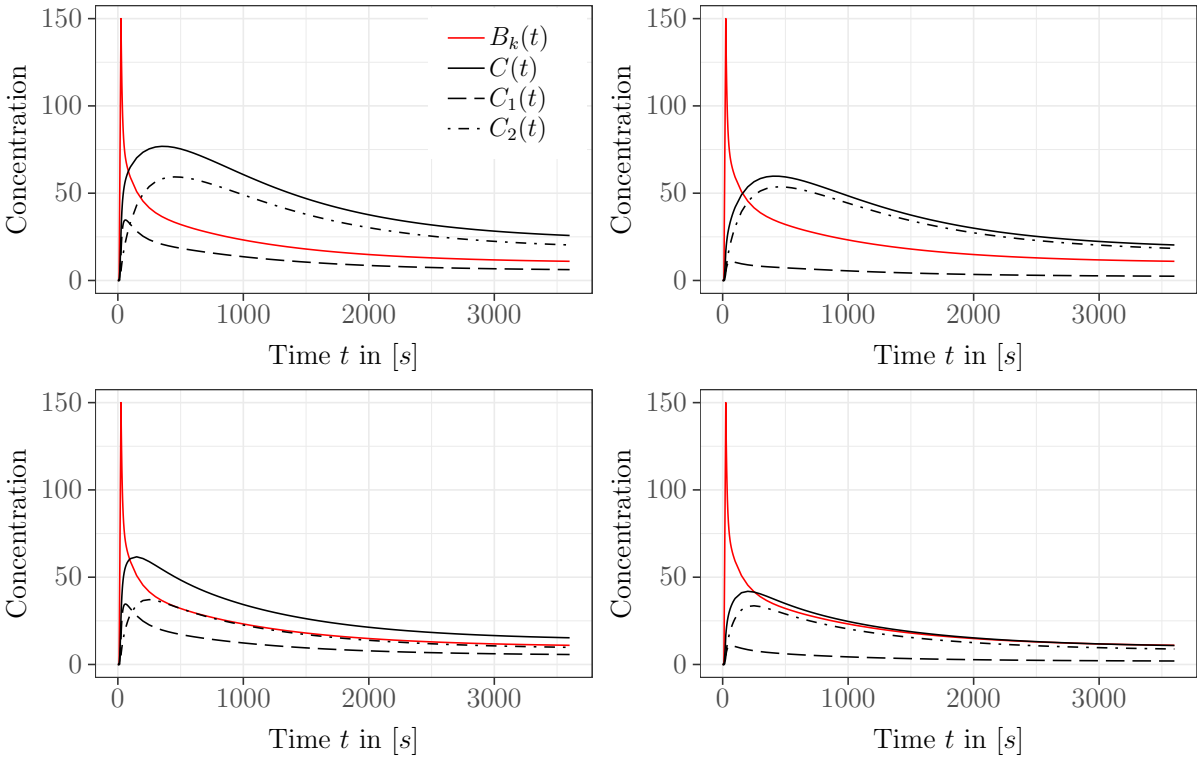


Figure 4.2: The dependence of the curve shape on the rate constants k_i . Upper left: $k_1 = 1, k_6 = 0.1$, upper right: $k_1 = 0.3, k_6 = 0.1$, lower left: $k_1 = 1, k_6 = 0.3$, lower right: $k_1 = 0.3, k_6 = 0.3$. The other rate constants are set $k_2 = 2, k_3 = 0.1, k_4 = 0.1, k_5 = 0.3$ for all graphics above. For this simulation a two-compartment model supplied by a kinky input function has been utilized.

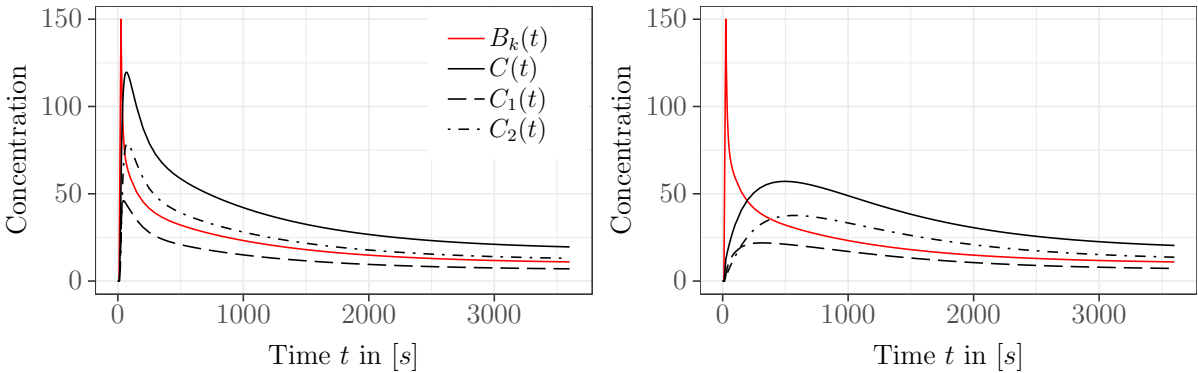


Figure 4.3: The impact of generally high and low rate constants k_i . In the left plot $k_1 = 2, k_2 = 1, k_3 = 4, k_4 = 1, k_5 = 1, k_6 = 2$ holds, whereas in the right plot all rate constants are divided by the factor 10. One can see that for high values of k_i the TAC rather follows the input function than for low values.

4.2 Generation of Information regarding Curve Shape

For the mathematical analysis of the TAC shape a PYTHON script has been written in order to read in the TAC data from the "TAC-simulation.csv" file. This data was then investigated for mathematical properties such as the relative slope and intercept of a linear best fit function before and after a certain time point, the relative integral, the curvature as well as the overall TAC shape described in section 3.2. Subsequently the information obtained together with the parameters of the underlying model has been written into a new csv-file named "information.csv". In table 4.1 a small selection of this csv-file is shown. Apart from the entries displayed, the "information.csv" file also contained information such as the residual sum of squares (RSS) for the linear regression lines used, which are not shown in the table. The so called RSS is defined by

$$\text{RSS} = \sum_{t=t_0}^{t_1} (\text{TAC}(t) - f(t))^2 \quad ,$$

where $\text{TAC}(t)$ is the exact value of the simulated TAC, $f(t)$ is the approximated value of the linear regression line and t_0 and t_1 being the boundaries of the time interval considered for the linear regression. The lower the RSS, the better the linear fit [47].

The first column in table 4.1 called Input describes the type of input function: this can either be a pointed, blunt or kinky input function. The second and third column specify the underlying kinetic model, so either open or trapped and a one-, a two- or a three-compartment model. For a three-compartment model the csv-file comprised of twelve rate constants k_i instead of six, which is not shown in the table.

For a two-compartment model 195 TACs and for a three-compartment model 209 have been generated for each input function type. In contrast, for a one-compartment model only 13 TACs have been simulated for each kind of input function due to the model's simplicity. By analyzing all these TACs, the generated "information.csv" file containing all mathematical properties mentioned in section 3.2 comprised of 1251 entries in total. This large number of data was necessary for increasing the significance and reliability when it comes to further analysis via statistical methods.

