

Martina Feilke

## Estimation and model selection for dynamic biomedical images





---

# Estimation and model selection for dynamic biomedical images

Martina Feilke

---



Dissertation  
an der Fakultät für Mathematik, Informatik und Statistik  
der Ludwig-Maximilians-Universität München

vorgelegt von  
Martina Feilke

München, den 11. März 2015

Erster Berichterstatter: Prof. Dr. Volker Schmid

Zweiter Berichterstatter: Prof. Dr. Matthias Schmid

Tag der Disputation: 15. Mai 2015

# Zusammenfassung

Zur Analyse dynamischer Bilder aus Medizin und Biologie werden häufig Kompartimentmodelle verwendet. Ausgehend von diesen Modellen können Differentialgleichungen hergeleitet werden, deren Lösungen nichtlineare parametrische Funktionen darstellen, anhand derer die Veränderung bestimmter Konzentrationen über einen Zeitraum hinweg beschrieben werden kann.

Oft ist die Anzahl der Kompartimente eines Modells unbekannt. Da die Kompartimentanzahl, also die Modellkomplexität, jedoch eine wichtige Information darstellt, soll sie zusätzlich zu den unbekanntem Parametern des Modells aus den Daten geschätzt werden. Somit werden Methoden zur Parameterschätzung und zur Modellwahl benötigt. Die Methoden in dieser Dissertation sind durch zwei Anwendungen aus dem Bereich der biomedizinischen Bildgebung motiviert.

In der ersten Anwendung, der quantitativen Analyse von Fluorescence recovery after photobleaching (FRAP)-Daten, werden Kompartimentmodelle verwendet, um Einblick in das Bindungsverhalten von Molekülen in lebenden Zellen zu gewinnen. Zum einen wurde ein Bayesianisches nichtlineares gemischtes Modell für die Analyse einer Serie von FRAP-Bildern entwickelt. Für die Parameter des Modells werden mixed-effect Prioris verwendet, was ein völlig neuer Ansatz ist. Das Modell liefert neben Parameterschätzern auch Informationen über die Variabilität zwischen den betrachteten Zellkernen. Die Evaluierung der Methode erfolgte anhand von Daten aus FRAP-Experimenten, in denen eine Zellkernhälfte gebleicht wurde, im Vergleich mit Modellen ohne zufällige Effekte. Zum anderen wurde eine Methode für die Analyse von FRAP-Daten auf Pixelebene entwickelt, bei der Information aus den Nachbarpixeln für das Modell jedes Pixels eines Bildes genutzt wird. Diese Methode ist innovativ, da bisher existierende Modelle nur für die Analyse von Bildbereichen, welche mehrere Pixel umfassen, geeignet sind.

In der zweiten Anwendung, der quantitativen Analyse dynamischer kontrastmittelgestützter Magnetresonanztomographie der Brust, wird ein Kompartimentmodell verwendet, das den Austausch von Blut zwischen verschiedenen Kompartimenten beschreibt. Die Anzahl der Kompartimente eines Modells lässt hierbei auf die Heterogenität des untersuchten Tumorgewebes schließen. Es wurde eine Boosting-Methode zur Schätzung der Kompartimentanzahl sowie der Parameter eines Modells auf Voxel Ebene entwickelt. Da Boosting bisher zwar für additive Regression unter Verwendung von Glättung, jedoch nicht für nichtlineare parametrische Regression beschrieben wurde, ist das ein vollkommen neuer Ansatz. Um die räumliche Struktur eines Bildes zu berücksichtigen, werden in einer Erweiterung der Methode Abweichungen von Parameterschätzern benachbarter Voxel bestraft. Die Evaluierung der Methode erfolgte in Simulationsstudien sowie in der Anwendung auf Daten aus einer Brustkrebsstudie.

Der Großteil des Programmcodes, der in den drei Methoden verwendet wird, wurde in den Programmiersprachen R und C neu entwickelt. Darauf basierend entstanden zwei R Pakete.



# Abstract

Compartment models are a frequently used tool for imaging data gained with medical and biological imaging techniques. The solutions of the differential equations derived from a compartment model provide nonlinear parametric functions, based on which the behavior of a concentration of interest over time can be described.

Often, the number of compartments in a compartment model is unknown. As the model complexity itself, which is, the number of compartments, is certainly an important information, it is desirable to estimate it from the observed data. Additionally, the unknown parameters have to be estimated. Therefore, methods dealing with both the parameter estimation and model selection in compartment models are needed. The methods proposed in this thesis are motivated by two applications from the field of medical and biological imaging.

In the first application, the quantitative analysis of Fluorescence recovery after photobleaching (FRAP) data, compartment models are used in order to gain insight into the binding behavior of molecules in living cells. As a first approach, we developed a Bayesian nonlinear mixed-effects model for the analysis of a series of FRAP images. Mixed-effect priors are defined on the parameters of the nonlinear model, which is a novel approach. With the proposed model, we get parameter estimates and additionally gain information about the variability between nuclei, which has not been studied so far. The proposed method was evaluated on half-nucleus FRAP data, also in comparison with different kinds of fixed-effects models. As a second approach, a pixelwise analysis of FRAP data is proposed, where information from the neighboring pixels is included into the nonlinear model for each pixel. This is innovative as the existing models are suitable for the analysis of FRAP data for some regions of interest only.

For the second application, the quantitative analysis of dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) of the breast, we use a compartment model which describes the exchange of blood between different, well-mixed compartments. In the analysis of such data, the number of compartments allows conclusions about the heterogeneity of cancerous tissue. Therefore, an estimation and model selection approach based on boosting, with which the number of compartments and the unknown parameters can be estimated at the voxel level, is proposed. In contrast to boosting for additive regression, where smoothing approaches are used, boosting in nonlinear parametric regression as described in this thesis is a novel approach. In an extension of this approach, the spatial structure of an image is taken into account by penalizing the differences in the parameter estimates of neighboring voxels. The evaluation of the method was done in simulation studies, as well as in the application to data from a breast cancer study.

The majority of the program code used in the three approaches was newly developed in the programming languages R and C. Based on that code, two R packages were built.



# Danksagung

There are more people involved in telling a story than the writer.

---

(Cecelia Ahern, One Hundred Names)

Diese Worte treffen auch auf die vorliegende Dissertation zu. Deshalb bedanke ich mich an dieser Stelle herzlich bei allen, die mich bei der Arbeit an dieser Dissertation unterstützt haben.

Besonderer Dank gilt Volker Schmid, meinem Doktorvater, der mir diese Dissertation ermöglicht hat. Ich danke ihm für die Betreuung während meiner Zeit als wissenschaftliche Mitarbeiterin und Doktorandin in der AG Bioimaging, seine Unterstützung und viele hilfreiche Gespräche.

Vielen Dank an Matthias Schmid dafür, dass er sich als Zweitgutachter für diese Dissertation bereit erklärt hat.

Jan Gertheiss, Bernd Bischl und Katrin Schneider danke ich für die ergiebige, angenehme und freundschaftliche Zusammenarbeit, aus der zwei Kapitel dieser Dissertation entstanden sind.

Meinen Kollegen am Institut für Statistik – zu nennen sind hier insbesondere Linda Schulze Waltrup, Gunther Schauburger, Verena Maier und Stephanie Thiemichen – und meinen Freunden außerhalb des Instituts danke ich für die schöne Zeit, die wir gemeinsam verbracht haben.

Julia Sommer danke ich für die schöne Zeit, in der wir ein Büro geteilt haben. Sie hat mir durch die Beantwortung zahlreicher Fragen und ihre Unterstützung einen angenehmen Einstieg als Doktorandin in der AG Bioimaging ermöglicht.

Ich bedanke mich bei Ramona Feilke und Clara Happ für das Korrekturlesen von Teilen dieser Dissertation und bei Marianne Feilke für die Unterstützung bei der Erstellung von Grafiken für diese Dissertation.

Ich danke Anwar Padhani, Heinrich Leonhardt und Lothar Schermelleh für die Bereitstellung von Daten, welche die Basis für die Anwendungen in dieser Dissertation darstellen.

Der DFG danke ich für die Finanzierung eines Großteils meiner Arbeit.

Bei Alexander Kopp bedanke ich mich herzlich für seinen Beistand während der Entstehung dieser Dissertation.

Meinen Eltern Peter und Marianne Feilke, meiner Schwester Ramona Feilke und meinem Schwager Christian Schwalb danke ich dafür, dass sie mich immer bedingungslos unterstützen. In meiner Familie – wobei ich hier insbesondere auch meine Großeltern sowie meine Cousinen und Cousins erwähnen möchte – finde ich stets den nötigen Rückhalt.

München, im Mai 2015

Martina Feilke



# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b>  |
| 1.1      | Outline . . . . .   | 3         |
| 1.2      | Contributions . . . . .   | 4         |
| 1.3      | Software . . . . .  | 5         |
| <b>2</b> | <b>Compartment models and differential equations</b>                            | <b>7</b>  |
| 2.1      | Introduction to compartment models . . . . .                                    | 8         |
| 2.2      | A general compartment model . . . . .   | 8         |
| 2.3      | Compartment models in this thesis . . . . .                                     | 9         |
| <b>3</b> | <b>Nonlinear regression</b>   | <b>15</b> |
| 3.1      | Introduction to regression analysis . . . . .                                   | 16        |
| 3.2      | Regression models in this thesis . . . . .                                      | 17        |
| <b>4</b> | <b>Bayesian data analysis</b>   | <b>19</b> |
| 4.1      | Bayesian models in this thesis . . . . .  | 20        |
| 4.2      | Prior distributions . . . . .   | 20        |
| 4.3      | The Bayesian principle . . . . .  | 21        |
| 4.4      | Introduction to MCMC inference . . . . .  | 21        |
| <b>5</b> | <b>Bayesian mixed-effects model for the analysis of a series of FRAP images</b> | <b>23</b> |
| 5.1      | Introduction . . . . .  | 25        |
| 5.2      | FRAP experiments . . . . .  | 26        |
| 5.2.1    | Data . . . . .  | 26        |
| 5.2.2    | Nonlinear recovery model . . . . .  | 27        |
| 5.3      | Bayesian nonlinear mixed-effects model . . . . .                                | 31        |
| 5.3.1    | Data model . . . . .  | 31        |
| 5.3.2    | Prior model . . . . .   | 31        |
| 5.3.3    | Hyper prior model . . . . .   | 32        |
| 5.3.4    | Posterior distribution and MCMC inference . . . . .                             | 33        |
| 5.4      | Results . . . . .   | 37        |
| 5.4.1    | Mixed-effects model . . . . .   | 37        |
| 5.4.2    | Comparison between the mixed-effects model and the fixed-effects models         | 41        |
| 5.5      | Discussion . . . . .  | 48        |

|          |  |            |
|----------|--|------------|
| <b>6</b> | <b>Analysis of FRAP images at the pixel level</b>                                  | <b>51</b>  |
| 6.1      | Introduction . . . . .   | 53         |
| 6.2      | Nonlinear recovery model . . . . .   | 53         |
| 6.3      | Bayesian nonlinear model . . . . .   | 56         |
| 6.3.1    | Data model . . . . .   | 57         |
| 6.3.2    | Prior model . . . . .  | 57         |
| 6.3.3    | Hyper prior model . . . . .  | 58         |
| 6.3.4    | Posterior distribution and MCMC inference . . . . .                                | 58         |
| 6.4      | Simulation . . . . .   | 61         |
| 6.5      | Conclusions and Discussion . . . . .   | 62         |
| <b>7</b> | <b>Boosting in nonlinear regression models with an application to DCE-MRI data</b> | <b>65</b>  |
| 7.1      | Introduction . . . . .   | 67         |
| 7.2      | Methods . . . . .  | 69         |
| 7.2.1    | Compartment model . . . . .  | 69         |
| 7.2.2    | Nonlinear regression model . . . . .   | 71         |
| 7.2.3    | Gradient boosting . . . . .  | 71         |
| 7.2.4    | Boosting algorithm . . . . .   | 72         |
| 7.2.5    | Voxelwise estimation . . . . .   | 76         |
| 7.2.6    | Spatially regularized estimation . . . . .   | 76         |
| 7.2.7    | Refit . . . . .  | 77         |
| 7.2.8    | Competing Methods . . . . .  | 77         |
| 7.3      | Simulation studies . . . . .   | 78         |
| 7.3.1    | Simulation setup . . . . .   | 78         |
| 7.3.2    | Results . . . . .  | 79         |
| 7.4      | Application to DCE-MRI data . . . . .  | 80         |
| 7.4.1    | Description of the data . . . . .  | 80         |
| 7.4.2    | Results . . . . .  | 82         |
| 7.5      | Assessment of therapy success . . . . .  | 91         |
| 7.6      | Conclusions and Discussion . . . . .   | 91         |
| <b>8</b> | <b>Conclusion</b>  | <b>95</b>  |
| 8.1      | Summary . . . . .  | 96         |
| 8.2      | Discussion . . . . .   | 97         |
| 8.3      | Outlook . . . . .  | 99         |
| <b>A</b> | <b>Appendix to Chapter 5</b>   | <b>101</b> |
| <b>B</b> | <b>Appendix to Chapter 6</b>   | <b>105</b> |
| <b>C</b> | <b>Appendix to Chapter 7</b>   | <b>107</b> |
|          | <b>Bibliography</b>  | <b>137</b> |

# **Chapter 1**

---

## **Introduction**

---

In medicine and biology, various *in vivo* imaging techniques are used today. Among established clinical imaging modalities are for example ultrasound, computed tomography, magnetic resonance imaging (MRI), and positron emission tomography. In the field of biology, fluorescence microscopy techniques are commonly used in order to visualize the behavior of molecules of interest in living cells.

This thesis focuses on the analysis of dynamic biomedical images, i.e., images produced by imaging techniques in medicine and biology that show for example some tissue of interest or some cell nucleus of interest over time. The two applications that motivated the development of the novel statistical methods presented in this thesis are Fluorescence recovery after photobleaching (FRAP) and dynamic contrast-enhanced (DCE)-MRI. Data from FRAP experiments on nuclei of living cells can be analyzed in order to investigate the binding dynamics of molecules of interest. With DCE-MRI, medical tracer experiments can be done by which the uptake of tracer in body regions of interest can be observed. The DCE-MR images analyzed in this thesis were recorded in the framework of a breast cancer study. Thus, the methods presented in this thesis promote image analysis in fluorescence microscopy and oncology.

The analysis of dynamic biomedical images in this thesis is based on compartment models representing the exchange of fluorescent molecules between different compartments in the FRAP application and the exchange of tracer between different compartments in the DCE-MRI application. These compartment models can be described by ordinary differential equations (ODEs). The solution of such ODEs is nonlinear and a function of kinetic parameters and time, and a nonlinear parametric regression model can be formulated based upon it. As the only observable in the regression model is time, the kinetic parameters describing the exchange of fluorescent molecules or tracer between the compartments as well as parameters related to the size of the compartments have to be estimated. Moreover, often, the number of compartments of a compartment model is unknown and has to be estimated in addition. Therefore, novel methods dealing with both the parameter estimation and model selection in compartment models are presented in this thesis.

For the FRAP application, two different methods for parameters estimation are proposed. The first method is a nonlinear mixed-effects model for the analysis of a series of FRAP images, with which parameter estimates are obtained. The definition of mixed-effect priors on the parameters of a nonlinear model in a Bayesian framework as done in this thesis is a novel approach. With this method, in addition to the parameter estimates, information about the variability between several similar cell nuclei is gained, which has not been studied so far. As a second method, a nonlinear model at the pixel level is proposed into which information from the neighboring pixels is included. The analysis of FRAP data at the pixel level is innovative, as FRAP data are usually analyzed for some regions of interest which comprise lots of pixels. For the DCE-MRI application, an estimation and model selection method based on boosting is proposed. With this method, the number of compartments and the unknown parameters can be estimated at the voxel level. The spatial structure of an image is additionally taken into account in an extension of this method, where differences in the parameter estimates of neighboring voxels are penalized. To the best of our knowledge, boosting used in nonlinear parametric regression is described for the first time in Feilke, Bischl, Schmid, and Gertheiss (2015) and in the corresponding chapter of this thesis.

## 1.1 Outline

This thesis is organized as follows. In Chapter 2, an introduction to compartment models and differential equations is given. A selection of application areas including the ones in this thesis is presented. Moreover, a general compartment model and the common structure of the compartment models in this thesis as well as assumptions commonly made in compartmental modeling are introduced. Chapter 3 provides an introduction to nonlinear regression. It starts with a general introduction to regression analysis and then introduces the common characteristics of the nonlinear regression models in this thesis. Moreover, approaches for the parameter estimation in linear and nonlinear regression models are discussed. Chapter 4 gives a short introduction to Bayesian data analysis. The common structure of the Bayesian models in this thesis is described. This is followed by a section about prior distributions in general and in this thesis. Moreover, the Bayesian principle as well as Markov chain Monte Carlo inference are introduced. In Chapter 5, a Bayesian nonlinear mixed-effects model for the analysis of a series of FRAP images, i.e., for FRAP images of several similar cell nuclei, is proposed. After a short introduction, the data analyzed in the application in this chapter is described. Then, the nonlinear recovery model that is the basis of the Bayesian nonlinear mixed-effects model is introduced together with the Bayesian model itself. This is followed by a presentation of the results of the evaluation of the proposed model using FRAP data. The chapter ends with a discussion. In Chapter 6, a Bayesian nonlinear model for the analysis of FRAP images at the pixel level is suggested. The introduction is followed by a presentation of the Bayesian model and the nonlinear recovery model that serves as a basis for it. After that, the setup of a simulation study for the evaluation of the Bayesian nonlinear model is described. The chapter closes with conclusions and a discussion. In Chapter 7, boosting in nonlinear regression models for the analysis of DCE-MRI data is proposed. After an introduction, the compartment model and the nonlinear regression model derived from it are described. Thereafter, gradient boosting is introduced. This is followed by an introduction of the novel boosting algorithm together with two estimation procedures, voxelwise and spatially regularized estimation, as well as an introduction of the refit procedure and the competing methods that are used for the evaluation of the new boosting algorithm. Subsequently, the evaluation of the boosting approach in comparison with the competing methods in the framework of a simulation study as well as in the application to data from a breast cancer study is described. After that, there is a section focusing on how the proposed method can be used for the assessment of therapy success in a breast cancer study. The chapter ends with conclusions and a discussion. Chapter 8 gives a summary of the thesis, as well as a discussion on some general topics concerning the applications in this thesis, and an outlook.