Model type			Rate Constants						Shape			Before		After		I_{rel}	C
Input	#C	t/o	k_1	k_2	k_3	k_4	k_5	k_6	max ₁	max ₂	fall	m_{rel}	b_{rel}	m_{rel}	b_{rel}		
p	2	o	0.2	0.5	0.1	0.05	0.1	0.6	×		×	0.084	0.095	-0.021	0.063	0.142	-14.488
p	2	o	0.5	0.3	0.1	0.5	0.03	0	×		×	0.112	0.130	-0.106	0.061	0.221	-19.370
p	2	t	0.2	0.1	0	0	0.06	0		×		0.081	0.090	-0.039	0.051	0.138	-11.876
b	2	o	0.2	0.5	0.1	0.05	0.1	0.02	×		×	0.085	0.095	-0.380	0.014	0.132	-8.680
b	2	t	1	0.7	0	0	0.05	0	×	×		0.169	0.205	-1.216	-0.033	0.320	-17.892
k	2	o	0.2	0.6	0.1	0.8	0.5	0.4	×		×	0.131	0.155	-0.385	0.037	0.253	-16.885
k	2	t	0.2	1	0	0	0.05	0			×	0.078	0.087	-0.081	0.051	0.118	-10.077
k	2	t	0.1	0.01	0.2	0	0	0			×	0.062	0.065	-0.033	0.042	0.082	-6.213

Table 4.1: A small selection of the "information.csv" file written by the PYTHON script. Besides the mathematical properties (relative slope m_{rel} and intercept b_{rel} before and after a certain time point, relative integral I_{rel} , curvature C) and the overall TAC shape (max₁, max₂, fall off) the model parameters are recorded as well. This table is valid for a two-compartment model. For a three-compartment model there would be twelve rate constants k_i and for a one-compartment model only k_1 and k_2 would be present.

4.3 Evaluation of obtained Data

For the examination of significant relations between the rate constants k_i and the calculated mathematical properties, the corresponding correlation coefficients according to equation (26) had to be calculated. This was performed for all the one-, the two- and the three-compartment model. The mathematical properties considered have been the relative slope m_{rel} and intercept b_{rel} before and after a certain time point, the relative integral I_{rel} as well as the curvature C . These properties were set into relation to all rate constants k_i . In table 4.2 the correlation coefficients r_{xy} for all rate constants k_i and mathematical properties are shown. Since an in amount high coefficient $|r_{xy}| \approx 1$ corresponds to a strong dependence, for a better overview the coefficients of high significance are color coded. For $|r_{xy}| \geq 0.95$ the cells are marked as dark-gray, for $|r_{xy}| \geq 0.75$ the cells are highlighted as light-gray.

With table 4.2 one can see that there are high correlations between the first two rate constants and specific mathematical properties. Especially for k_1 there is a strong relation between the relative slope m_{rel} , intercept b_{rel} and curvature C . Furthermore, it is apparent that for all other rate constants there is no significant relation.

This result justifies the estimation of k_1 and k_2 by using these mathematical properties for back-calculation. The other rate constants had to be estimated by alternative methods, which leads to the Student's t -test.

As described in section 3.3.2, first it was necessary to sort the "information.csv" file shown in table 4.1 by different characteristics, namely the type of input function $B(t)$, the number

of compartments $\#C$, the presence of an open or trapped model as well as the TAC shape. All the considered classifications are listed in table 4.3. For classifying the data a two-step classification was performed, meaning first the dataset has been sorted by a primary characteristic and then optionally these classes could be further divided into different subclasses. For example, with $\#C$ being the main class and $B(t)$ being the secondary class, the datasets between different input function types within either a one-, a two- or a three-compartment model were compared. By that, it could be determined whether or not there is a difference between a pointed, a blunt and a kinky input function for a certain compartment model. If table 4.3 shows a \times in the secondary class, it means that only a first classification has been performed. Since there are many different combinations for the input function, the number of compartments, the existence of an open or trapped model and the TAC shape, the t -test had to be conducted for all mathematical properties for all feasible combinations of these classifications.

By means of the t -test one was able to determine if there is a significant difference in mathematical values between two different classifications. If so, these classes seemed to have a considerable influence on the specific mathematical properties.

The unpaired, two-sided t -test has been performed with the spreadsheet software EXCEL. The output of the t -test hereby is the p -value described in equation (27). In table 4.4 the outcome of the t -test for the first class being the number of compartments $\#C$ and the secondary being the type of input function $B(t)$ is shown. The fields with high significance are highlighted: light-gray for $p < 0.05$ and dark-gray for $p < 0.01$. For a low p -value the null hypothesis \mathcal{H}_0 describing the equality of datasets between two classes could be rejected. It should be underlined that for the relative slope m_{rel} and the relative intercept b_{rel} before a certain time point the p -value equals 1 for different input functions, meaning there is no difference between them at all. This result was due to the fact that m_{rel} and b_{rel} were calculated for a regression line within the time interval $[t_{\text{start}}, t_{\text{peak}}]$, where the input functions are the same (see the beginning term $A(t)$ in section 3.2.1). The TAC values within this time region therefore were the same as well, leading to equal relative slopes and intercepts before t_{peak} .

Apart from the one-compartment model, where only two rate constants k_i are present, for the two- and the three-compartment model table 4.4 consists of many fields with a low p -value. Even though the relative integral I_{rel} did not have as small p -values as other mathematical properties, the different classes seemed to have an influence on the relative integral

	I_{rel}	Before		After		C
		m_{rel}	b_{rel}	m_{rel}	b_{rel}	
k_1	0.859	0.958	0.954	0.068	0.906	-0.928
k_2	0.647	0.799	0.792	0.219	0.849	-0.786
k_3	-0.053	-0.065	-0.065	-0.010	-0.061	0.055
k_4	-0.203	-0.221	-0.220	-0.011	-0.209	0.219
k_5	0.135	0.048	0.052	-0.252	-0.118	0.010
k_6	0.168	0.115	0.119	-0.134	0.032	-0.111

Table 4.2: Correlation coefficients r_{xy} between the rate constants k_i and the calculated mathematical properties. The correlations with $|r_{xy}| \geq 0.95$ are highlighted as dark-gray, signaling a high relationship. In light-gray the cells are marked where $|r_{xy}| \geq 0.75$ holds.

1 st class	2 nd class	1 st class	2 nd class	1 st class	2 nd class	1 st class	2 nd class
$B(t)$	× # C TAC shape t/o	# C	× $B(t)$ TAC shape t/o	TAC shape	× $B(t)$ # C t/o	t/o	× $B(t)$ # C TAC shape

Table 4.3: Possible classifications of the "information.csv" file for further analysis via a Student's t -test. The considered classes were the type of input function $B(t)$, the number of compartments # C , the presence of either an open or trapped model and the general TAC shape. Each of them could either be used as a main class or as a secondary one. A × illustrates that no secondary classification has been performed.

anyway. Such a table comprising of p -values for the various mathematical properties had to be generated for all feasible combinations of classes (see table 4.3). Thereby, one got a good overview of the importance of the performed subdivision. Since for all classifications there were many fields comprising of low p -values, there seemed to be a large dependence of the mathematical properties on the type of input function $B(t)$, the number of compartments # C , the presence of an open or trapped model as well as the TAC shape. This result clearly stated that it is necessary to look at all classes one by one.

4.4 Programming the Algorithm

In the previous section the results pointed out that the investigations via statistical methods are less informative than expected. Thus, the data needed to be classified by its input function, the model type and the TAC shape. By that the algorithm has to output values and feasible intervals for the rate constants on basis of the underlying characteristic, leading to a more complex structure of the algorithm than initially presumed. Another drawback is the limitation of the algorithm on the rate constants k_1 and k_2 for a one-compartment model, k_1

1 st class	2 nd class	I_{rel}	Before		After		C
			m_{rel}	b_{rel}	m_{rel}	b_{rel}	
1C	$B_p(t) \leftrightarrow B_b(t)$	0.817	1	1	0.003	0.451	0.711
	$B_p(t) \leftrightarrow B_k(t)$	0.913	1	1	0.185	0.819	0.672
	$B_b(t) \leftrightarrow B_k(t)$	0.903	1	1	0.035	0.604	0.969
2C	$B_p(t) \leftrightarrow B_b(t)$	0.077	1	1	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-3}$
	$B_p(t) \leftrightarrow B_k(t)$	0.415	1	1	$< 10^{-5}$	0.010	0.002
	$B_b(t) \leftrightarrow B_k(t)$	0.336	1	1	$< 10^{-5}$	$< 10^{-5}$	0.389
3C	$B_p(t) \leftrightarrow B_b(t)$	0.056	1	1	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-3}$
	$B_p(t) \leftrightarrow B_k(t)$	0.374	1	1	$< 10^{-5}$	0.018	0.001
	$B_b(t) \leftrightarrow B_k(t)$	0.303	1	1	$< 10^{-5}$	$< 10^{-5}$	0.562

Table 4.4: Calculation of the p -value of a Student's t -test for the data with EXCEL. As an example, the first classification is the number of compartments $\#C$, the secondary being the type of input function $B(t)$. Highlighted as light-gray and dark-gray are the fields of high significance where $p < 0.05$ and $p < 0.01$, respectively, holds.

to k_6 for a two-compartment model and k_1 to k_4 , k_7 and k_8 in a three-compartment model due to the poor information obtained by statistical methods such as the Student's t -test and the correlation coefficient. The other six rate constants appearing in a three-compartment model remain unconsidered.

As already described in section 3.4, the first task of the algorithm is to determine the type of input function as well as examining the measured TAC for its shape. In the next step the PYTHON script analyzes the TAC for various mathematical properties such as the relative slope m_{rel} and intercept b_{rel} before and after a certain time point, the relative integral I_{rel} as well as its curvature C .

Since there happened to be a high correlation between the first two rate constants k_1, k_2 and various mathematical characteristics, the attempt was made to compute these rate constants by back-calculation. In figure 4.4 the data points with their linear regression lines are shown for an open two-compartment model supplied by a kinky input function. As one can see, in this case the two mathematical properties showing the best correlation coefficients were the relative slope before and the relative intercept after a certain time point for the first rate constant k_1 (left plot). For the second constant k_2 (right plot) the best correlating properties were the relative intercept after a certain time point as well as the curvature. In the case of some other compartment model and input function type the linear regression were different. All these linear regression lines have been used for determining k_1 and k_2 via back-calculation according to equation (29).

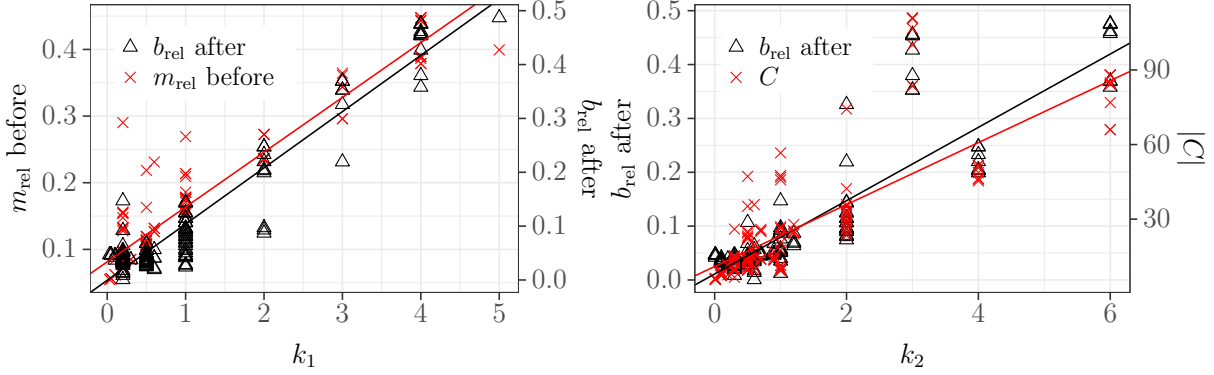


Figure 4.4: Linear regression lines for an open two-compartment model supplied by a kinky input function: the two mathematical properties having the best correlation with k_1 (left) and k_2 (right) are shown.

As one can see in table 4.2, the other rate constants did not comprise of such a good correlation coefficient. Thus, a linear fit of the data points and the subsequent back-calculation were not expected to yield trustworthy values for these rate constants. Nevertheless, to make an estimation for the rate constants k_3 to k_6 in a two-compartment model and k_3, k_4, k_7 and k_8 in a three-compartment model, two-sided 95% confidence intervals have been calculated. These intervals provide a good first assessment for the values of these rate constants.

As described in section 3.4, to generate such confidence intervals, the data has been classified and afterwards the confidence interval to the significance level $\alpha = 5\%$ was computed. These intervals have then been intersected to obtain feasible intervals for each rate constant for all types of input functions, all compartment models and all TAC shapes as well. In tables 4.5, 4.6 and 4.7 all the feasible intervals are shown. These intervals were imported into the algorithm in order to provide the estimated values for k_1 and k_2 as well as the feasible intervals for the other rate constants. An example for an output of the algorithm is shown in figure 4.5.

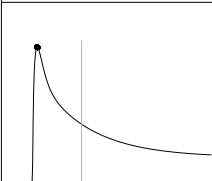
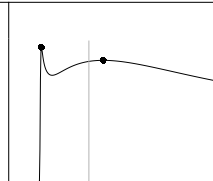
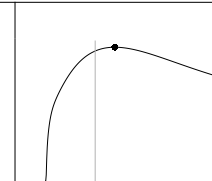
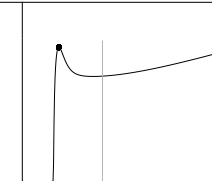
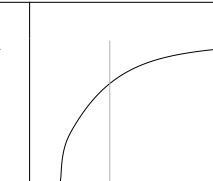
Pointed Input Function					
					
	max ₁ no max ₂ fall off	max ₁ max ₂ fall off	no max ₁ max ₂ fall off	max ₁ no max ₂ no fall off	no max ₁ no max ₂ no fall off
1C	$0 < k_1 < 3.257$ $0 < k_2 < 3.226$	$0 < k_1 < 3.519$ $0 < k_2 < 3.494$	$0 < k_1 < 5.274$ $0 < k_2 < 5.235$	$0.042 < k_1 < 0.178$ $0 < k_2 < 0.017$	$0.042 < k_1 < 0.178$ $0 < k_2 < 0.017$
2C open	$0.878 < k_1 < 1.190$ $1.199 < k_2 < 1.634$ $0.458 < k_3 < 0.633$ $0.325 < k_4 < 0.446$ $0.091 < k_5 < 0.162$ $0.361 < k_6 < 0.487$	$0.878 < k_1 < 1.257$ $1.199 < k_2 < 1.711$ $0.458 < k_3 < 0.634$ $0.325 < k_4 < 0.446$ $0.091 < k_5 < 0.174$ $0.361 < k_6 < 0.527$	$0.458 < k_3 < 0.718$ $0.325 < k_4 < 0.488$	$0.878 < k_1 < 1.190$ $1.199 < k_2 < 1.634$ $0.426 < k_3 < 0.483$ $0.310 < k_4 < 0.446$	$0.426 < k_3 < 0.483$ $0.310 < k_4 < 0.490$
2C trapped	$1.101 < k_1 < 1.792$ $1.497 < k_2 < 2.611$ $0.042 < k_3 < 0.170$ $0.012 < k_5 < 0.029$	$1.101 < k_1 < 2.044$ $1.497 < k_2 < 3.003$ $0.035 < k_3 < 0.183$ $0.011 < k_5 < 0.029$	$0.147 < k_1 < 1.253$ $0.041 < k_2 < 0.247$ $0 < k_3 < 0.198$ $0 < k_5 < 0.062$	$1.101 < k_1 < 1.792$ $1.497 < k_2 < 2.602$ $0.043 < k_3 < 0.170$ $0.012 < k_5 < 0.029$	$0.043 < k_3 < 0.170$ $0.012 < k_5 < 0.030$
3C open	$1.049 < k_1 < 1.255$ $1.599 < k_2 < 2.015$ $0.515 < k_3 < 0.605$ $0.327 < k_4 < 0.472$ $0.058 < k_7 < 0.099$ $0.183 < k_8 < 0.348$	$1.049 < k_1 < 1.255$ $1.599 < k_2 < 2.015$ $0.515 < k_3 < 0.606$ $0.327 < k_4 < 0.472$ $0.058 < k_7 < 0.100$ $0.183 < k_8 < 0.350$	$0.779 < k_1 < 1.255$ $1.278 < k_2 < 2.015$ $0.515 < k_3 < 0.756$ $0.304 < k_4 < 0.472$ $0.047 < k_7 < 0.113$ $0.153 < k_8 < 0.412$	$1.107 < k_1 < 1.407$ $1.614 < k_2 < 2.127$ $0.457 < k_3 < 0.546$ $0.327 < k_4 < 0.497$ $0.058 < k_7 < 0.099$ $0.183 < k_8 < 0.348$	$1.107 < k_1 < 1.407$ $1.614 < k_2 < 2.127$ $0.457 < k_3 < 0.546$ $0.325 < k_4 < 0.497$ $0.058 < k_7 < 0.099$ $0.182 < k_8 < 0.348$
3C trapped	$1.092 < k_1 < 2.103$ $1.284 < k_2 < 2.654$ $0.219 < k_3 < 0.681$ $0.131 < k_4 < 0.565$ $0.005 < k_7 < 0.052$ $0.091 < k_8 < 0.869$	$1.092 < k_1 < 2.317$ $1.284 < k_2 < 2.942$ $0.201 < k_3 < 0.712$ $0.131 < k_4 < 0.565$ $0.005 < k_7 < 0.058$ $0.091 < k_8 < 0.972$	$0 < k_3 < 1.691$ $0 < k_4 < 7.070$ $0 < k_8 < 0.210$	$1.092 < k_1 < 2.103$ $1.284 < k_2 < 2.654$ $0.219 < k_3 < 0.681$ $0.131 < k_4 < 0.565$ $0.005 < k_7 < 0.052$ $0.091 < k_8 < 0.869$	$0.219 < k_3 < 0.681$ $0.111 < k_4 < 0.812$ $0.004 < k_7 < 0.052$ $0.087 < k_8 < 0.210$