## 1.2 Contributions

Parts of this thesis are published or submitted as peer-reviewed journal articles. The outline given below lists the titles of the articles, gives a short summary of the content, and highlights the contributions of the authors and additional contributors.

- Chapter 5 is mainly based on

Feilke, M., K. Schneider, and V.J. Schmid (2015). Bayesian mixed-effects model for the analysis of a series of FRAP images. *Statistical Applications in Genetics and Molecular Biology* 14(1), 35-51. (Feilke, Schneider, and Schmid, 2015)

In this manuscript, we suggest the usage of a hierarchical Bayesian mixed-effects model for the analysis of a series of FRAP images, i.e., for FRAP images of several similar cell nuclei. Martina Feilke and Volker J. Schmid developed the model. Martina Feilke conducted all analyses and drafted the manuscript. Volker J. Schmid and Katrin Schneider revised the manuscript. Heinrich Leonhardt (Biocenter Martinsried, LMU Munich) and Lothar Schermelleh (Department of Biochemistry, University of Oxford) provided the data. Joseph W. Sakshaug proofread the manuscript, as he is a native English speaker. Martina Feilke and Volker J. Schmid were supported by Deutsche Forschungsgemeinschaft (DFG SCHM 2747/1-1). Katrin Schneider was supported by the International Max Planck Research School for Molecular and Cellular Life Sciences (IMPRS-LS).

- Chapter 7 is mainly based on

Feilke, M., B. Bischl, V.J. Schmid, and J. Gertheiss. Boosting in nonlinear regression models with an application to DCE-MRI data. *Methods of Information in Medicine*. Accepted for publication on May 26, 2015. (Feilke, Bischl, Schmid, and Gertheiss, 2015)

In this manuscript, we introduce boosting in nonlinear regression models for the analysis of DCE-MRI data. Martina Feilke, Jan Gertheiss, and Volker J. Schmid developed the method. Martina Feilke conducted all analyses and drafted the manuscript. Bernd Bischl supported the analyses by providing R code and technical advice for the benchmark study in Chapter 7 and the parallelization of R code. Volker J. Schmid, Jan Gertheiss, and Bernd Bischl revised the manuscript. Anwar Padhani, PSSC, Mount Vernon Hospital, Northwood, U.K, provided the clinical data. Martina Feilke and Volker J. Schmid were supported by Deutsche Forschungsgemeinschaft (DFG SCHM 2747/1-1).

## 1.3 Software

The majority of the program code used for the analyses in this thesis was newly developed in the programming languages R (R Core Team, 2013) and C. The R packages used for the various analyses are given in each chapter. Two R packages were developed within the framework of this thesis:

- `frapmm`: Bayesian nonlinear mixed-effects model for a series of FRAP images.
- `dcmriboost`: Boosting for DCE-MRI data.

They are both available at <https://github.com/feilke>. Version 1.0 of the R packages was used for the applications in this thesis. With the R package `frapmm`, a nonlinear mixed model can be fitted to data from a series of FRAP images, i.e., FRAP images of various similar cell nuclei, as described in Chapter 5 of this thesis. An accompanying file that is also available at GitHub provides a minimal working example including data simulation. With the R package `dcmriboost`, voxelwise and spatially regularized boosting for DCE-MRI data as described in Chapter 7 of this thesis can be performed. This package contains the simulated data for one of the simulation settings introduced in Chapter 7. Moreover, we uploaded two accompanying R files to GitHub that provide code for performing voxelwise and spatially regularized boosting for exactly this simulation setting.



## **Chapter 2**

---

# **Compartment models and differential equations**

---

This chapter gives an introduction to compartment models and differential equations. As in many fields of research, compartment models are commonly used to describe processes of interest, a selection of application areas – including those motivating the methods proposed in this thesis – is presented in the first section. In the following sections, a general compartment model and the common structure of the compartment models in this thesis as well as assumptions usually made in compartmental modeling are introduced.

## 2.1 Introduction to compartment models

Compartment models can for example be used to describe the drug absorption in pharmacokinetic studies (see for example Savic et al., 2007; Yu and Amidon, 1999), and they are also applied in epidemiology, for example to describe transmission processes. Among others, Lipsitch et al. (2000) use a compartment model to represent the bacterial transmission within a hospital, and Legrand et al. (2008) model the dynamics of tuberculosis in prison based on a compartment model. In physiology, compartment models are for example used to model the exchange of oxygen between organs (for example in Jacquez, 1972) or the exchange of pulmonary nitric oxide (for example in Tsoukias and George, 1998). Compartment models are moreover applied in the social sciences (for example in Herbst, 1963) and with ecological systems, where the exchange of energy is described (for example in Eriksson, 1971).

In this thesis, the focus is on compartment models used in the analysis of data from the field of biomedical imaging. We use compartmental modeling for the analysis of data from biochemical experiments that are done in order to investigate the binding dynamics of molecules of interest in living cells (Sprague and McNally, 2005). These experiments are done with the FRAP technique. Moreover, we use compartmental modeling for the analysis of data from medical tracer experiments. In these experiments, the uptake of tracer in body regions of interest is observed using imaging techniques such as MRI. We focus on a DCE-MRI application, where the exchange of tracer between different compartments is modeled.

## 2.2 A general compartment model

A compartment model is a tool that is used in order to model a dynamic system consisting of a certain number of compartments exchanging material with each other. In addition to the exchange of material between the compartments, there may also be input of material from the outside into some compartments and excretion of material from some compartments to the outside environment (Anderson, 1983). In general, if there is no input of material from the outside and no excretion of material to the outside, a system is called a closed system.

According to Anderson (1983), the general compartment model is a compartment model with  $i$  compartments,  $i = 1, \dots, n$ , which can be described by the following differential equation:

$$\frac{d}{dt}q_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^n f_{ij}(q(t), t, \alpha)q_j(t) - \sum_{\substack{j=0 \\ j \neq i}}^n f_{ji}(q(t), t, \alpha)q_i(t) + I_i(t). \quad (2.1)$$

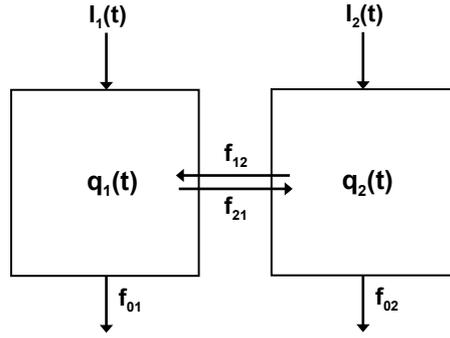


Figure 2.1: General two compartment model.

We use the same notation as in Anderson (1983). The quantity of material in compartment  $i$  at time  $t$  is nonnegative and is denoted by  $q_i(t)$ .  $f_{ij}q_j$  is the rate of transfer of material from compartment  $j$  to compartment  $i$ , where  $i \neq j$ .  $f_{ij} \geq 0$  is called ‘fractional transfer coefficient’. It may be a function of  $q = (q_1, \dots, q_n)'$ , time  $t$ , and a vector of parameters  $\alpha = (\alpha_1, \dots, \alpha_\nu)'$ . The rate of input of material from the outside into compartment  $i$  is denoted by  $I_i(t)$ , and  $f_{0i}q_i$  is the rate of excretion of material from compartment  $i$  to the outside, with  $f_{0i}$  the fractional excretion coefficient.

In Figure 2.1, a block diagram for a general two compartment model, which is a special case of the general compartment model, is given. In this model, there are two compartments that exchange material with each other with fractional transfer coefficients  $f_{12}, f_{21}$ . For both compartments, there can additionally be input of material from the outside ( $I_1(t), I_2(t)$ ) as well as excretion of material to the outside environment ( $f_{01}, f_{02}$ ).

## 2.3 Compartment models in this thesis

In the applications in this thesis, we assume that the fractional transfer coefficients are constant over time, i.e.,  $f_{ij}(q(t), t, \alpha) = f_{ij}$  and  $f_{ji}(q(t), t, \alpha) = f_{ji}$ , and therefore also call them ‘rate constants’. With that, Equation (2.1) is simplified to

$$\frac{d}{dt}q_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^n f_{ij}q_j(t) - \sum_{\substack{j=0 \\ j \neq i}}^n f_{ji}q_i(t) + I_i(t), \quad (2.2)$$

which is also called a nonhomogeneous linear system of ODEs with constant coefficients (Heuser, 1995). With

$$f_{ii} := - \sum_{\substack{j=0 \\ j \neq i}}^n f_{ji},$$

the system in Equation 2.2 can also be written as

$$\begin{aligned} \frac{d}{dt}q_1(t) &= f_{11}q_1(t) + f_{12}q_2(t) + \dots + f_{1n}q_n(t) + I_1(t) \\ \frac{d}{dt}q_2(t) &= f_{21}q_1(t) + f_{22}q_2(t) + \dots + f_{2n}q_n(t) + I_2(t) \\ &\vdots \\ \frac{d}{dt}q_n(t) &= f_{n1}q_1(t) + f_{n2}q_2(t) + \dots + f_{nn}q_n(t) + I_n(t). \end{aligned} \quad (2.3)$$

When formulating the system in 2.3 in matrix notation, we get

$$\frac{d}{dt}q(t) = Fq(t) + I(t). \quad (2.4)$$

By solving the system in Equation 2.4 with boundary condition  $q(0) = q_0$ , where  $q_0$  is fixed, what is equivalent to solving the system of ODEs given in Equations 2.2 and 2.3 with boundary conditions  $q_i(0) = q_{i0}$ ,  $i = 1, \dots, n$ , where  $q_{i0}$  are fixed values, a unique solution of the form

$$q(t) = \exp(Ft)q_0 + \int_0^t \exp(F(t-\tau))I(\tau)d\tau \quad (2.5)$$

is obtained (Heuser, 1995).

As explained in Sommer (2013), the spectral decomposition  $F = S\Lambda S^{-1}$  has to be computed in order to compute the matrix valued exponentials in Equation 2.5, where  $S$  consists of eigenvectors of  $F$  and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  is a diagonal matrix containing the corresponding eigenvalues. According to Seber and Wild (1989), the following equation applies for the integral in Equation 2.5:

$$\int_0^t \exp(F(t-\tau))I(\tau)d\tau = S \int_0^t \exp(\Lambda(t-\tau))S^{-1}I(\tau)d\tau,$$

with

$$\left[ \int_0^t \exp(\Lambda(t-\tau))S^{-1}I(\tau)d\tau \right]_j = \sum_{i=1}^n s^{ji} \int_0^t \exp(\lambda_j(t-\tau))I_i(\tau)d\tau = \sum_{i=1}^n s^{ji}(I_i * \exp(\lambda_j t)), \quad (2.6)$$

where  $s^{ji}$  are the elements of the matrix  $S^{-1}$ . In Equation 2.6,  $*$  denotes the convolution operator, for which  $(f * g)(t) = \int_0^t f(\tau)g(t-\tau)d\tau$  applies. The exponential rates in this equation are the eigenvalues of the connectivity matrix  $F$  (Sommer, 2013).

In this thesis, special cases of the compartment model in Equation 2.2 are considered. A block diagram showing the common structure of these special cases is given in Figure 2.2. We assume that there is one base compartment which exchanges material with up to  $n-1$  other compartments. At time  $t$ , the base compartment contains the quantity of material  $q_1(t)$  and the other compartments contain the quantities of material  $q_2(t), \dots, q_n(t)$ . The exchange between the base compartment and the other compartments happens with rate constants  $f_{1k}$  and  $f_{k1}$ ,  $k =$

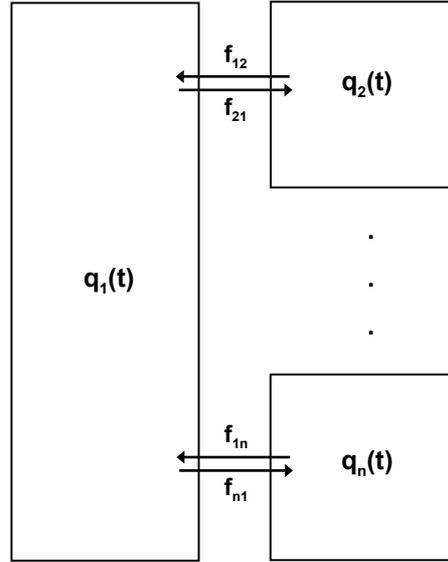


Figure 2.2: Common structure of the compartment models in this thesis.

$2, \dots, n$ , and depends on the current quantities of material in the compartments. Input of material from the outside is optional, and is therefore not depicted in the block diagram in Figure 2.2.

The change of the quantity of material in compartment  $k$ ,  $k = 2, \dots, n$ , can be described by the differential equation

$$\frac{d}{dt}q_k(t) = f_{k1}q_1(t) - f_{1k}q_k(t), \quad (2.7)$$

which is a special case of Equation 2.4 with  $F = \text{diag}(-f_{12}, \dots, -f_{1n})$  and the base compartment playing the role of an external source of material with  $I(t) = (f_{21}q_1(t), \dots, f_{n1}q_1(t))'$ . With boundary condition  $q_k(0) = q_{k0}$ , the solution of Equation 2.7 is

$$q_k(t) = q_{k0} \exp(-f_{1k}t) + f_{k1}q_1 * \exp(-f_{1k}t). \quad (2.8)$$

For the two imaging applications in this thesis, FRAP and DCE-MRI, however, it is not possible to observe the quantity of material for each compartment individually. We can only observe the total quantity of material  $q(t)$ , that is, the sum of the quantities of material in all compartments.

For the FRAP application in Chapter 5 of this thesis, the sum of the quantities of material in all compartments in the notation used above is

$$q(t) = \sum_{i=1}^n q_i(t). \quad (2.9)$$

Here,  $q_1(t)$  denotes the quantity of material in the base compartment, and the quantities of material in the remaining compartments are denoted by  $q_k(t)$ ,  $k = 2, \dots, n$ . In this application, the observed and modeled quantity is the total fluorescence intensity in the bleached area of a

cell nucleus. It is denoted by  $\text{total}(t)$  and corresponds to  $q(t)$  in Equation 2.9.  $\text{total}(t)$  is the sum of the concentration of free unbleached molecules in the base compartment denoted by  $f(t)$  and the concentration of bound unbleached molecules in the remaining  $k$  compartments  $a_k(t)$ ,  $k = 0, \dots, K$ , and an error.  $f(t)$  corresponds to  $q_1(t)$ , and  $a_k(t)$  to the remaining quantities of material  $q_2(t), \dots, q_n(t)$  in Equation 2.9.

As in the FRAP application in Chapter 5, the concentration of free unbleached molecules in the base compartment  $f(t)$  is assumed to be constant over time, the convolution in Equation 2.8 is simple, and the total quantity of material  $\text{total}(t)$  can be described by a term containing a sum of exponentials instead of a sum of convolved exponentials (Sommer, 2013).

For the application in Chapter 6 of this thesis, the observed and modeled quantity is the total fluorescence intensity in each pixel of an image of a cell nucleus in a FRAP experiment. This quantity is again denoted by  $\text{total}(t)$  and corresponds to  $q(t)$  in Equation 2.9. In the application in Chapter 6, similar to the application in Chapter 5, it is the sum of the concentration of free unbleached molecules in the base compartment  $f(t)$  and the concentration of bound unbleached molecules in a second compartment  $a(t)$ , plus an error, where  $f(t)$  corresponds to  $q_1(t)$ , and  $a(t)$  to  $q_2(t)$  (with  $n = 2$ ) in Equation 2.9. In contrast to the application in Chapter 5 we, however, do not assume that  $f(t)$  is constant, and therefore, we have to deal with convolved exponentials in Chapter 6.

For the DCE-MRI application in Chapter 7 of this thesis, the fractional volumes of the compartments have to be considered in addition to the quantities of material in the compartments.

The fractional volume of the base compartment denoted by  $v_1$  is assumed to be negligible, and is therefore set zero. Hence, the total quantity of material  $q(t)$  in the general notation used above is

$$q(t) = \sum_{k=2}^n v_k q_k(t), \quad (2.10)$$

where  $v_k$  is the fractional volume of the  $k$ -th compartment for  $k = 2, \dots, n$ . The observed quantity in this application is the total concentration of contrast agent in a voxel of an DCE-MR image. It is denoted by  $C(t)$ , what corresponds to  $q(t)$  in Equation 2.10. It is the sum of the products of the fractional volumes of the tissue compartments  $v_{e_k}$ ,  $k = 1, \dots, q$ , and the concentrations of contrast agent in the tissue compartments  $C_{e_k}(t)$ , where  $v_{e_k}$  corresponds to  $v_k$ , and  $C_{e_k}(t)$  to  $q_k(t)$  in Equation 2.10. Using Equation 2.8 and Equation 2.10, we get

$$q(t) = q_1 * \sum_{k=2}^n v_k f_{k1} \exp(-f_{1k}t) \quad (2.11)$$

and therefore, we have to deal with a sum of convolved exponentials in Chapter 7.