Table 4.5: Feasible intervals for a pointed input function: for each model type and each TAC shape there are different intervals.

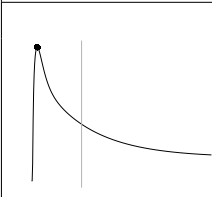
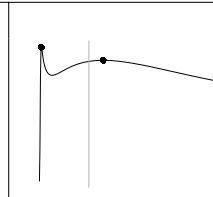
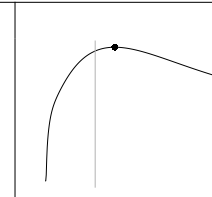
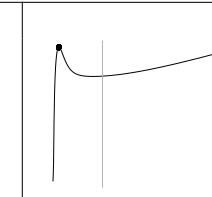
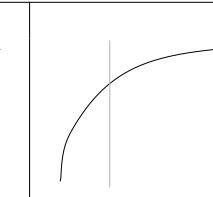
Blunt Input Function					
					
	max ₁ no max ₂ fall off	max ₁ max ₂ fall off	no max ₁ max ₂ fall off	max ₁ no max ₂ no fall off	no max ₁ no max ₂ no fall off
1C	$0 < k_1 < 5.274$ $0 < k_2 < 5.235$	$0 < k_1 < 5.274$ $0 < k_2 < 5.235$	$0.042 < k_1 < 0.178$ $0 < k_2 < 0.017$	$0.042 < k_1 < 0.178$ $0 < k_2 < 0.017$	$0.042 < k_1 < 0.178$ $0 < k_2 < 0.017$
2C open	$0.847 < k_1 < 1.190$ $1.146 < k_2 < 1.634$ $0.442 < k_3 < 0.633$ $0.310 < k_4 < 0.470$ $0.082 < k_5 < 0.162$ $0.292 < k_6 < 0.487$	$0.847 < k_1 < 1.204$ $1.146 < k_2 < 1.666$ $0.442 < k_3 < 0.635$ $0.309 < k_4 < 0.470$ $0.082 < k_5 < 0.168$ $0.292 < k_6 < 0.503$	$0.442 < k_3 < 0.661$ $0.309 < k_4 < 0.490$	$0.847 < k_1 < 1.190$ $1.146 < k_2 < 1.634$ $0.426 < k_3 < 0.523$ $0.310 < k_4 < 0.470$	$0.426 < k_3 < 0.523$ $0.310 < k_4 < 0.490$
2C trapped	$1.107 < k_1 < 1.792$ $1.461 < k_2 < 2.611$ $0.042 < k_3 < 0.070$ $0.015 < k_5 < 0.030$	$1.107 < k_1 < 2.072$ $1.461 < k_2 < 3.021$ $0.021 < k_3 < 0.070$ $0.015 < k_5 < 0.036$	$0.041 < k_1 < 1.534$ $0.065 < k_2 < 0.837$ $0.068 < k_3 < 0.734$	$1.107 < k_1 < 1.792$ $1.461 < k_2 < 2.602$ $0.043 < k_3 < 0.070$ $0.015 < k_5 < 0.030$	$0.068 < k_3 < 0.170$ $0.012 < k_5 < 0.030$
3C open	$1.051 < k_1 < 1.399$ $1.598 < k_2 < 2.127$ $0.457 < k_3 < 0.597$ $0.325 < k_4 < 0.484$ $0.064 < k_7 < 0.099$ $0.184 < k_8 < 0.310$	$1.051 < k_1 < 1.399$ $1.598 < k_2 < 2.147$ $0.452 < k_3 < 0.597$ $0.314 < k_4 < 0.484$ $0.064 < k_7 < 0.100$ $0.184 < k_8 < 0.310$	$1.559 < k_2 < 1.997$ $0.435 < k_3 < 0.597$ $0.308 < k_4 < 0.504$ $0.064 < k_7 < 0.112$ $0.153 < k_8 < 0.310$	$1.051 < k_1 < 1.407$ $1.598 < k_2 < 2.127$ $0.457 < k_3 < 0.601$ $0.325 < k_4 < 0.484$ $0.058 < k_7 < 0.081$ $0.184 < k_8 < 0.348$	$1.581 < k_2 < 1.997$ $0.457 < k_3 < 0.605$ $0.325 < k_4 < 0.497$ $0.058 < k_7 < 0.081$ $0.182 < k_8 < 0.348$
3C trapped	$0.959 < k_1 < 2.103$ $1.108 < k_2 < 2.654$ $0.219 < k_3 < 0.681$ $0.111 < k_4 < 0.812$ $0.004 < k_7 < 0.052$ $0.087 < k_8 < 0.869$	$0.959 < k_1 < 2.103$ $1.108 < k_2 < 2.654$ $0.219 < k_3 < 0.681$ $0.111 < k_4 < 0.812$ $0.004 < k_7 < 0.052$ $0.087 < k_8 < 0.869$		$0.959 < k_1 < 1.989$ $1.108 < k_2 < 2.440$ $0.230 < k_3 < 0.681$ $0.117 < k_4 < 0.812$ $0.004 < k_7 < 0.049$ $0.087 < k_8 < 0.869$	$0.959 < k_1 < 1.989$ $1.108 < k_2 < 2.440$ $0.230 < k_3 < 0.681$ $0.117 < k_4 < 0.812$ $0.004 < k_7 < 0.049$ $0.087 < k_8 < 0.869$

Table 4.6: Feasible intervals for a blunt input function: for each model type and each TAC shape there are different intervals.