In each application in this thesis, the solution of the differential equations describing the used compartment model is nonlinear in the coefficients. Based on it, a nonlinear regression model can be formulated with which the data of interest can be modeled. This means, the evolution of the quantity of some material over time can be described with a known parametric function. This function depends on the structure of the compartment model, i.e., on the number of compartments and the exchange structure between the compartments, as well as on the parameters of the compartment model. Usually, the quantity of material is observed over time, and the parameters

of the model have to be estimated for a compartment model with a given structure. In Chapters 5 and 6 of this thesis, for example, compartment models with fixed numbers of compartments are assumed. However, there are cases in which the structure of a compartment model is not or only partially given *a priori*, and therefore, it has to be (finally) determined in addition to the parameter estimation. In Chapter 7 of this thesis, for example, the number of compartments in the assumed compartment model is not fixed *a priori*, and is therefore estimated from the data.

Several general assumptions are usually made in compartmental modeling. Two assumptions applying for all compartment models considered in this thesis are time invariance and the assumption of well-mixed and homogeneous compartments. If time invariance is assumed, the parameters in a compartment model are assumed to be constant during the time of data acquisition (Tofts, 1997; Sommer, 2013). The second assumption made for all compartment models considered in this thesis implies that the material in the compartments of a compartment model is well-mixed and homogeneous (Tofts, 1997; Sommer, 2013). A discussion of these two general assumptions regarding the applications in this thesis can be found in Chapter 8. More assumptions that can be made in a compartmental modeling approach that are not discussed in detail here can be found for example in Tofts (1997) and in Sommer (2013).



# **Chapter 3**

---

## **Nonlinear regression**

---

In this chapter, an introduction to regression analysis is given. In the first section, the classical linear model and the method of parameter estimation in the linear model are introduced. After that, the common characteristics of the nonlinear regression models considered in this thesis are given. This is followed by an introduction of approaches for parameter estimation in nonlinear regression models.

### 3.1 Introduction to regression analysis

In general, in a regression framework, the goal is to model the effects of one or more explanatory variables on a response variable  $y$ . A crucial point for the choice of a suitable model to do this is the type of the response variable. It can be continuous, binary, categorical, or counts (Fahrmeir et al., 2013). Another important information for the choice of a suitable model are the types of the explanatory variables, which can be continuous, binary, or categorical (Fahrmeir et al., 2013). For an extensive overview of regression models, the linear model introduced in the following as well as extensions and generalizations of this model see for example Fahrmeir et al. (2013).

In a regression model, the response variable  $y$  is assumed to be a random variable, with its distribution depending on the explanatory variables. This means that the relationship between the response variable  $y$  and the explanatory variables is not deterministic, but shows random errors.

If linear dependency of  $y$  on the explanatory variables is assumed, this results in the classical linear model. In this model, it is moreover assumed that the random errors are independent and identically normally distributed with mean zero and variance  $\sigma^2$ . With only one explanatory variable  $x$ , the linear model can be written as

$$y = \beta_0 + \beta_1 x + \varepsilon,$$

with an error term  $\varepsilon$ . If we consider observations  $i = 1, \dots, n$  of  $x$  and  $y$ , we get

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

with Gaussian error terms  $\varepsilon_i \sim N(0, \sigma^2)$ .

The unknown parameters of the model,  $\beta_0$  and  $\beta_1$ , have to be estimated for given data  $(y_i, x_i)$ . In the linear model, they are estimated according to the least squares method, i.e., the residual sum of squares (RSS)

$$RSS(\beta) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - x_i' \beta)^2$$

is to be minimized with respect to  $\beta = (\beta_0, \beta_1)'$ . This results in the least squares estimate

$$\hat{\beta} = (X'X)^{-1}X'y,$$

which is in the case of the linear model identical to the maximum likelihood estimator of  $\beta$ .

## 3.2 Regression models in this thesis

In this thesis, we focus on continuous response variables. The only explanatory variable considered in the regression models in this thesis is time, which is continuous as well. As the regression models describing the relationship between the continuous response variables and the explanatory variable time are derived from nonlinear solutions of systems of ODEs, they are nonlinear regression models. This means that a relationship that is more complex than a linear relationship has to be modeled between the response variable and the explanatory variable. But we still assume independent Gaussian error terms in the regression models.

All regression models considered in this thesis contain terms of the form

$$\alpha \exp(-\beta x),$$

i.e., nonlinear combinations of the unknown parameters  $\alpha$  and  $\beta$  that have to be estimated and a given explanatory variable  $x$ . Consider now for example a simple regression model of the form

$$y = \alpha \exp(-\beta x) + \varepsilon,$$

with response  $y$  and error term  $\varepsilon$ . If we aimed to transform this nonlinear model into a linear model, we could think about applying a logarithmic transformation to this model in order to get rid of the exponential function. However, the response variable as well as the error term would then no longer follow a Gaussian distribution. This means, in particular, that the error structure of the model would change. As an alternative, we could use the fact that in some cases where the solution of the differential equations for a compartment model contains only one exponential function, a linear solution can be obtained by integrating the differential equation describing the kinetic behavior of some material (Murase, 2004). In this thesis, however, we throughout consider models containing sums of exponentials, what corresponds to

$$y = \sum_{k=1}^n \alpha_k \exp(-\beta_k x) + \varepsilon$$

in this example. A logarithmic transformation is obviously not constructive for models of this kind in order to obtain a linear model, as it would result in a multiplicative structure rather than in an additive structure as present in a linear model. Therefore, the models presented in this thesis will not be transformed into linear models and we will have to deal with nonlinear models.

In a frequentist framework, nonlinear least squares methods are used in order to estimate the parameters in a nonlinear model. In contrast to the linear model, where the least squares estimate can be analytically determined, this is not the case for the nonlinear model. Therefore, iterative methods have to be used in order to solve the nonlinear optimization problem and thus to determine the nonlinear least squares estimate. Among the various algorithms that have been proposed for the determination of the nonlinear least squares estimate are for example the Gauss-Newton method, the Nelder-Mead method, and the Levenberg Marquart algorithm. The latter is used for the curve-fitting problem in Chapter 7 of this thesis.

However, consistency problems by specifying starting values and convergence issues can occur with algorithms for nonlinear model fitting. Therefore, the model is usually fitted several

times using for example random starting values or a grid of starting values. The most suitable model is then determined according to an information criterion like the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). This may result in a high computational burden, in particular if a large number of starting values is used.

To avoid the consistency problems and convergence issues mentioned above, as an alternative to the nonlinear least squares method, the parameters can also be estimated in a Bayesian framework, where the algorithm usually converges and the parameter estimates do not depend on any starting values. Bayesian inference is introduced in Chapter 4, and Bayesian approaches for the estimation of the parameters in a nonlinear model are described in Chapters 5 and 6 of this thesis.

For details on nonlinear regression and nonlinear optimization see for example Bates and Watts (1988), Ratkowsky (1990), and Seber and Wild (1989).

# **Chapter 4**

---

## **Bayesian data analysis**

---

In this chapter, the structure of the Bayesian models in the applications in Chapters 5 and 6 of this thesis is described. This is followed by a section about prior distributions, the introduction of the Bayesian principle, and an introduction to Markov chain Monte Carlo inference.

## 4.1 Bayesian models in this thesis

In the applications in Chapters 5 and 6 of this thesis, we use hierarchical Bayesian models consisting of three levels:

1. Data model; derived from a nonlinear recovery model,
2. Prior model; prior assumptions on all unknown parameters in level 1,
3. Hyper prior model; prior assumptions on all unknown parameters in level 2.

This structure reflects that in the Bayesian framework probability density functions (PDFs) have to be assigned not only to the observables of a statistical model, but to all unknown parameters in a model, as there is no fundamental distinction between parameters and observables of a statistical model (Gilks et al., 1996; Carlin and Louis, 2009).

## 4.2 Prior distributions

The choice of the prior distribution for all unknown parameters in a model is a fundamental task in the Bayesian framework. In Chapters 5 and 6 of this thesis, we assume that the data follows a Gaussian distribution. The priors for the remaining unknown parameters have to be chosen according to the prior knowledge we have about these parameters.

Two commonly used types of prior distributions, which are not contradictory, are conjugate priors and uninformative priors (Carlin and Louis, 2009). If a prior is conjugate with another distribution, it means that the prior distribution belongs to the same distributional family as the posterior distribution (Carlin and Louis, 2009). Uninformative priors are used in cases where we have no reliable prior information about a parameter (Carlin and Louis, 2009).

In the application in Chapter 5 of this thesis, mixed-effect priors are defined on the parameters of a nonlinear model. A Gaussian distribution is defined on the random effect of one of the parameters that may take negative as well as positive values. The log-normal distribution can generally be used for a parameter in cases where we know that the parameter must be non-negative. Therefore, this distribution is used for the random-effects of the transfer rates in the application in Chapter 5 and for the transfer rates in the application in Chapter 6 of this thesis. In Chapter 5, we assume Inverse Gamma distributions for the variance parameters in level 3 of the hierarchical model, as this distribution is known as a conjugate prior for the Gaussian distribution with known mean. Uniform priors are used for the fixed effects in Chapter 5 and for the parameters describing the concentration of free and bound unbleached molecules at the time of the bleaching of a cell nucleus in Chapter 6. In Chapter 5, the uniform prior distributions can be regarded as uninformative, which means that they do not contain any relevant information. In

Chapter 6, they incorporate the information that the concentration of free and bound unbleached molecules at the time of the bleaching of a cell nucleus can each be at most equal to the sum of these concentrations.

### 4.3 The Bayesian principle

In the Bayesian framework, all conclusions are drawn from the posterior distribution (Gilks et al., 1996), which can be computed via Bayes' theorem (Carlin and Louis, 2009):

$$p(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\tilde{\theta})\pi(\tilde{\theta})d\tilde{\theta}},$$

with  $f(y|\theta)$  the PDF of the data distribution, and  $\pi(\theta)$  the product of the prior distributions. In the applications in Chapters 5 and 6 of this thesis, we assume *a priori* independence of all unknown parameters in a model. As in most cases the posterior distribution is only known up to the normalization constant  $\int f(y|\tilde{\theta})\pi(\tilde{\theta})d\tilde{\theta}$ , for example Markov chain Monte Carlo (MCMC)-methods have to be applied in order to derive the full conditional distributions (FCs) of the parameters in a model from the posterior distribution (Carlin and Louis, 2009).

### 4.4 Introduction to MCMC inference

In the applications in Chapters 5 and 6 of this thesis, hybrid MCMC algorithms with Gibbs- and Metropolis-Hastings (MH)-update steps are used in order to obtain samples from the FCs of the parameters in a model. If the FC of a parameter  $\theta$  is a familiar distribution, as for example the Gaussian distribution or the gamma distribution, the parameter can be sampled directly from this distribution. That means, the proposal distribution for the update of the parameter  $\theta$  is exactly its FC, and the proposed value is accepted with an acceptance probability of 1 (Gilks et al., 1996). Otherwise, an MH algorithm has to be used. Then, a proposal distribution for the update of the parameter  $\theta$ , that should preferably be similar to the FC of this parameter, has to be chosen (Schmid, 2004). Let  $p(\theta | \cdot)$  denote the full conditional distribution of  $\theta$ . Then, the MH algorithm for the update of  $\theta$  with a notation similar to the notation used in Schmid (2004) is given in Figure 4.1.

For the MH algorithm, the proposal distribution  $q(\theta^* | \theta^{(n-1)})$  has to be specified. In the applications in Chapters 5 and 6 of this thesis, random walk proposals of the form  $q(\theta^* | \theta^{(n-1)}) = N(\theta^{(n-1)}, \sigma_\theta^2)$  are used. That means, the proposal  $\theta^*$  is drawn from a Gaussian distribution with mean  $\theta^{(n-1)}$  and variance  $\sigma_\theta^2$ . As the proposal distribution is symmetric in that case, i.e.,  $q(\theta^{(n-1)} | \theta^*) = q(\theta^* | \theta^{(n-1)})$ , the acceptance probability in step 4 of the MH algorithm in Figure 4.1 reduces to

$$\alpha = \min \left\{ \frac{p(\theta^* | \cdot)}{p(\theta^{(n-1)} | \cdot)}, 1 \right\}.$$

1. Choose a starting value  $\theta^{(0)}$  and the number of iterations  $N$ .
2. Set  $n = 1$ .
3. Draw a proposal  $\theta^*$  according to the proposal distribution  $q(\theta^* | \theta^{(n-1)})$ .
4. Accept  $\theta^*$  with acceptance probability

$$\alpha = \min \left\{ \frac{p(\theta^* | \cdot)q(\theta^{(n-1)} | \theta^*)}{p(\theta^{(n-1)} | \cdot)q(\theta^* | \theta^{(n-1)})}, 1 \right\},$$

otherwise set  $\theta^{(n)} = \theta^{(n-1)}$ .

5. Set  $n = n + 1$ .
6. Iterate steps 3-5 until  $n = N$ , the number of iterations.

Figure 4.1: Metropolis-Hastings algorithm.

This is a special case of the MH algorithm which is called Metropolis algorithm (Carlin and Louis, 2009). For more details on MCMC inference see for example Gilks et al. (1996) or Brooks et al. (2011).

# Chapter 5

---

## Bayesian mixed-effects model for the analysis of a series of FRAP images

---

---

This chapter is mainly based on Feilke, Schneider, and Schmid (2015). Some of the Tables and Figures of the Appendix in Feilke, Schneider, and Schmid (2015) are additionally contained in this chapter and the content of this chapter is accordingly changed at some points compared to Feilke, Schneider, and Schmid (2015). Some parts that are common for this chapter and Chapter 6 are explained in Chapter 4 of this thesis and therefore, the content of this chapter is at some points accordingly changed compared to Feilke, Schneider, and Schmid (2015). This chapter moreover contains more detailed information on data preparation (in Section 5.2) than Feilke, Schneider, and Schmid (2015) and additionally provides information about an R package that was built based on the code used for the analyses in this chapter.

### Abstract

The binding behavior of molecules in nuclei of living cells can be studied through the analysis of images from fluorescence recovery after photobleaching experiments. However, there is still a lack of methodology for the statistical evaluation of FRAP data, especially for the joint analysis of multiple dynamic images.

We propose a hierarchical Bayesian nonlinear model with mixed-effect priors based on local compartment models in order to obtain joint parameter estimates for all nuclei as well as to account for the heterogeneity of the nuclei population. We apply our method to a series of FRAP experiments of DNA methyltransferase 1 tagged to green fluorescent protein expressed in a somatic mouse cell line and compare the results to the application of three different fixed-effects models to the same series of FRAP experiments.

With the proposed model, we get estimates of the off-rates of the interactions of the molecules under study together with credible intervals, and additionally gain information about the variability between nuclei. The proposed model is superior to and more robust than the tested fixed-effects models. Therefore, it can be used for the joint analysis of data from FRAP experiments on various similar nuclei.

## 5.1 Introduction

Fluorescence recovery after photobleaching is an imaging technique to investigate the binding behavior of molecules inside organisms, cells or cellular sub-compartments *in vivo* (Sprague et al., 2004; Meyvis et al., 1999; Reits and Neefjes, 2001). To analyze, for example, the dynamic properties of proteins of interest within the cell nucleus, the proteins are genetically tagged to a fluorescent protein (e.g. green fluorescent protein (GFP)) and expressed in cells of interest. A part of the molecules in the cell nucleus is bleached by a focused laser beam and the recovery in the bleached part of the nucleus is observed by capturing images of the nucleus in predefined time intervals (McNally, 2008). Typically, such analyses are done on a couple of similar nuclei and the resulting rate estimators are summarized afterwards. However, the results often differ between nuclei. This variability is however not only due to random observation noise and the randomness of the underlying diffusion process, but also caused by cell-to-cell variation or the cellular status of the examined cells. To this end, we propose to analyze all nuclei together and account for the variance between nuclei by using mixed-effects models.

To date, for the analysis of data from FRAP experiments for the same molecule on multiple similar cell nuclei with a mathematical model, either the data of each recovery curve is analyzed separately, and the results for all cell nuclei are regarded together (e.g. Schneider et al., 2013), or the data of the different cell nuclei are pooled, averaged and then analyzed (e.g. McNally, 2008). In the second case, the recovery curves of the cell nuclei are averaged to obtain a smooth curve that can then be analyzed by the same mathematical model that is usually used for the analysis of one recovery curve. It is, however, vital, that only data of comparable fluorescent intensities are averaged (Sprague et al., 2004).

Random effects are frequently used in linear models for longitudinal data. They account for the fact that subjects are sampled randomly from a heterogeneous population (Pinheiro and Bates, 2000). Typically, random effects are combined with fixed effects, i. e., the usual effects in a linear model, resulting into mixed-effects models. Mixed-effects models are used in many applications including agriculture, pharmacokinetics, and geophysics (Pinheiro and Bates, 2000), as well as clinical trials (Brown and Prescott, 1999). In mixed-effects models, the relationships between a response variable and covariates, which are grouped by one or several factors, are described (Pinheiro and Bates, 2000). In our approach, fixed effects are parameters that are associated with the recovery curves of all cell nuclei, while random effects are parameters associated with the recovery curves of the individual cell nuclei.