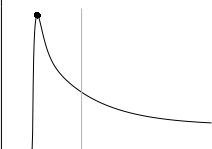
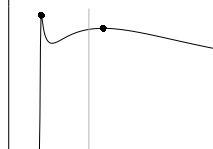
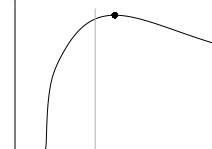
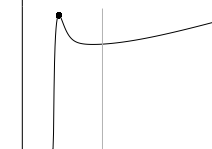
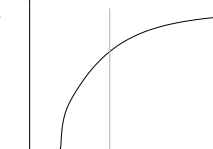
Kinky Input Function					
					
	max ₁ no max ₂ fall off	max ₁ max ₂ fall off	no max ₁ max ₂ fall off	max ₁ no max ₂ no fall off	no max ₁ no max ₂ no fall off
1C	$0 < k_1 < 4.181$ $0 < k_2 < 4.178$	$0 < k_1 < 0.713$ $0 < k_2 < 0.331$	$0.049 < k_1 < 0.385$ $0 < k_2 < 0.139$	$0.048 < k_1 < 0.127$ $0 < k_2 < 0.010$	$0.049 < k_1 < 0.127$ $0 < k_2 < 0.010$
2C open	$0.963 < k_1 < 1.226$ $1.281 < k_2 < 1.681$ $0.431 < k_3 < 0.638$ $0.321 < k_4 < 0.459$ $0.087 < k_5 < 0.168$ $0.352 < k_6 < 0.503$	$1.281 < k_2 < 1.412$ $0.425 < k_3 < 0.557$ $0.298 < k_4 < 0.459$	$0.425 < k_3 < 0.557$ $0.297 < k_4 < 0.471$	$0.431 < k_3 < 0.678$ $0.321 < k_4 < 0.459$	$0.431 < k_3 < 0.600$ $0.321 < k_4 < 0.499$
2C trapped	$1.645 < k_1 < 1.857$ $2.136 < k_2 < 2.714$	$1.645 < k_1 < 2.867$ $2.136 < k_2 < 3.442$ $0.007 < k_5 < 0.012$	$0.380 < k_2 < 0.734$ $0.008 < k_5 < 0.012$	$1.645 < k_1 < 1.748$ $2.136 < k_2 < 2.714$ $0.052 < k_3 < 0.083$ $0.013 < k_5 < 0.030$	$0.052 < k_3 < 0.177$ $0.013 < k_5 < 0.031$
3C open	$1.175 < k_1 < 1.399$ $1.779 < k_2 < 2.138$ $0.476 < k_3 < 0.608$ $0.321 < k_4 < 0.423$ $0.061 < k_7 < 0.104$ $0.189 < k_8 < 0.349$	$1.175 < k_1 < 1.399$ $1.779 < k_2 < 2.138$ $0.458 < k_3 < 0.543$ $0.321 < k_4 < 0.423$ $0.061 < k_7 < 0.104$ $0.188 < k_8 < 0.279$	$0.458 < k_3 < 0.543$ $0.321 < k_4 < 0.498$ $0.061 < k_7 < 0.106$ $0.184 < k_8 < 0.279$	$1.175 < k_1 < 1.429$ $1.779 < k_2 < 2.153$ $0.476 < k_3 < 0.613$ $0.308 < k_4 < 0.423$ $0.189 < k_8 < 0.376$	$0.476 < k_3 < 0.643$ $0.308 < k_4 < 0.447$ $0.189 < k_8 < 0.304$
3C trapped	$1.340 < k_1 < 2.333$ $1.420 < k_2 < 2.877$ $0.006 < k_7 < 0.034$ $0.091 < k_8 < 0.155$	$1.340 < k_1 < 2.082$ $1.420 < k_2 < 3.466$ $0.095 < k_4 < 0.105$ $0 < k_7 < 0.034$ $0.086 < k_8 < 0.155$	$0.135 < k_1 < 0.615$ $0.079 < k_2 < 1.196$ $0.091 < k_3 < 0.114$ $0 < k_4 < 0.105$ $0 < k_7 < 0.130$ $0 < k_8 < 0.155$	$1.340 < k_1 < 1.984$ $1.420 < k_2 < 2.349$ $0.262 < k_3 < 0.731$ $0.139 < k_4 < 0.589$ $0.006 < k_7 < 0.034$ $0.091 < k_8 < 1.012$	$1.178 < k_2 < 1.196$ $0.262 < k_3 < 0.634$ $0.139 < k_4 < 0.589$ $0.006 < k_7 < 0.050$ $0.091 < k_8 < 0.241$

Table 4.7: Feasible intervals for a kinky input function: for each model type and each TAC shape there are different intervals.

Max1 exists: True	2C-model (open):	3C-model (open):
Max2 exists: False	k1 = 0.878	k1 = 1.107
Fall at end: False	k2 = 1.199	k2 = 1.614
Blood curve: Sharp	k3 : [0.426 , 0.483]	k3 : [0.457 , 0.546]
a = 39.024	k4 : [0.31 , 0.446]	k4 : [0.327 , 0.497]
b = 0.098	k5 : [0 , inf]	k7 : [0.058 , 0.099]
c = 3.14	k6 : [0 , inf]	k8 : [0.183 , 0.348]
1C-model:	2C-model (trapped):	3C-model (trapped):
k1 = 0.042	k1 = 1.101	k1 = 1.092
k2 = 0.0	k2 = 1.497	k2 = 1.284
	k3 : [0.043 , 0.17]	k3 : [0.219 , 0.681]
	k5 : [0.012 , 0.029]	k4 : [0.131 , 0.565]
		k7 : [0.005 , 0.052]
		k8 : [0.091 , 0.869]

Figure 4.5: Example of an output of the algorithm. Besides the TAC shape (\max_1 , \max_2 , fall off), the type of input function is presented as well. The algorithm provides values and feasible intervals for the specific rate constants for all model types (1C, 2C open, 2C trapped, 3C open, 3C trapped) separately.

4.5 Testing the Algorithm

The last step was to test the algorithm's reliability. For this, the algorithm was fed with different real measurement curves containing concentration values for both the input function and the TAC. This is an important step, since the algorithm was built up on basis of ideal input functions and TACs as well as on a perfect exchange between the individual compartments. In contrast to the simulated TACs, the ones obtained from real measurements comprise of large fluctuations. These fluctuations in particular can lead to a wrong assumption of the existence of a \max_1 and a \max_2 .

While the relative slope m_{rel} and intercept b_{rel} before and after a certain time point, the relative integral I_{rel} , the curvature C as well as the property of falling down at the end are calculated over a certain time range, the existence of a \max_1 and \max_2 is determined on basis of a maximum appearing at a single time point. Hence, fluctuations in the TAC are expected of having a possible impact on the existence of the two maximums. This is due to the fact that an outlier in the TAC concentration can be thought of as being a maximum although in reality it only appears due to noise and measurement uncertainties. In contrast, for calculations over a certain time period the fluctuations are supposed to have a less important impact on the specific mathematical properties.