The aim of a FRAP experiment is to infer the binding behavior of the unbleached molecules in the cell nucleus from their speed of movement. Because the bleached and unbleached molecules are assumed to behave identically, we can infer from that the binding behavior of all – bleached and unbleached – molecules of interest in the cell nucleus. See Sprague and McNally (2005) for more information on FRAP experiments. In this chapter, we concentrate on half-nucleus FRAP (as opposed to circle FRAP or strip FRAP (Mueller et al., 2008; Sprague et al., 2004)), which should cover representative fractions of heterogeneously distributed binding sites in all cell cycle stages (Schneider et al., 2013). An example for such data is given in Figure 5.1. In the first post-bleach image (second image from left, after 0.15 seconds), it is apparent that one half of

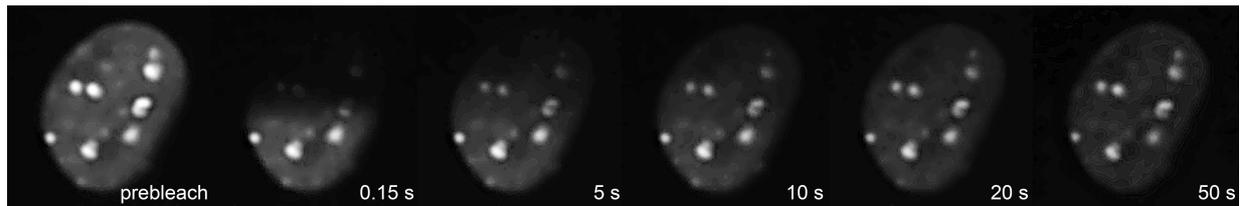


Figure 5.1: Fluorescence recovery after photobleaching one half of the nucleus of a mouse C2C12 cell expressing GFP-Dnmt1 in late S phase. Images of a cell nucleus in a FRAP experiment: In the prebleach image, the complete cell nucleus is visible because all molecules are fluorescent. In the first postbleach image (acquired after 0.15 seconds), it is obvious that one half of the nucleus has been bleached. The subsequent images show the recovery of the fluorescence in the nucleus after 5, 10, 20 and 50 seconds.

the cell nucleus has been bleached. The recovery of fluorescence in this half can be tracked over time in the subsequent images.

We propose a Bayesian nonlinear regression with mixed-effect priors for the simultaneous analysis of all recovery curves resulting from a series of FRAP experiments. In the following section, the data used in our analyzes is described and the compartment model as well as the differential equations associated with it are introduced. Thereafter, the model equation, which is based on the solution to the differential equations, is presented. In the subsequent section, the Bayesian nonlinear mixed-effects model, which consists of the data model, the prior model and the hyper priors, is described. Then, the parameter estimation procedure is presented, and the three different fixed-effects models with which the mixed-effects model is compared, are introduced, together with the information criterion used for the comparison. In the Results section, we present the results from the application of the proposed mixed-effects model to a series of FRAP experiments of GFP-tagged DNA methyltransferase 1 (GFP-Dnmt1) expressed in a somatic mouse cell line. Moreover, the comparison between the mixed-effects model and the fixed-effects models is done. The chapter ends with a discussion of the proposed approach. With the R package `frapmm`, which is available on GitHub (<https://github.com/feilke/frapmm>), the Bayesian nonlinear mixed-effects model can be fitted to data from a series of FRAP images. An accompanying file that is also available at GitHub provides a minimal working example including data simulation.

## 5.2 FRAP experiments

### 5.2.1 Data

In this chapter, we use FRAP data sets of GFP-Dnmt1 expressed in mouse C2C12 myoblast cells (Schneider et al., 2013), which were obtained from multiple cell nuclei and can therefore be utilized to illustrate how our nonlinear regression model can be used to fit all available data at once. DNA methylation at position 5 of cytosines within CpG dinucleotide sequences is an important biochemical process for the stable epigenetic gene silencing in vertebrates (Bird,

2002; Spada et al., 2006). The maintenance methyltransferase Dnmt1 reestablishes methylation of hemi-methylated CpG sites generated during DNA replication in S phase and thus ensures propagation of genomic methylation pattern over many cell divisions.

In order to study the cell cycle dependent binding behavior of wild type Dnmt1, we analyzed data from cells in different cell cycle stages as identified by the nuclear distribution pattern of GFP-Dnmt1 (Schneider et al., 2013): 12 cells with diffuse nuclear distribution (mostly G1 phase and possibly also late G2 phase), 26 cells in early S phase with Dnmt1 association at early replication foci, and 11 cells in late S phase with Dnmt1 associating with late replicating heterochromatin clusters. For each cell, the concentration of unbleached GFP-Dnmt1 in the bleached half of the cell nucleus was documented every 0.15 seconds up to 779 times after the bleaching. For the cells with diffuse nuclear distribution, the concentration was measured 778 times for 10 cells. For the two remaining cells with diffuse nuclear distribution, 390 and 480 measurements, respectively, were available. 778 measurements were available for 24 cells in early S phase. For the other two cells in early S phase, 774 and 754 measurements, respectively, were available. For the cells in late S phase, 779 measurements were available for 10 cells, whereas 777 measurements were available for the eleventh cell. The original FRAP data has been normalized by a triple normalization procedure described in detail in Dargatz (2010) and Schneider et al. (2013), converted to 8-bit, and Gauss-filtered with a two pixel radius (Schneider et al., 2013). Image registration has been done in order to correct for lateral movement of cells. In the normalization procedure, more specifically, it has been corrected for the intensity of the background of the images. In addition, it has been accounted for the variation in fluorescence during postbleach image acquisition as well as for cell-to-cell differences in the bleaching depth. Furthermore, the fluorescence loss due to half nucleus bleaching has been taken into account.

The major part of the FRAP data analyzed in this chapter has previously been published and analyzed (Schneider et al., 2013). The goal of the experiment was to identify the contribution of two different kinds of interactions Dnmt1 is involved in in different cell cycle phases. The interactions can be attributed to two subdomains of Dnmt1, first, the proliferating cell nuclear antigen (PCNA)-binding domain (Schermelleh et al., 2007; Schneider et al., 2013), and second, the targeting sequence domain, which targets Dnmt1 to the replication sites in S phase (Schneider et al., 2013). A figure showing the domain structure of GFP-Dnmt1 constructs can be found in (Schneider et al., 2013). Schneider et al. (2013) use the term “mobility classes” instead of the term “binding partners,” because all interactions with similar on- and off-rates can not be distinguished (Schneider, 2009), and, hence, build one mobility class (MC). Moreover, processes like anomalous diffusion, which are not related to binding, can also be represented by a MC.

### 5.2.2 Nonlinear recovery model

The movement of a molecule of interest in a cell nucleus is influenced by diffusion and by interactions, including binding reactions, the molecule is involved in (van Royen et al., 2011; Hemmerich et al., 2011; Mueller et al., 2010; Mazza et al., 2012). It is possible to model this process by using the full reaction-diffusion equations (Sprague et al., 2004; Carrero et al., 2004; Beaudouin et al., 2006). As we strive for an analytical solution of the equations describing the movement of the molecule of interest, we use a simplification of the full reaction-diffusion

equations. Usually, one of the following three simplifications is employed: the pure-diffusion scenario, the effective diffusion scenario or the reaction dominant scenario (Sprague et al., 2004). A pure-diffusion dominant scenario (Sprague et al., 2004) is present when most of the fluorescent molecules are free and interactions can be ignored. An effective diffusion scenario (Sprague et al., 2004; Mueller et al., 2008; Beaudouin et al., 2006; van Royen et al., 2009) occurs “when the reaction process is much faster than diffusion” (Sprague et al., 2004). A reaction dominant scenario is present, when diffusion is very fast compared to the timescale of the image acquisition and to the reaction process (Sprague et al., 2004).

The interactions Dnmt1 is involved in are described by on- and off-rates. In Schneider et al. (2013), where a correction value for diffusion was used, it was found that, in S phase, Dnmt1 is involved in interactions with relatively small off-rates, which means in the case of binding reactions, that the molecules of interest have a relatively long residence time (about 10-20 seconds) at their binding sites. We have no sufficient information about the magnitude of the on-rate. For these reasons and because we aim to have an analytical solution to the ordinary differential equations describing the movement of Dnmt1, we assume a reaction dominant scenario for our data.

In a reaction dominant FRAP scenario, diffusion is very fast in comparison to reaction processes and the time scale of the FRAP measurement (Bulinski et al., 2001; Coscoy et al., 2002; Dundr et al., 2002) and the recovery curve in the bleached part of the cell nucleus can be modeled using a nonlinear regression model (Sprague et al., 2004).

Here, we regard cases with two or three MCs (Schneider et al., 2013). In all considered cell cycle phases, a MC with a very long residence time compared to the time of image acquisition is indicated (Schneider et al., 2013). For this MC, we estimate only one parameter, and it is later also referred to as “immobile fraction”. In cells with diffuse localization and in early S phase, one additional MC has been identified. For the late S phase, two additional MCs with different off-rates were found.

The binding sites to which the molecules of interest bind are assumed to be part of large complexes, which are relatively immobile on the time scale of the FRAP measurement and the molecular movement (Sprague et al., 2004; Carrero et al., 2004). A compartment model with two or three compartments (Figure 5.2; the immobile fraction is ignored in this representation) is used to describe the change of the concentration of unbleached molecules in the bleached part of the cell nucleus. In a compartment model with two compartments, the molecules can be either free or bound. Exchange between the compartment of the free and the compartment of the bound molecules occurs with rates  $b_1^{\text{on}*}$  and  $b_1^{\text{off}}$ . In a compartment model with three compartments, the molecules can be either free or bound in one of two discriminable binding states. Exchange between the compartment of the free molecules and the compartments of the bound molecules occurs with rates  $b_1^{\text{on}*}$  and  $b_1^{\text{off}}$  and  $b_2^{\text{on}*}$  and  $b_2^{\text{off}}$ , respectively. A similar procedure based on the reaction equation of a binding interaction was proposed by Sprague et al. (2004).

The on- and off-rates of the binding reaction are denoted by  $b_k^{\text{on}*}$  and  $b_k^{\text{off}}$ ,  $k = 0, \dots, K$ . As stated in Sprague et al. (2004),  $b_k^{\text{on}*}$  is actually a pseudo-on-rate. It is the product of the actual on-rate  $b_k^{\text{on}}$  and the concentration of vacant bindings sites belonging to MC  $k$ . It is constant during the entire recovery process, because we assume that the biological system is in equilibrium before the bleaching and because bleaching does not affect the number of vacant binding sites (Sprague

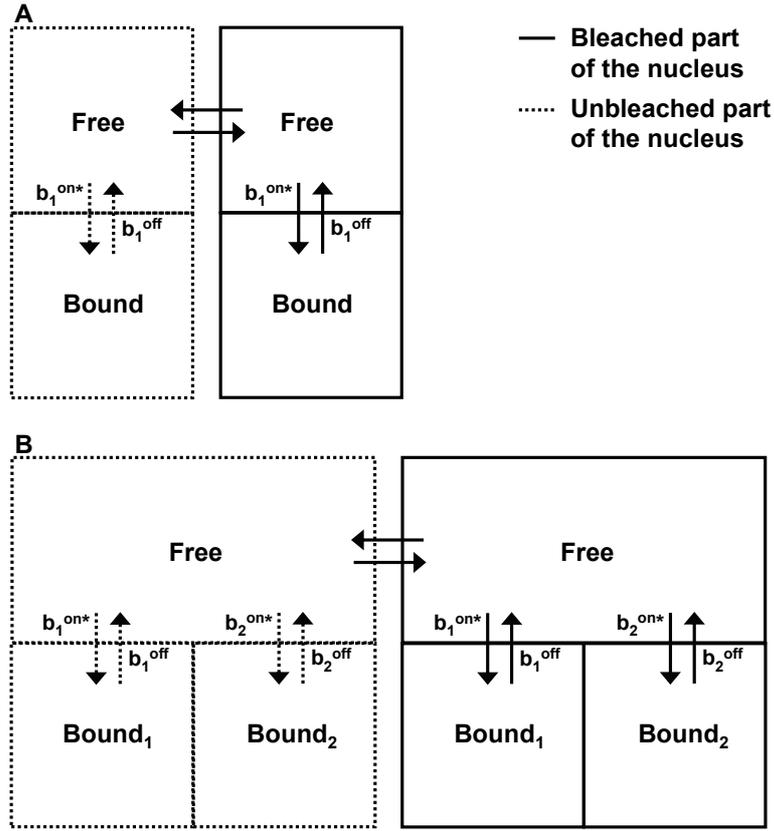


Figure 5.2: (A) Compartment model with two compartments and (B) compartment model with three compartments.

et al., 2004).

Let  $f(t) = [\text{Free}](t)$  denote the concentration of the free molecules and  $a_k(t) = [\text{Bound}_k](t)$  the concentration of the bound molecules in MC  $k$  at time  $t$ . We can describe the change of the concentration of the free and bound molecules based on the compartment model by the two differential equations

$$\frac{d}{dt}f(t) = \sum_{k=0}^K \left( -b_k^{\text{on}*} f(t) + b_k^{\text{off}} a_k(t) \right) + D_f \nabla^2 f(t), \quad (5.1)$$

with  $\nabla^2$  the Laplacian operator and  $D_f$  the diffusion coefficient for free proteins, and

$$\frac{d}{dt}a_k(t) = b_k^{\text{on}*} f(t) - b_k^{\text{off}} a_k(t). \quad (5.2)$$

The molecules in the cell nucleus are in equilibrium before the bleaching. In a diffusion-uncoupled FRAP scenario, the free molecules are moreover assumed to be in equilibrium again

immediately after the bleaching. Therefore  $f(t) = f_{\text{eq}}$ , a constant, and Equation (5.2) can be written as

$$\frac{d}{dt}a_k(t) = b_k^{\text{on}*} f_{\text{eq}} - b_k^{\text{off}} a_k(t). \quad (5.3)$$

Moreover, we do not have to model the change of the concentration of the free molecules, it suffices to model the change of concentration of the bound molecules, which means that Equation (5.1) can be ignored.

With boundary condition  $a_k(0) = 0$ , which means that at time  $t = 0$  (the time of the bleaching) the concentration of unbleached bound molecules in MC  $k$  in the bleached area equals zero, the solution of Equation (5.3) is

$$a_k(t) = \frac{b_k^{\text{on}*} f_{\text{eq}}}{b_k^{\text{off}}} - \frac{b_k^{\text{on}*} f_{\text{eq}}}{b_k^{\text{off}}} \exp(-b_k^{\text{off}} t). \quad (5.4)$$

As the system is in equilibrium before bleaching we have  $\frac{d}{dt}f(t) = 0$ ,  $\frac{d}{dt}a_k(t) = 0$  and constant steady-state intensities  $f_{\text{eq}}$ ,  $a_{k,\text{eq}}$ . Together with Equation (5.3) we get

$$a_{k,\text{eq}} = \frac{b_k^{\text{on}*} f_{\text{eq}}}{b_k^{\text{off}}},$$

and can therefore write Equation (5.4) as

$$a_k(t) = a_{k,\text{eq}}(1 - \exp(-b_k^{\text{off}} t)). \quad (5.5)$$

The observed value during FRAP recovery is the total fluorescence intensity in the bleached area. It can be described by the sum of the bound and the free unbleached molecules plus an error. The sum of the bound and the free unbleached molecules is denoted by  $\text{total}(t)$ :

$$\text{total}(t) = f_{\text{eq}} + \sum_{k=0}^K a_k(t).$$

For our analysis, in each cell nucleus, the fluorescence intensity has been averaged over the bleached part of the cell nucleus. Therefore, in our analysis,  $f_{\text{eq}}$  is the average of the intensity of the free fluorescent molecules in the bleached half, and  $a_k(t)$  is the average of the intensity of the bound fluorescent molecules in the bleached part of the nucleus. With Equation (5.5) we can then write

$$\text{total}(t) = f_{\text{eq}} + \sum_{k=0}^K a_{k,\text{eq}}(1 - \exp(-b_k^{\text{off}} t)).$$

With  $f_{\text{eq}} + \sum_{k=0}^K a_{k,\text{eq}} = 1$ , which holds because the concentration of the unbleached molecules has been normalized to one, we arrive at

$$\text{total}(t) = 1 - \sum_{k=0}^K a_{k,\text{eq}} \exp(-b_k^{\text{off}} t),$$

which is the deterministic approximation of the model with multiple mobility classes in Fuchs (2013), where stochastic differential equations are used to describe the dynamics of molecules in a cell nucleus.

## 5.3 Bayesian nonlinear mixed-effects model

In order to analyze all recovery curves from all nuclei simultaneously, we use a hierarchical Bayesian model consisting of the three levels described in Section 4.1 of Chapter 4.

### 5.3.1 Data model

The total observed concentration of unbleached molecules in the bleached part of the cell nucleus of cell  $j$ ,  $j = 1, \dots, J$ , at time  $t_i$ ,  $i = 1, \dots, T_j$ , is denoted by  $C_j(t_i)$ . We assume Gaussian noise for the observations

$$C_j(t_i) \sim N(\text{total}_j(t_i), \sigma^2).$$

The true concentration of unbleached molecules is modeled by the nonlinear model

$$\text{total}_j(t_i) = 1 - \sum_{k=0}^K a_{kj} \exp(-b_{kj}^{\text{off}} t_i).$$

Therefore, we fit the mixed-effects model

$$C_j(t_i) = 1 - \sum_{k=0}^K a_{kj} \exp(-b_{kj}^{\text{off}} t_i) + \varepsilon_{ij}$$

to the data of each cell cycle phase, where  $\varepsilon_{ij}$  are independent Gaussian noise terms with mean 0 and variance  $\sigma^2$ .

### 5.3.2 Prior model

In a Bayesian framework prior probability density functions have to be defined for all unknown parameters. Here, for the parameters  $a_{kj}$  and  $b_{kj}^{\text{off}}$ , we use a mixed-effect decomposition of the form

$$a_{kj} = a_k + \alpha_{kj}, \quad b_{kj}^{\text{off}} = \exp(f_k + \phi_{kj}) = \exp(f_k) \cdot \exp(\phi_{kj}),$$

with  $b_k^{\text{off}} = \exp(f_k)$  and  $\beta_{kj}^{\text{off}} = \exp(\phi_{kj})$ . So each of these parameters is split into a fixed effect, which represents a joint parameter for all recovery curves of all cell nuclei, and a random effect representing a curve-specific parameter. The prior for the parameter  $b_{kj}^{\text{off}}$  incorporates moreover the knowledge that transfer rates must be nonnegative (Schmid, Whitcher, Padhani, Taylor, and Yang, 2009). We do not assume nonnegativity for the parameter  $a_{kj}$ , as  $a_{0j}$  can be also negative. This is due to the fact that a triple normalization procedure has been applied to the data, which assumes that the equilibrium concentration of unbleached molecules in the bleached part of the cell nucleus is one. However, due to erroneous pre-processing,  $a_{0j}$  can also be smaller than zero, which leads to a equilibrium concentration bigger than one.

For the fixed effects, uniform priors of the form

$$p(a_k) = p(b_k^{\text{off}}) \propto \text{constant}$$

are used. These prior distributions are uninformative, which means that they do not contain any relevant information.

As prior distributions for the nuclei-specific random effects, we use Gaussian distributions and log-normal distributions, respectively, which are given by

$$\alpha_{kj} \sim N(0, \tau_{\alpha_k}^2), \beta_{kj}^{\text{off}} \sim LN(0, \tau_{\beta_k^{\text{off}}}^2),$$

where  $\tau_{\alpha_k}^2$  and  $\tau_{\beta_k^{\text{off}}}^2$  are unknown variance parameters.

### 5.3.3 Hyper prior model

Additional prior PDFs have to be defined for all other unknown parameters. As prior distributions for the unknown variance parameters, Inverse Gamma distributions, which are given by

$$\tau_{\alpha_k}^2 \sim IG(c_k, d_k), \tau_{\beta_k^{\text{off}}}^2 \sim IG(e_k, g_k),$$

are used. The Inverse Gamma distribution is known as a conjugate prior for the Gaussian distribution with known mean.

By using uninformative priors for the parameters  $a_k$  and  $b_k^{\text{off}}$ , we ensure that as much variance as possible is covered by the fixed effects. Only the variability that is not covered by the fixed effects is captured by the random effects. The definition of the hyperpriors with prudently chosen parameters on the variances of the parameters  $\alpha_{kj}$  and  $\beta_{kj}^{\text{off}}$  leads to a shrinkage of the random effects, so that they do not cover variance explained by the fixed effects (Schmid, Whitcher, Padhani, Taylor, and Yang, 2009).

If  $K = 1$ , which means that there is one MC in addition to the immobile fraction, we have to choose the parameters for the three Inverse Gamma distributions

$$\tau_{\alpha_0}^2 \sim IG(c_0, d_0), \tau_{\alpha_1}^2 \sim IG(c_1, d_1), \tau_{\beta_1^{\text{off}}}^2 \sim IG(e_1, g_1).$$

For the diffuse and the early S phase, we choose the parameters  $c_0 = c_1 = e_1 = 1$ ,  $d_0 = d_1 = g_1 = 10^{-4}$ .

When  $K = 2$ , which means that there exist two MCs in addition to the immobile fraction, we have the Inverse Gamma distributions

$$\tau_{\alpha_0}^2 \sim IG(c_0, d_0), \tau_{\alpha_1}^2 \sim IG(c_1, d_1), \tau_{\beta_1^{\text{off}}}^2 \sim IG(e_1, g_1),$$

$$\tau_{\alpha_2}^2 \sim IG(c_2, d_2), \tau_{\beta_2^{\text{off}}}^2 \sim IG(e_2, g_2).$$

The chosen parameters for the late S phase similar to above are  $c_0 = c_1 = c_2 = e_1 = e_2 = 1$ ,  $d_0 = d_1 = d_2 = g_1 = g_2 = 10^{-4}$ . By choosing the first parameter of the Inverse Gamma distribution to be 1 and the second parameter to be considerably smaller than 1, we perform shrinkage of the variances of the random effects. The smaller the second parameter is with respect to 1, the stronger is the shrinkage of the variance of the corresponding random effect, and, hence, the more variance is covered by the corresponding fixed effect.

As prior for the variance  $\sigma^2$  of the noise term  $\varepsilon_{ij}$ , we define the Inverse Gamma distribution  $\sigma^2 \sim IG(a, b)$  with  $a = b = 1$ . We assume *a priori* independence of all unknown parameters.

### 5.3.4 Posterior distribution and MCMC inference

The full conditional distributions of the parameters of the nonlinear regression model can be derived from the posterior distribution that is computed via Bayes' theorem. A Markov chain Monte Carlo algorithm with Gibbs- and Metropolis-Hastings-update steps is applied to obtain samples from the full conditional distributions. Therefore, in each iteration of the algorithm, a random sample from the conditional posterior distribution (given all other parameters and the data) is drawn for each parameter. The full conditional distributions of all parameters can be found in the Appendix (Part A).

The parameters  $a_k$  and  $\alpha_{kj}$  are drawn in Gaussian Gibbs steps, because their full conditional distributions are Gaussian distributions, from which one can sample directly. For the parameters  $\sigma^2$ ,  $\tau_{\alpha_k}^2$ , and  $\tau_{\beta_k^{\text{off}}}^2$ , Gamma Gibbs steps are used, because the full conditional distributions of the parameters are Inverse Gamma distributions. The parameters  $b_k^{\text{off}}$  and  $\beta_{kj}^{\text{off}}$  are drawn in MH-steps with random walk proposals, because their full conditional distributions are not standard distributions. For the MC which is present in all considered cell cycle phases and has a very long residence time compared to the time of image acquisition ( $k = 0$ ), the parameters  $b_0^{\text{off}}$  and  $\beta_{0j}^{\text{off}}$  are close to zero. Therefore, we set  $b_0^{\text{off}} = \beta_{0j}^{\text{off}} = 0$  and estimate only the parameter  $a_0$  for this immobile fraction (Schermelleh et al., 2007; Schneider et al., 2013; Sprague and McNally, 2005).

The random walk proposal of the Metropolis Hastings algorithm was tuned and resulted in acceptance rates between 35% and 52%. This is in accordance with recommendations for acceptance rates; for example Gilks et al. (1996) recommend acceptance rates between 15% and 50% (see also Gelman et al., 1996; Roberts et al., 1994). In our opinion, acceptance rates should be rather higher than lower, because with acceptance rates that are too low it might happen that part of the state space is never visited, what we intend to avoid.

We ran 10 parallel chains for each model (one model per cell cycle phase). For each parameter, a point estimate was obtained via the median of the sample formed by the observations of

| $d_0$     | $d_1$     | $g_1$     | $a_0$                   | $a_1$                  | $b_1^{\text{off}}$     | $\tau_{\alpha_0}^2$       | $\tau_{\alpha_1}^2$       | $\tau_{\beta_1}^2$        |
|-----------|-----------|-----------|-------------------------|------------------------|------------------------|---------------------------|---------------------------|---------------------------|
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.815,0.848) | 0.163<br>(0.152,0.174) | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) | 0.0107<br>(0.0052,0.0267) |
| $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | 0.007<br>(-0.002,0.016) | 0.831<br>(0.815,0.848) | 0.163<br>(0.152,0.173) | 0.0002<br>(0.0001,0.0006) | 0.0007<br>(0.0003,0.0018) | 0.0106<br>(0.0052,0.0263) |
| $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.813,0.850) | 0.163<br>(0.152,0.174) | 0.0001<br>(0.0000,0.0002) | 0.0009<br>(0.0004,0.0022) | 0.0107<br>(0.0052,0.0268) |
| $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.815,0.848) | 0.163<br>(0.152,0.174) | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) | 0.0109<br>(0.0053,0.0273) |
| $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.815,0.848) | 0.163<br>(0.152,0.175) | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) | 0.0106<br>(0.0052,0.0270) |
| $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.815,0.848) | 0.163<br>(0.152,0.174) | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) | 0.0106<br>(0.0052,0.0266) |
| $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | 0.006<br>(0.000,0.012)  | 0.831<br>(0.815,0.848) | 0.163<br>(0.153,0.174) | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) | 0.0107<br>(0.0052,0.0266) |

Table 5.1: Results of the sensitivity analysis for the diffuse phase.

| $d_0$     | $d_1$     | $g_1$     | $a_0$                  | $a_1$                  | $b_1^{\text{off}}$     | $\tau_{\alpha_0}^2$       | $\tau_{\alpha_1}^2$       | $\tau_{\beta_1}^2$        |
|-----------|-----------|-----------|------------------------|------------------------|------------------------|---------------------------|---------------------------|---------------------------|
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.030<br>(0.021,0.037) | 0.818<br>(0.799,0.835) | 0.094<br>(0.089,0.105) | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) | 0.0510<br>(0.0297,0.1003) |
| $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | 0.030<br>(0.020,0.037) | 0.818<br>(0.799,0.835) | 0.094<br>(0.088,0.105) | 0.0005<br>(0.0003,0.0008) | 0.0019<br>(0.0011,0.0034) | 0.0506<br>(0.0298,0.0994) |
| $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | 0.030<br>(0.022,0.037) | 0.817<br>(0.798,0.834) | 0.093<br>(0.087,0.103) | 0.0004<br>(0.0002,0.0007) | 0.0020<br>(0.0012,0.0036) | 0.0500<br>(0.0293,0.0960) |
| $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | 0.029<br>(0.023,0.036) | 0.817<br>(0.801,0.832) | 0.092<br>(0.087,0.100) | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) | 0.0491<br>(0.0292,0.0911) |
| $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | 0.030<br>(0.021,0.037) | 0.818<br>(0.799,0.835) | 0.094<br>(0.088,0.105) | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) | 0.0509<br>(0.0297,0.0992) |
| $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | 0.030<br>(0.022,0.037) | 0.817<br>(0.799,0.833) | 0.093<br>(0.087,0.103) | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) | 0.0500<br>(0.0294,0.0951) |
| $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | 0.030<br>(0.021,0.037) | 0.818<br>(0.799,0.834) | 0.094<br>(0.088,0.103) | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) | 0.0509<br>(0.0298,0.0981) |

Table 5.2: Results of the sensitivity analysis for the early S phase.

the converged parallel chains. Additionally, a 95% credible interval was calculated for each parameter. Approximate convergence of the parallel chains was diagnosed if the upper confidence limit of the potential scale reduction factor (Gelman and Rubin, 1992; Brooks and Gelman, 1998; Plummer et al., 2006) was smaller than or equal to 1.1. The number of burn-in iterations was determined by visual inspection of the sampling paths.

In order to check the sensitivity to the prior assumptions, i.e., to the choice of the parameter values of the hyper priors, we ran the mixed-effects model with different parameter values for the hyper priors. The parameter estimates together with 95% credible intervals can be found in Tables 5.1-5.4. We found that the point estimates stay the same for the different parameters of the hyper priors. Changes take place at most at the third decimal place of the estimates. The width of the credible intervals varies slightly. Overall, we found that the parameter estimates are not very sensitive to the choice of the parameter values of the hyper priors.

| $d_0$     | $d_1$     | $d_2$     | $g_1$     | $g_2$     | $a_0$                   | $a_1$                  | $a_2$                  | $b_1^{\text{off}}$     | $b_2^{\text{off}}$     |
|-----------|-----------|-----------|-----------|-----------|-------------------------|------------------------|------------------------|------------------------|------------------------|
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.602) | 0.217<br>(0.183,0.273) | 0.043<br>(0.038,0.048) |
| $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(-0.001,0.027) | 0.309<br>(0.260,0.353) | 0.557<br>(0.511,0.602) | 0.214<br>(0.180,0.263) | 0.043<br>(0.038,0.049) |
| $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.002,0.024)  | 0.308<br>(0.260,0.353) | 0.557<br>(0.510,0.602) | 0.215<br>(0.180,0.267) | 0.044<br>(0.038,0.048) |
| $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.601) | 0.217<br>(0.183,0.267) | 0.043<br>(0.038,0.048) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | 0.013<br>(0.001,0.023)  | 0.309<br>(0.260,0.353) | 0.559<br>(0.511,0.602) | 0.217<br>(0.180,0.273) | 0.044<br>(0.038,0.048) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.602) | 0.218<br>(0.182,0.270) | 0.044<br>(0.038,0.048) |
| $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.004,0.024)  | 0.308<br>(0.259,0.351) | 0.553<br>(0.508,0.603) | 0.212<br>(0.176,0.247) | 0.044<br>(0.038,0.049) |
| $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.351) | 0.558<br>(0.511,0.601) | 0.218<br>(0.182,0.282) | 0.043<br>(0.038,0.048) |
| $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.601) | 0.218<br>(0.182,0.282) | 0.043<br>(0.039,0.048) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.511,0.601) | 0.218<br>(0.182,0.275) | 0.044<br>(0.039,0.049) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | 0.013<br>(0.001,0.025)  | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.601) | 0.217<br>(0.183,0.265) | 0.043<br>(0.038,0.048) |

Table 5.3: Results of the sensitivity analysis for the late S phase - fixed parameters.

To evaluate the model fit, we compared the mixed-effects model – which was fitted to the whole of the data resulting from the FRAP experiments – to

1. a model without random effects fitted to the whole of the data,
2. a model without random effects fitted to the individual recovery curves,
3. a model without random effects fitted to the averaged recovery curves (all recovery curves of the same phase were averaged).

For each of these three scenarios, we fitted the following fixed-effects model to the data of each cell cycle phase:

$$C(t_i) = 1 - \sum_{k=0}^K a_k \exp(-b_k^{\text{off}} t_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2).$$

As for the mixed-effects model, we ran 10 parallel chains for each modeling alternative. Again, we calculated the medians and the 95% credible intervals of the samples formed by the observations of the converged chains (upper confidence limit of the potential scale reduction factor  $\leq 1.1$ ). The number of burn-in iterations was again determined by visual inspection of the sampling paths.

The Deviance Information Criterion (DIC) served as a measure of the model fit for the comparison of the mixed-effects model to the fixed-effects models 1 and 2. It is a suitable information criterion for model selection in hierarchical models, where parameters may outnumber observations and measures like the Akaike information criterion or Bayesian information criterion cannot be directly applied (Spiegelhalter et al., 2002). The DIC itself is not an absolute measure,

| $d_0$     | $d_1$     | $d_2$     | $g_1$     | $g_2$     | $\tau_{\alpha_0}^2$       | $\tau_{\alpha_1}^2$       | $\tau_{\alpha_2}^2$       | $\tau_{\beta_1}^2$        | $\tau_{\beta_2}^2$        |
|-----------|-----------|-----------|-----------|-----------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0144) | 0.0669<br>(0.0292,0.1921) | 0.0443<br>(0.0210,0.1140) |
| $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0005<br>(0.0002,0.0013) | 0.0051<br>(0.0023,0.0135) | 0.0054<br>(0.0024,0.0143) | 0.0671<br>(0.0296,0.1861) | 0.0462<br>(0.0217,0.1197) |
| $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0050<br>(0.0023,0.0134) | 0.0056<br>(0.0025,0.0147) | 0.0663<br>(0.0290,0.1891) | 0.0450<br>(0.0212,0.1181) |
| $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0144) | 0.0666<br>(0.0292,0.1860) | 0.0445<br>(0.0210,0.1144) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0053<br>(0.0024,0.0139) | 0.0055<br>(0.0025,0.0143) | 0.0675<br>(0.0295,0.1922) | 0.0444<br>(0.0211,0.1139) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-3}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0144) | 0.0671<br>(0.0294,0.1895) | 0.0445<br>(0.0211,0.1145) |
| $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0001,0.0008) | 0.0049<br>(0.0022,0.0129) | 0.0054<br>(0.0024,0.0144) | 0.0634<br>(0.0281,0.1724) | 0.0449<br>(0.0211,0.1156) |
| $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0135) | 0.0055<br>(0.0025,0.0143) | 0.0680<br>(0.0293,0.2018) | 0.0443<br>(0.0210,0.1139) |
| $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0135) | 0.0055<br>(0.0025,0.0144) | 0.0675<br>(0.0294,0.2016) | 0.0443<br>(0.0210,0.1138) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-5}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0050<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0143) | 0.0675<br>(0.0293,0.1933) | 0.0445<br>(0.0211,0.1145) |
| $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | $10^{-4}$ | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0144) | 0.0663<br>(0.0290,0.1836) | 0.0441<br>(0.0209,0.1134) |

Table 5.4: Results of the sensitivity analysis for the late S phase - variances.

that is, the absolute values cannot be interpreted, but can be compared relatively between models. We did not compare the mixed-effects model and fixed-effects model 3 on the basis of the DIC, because these two models were fit to different kinds of data. The fixed-effects model 3 in contrast to the mixed-effects model was not fitted to the whole of all recovery curves but to the averaged recovery curve per phase.

The DIC can be calculated by the deviance of the medians  $D(\theta_{\text{med}})$  plus two times the effective number of parameters  $p_D$  (Spiegelhalter et al., 2002):

$$DIC = D(\theta_{\text{med}}) + 2p_D.$$

The deviance is a measure of the fit of a model and is calculated by

$$D(\theta) = -2l(\theta),$$

where  $l(\theta)$  is the log-likelihood. The effective number of parameters is a measure of the complexity of the model. It is the median deviance minus the deviance of the medians and is calculated by

$$p_D = \text{median}(D(\theta)) - D(\theta_{\text{med}}).$$

The effective number of parameters is high for models with a high effective model complexity. When comparing two models on the basis of their DIC, the model with the lower DIC is to be favored.

All software was written in the programming languages R and C.

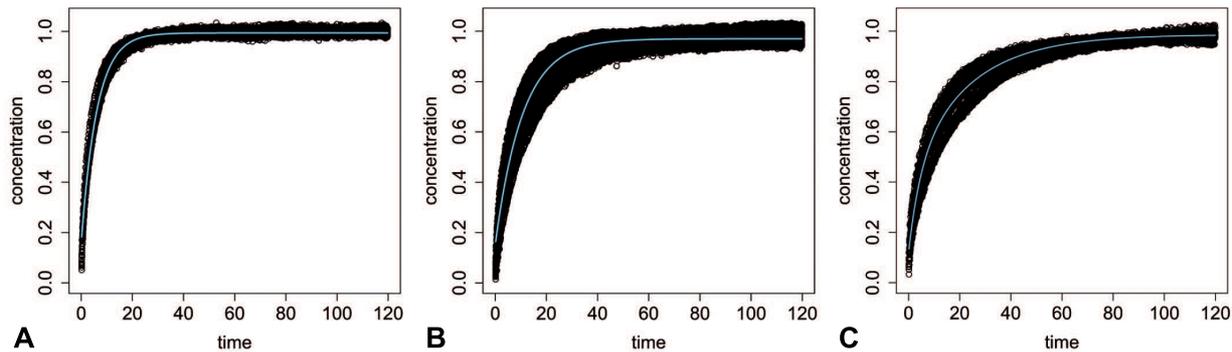


Figure 5.3: Normalized data together with estimated joint recovery curve. The estimated joint recovery curves for all cell nuclei using the posterior medians of the MCMC-samples of the fixed effects are shown together with the normalized data for all three reviewed cell cycle phases – (A) diffuse, (B) early S, (C) late S phase.