Another crucial factor to keep in mind were all the external parameters. For adjusting the

algorithm to real recorded TACs, one had to consider these parameters, since they have a great influence on the outcome of the algorithm. It turned out that the best results of the algorithm are obtained with the following choice of external parameters:

- *threshold level* $c_{\text{th}} = 0.03$: the measurement starts as soon as the concentration exceeds 3% of the input function peak. This level seems to be high enough to ignore noise in the beginning. On the other hand it is low enough so that the start of the measurement will be recognized properly.
- *slope range* $r_s = 3$: the length of the time window for calculating the relative slope and intercept after a certain time point is three times the interval length $[t_{\text{start}}, t_{\text{peak}}]$.
- *integral range* $r_I = 2$: for the relative integral, the considered time window is twice as long as the interval $[t_{\text{start}}, t_{\text{peak}}]$.
- *curvature range* $r_C = 4$: the time window for computing the curvature is four times as wide as the time interval $[t_{\text{start}}, t_{\text{peak}}]$. This relatively large range is necessary since for shorter time windows the calculated curvature is hardly reliable (see figure 3.3).
- *fall off range* $r_{\text{fall off}} = 0.2$: for examining the TAC for a fall off at the end, the last 20% of the data points are taken into account. This range turned out to be a good compromise between considering only a small part of data points for determining a fall off, which increases its reliability and on the other hand taking into account enough points to neglect TAC fluctuations which could alter the result of a fall off.
- *limit range* $r_{\text{limit}} = 35$: the time point distinguishing a max_1 from a max_2 is located at 35 times the window length of $[t_{\text{peak}}, t_{\text{start}}]$.
- *percentage step* $p = 0.05$: with each iteration in finding a maximum (both max_1 and max_2), the considered time window gets shortened by 5% of the initial window length (see section 3.2.6). A smaller *percentage step* would be more precise while increasing the computing time. In contrast, a larger p would be faster but less accurate. A value of 5% is sufficiently precise while keeping the calculation time low.
- *concentration difference* $\varepsilon = 0.03$: the concentration of a max_2 has to exceed 3% of the concentrations at the boundaries of the considered time window in order to be designated as max_2 .

With this choice of the external parameters the algorithm seemed to yield the most reliable results for the rate constants k_i . As aforementioned, one has to ensure that the existence of a maximum is associated with a real maximum and not only due to concentration fluctuations.

Therefore for future applications it might be helpful to adapt ε for each measurement in order to obtain better results. The *concentration difference* ε turned out to be the external parameter influencing the algorithm most.

With all the external parameters being adapted, the algorithm finally can be used for estimating the rate constants of a kinetic model. In section 3.2 all these external parameters with their application and implementation is described more precisely.

To sum up, the algorithm outputs both the TAC shape (\max_1 , \max_2 , fall off at the end) and the parameters and type of the fitted input function $B_{\text{app}}(t)$ (pointed, blunt or kinky) as well as initial values and feasible intervals for the rate constants for each type of compartment model (see figure 4.5). Due to the lack of information the algorithm itself is not able to propose a certain type best fitting the measured TAC. Hence, the algorithm outputs estimated values for the rate constant for each type of compartment model, which leaves it up to the operator to decide which model to use. As already mentioned, the values and intervals provided by the algorithm are not supposed to be exact values representing the measured TAC. They rather should present good initial values which can be adapted by an optimization tool such as the Nelder-Mead simplex method. By that, rate constants providing a good accordance between the calculated and the real TAC can be obtained.

5 Discussion

The goal in kinetic modeling is to describe the temporal behavior of a tracer within a VOI. In compartmental modeling, the quantification takes place by using mathematical models which are based on compartments as well as on their tracer exchange governed by rate constants. The aim is now to get a good estimation of these rate constants. In order to do so, simulated TACs are adapted to real measured ones by applying an optimization tool such as the Nelder-Mead simplex method. The accordance between the simulated and the measured TAC thereby is gradually increased by altering the rate constants, until a local best fit is reached.

Since all these optimization tools strongly depend on the initial values given, it is very important to start with proper rate constants. The algorithm programmed for this master thesis should provide such good initial values to enhance the general outcome after applying an optimization method. Nevertheless, there were some issues coming up during the programming of this algorithm.

The set of differential equations (1), (2) and (3) can either be solved analytically or numerically. Since the analytical solution turned out to be very complex, even for a simplified two-compartment model (see section 2.1), the differential equations have been solved by numerical means. In section 3.1.2 the Euler method with step size $h = 1$ s was performed. Even though an analytical solution would be more precise than applying numerical methods, the errors produced by the Euler method are small and therefore negligible. Besides, with solving the set of differential equations numerically, the implementation of the algorithm can be extended to other problems as well (for example when using a constant input function). In the case of analytical solutions, one would have to start all over again when changing external conditions having an influence on the TAC behavior.

Apart from the Euler method there are other numerical techniques such as the second-order Runge-Kutta methods, by which the sets of differential equations could be solved (see section 2.3). Nevertheless, the more precise a numerical approach, the more complex and hence unsuitable for programming the algorithm. The Euler method hereby turned out to yield good simulated TACs while keeping low the complexity.

Some other issues arising when numerically calculating various mathematical properties by the PYTHON script are the relative integral and the TAC's curvature. For the relative in-

tegral I_{rel} the sum of the discrete TAC values within a certain time window is taken and divided by that of the input function. This can be seen as a quadrature formula using polynomial functions of order zero (constant functions) between the individual time points. A more precise approximation could be obtained by using functions of higher order [48]. Another possibility would be to use a spline interpolation for approximating the TAC by a polynomial function and subsequently computing its integral [49].

The same could be done for the calculation of the curvature C : instead of estimating the curvature by calculating the difference quotients twice (see equation (22)), again for better results one could approximate the TAC. Nevertheless, since the main errors appearing during programming the algorithm were caused by statistical uncertainties, the calculation methods for the relative integral and the curvature are rather unimportant. The use of more complex formulas would not considerably enhance the reliability of the obtained results, but the calculation would get more complex. Thus, the most simple techniques for numerically calculating these mathematical properties were performed.

Two other mathematical values of interest were the relative slope and intercept before a certain time point. As expected, they turned out to show high correlations with the first two rate constants k_1 and k_2 , wherefore these values have been used for back-calculation (see section 3.4.4). Nevertheless, table 4.4 shows that their p -values were equal to 1, stating that for these mathematical properties there is no difference at all between different types of input function. This is due to the beginning term $A(t)$ for the various input functions (see equation (17)). $A(t)$ was the same for all types, hence the calculation of k_1 and k_2 using the relative slope and intercept before a certain time point actually was based on the same linear rising term. Hence, when estimating the first two rate constants via these mathematical properties, the different types of input function had no effect, which reduces the reliability of the obtained values for k_1 and k_2 .

A more important drawback appearing in this master thesis was the statistical uncertainty. After generating the TACs they have been classified by the input function $B(t)$, the number $\#C$ of compartments, the overall TAC shape and whether the underlying model is trapped or open. Subsequently a t -test was performed to see if there is a noteworthy difference among these classifications. Since most of the p -values turned out to be very low (see table 4.4), the classifications were supposed to have a huge impact on the various mathematical properties. The reason for performing these classifications was to increase the reliability of the obtained confidence intervals and in consequence to get a better estimation for the rate constants.

However, due to these classifications, the sample sizes n of the individual classifications were drastically reduced, sometimes even $n < 10$. Therefore the reliability of the p -values was reduced as well. To ensure a better significance and to reduce the statistical uncertainties, many more TACs would have to be simulated in the beginning, as the original number of 1251 generated TACs turned out to be too low to really provide reliable values.

Another issue in programming the algorithm are the input functions used for simulation. As explained in section 3.1.1 there have been three different types of input functions which were used for generating the TACs. Nevertheless, there are other possible types as well. For instance there is the constant input function, which can be used for tracer delivery but which is not covered in this master thesis. Using such a constant input function would have an enormous impact on the TAC shape and thus would alter the obtained rate constants tremendously.

The last issue discussed is the algorithm's outcome and the still needed operator's decision on which model to use best. Originally, the aim of this master thesis was to program an algorithm which was supposed to determine both the input function type and the underlying model (one-, two- or three-compartment model and open or trapped) and to provide good initial values for the rate constants. Thereby the assessment of the rate constants should not include the operator's decision at all. However, due to the lack of statistical information, one had to abandon this idea. Instead, the algorithm now provides initial values for each model type (see figure 4.5), which leaves it up to the operator to decide which model to use. So the estimation of the rate constants still includes a subjective assessment.