## 5.4 Results

### 5.4.1 Mixed-effects model

By using the Bayesian regression model with mixed-effect priors we gain common parameter estimates for all cell nuclei through the estimation of the fixed effects, as well as curve specific parameter estimates through the estimation of the random effects, and estimates for the variances of the random effects.

In Figure 5.3, for each phase (diffuse, early S, and late S phase), the estimated joint recovery curve for all cell nuclei is shown together with the normalized data. The joint recovery curve is computed using the posterior medians of the MCMC-samples of the fixed effects.

The random effects take into account the variability resulting from the joint analysis of data of multiple cell nuclei, which is not covered by the fixed effects. In Figure 5.4, the estimated joint recovery curve for all cell nuclei (black, solid line) is shown together with the cell nuclei-specific curves (colored, dashed lines), which are computed using the posterior medians of the MCMC-samples of the curve-specific random effects.

The posterior medians of the fixed parameters  $a_k$  and  $b_k^{\text{off}}$  together with 95%-credible intervals can be found in Table 5.5. Density plots of the posterior distributions of the fixed parameters for the diffuse phase can be found in Figure 5.5. The density plots for the remaining cell cycle phases can be found in the Appendix of Feilke, Schneider, and Schmid (2015). All parameters could be estimated with small variance. For the diffuse phase, the posterior median of the fixed effect of the off-rate is denoted by  $b_1^{\text{off}}$  and equals 0.163 (0.152, 0.174). For the early S phase,  $b_1^{\text{off}}$  equals 0.094 (0.089, 0.105). In both cases, we assumed that there is only one MC in addition to the immobile fraction ( $K = 1$ ), based on Schneider et al. (2013). In the presence of binding, the off-rate is the rate of the unbinding reaction where a protein is unsoldered from its binding site (Sprague and McNally, 2005), and its inverse is the residence time, i.e., the time a protein remains at a binding site (McNally, 2008). In early S phase, binding of Dnmt1 to immobilized

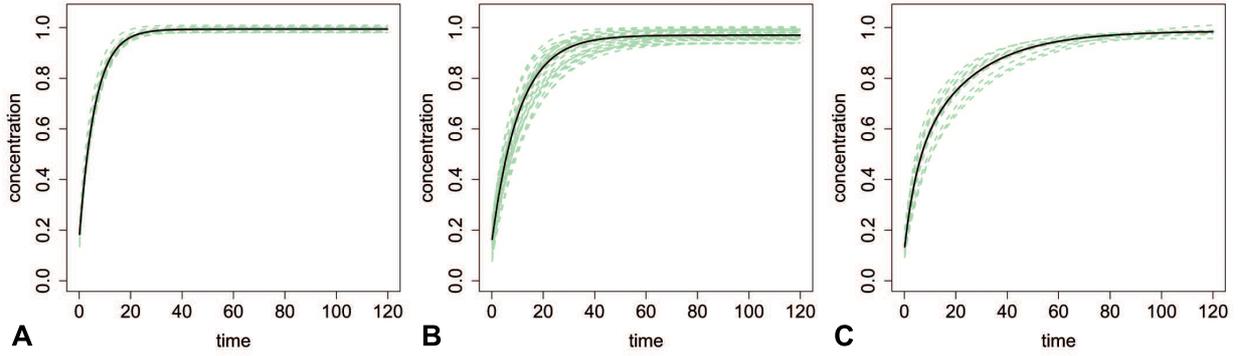


Figure 5.4: Estimated joint recovery curve together with cell nuclei-specific recovery curves. The cell nuclei-specific recovery curves using the posterior medians of the MCMC-samples of the random effects for all three reviewed cell cycle phases – (A) diffuse, (B) early S, (C) late S phase – are shown together with the estimated joint recovery curve.

| phase   | number of recovery curves | $a_0$                  | $a_1$                  | $a_2$                  | $b_1^{\text{off}}$     | $b_2^{\text{off}}$     |
|---------|---------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| diffuse | 12                        | 0.006<br>(0.000,0.012) | 0.831<br>(0.815,0.848) |                        | 0.163<br>(0.152,0.174) |                        |
| early S | 26                        | 0.030<br>(0.021,0.037) | 0.818<br>(0.799,0.835) |                        | 0.094<br>(0.089,0.105) |                        |
| late S  | 11                        | 0.013<br>(0.001,0.025) | 0.309<br>(0.262,0.352) | 0.558<br>(0.512,0.602) | 0.217<br>(0.183,0.273) | 0.043<br>(0.038,0.048) |

Table 5.5: Mixed-effects model: Fixed effects - median plus 95% credible interval.

PCNA trimetric rings at replication forks takes place (Schneider et al., 2013; Schermelleh et al., 2007; Sporbert et al., 2005). The median residence time of Dnmt1 at this binding site is about 11 seconds. In the diffuse phase, the off-rate can not be interpreted in the same way, because in this phase, there is no specific binding partner present and the additional MC is probably due to anomalous diffusion behavior (Schneider et al., 2013).

For the late S phase, we assumed that there are two distinctive MCs in addition to the immobile fraction ( $K = 2$ ) (Schneider et al., 2013). The posterior medians of the fixed effect of the off-rates are  $b_1^{\text{off}} = 0.217$  (0.183,0.273) and  $b_2^{\text{off}} = 0.043$  (0.038,0.048), which corresponds to median residence times of about 5 seconds and about 23 seconds, respectively. This is in compliance with the finding that the protein Dnmt1 is involved in two distinctive interactions in the late S phase (Schneider et al., 2013).

As it is of essential interest how much variance is captured by the random effects, point estimates for the variances of the random effects plus 95%-credible intervals were calculated and can be found in Table 5.6. Density plots of the posterior distributions of the variances for the diffuse phase can be found in Figure 5.6 and the density plots of the variances for the remaining cell cycle phases can be found in the Appendix of Feilke, Schneider, and Schmid (2015).

Each of the credible intervals in Tables 5.5 and 5.6 embodies the true parameter with a probability of 95%. To give an impression about the variation of the off-rates in the population of cell nuclei, for each cell cycle phase, we calculated the minimum and the maximum of the products

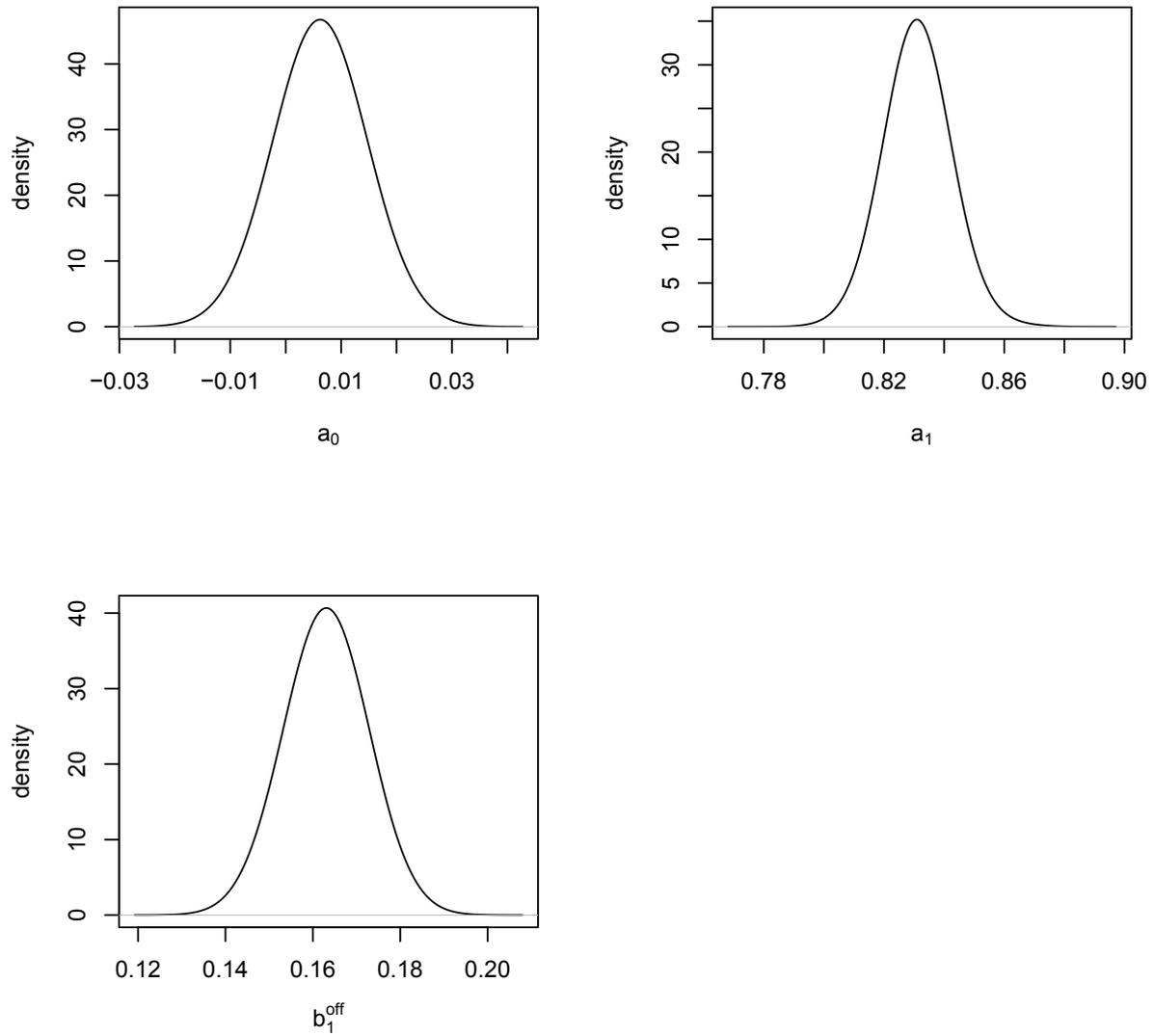


Figure 5.5: Mixed-effects model: Density plots of the posterior distributions of the fixed parameters for the diffuse phase.

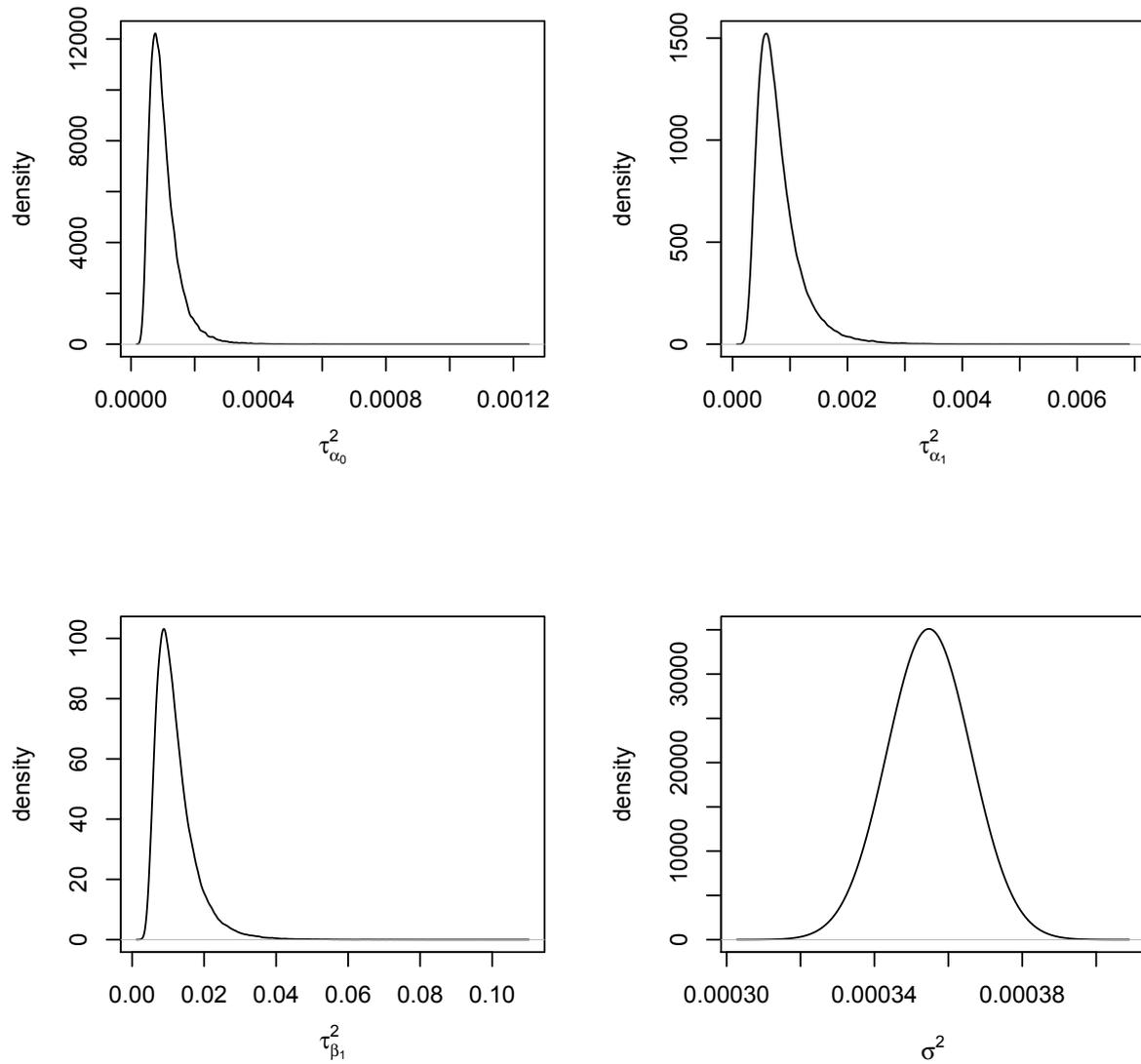


Figure 5.6: Mixed-effects model: Density plots of the posterior distributions of the variances for the diffuse phase.

| phase   | number of recovery curves | $\tau_{a_0}^2$            | $\tau_{a_1}^2$            | $\tau_{a_2}^2$            | $\tau_{\beta_1}^2$        | $\tau_{\beta_2}^2$        |
|---------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| diffuse | 12                        | 0.0001<br>(0.0000,0.0002) | 0.0007<br>(0.0003,0.0018) |                           | 0.0107<br>(0.0052,0.0267) |                           |
| early S | 26                        | 0.0004<br>(0.0002,0.0007) | 0.0019<br>(0.0011,0.0034) |                           | 0.0510<br>(0.0297,0.1003) |                           |
| late S  | 11                        | 0.0003<br>(0.0002,0.0009) | 0.0051<br>(0.0023,0.0134) | 0.0055<br>(0.0025,0.0144) | 0.0669<br>(0.0292,0.1921) | 0.0443<br>(0.0210,0.1140) |

Table 5.6: Mixed-effects model: Variances of random effects - median plus 95% credible interval.

| phase   | number of recovery curves | $a_0$                  | $a_1$                  | $a_2$                  | $b_1^{\text{off}}$     | $b_2^{\text{off}}$     |
|---------|---------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| diffuse | 12                        | 0.005<br>(0.005,0.006) | 0.829<br>(0.824,0.833) |                        | 0.159<br>(0.158,0.161) |                        |
| early S | 26                        | 0.032<br>(0.031,0.033) | 0.809<br>(0.805,0.813) |                        | 0.086<br>(0.085,0.087) |                        |
| late S  | 11                        | 0.015<br>(0.013,0.018) | 0.316<br>(0.291,0.343) | 0.546<br>(0.516,0.573) | 0.202<br>(0.179,0.231) | 0.042<br>(0.040,0.044) |

Table 5.7: Fixed-effects model 1 (fitted to the whole of the data): Fixed effects - median plus 95% credible interval.

of the median of the fixed effect of the off-rate and the medians of the nuclei-specific off-rates to get approximate limits in which the off-rates of the different cell nuclei lie. According to that, the off-rate for the diffuse phase varies approximately between 0.141 and 0.195 over the different cell nuclei, whereas for the early S phase, it lies approximately between 0.056 and 0.143 (residence time: about 7 - 18 seconds). For the late S phase,  $b_1^{\text{off}}$  varies between 0.129 and 0.293 (residence time: about 3 - 8 seconds), and  $b_2^{\text{off}}$  varies between 0.026 and 0.059 (residence time: about 17 - 38 seconds). Thus, we do not only get a joint point estimate of the off-rate for the population of the cell nuclei for each cell cycle phase, but we also gain information about the variation of the off-rate in the population of nuclei.