6 Conclusion

In kinetic modeling the aim is to quantify the dynamic behavior of tracers administered to patients. For a compartment model the mathematical description is based on different compartments including respective rate constants. To estimate these rate constants, various curve fitting programs exist and are already in use. Nevertheless, a general problem hereby is the dependence of these programs' reliability on the given initial values of the rate constants.

In this master thesis an algorithm was programmed which aim is to asses such rate constants. For this purpose the algorithm is based on a previously calculated database of linear regression lines and feasible intervals for different compartment models. By that the algorithm reads in a measured TAC, collates it with the simulated ones and outputs proper values which in turn can be used for further curve fitting.

By having good initial values, the examination of physiological processes can be facilitated and the understanding of the physical, biological and chemical behavior of the tracer within a VOI can be improved as well. Hence, a first assessment of rate constants is imperative when it comes to PET data quantification using compartment models. Plus, the investigation and study of new tracers applicable in nuclear medicine can be supported by using an algorithm suitable for parameter estimation.

List of Abbreviations

^{18}F FDG	Fludeoxyglucose ^{18}F
^{68}Ga -PSMA	Prostate Specific Membrane Antigen labeled with ^{68}Ga
$^{99\text{m}}\text{Tc}$ -MAG3	Mercaptoacetyltriglycine labeled with $^{99\text{m}}\text{Tc}$
CT	Computed Tomography
EC	Electron Capture
FUR	Fractional Uptake Rate
IT	Isomeric Transition
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
PVE	Partial Volume Effect
ROI	Region of Interest
RSS	Residual Sum of Squares
SPECT	Single Photon Emission Computed Tomography
SUV	Standardized Uptake Value
TAC	Time Activity Curve
VOI	Volume of Interest

Nomenclature

$\#C$	Number of Compartments
$A(t)$	Beginning Term in the Input Function
$B(t)$	Input Function in Kinetic Modeling
$B_{\text{app}}(t)$	Approximated Input Function
$B_p(t), B_b(t), B_k(t)$	Pointed, Blunt and Kinky Input Function considered in this Master Thesis
$b_{\text{TAC}}, b_B, b_{\text{rel}}$	Intercept of the Linear Regression Line for the Time Activity Curve, the Input Function and their Ratio
C	Curvature
$C_i(t)$	Concentration in the i -th Compartment
c_{th}	Threshold Level to disregard Measurement Noise
ε	Concentration Difference
I_{rel}	Relative Integral
k_i	Rate Constants
$m_{\text{TAC}}, m_B, m_{\text{rel}}$	Slope of the Linear Regression Line for the Time Activity Curve, the Input Function and their Ratio
$\text{max}_1, \text{max}_2$	TAC Maximums before and after t_{limit}
p	Percentage Step
$r_C, r_{\text{fall off}}, r_I, r_s$	Curvature Range, Fall Off Range, Integral Range and Slope Range
t_{end}	Both the End Time for a certain Time Window and the Measurement End Time
$t_{\text{fall off start}}$	Starting Time for the Fall Off Interval
t_{limit}	Limit Time Distinguishing a max_1 from a max_2

t_{mean}	Arithmetic Mean Value of t_{start} and t_{end}
t_{peak}	Peak Time of the Input Function
$\theta(t), \theta'(t)$	Heaviside Function, Half-Maximum Convention of the Heaviside Function
w	Weighting Factor for Calculating the TAC
$X(t), \bar{X}(t)$	Response Function, Simplified Response Function

References

- [1] Rachel A. Powsner, Matthew R. Palmer, and Edward R. Powsner. *Essentials of Nuclear Medicine Physics and Instrumentation*. John Wiley & Sons, third edition, 2013.
- [2] Wolfgang Schlegel, Christian P. Karger, and Oliver Jäkel. *Medizinische Physik: Grundlagen-Bildgebung-Therapie-Technik*. Springer, 2018.
- [3] RadiologyInfo. *General Nuclear Medicine*. <https://www.radiologyinfo.org/en/info.cfm?pg=gennuclear>, 2018. Accessed on March 13th, 2019.
- [4] Giovanni Lucignani. SUV and Segmentation: Pressing Challenges in Tumour Assessment and Treatment. *European Journal of Nuclear Medicine and Molecular Imaging*, 36(4):715–720, 2009.
- [5] Alexander Stahl, Katja Ott, Markus Schwaiger, and Wolfgang A. Weber. Comparison of Different SUV-Based Methods for Monitoring Cytotoxic Therapy with FDG PET. *European Journal of Nuclear Medicine and Molecular Imaging*, 31(11):1471–1478, 2004.
- [6] Peter Morris and Alan Perkins. Physics and Medicine 2: Diagnostic imaging. *The Lancet*, 379(9825):1525–1533, 2012.
- [7] Science Encyclopedia. *Nuclear Medicine - Radionuclides and Radiopharmaceuticals*. <https://science.jrank.org/pages/4738/Nuclear-Medicine-Radionuclides-radiopharmaceuticals.html>, 2019. Accessed on March 8th, 2019.
- [8] Richard E. Carson. *Positron Emission Tomography*. Springer, 2005.
- [9] Michael E. Phelps, John C. Mazziotta, and Heinrich R. Schelbert. *Positron Emission Tomography*. Los Alamos National Laboratory, 1988.
- [10] Victor I Mikla and Victor V Mikla. *Medical Imaging Technology*. Elsevier, 2013.
- [11] Luis W. Alvarez. The Capture of Orbital Electrons by Nuclei. *Physical Review*, 54(7):486–497, 1938.
- [12] Valerii P. Chechev. An Exotic Type of Nuclear Isomeric Transition. *Bulletin of the Russian Academy of Sciences: Physics*, 78(7):634–635, 2014.

- [13] University of Washington, Division of Nuclear Medicine. *Introduction to PET Physics*. https://depts.washington.edu/nucmed/IRL/pet_intro/intro_src/section6.html, 1999. Accessed on March 29th, 2019.
- [14] Josef A. Weber, Hubert Ceeh, and Christoph Hugenschmidt. *Messung des Fermi-Impulses durch die Winkelkorrelation von Gamma-Strahlung aus der Annihilation von Elektron-Positron-Paaren*. <https://www.ph.tum.de/academics/org/labs/fopra/docs/userguide-86.de.pdf>, 2015. Accessed on March 7th, 2019.
- [15] Scott Britz-Cunningham and Victor H. Gerbaudo. *A Case-Based Approach to PET/CT in Oncology*. Cambridge University Press, 2012.
- [16] Sandip Basu, Thomas C. Kwee, Suleman Surti, Esma A. Akin, Don Yoo, and Abass Alavi. Fundamentals of PET and PET/CT Imaging. *Annals of the New York Academy of Sciences*, 1228(1):1–18, 2011.
- [17] Abass Alavi and Steve S. Huang. *Cancer Imaging*. Elsevier, 2008.
- [18] Kin Leung. *Molecular Imaging and Contrast Agent Database (MICAD)*. <https://www.ncbi.nlm.nih.gov/books/NBK23549/>, 2004. Accessed on March 10th, 2019.
- [19] Kazuo Itoh. 99m Tc-MAG3: Review of Pharmacokinetics, Clinical Application to Renal Diseases and Quantification of Renal Function. *Annals of Nuclear Medicine*, 15(3):179–190, 2001.
- [20] Science Medicines Health European Medicine Agencies. *Zusammenfassung des EPAR für die Öffentlichkeit - DaTSCAN*. https://www.ema.europa.eu/en/documents/overview/datscan-epar-summary-public_de.pdf, 2010. Accessed on March 10th, 2019.
- [21] Matthias R. Benz, Vladimir Evilevitch, Martin S. Allen-Auerbach, Fritz C. Eilber, Michael E. Phelps, Johannes Czernin, and Wolfgang A. Weber. Treatment Monitoring by 18F-FDG PET/CT in Patients with Sarcomas: Interobserver Variability of Quantitative Parameters in Treatment-Induced Changes in Histopathologically Responding and Nonresponding Tumors. *Journal of Nuclear Medicine*, 49(7):1038–1046, 2008.
- [22] M’hamed Bentourkia and Habib Zaidi. Tracer Kinetic Modeling in PET. *PET Clinics*, 2(2):267–277, 2007.