## 5.4.2 Comparison between the mixed-effects model and the fixed-effects models

Table 5.7 provides the posterior medians of the fixed effects  $a_k$  and  $b_k^{\text{off}}$  together with 95%-credible intervals for fixed-effects model 1, where regression model (5.6) was fitted to the whole of the data resulting from the FRAP experiments. Table 5.8 provides the posterior medians, minima, and maxima of the fixed effects  $a_k$  and  $b_k^{\text{off}}$  for fixed-effects model 2, where regression model (5.6) was fitted to the individual recovery curves of all cell nuclei. In Table 5.9, the posterior medians of the fixed effects  $a_k$  and  $b_k^{\text{off}}$  together with 95%-credible intervals for fixed-effects model 3, where regression model (5.6) was fitted to the averaged recovery curves, are shown. All three tables provide estimates for each cell cycle phase under review. Density plots of the posterior distributions of the inferred parameters for the diffuse phase for the fixed-effects models 1-3 can be found in Figures 5.7-5.10. The density plots for the remaining cell cycle phases can be found in the Appendix of Feilke, Schneider, and Schmid (2015). For fixed-effects

| phase   | number of recovery curves | $a_0$                   | $a_1$                  | $a_2$                  | $b_1^{\text{off}}$     | $b_2^{\text{off}}$     |
|---------|---------------------------|-------------------------|------------------------|------------------------|------------------------|------------------------|
| diffuse | 12                        | 0.004<br>[-0.009,0.020] | 0.820<br>[0.793,0.889] |                        | 0.153<br>[0.139,0.194] |                        |
| early S | 26                        | 0.034<br>[-0.003,0.063] | 0.824<br>[0.727,0.892] |                        | 0.087<br>[0.056,0.144] |                        |
| late S  | 11                        | 0.014<br>[-0.048,0.043] | 0.293<br>[0.220,0.450] | 0.503<br>[0.455,0.686] | 0.212<br>[0.102,0.333] | 0.043<br>[0.022,0.060] |

Table 5.8: Fixed-effects model 2 (fitted to the individual recovery curves): Fixed effects - median [min,max].

| phase   | number of recovery curves | $a_0$                  | $a_1$                  | $a_2$                  | $b_1^{\text{off}}$     | $b_2^{\text{off}}$     |
|---------|---------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| diffuse | 1                         | 0.005<br>(0.001,0.009) | 0.826<br>(0.794,0.861) |                        | 0.158<br>(0.149,0.169) |                        |
| early S | 1                         | 0.032<br>(0.027,0.037) | 0.806<br>(0.783,0.832) |                        | 0.086<br>(0.081,0.090) |                        |
| late S  | 1                         | 0.015<br>(0.002,0.024) | 0.329<br>(0.230,0.468) | 0.533<br>(0.383,0.636) | 0.190<br>(0.120,0.345) | 0.041<br>(0.031,0.049) |

Table 5.9: Fixed-effects model 3 (fitted to the averaged recovery curves): Fixed effects - median plus 95% credible interval.

model 2, exemplary density plots of the inferred parameters are shown for two cell nuclei of each cell cycle phase. In Figure 5.11, the point estimates and the 95%-credible intervals for the mixed-effects model and fixed-effects models 1 and 3 are displayed.

Table 5.10 contains the DIC, the effective number of parameters ( $p_D$ ) and the deviance of the medians ( $D(\theta_{\text{med}})$ ) for the proposed mixed-effects model and the fixed-effects models 1 and 2 introduced in Section 5.3.4.

Regarding the point estimates in Table 5.7, one sees that the point estimates for the fixed effects provided by fixed-effects model 1 differ from the estimates provided by the mixed-effects model (Table 5.5). Figure 5.11 and the comparison of Table 5.7 with Table 5.5 reveal that the 95%-credible intervals for the fixed parameters resulting from fitting fixed-effects model 1 to the whole of the data are considerably smaller than the credible intervals for the fixed parameters resulting from fitting the proposed mixed-effects model to the whole of the data. Therefore, it could be erroneously concluded that the estimation of the fixed parameters is more exact by using the fixed-effects model 1 than by using the proposed mixed-effects model. This is not the case, as for the fixed-effects model 1 we assume many independent observations and, hence, ignore the structure of the data, i.e., the observations come from different cells and are not all independent. Due to this assumption, the variance of the fixed effects is underestimated.

For all three cell cycle phases (diffuse, early S, and late S phase), when comparing the DIC of the mixed-effects model to the DIC of fixed-effects model 1, it is obvious that the DIC is considerably lower for the mixed-effects model. The clearest result can be found for the late S phase, where the DIC of the mixed-effects model is almost two times lower than the DIC of fixed-effects model 1. This means, that the proposed mixed-effects model is superior to fixed-effects model 1 for all three cell cycle phases.

Regarding the point estimates resulting from fitting fixed-effects model 2 to the individual recovery curves (Table 5.8), we observe that most of them differ from the point estimates result-

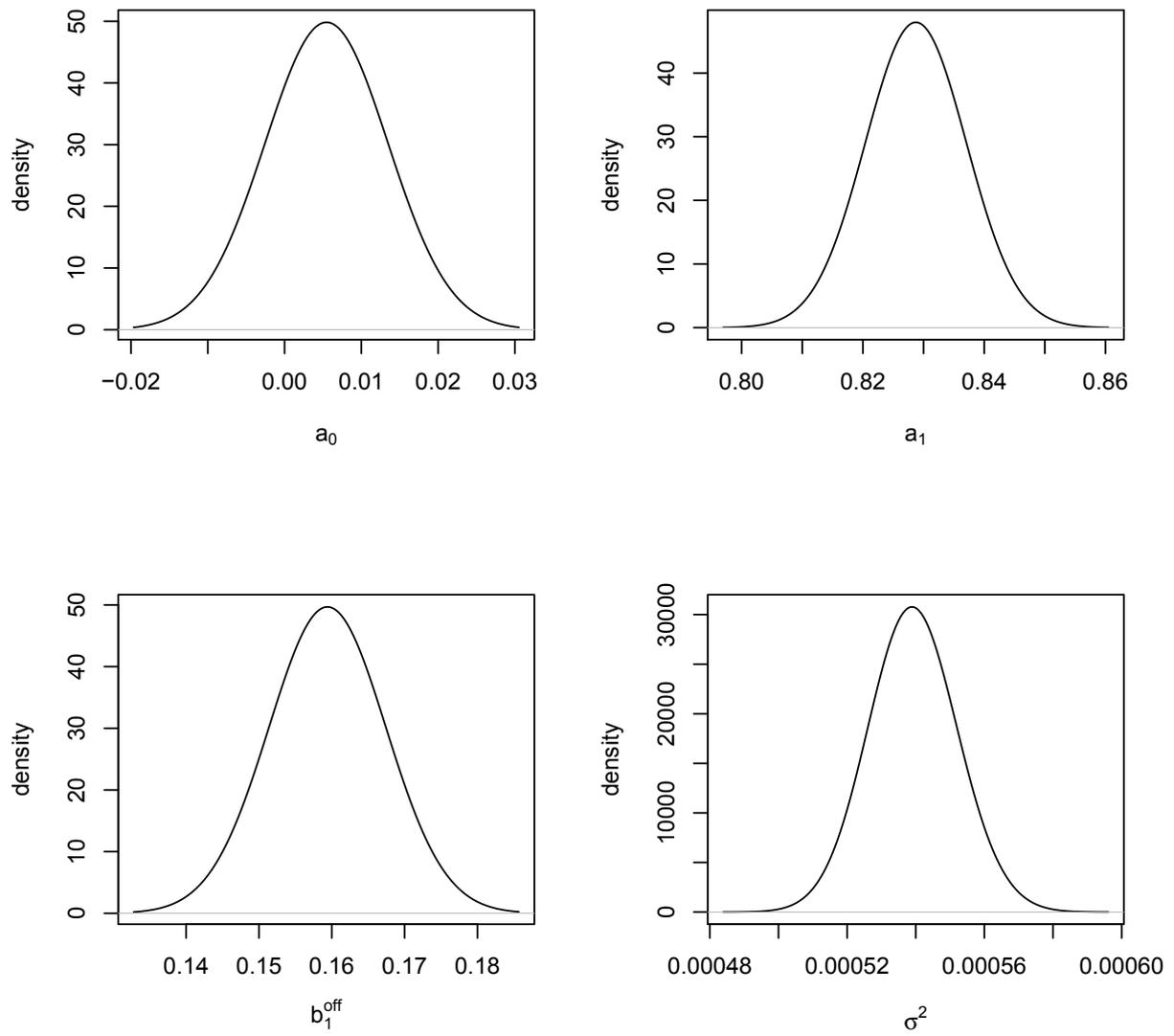


Figure 5.7: Fixed-effects model 1: Density plots of the posterior distributions of the parameters for the diffuse phase.

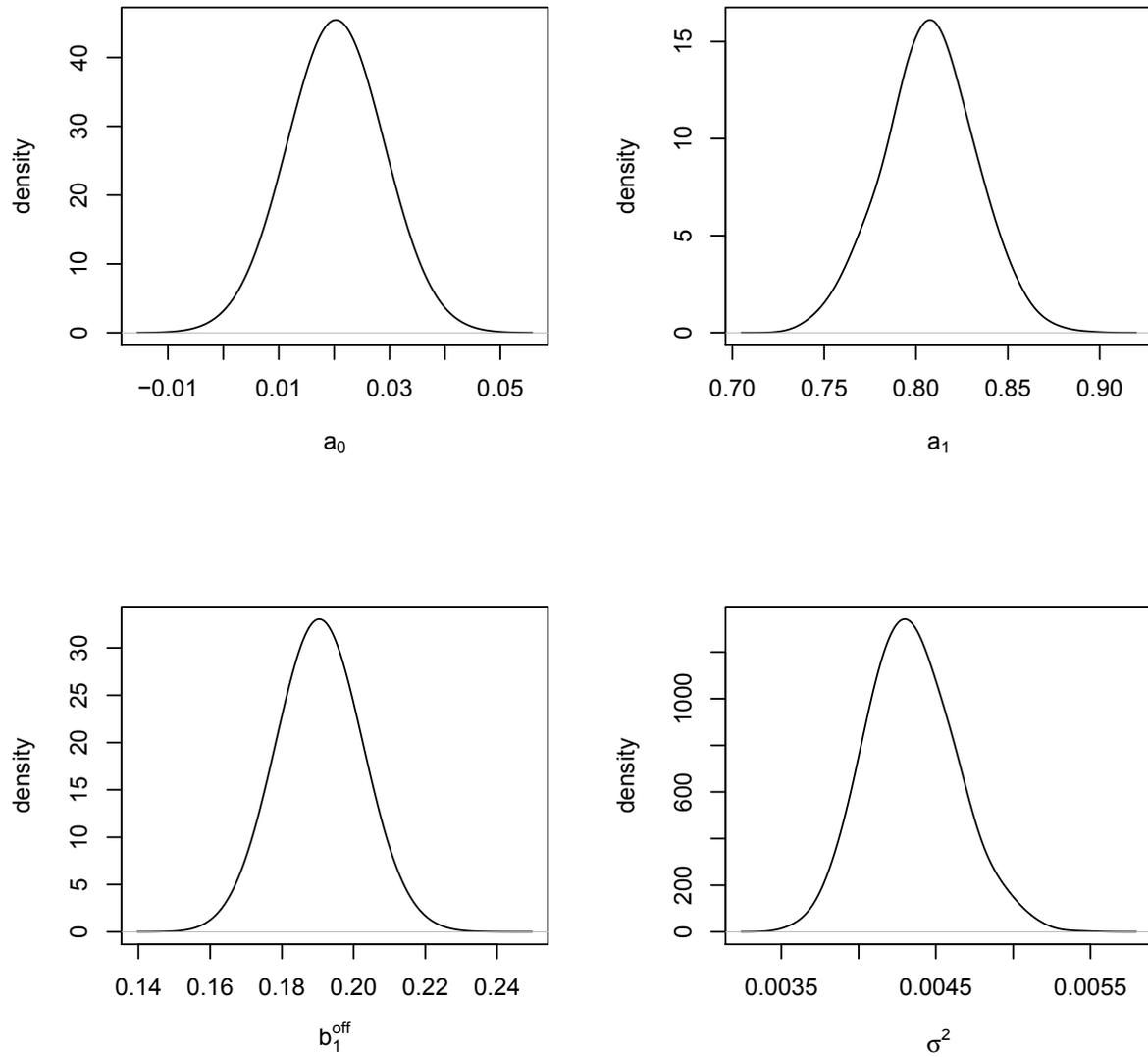


Figure 5.8: Fixed-effects model 2: Density plots of the posterior distributions of the parameters for the diffuse phase (curve 4).

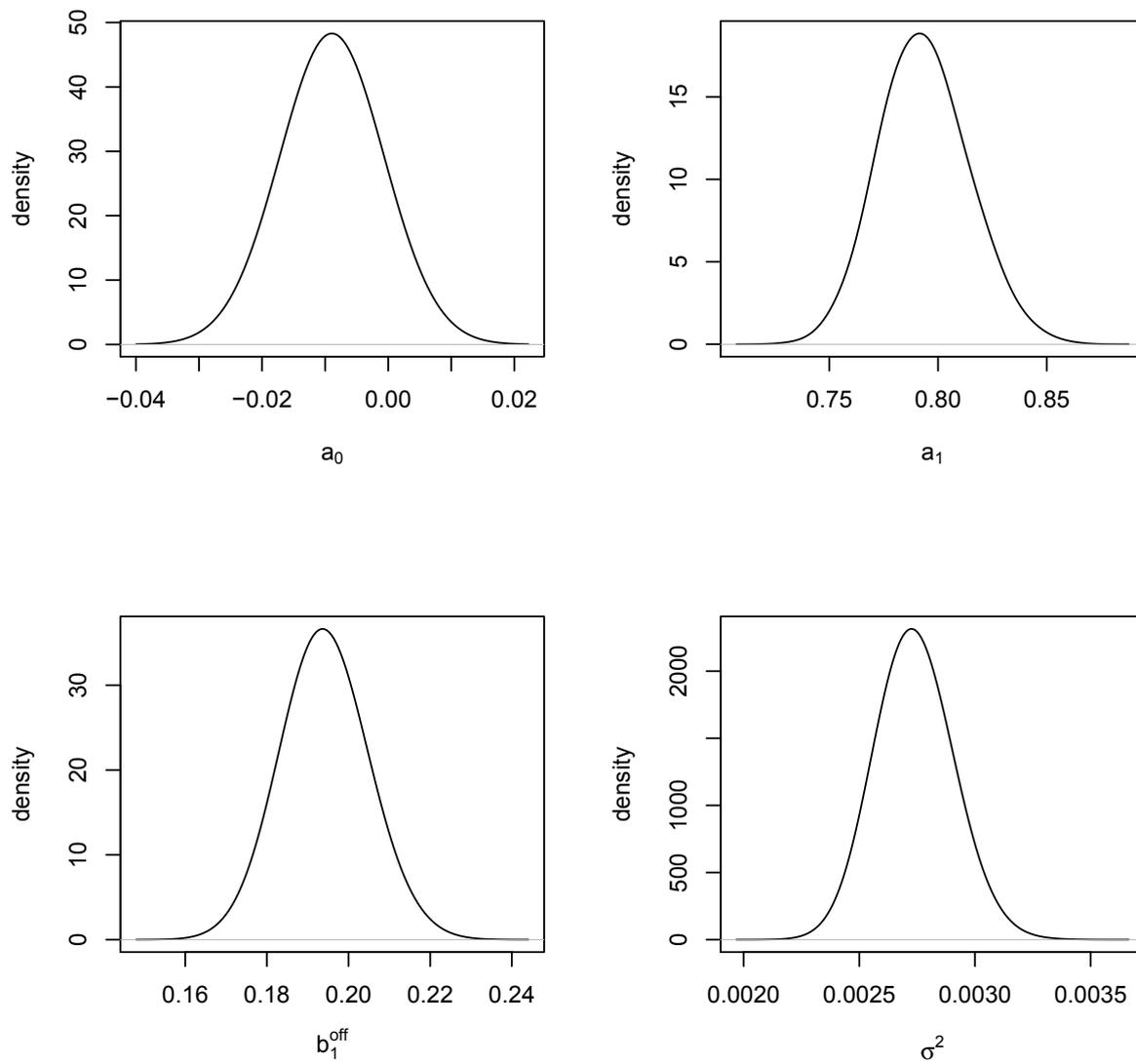


Figure 5.9: Fixed-effects model 2: Density plots of the posterior distributions of the parameters for the diffuse phase (curve 11).

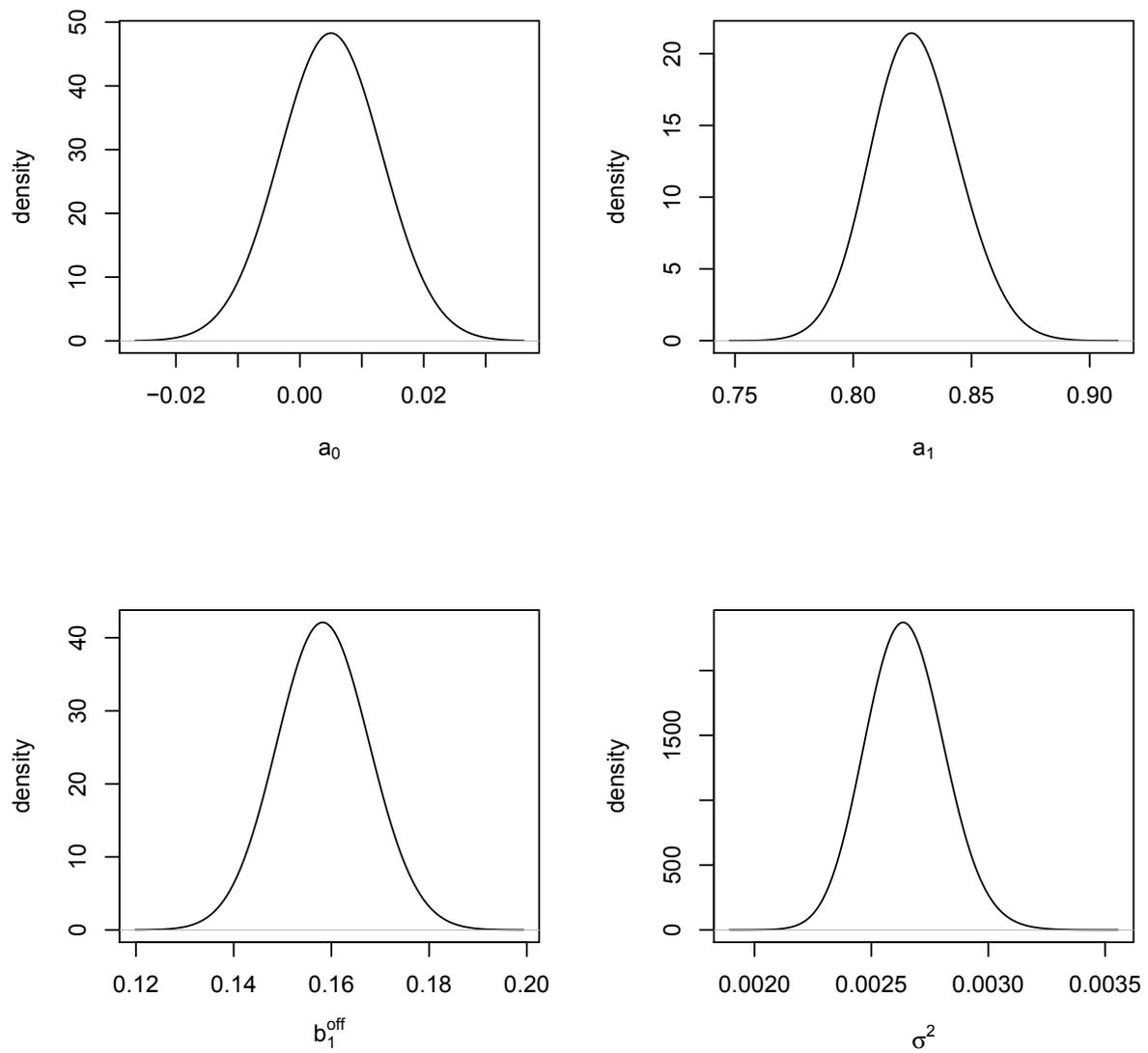


Figure 5.10: Fixed-effects model 3: Density plots of the posterior distributions of the parameters for the diffuse phase.