- [23] Koichi Ishizu and Yoshiharu Yonekura. Clarification of a Fractional Uptake Concept - Reply. *Journal of Nuclear Medicine*, 36(4):712–712, 1995.
- [24] Evan D. Morris, Christopher J. Endres, Kathleen C. Schmidt, Bradley T. Christian, Raymond F. Muzic, and Ronald E. Fisher. *Kinetic Modeling in Positron Emission Tomography*. Elsevier, 2004.
- [25] Richard Laforest, Terry L. Sharp, John A. Engelbach, Nicole M. Fetting, Pilar Herrero, Joonyoung Kim, Jason S. Lewis, Douglas J. Rowland, Yuan-Chuan Tai, and Michael J. Welch. Measurement of Input Functions in Rodents: Challenges and Solutions. *Nuclear Medicine and Biology*, 32(7):679–685, 2005.
- [26] Kazuhiro Koshino, Kazuhito Fukushima, Masaji Fukumoto, Yuki Hori, Tetsuaki Moriguchi, Tsutomu Zeniya, Yoshihiro Nishimura, Keisuke Kiso, and Hidehiro Iida. Quantification of Myocardial Blood Flow using 201Tl SPECT and Population-Based Input Function. *Annals of Nuclear Medicine*, 28(9):917–925, 2014.
- [27] Paolo Zanotti-Fregonara, Jussi Hirvonen, Chul H. Lyoo, Sami S. Zoghbi, Denise Rallis-Frutos, Marilyn A. Huestis, Cheryl Morse, Victor W. Pike, and Robert B. Innis. Population-Based Input Function Modeling for 18F FMPEP-d2, an Inverse Agonist Radioligand for Cannabinoid CB1 Receptors: Validation in Clinical Studies. *PLoS One*, 8(4):e60231, 2013.
- [28] Marine Soret, Stephen L. Bacharach, and Irene Buvat. Partial-Volume Effect in PET Tumor Imaging. *Journal of Nuclear Medicine*, 48(6):932–945, 2007.
- [29] Nuklearmedizinische Klinik und Poliklinik, Technische Universität München. *Positronen-Emissions-Tomographie (PET) bei onkologischen Fragestellungen*. https://web.archive.org/web/20070929130844/http://www.nuk.med.tu-muenchen.de/fachinfo/pdf/protokoll_pet1.pdf, 2007. Accessed on April 1st, 2019.
- [30] Anton Deitmar. *Analysis*. Springer, 2017.
- [31] Andy C. King, John Billingham, and Stephen R. Otto. *Differential Equations: Linear, Nonlinear, Ordinary, Partial*. Cambridge University Press, 2003.
- [32] Martin Brokate, Norbert Henze, Frank Hettlich, Andreas Meister, Gabriela Schranz-Kirlinger, and Thomas Sonar. *Grundwissen Mathematikstudium*. Springer, 2016.

- [33] Technische Universität Graz. *Variation der Konstanten, Wronski Determinante*. https://www.math.tugraz.at/~ganster/lv_analysis_2/20_variation_der_konstanten_wronsky.pdf, 2009. Accessed on February 22nd, 2019.
- [34] Jaan Kiusalaas. *Numerical Methods in Engineering with MATLAB*. Cambridge University Press, 2009.
- [35] Donald M. Olsson and Lloyd S. Nelson. The Nelder-Mead Simplex Procedure for Function Minimization. *Technometrics*, 17(1):45–51, 1975.
- [36] Fuchang Gao and Lixing Han. Implementing the Nelder-Mead Simplex Algorithm with Adaptive Parameters. *Computational Optimization and Applications*, 51(1):259–277, 2012.
- [37] Louis G. Birta and Arbez Gilbert. *Modelling and Simulation: Exploring Dynamic System Behaviour*. Springer, 2013.
- [38] Gregory J. Gbur. *Mathematical Methods for Optical Physics and Engineering*. Cambridge University Press, 2011.
- [39] John C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, second edition, 2016.
- [40] Kendall Atkinson, Weimin Han, and David E. Stewart. *Numerical Solution of Ordinary Differential Equations*. John Wiley & Sons, 2011.
- [41] Arieh Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge University Press, second edition, 2009.
- [42] Nuklearmedizinische Klinik und Poliklinik, Universitätsklinikum Bonn. *Ga-68-PSMA PET/CT - mehr Sensibilität für Männer mit Prostatakarzinom*. <https://www.nuklearmedizin-ukbonn.de/startseite/untersuchungen/psma-pet-ct/>, 2013–2019. Accessed on April 3rd, 2019.
- [43] Norbert Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie: Eine Einführung*. Springer, 2014.
- [44] Graeme D. Ruxton. The Unequal Variance t-Test is an Underused Alternative to Student's t-Test and the Mann-Whitney U Test. *Behavioral Ecology*, 17(4):688–690, 2006.

- [45] Michael Allwood. The Satterthwaite Formula for Degrees of Freedom in the Two-Sample t-Test. *The College Board*, 2008.
- [46] David W. Gaylor and F. N. Hopper. Estimating the Degrees of Freedom for Linear Combinations of Mean Squares by Satterthwaite's Formula. *Technometrics*, 11(4):691–706, 1969.
- [47] Robert A. Donnelly and Fatma Abdel-Raouf. *Statistics*. Idiot's guides: as easy as it gets! Alpha, third edition, first american edition, 2016.
- [48] Hans-Rudolf Schwarz and Norbert Köckler. *Numerische Mathematik*. Springer Verlag, 2013.
- [49] Roland W. Freund and Ronald W. Hoppe. *Stoer/Bulirsch: Numerische Mathematik 1*. Springer Verlag, 2007.

List of Figures

1.1	Imaging via PET/CT	5
1.2	Kinetic model for a one-compartment model	14
1.3	Kinetic model for a two-compartment model	15
1.4	Kinetic model for a three-compartment model	15
2.1	Dependence of ξ on the rate constants k_2 and k_3	24
2.2	Different scenarios for the response function $X(t)$	26
3.1	Shape of the three different input functions $B(t)$ considered in this master thesis	32
3.2	Calculation of the relative slope m_{rel} and intercept b_{rel} before and after a certain time point	36
3.3	Approach for calculating the curvature C	39
3.4	Procedure for analyzing the TAC for the existence of a \max_1	42
3.5	All five possible TAC shapes with respect to a \max_1 , a \max_2 and a fall off at the end	43
3.6	Real TACs and their input functions	54
4.1	Simulated TACs for all types of input functions $B(t)$ and for both a two- compartment and a three-compartment model	57
4.2	Dependence of the TAC shape on the rate constants k_i	58
4.3	Different impacts on the TAC shape between high and low rate constants . .	58
4.4	Example for the linear best fit functions for both k_1 and k_2 and their two best correlating mathematical properties	64
4.5	Example of an output of the algorithm	68

List of Tables

1.1	Common radionuclides in nuclear medicine with their half-lives, decay modes and decay products	7
4.1	A small selection of the "information.csv" file written by the PYTHON script	60
4.2	Correlation coefficients r_{xy} between the rate constants k_i and the calculated mathematical properties	62
4.3	All possible classifications of the data in the "information.csv" file	62
4.4	The p -values of a Student's t -test for a specific classification	63
4.5	Feasible intervals for a pointed input function	65
4.6	Feasible intervals for a blunt input function	66
4.7	Feasible intervals for a kinky input function	67