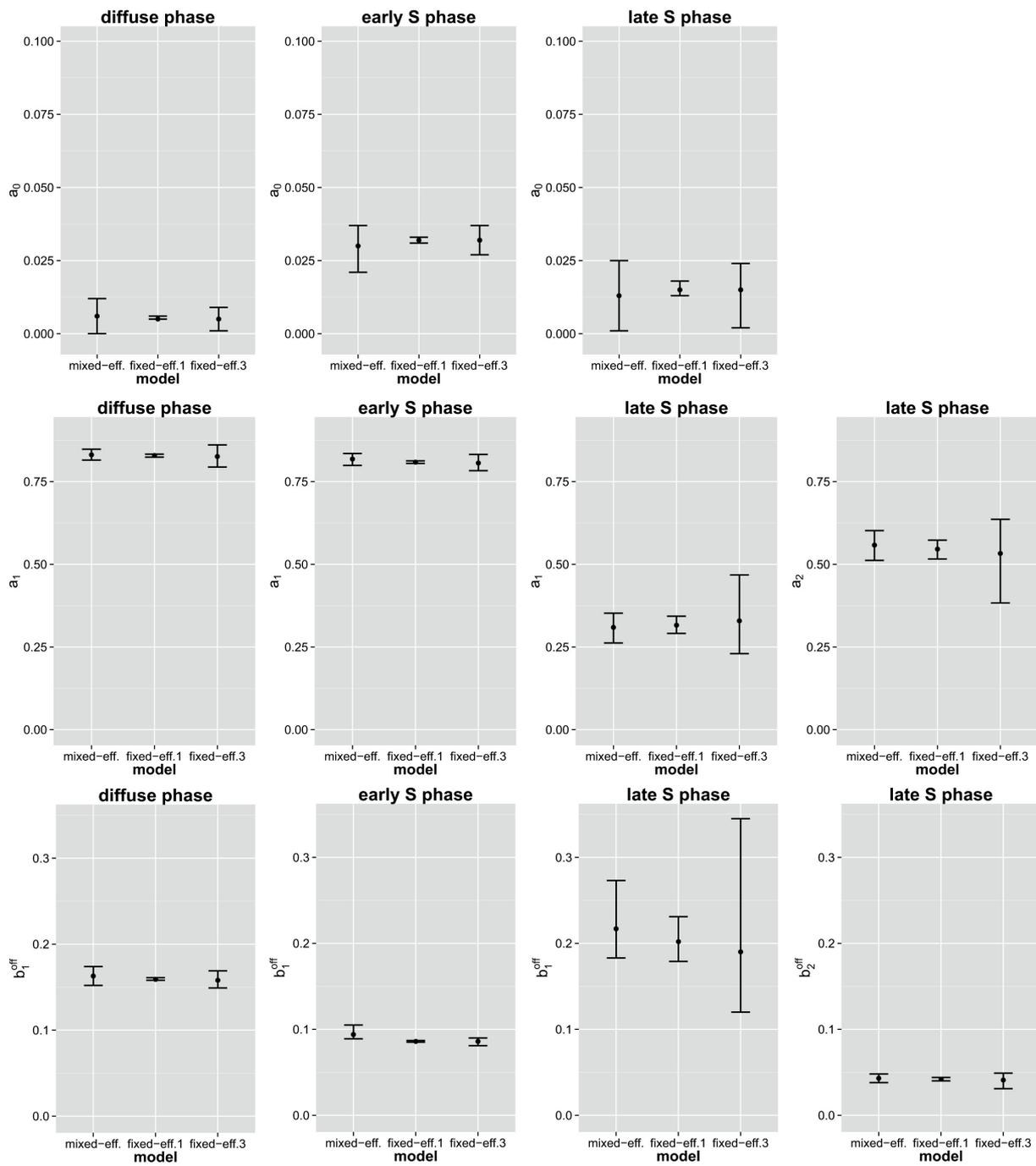


Figure 5.11: Point estimates and 95%-credible intervals for the fixed effects of the mixed-effects model and fixed-effects models 1 and 3. The posterior medians of the fixed effects together with 95%-credible intervals are shown for the proposed mixed-effects model and fixed-effects models 1 and 3 for all three reviewed cell cycle phases.

|         | Mixed-effects model      |        |            | Fixed-effects model 1    |       |           | Fixed-effects model 2    |         |            |
|---------|--------------------------|--------|------------|--------------------------|-------|-----------|--------------------------|---------|------------|
|         | $D(\theta_{\text{med}})$ | $p_D$  | DIC        | $D(\theta_{\text{med}})$ | $p_D$ | DIC       | $D(\theta_{\text{med}})$ | $p_D$   | DIC        |
| diffuse | -53393.32                | 99.66  | -53193.01  | -45393.75                | 4.19  | -45385.38 | -55231.80                | 1084.99 | -53061.82  |
| early S | -115018.28               | 114.67 | -114788.94 | -73488.90                | 2.59  | -73483.71 | -119243.50               | 1212.22 | -116819.10 |
| late S  | -60694.44                | 278.22 | -60138.01  | -34218.33                | 5.43  | -34207.48 | -61334.80                | 2515.63 | -56303.55  |

Table 5.10: Deviance of the medians ( $D(\theta_{\text{med}})$ ), effective number of parameters ( $p_D$ ) and DIC of the mixed-effects model and fixed-effects models 1 and 2.

ing from fitting the proposed mixed-effects model to the whole of the data (Table 5.5). When comparing the DIC of the mixed-effects model to the DIC of fixed-effects model 2, it can be seen that the DIC of the mixed-effects model is lower than the DIC of fixed-effects model 2 for the diffuse and late S phase. Only for the early S phase, the DIC of fixed-effects model 2 is lower than the DIC of the mixed-effects model. But overall, the DIC is approximately of the same magnitude for both models.

Figure 5.11 reveals that most of the credible intervals for the fixed effects  $a_k$  and  $b_k^{\text{off}}$  resulting from fitting fixed-effects model 3 to the averaged recovery curves for the fixed effects  $a_k$  and  $b_k^{\text{off}}$  (Table 5.9) are broader or of approximately the same size as the corresponding credible intervals resulting from the proposed mixed-effects model (Table 5.5). Only for the fixed effect  $a_0$ , it is converse. This is due to the difference in the number of data points between the two models. For fixed-effects model 3, we have only one recovery curve per cell cycle phase because of the averaging of the data. Therefore, the estimation of the fixed effects is more exact when using the proposed mixed-effects model. Moreover, we are of the opinion that averaging the recovery curves induces a loss of information because not all available data is used and the variability contained in the data is not appropriately quantified, which is why we favor the proposed mixed-effects model over fixed-effects model 3.

Overall, we conclude that the mixed-effects model is superior to the fixed-effects models 1-3, because the DIC of the mixed-effects model is lower in almost all considered scenarios. Moreover, it adequately reflects the heterogeneity of the data caused by cell-to-cell variability through the estimation of the variances of the random effects. The heterogeneity of the data is also taken into account by fixed-effects model 2, which gives point estimates of the fixed effects for each curve per cell cycle phase. However, estimating the variance through a mixed-effects model is the more appropriate and comfortable way to quantify the cell-to-cell variability. Moreover, the proposed mixed-effects model is more robust than the fixed-effects models 1-3 because it uses more information.

## 5.5 Discussion

Our objective was to develop an approach with which data from FRAP experiments on various similar cell nuclei can be analyzed in one model, taking into account the variability contained in the data. The variability can only be assessed by considering data of several cell nuclei in a joint model.

Using the proposed Bayesian nonlinear regression model with mixed-effect priors, we are able to do a joint analysis of the recovery curves of all available cell nuclei per cell cycle phase.

So all available data resulting from different FRAP experiments can be used for the estimation of the parameters of interest and no data is ignored. Hence, a distinct benefit of the proposed model is that we fit only one model to the whole of the data arising from all available cell nuclei, which is more comfortable than fitting one model per cell nucleus and analysing the results afterwards.

Curve-specific effects are taken into account by the use of random effects. These are, however, shrunk towards zero, so that most variability in the data is captured by the fixed effects. The variability of the parameters of interest can however be quantified through the estimation of the variance of the random effects. Hence, the proposed method allows not only to gain estimates of the parameters of interest, that is, the binding rates, but it also allows to gain insight into the variability between cells. That is, the proposed method allows to decompose the total variability in the data into the variability between cells and the remaining variability for example due to noise. The Bayesian technique gives complete posterior distributions for the binding rates in each phase, allowing to compute credible intervals for these binding rates, and, hence, showing the precision of the binding rate point estimates. The mixed-effect approach allows to quantify the variability between cells, which, so far, has not been studied. For example, in our data we see a similar variability between cells in the off-rates in early S phase and late S phase, but the variability in the concentration of bound molecules at equilibrium is higher in the late S phase compared to the early S phase.

Algorithms for nonlinear model fitting have consistency problems by specifying starting values and have convergence issues. Therefore, the model is typically fitted several times using a grid of starting values or random starting values, and the best model is determined using an information criterion like AIC or BIC. This results in a high computational burden. Using a Bayesian approach, the algorithm is guaranteed to converge and the resulting parameter estimates are not dependent on any starting values, which reduces computation time. Moreover, the regression model is very flexible. Mixed-effect priors on the nonlinear parameters can be incorporated easily into the nonlinear regression, which is a novel approach. In addition, the proposed technique allows to analyze all data jointly. For our data, the whole analysis of all data took 42 minutes (diffuse phase), 127 minutes (early S phase), and 138 minutes (late S phase), respectively.

In our approach, the number of MCs for the molecule and the cell cycle phase is a fixed parameter that was adopted from a previous study using a refined compartmental approach (Schneider et al., 2013). To make sure that the number of MCs is still valid with the mixed-effects model, we additionally fitted the mixed-effects model with two MCs for the cell cycle phases diffuse and early S, and the mixed-effects model with one MC for the cell cycle phase late S. For the cell cycle phases diffuse and early S, when fitting the model with two MCs, convergence and redundancy issues arise, which give a hint that the model with two MCs is not the appropriate one. Redundancy issues may for example occur when the exponential rates are too similar. In Reich (1981), a redundancy measure has been used to show that parameters in a sum of two exponentials model are highly redundant if the exponential rates differ by less than a factor of five (Sommer, 2013). The mixed-effects model with one MC for the cell cycle phase late S can be fitted without convergence problems. The resulting DIC is  $-47057.77$  ( $p_D = 63.26$ ,  $D(\theta_{\text{med}}) = -47184.30$ ), which is bigger than the DIC for the mixed-effects model with two MCs ( $DIC = -60138.01$ ). Hence, we conclude that according to the DIC, the model with two MCs is more suitable for the data of the late S phase than the model with one MC. The number of MCs is

moreover biologically sound, as described in Schneider et al. (2013).

We can conclude, that the DIC of the mixed-effects model is lower or in approximately the same range as the DIC of all considered models without random effects for all three cell cycle phases. With the mixed-effects model, we additionally gain precious insight into the variability in the population of cell nuclei in the different cell cycle phases through the estimation of random effects and their variances.

With the proposed mixed-effects model, estimates of the off-rates of the interactions the molecules of interest are involved in, and of the variances of the random effects are attained. Therefore, the model is useful for the analysis of data from FRAP experiments on various similar cell nuclei. With that model, it is no longer necessary to analyze each recovery curve belonging to an experiment on one cell nucleus separately and summarize the results afterwards, or to pool and average the data of experiments on multiple similar cell nuclei to be able to analyze it. The data of FRAP experiments on different cell nuclei can rather be analyzed simultaneously by one single model.

The main goal of this study is to show that the proposed technique can be used for the joint analysis of the data of many cells at once, furthermore providing insight into the variation of the off-rates in the population of cell nuclei. This is a novel approach in the field of FRAP analysis. Although we use a simplified kinetic model here, the approach can easily be adapted to other FRAP experiments and any kinetic model for such FRAP experiments.

# **Chapter 6**

---

## **Analysis of FRAP images at the pixel level**

---

### Abstract

To date, there is still a lack of methodology for the statistical analysis of Fluorescence recovery after photobleaching data. With the imaging technique FRAP, the binding behavior of molecules inside organisms, cells or cellular sub-compartments can be studied *in vivo*. But today, there is still no method for the analysis of FRAP images at the pixel level.

In this chapter, we propose a Bayesian nonlinear model to study the binding behavior of molecules in nuclei of living cells at the pixel level. This model is advantageous in cases where the spatial structure of the nuclei is of interest, that is, if the binding behavior of molecules should not be studied only for some regions of interest, but for each pixel of an image. With such a model, for example, binding sites of molecules in a nucleus could be located.

We present the setup of a simulation study for the evaluation of the model. Program code for the fit of the simulated data was implemented in the programming languages R and C. In order to reduce the runtime, part of the C code was parallelized. However, 1000 iterations of the MCMC algorithm still took approximately five days using 9 cores on a linux server with 64 cores and 512 GB memory, which is rather long. The MCMC algorithm was run with up to 15000 iterations and different settings for the proposal variances of the MH-updates in the MCMC algorithm, but the results were not satisfactory regarding the convergence of the MCMC algorithm.

As manual tuning of the proposal variances for the MH-updates can not be realized within a reasonable time frame and an adaptive MCMC algorithm as an alternative can as well not be realized within a reasonable time frame for our problem, the implementation of an algorithm that tackles the problem described in this chapter in a reasonable time frame is left for future research.

## 6.1 Introduction

To date, FRAP data has been analyzed only for some regions of interest, not for each pixel of an image. Half nucleus FRAP based on ordinary differential equations was performed for example in Feilke, Schneider, and Schmid (2015), Schneider et al. (2013), and Phair et al. (2004). In Beaudouin et al. (2006), the nuclei of the cells of interest were divided into six parts, and an analysis based on partial differential equations was carried out for each of these parts. In this chapter, we introduce a model that can be used for a pixelwise analysis of FRAP images. The model applied to the data of each pixel incorporates information from the neighboring pixels, i.e., we account for diffusion by replacing the term  $D_f \nabla^2 f(t)$  in Equation (5.1) of Chapter 5 by a term describing the exchange of free unbleached molecules between one pixel and its neighbors. Moreover, we assume one binding reaction per pixel.

The remainder of this chapter is organized as follows. In Section 6.2, the nonlinear regression model and the underlying compartment model are introduced. In Section 6.3, the Bayesian nonlinear model consisting of the data model, the prior model and the hyper prior model, is presented. This is followed by a description of the parameter estimation procedure. The setup of a simulation study for the evaluation of the proposed model is described in Section 6.4. Section 6.5 concludes and gives a short discussion.

## 6.2 Nonlinear recovery model

For the model at the pixel level, we use a two-dimensional neighborhood structure where adjacent pixels are neighbors. Therefore, each pixel has up to four neighbors. Pixels that are located at the edge of an image have less than four neighbors, whereas all other pixels have four neighbors. We assume one single mobility class per pixel. The interactions the molecule of interest is involved in are described by on- and off-rates denoted by  $b_1^{\text{on}*}$  and  $b_1^{\text{off}}$ . As stated in Sprague et al. (2004),  $b_1^{\text{on}*}$  is actually a pseudo-on-rate. It is the product of the actual on-rate  $b_1^{\text{on}}$  and the concentration of vacant bindings sites. As we assume that the biological system is in equilibrium before the bleaching and because bleaching does not affect the number of vacant binding sites (Sprague et al., 2004), it is constant during the entire recovery process. To account for diffusion, we include the concentrations of free unbleached molecules in the neighboring pixels into the model. In Figure 6.1, the assumed interaction structure for a pixel with four neighbors is shown. Free molecules are exchanged between the considered pixel and its neighbors at diffusion rates  $b_2^{\text{in}}$  and  $b_2^{\text{out}}$ . Moreover, one interaction the molecule of interest is involved in is illustrated. We assume that bleached and unbleached molecules have the same properties regarding the binding and diffusion behavior. Therefore, in Figure 6.1 and in the following, we concentrate only on the unbleached molecules.

Let  $f(t) = [\text{Free}](t)$  denote the concentration of the free molecules,  $a(t) = [\text{Bound}](t)$  the concentration of the bound molecules, and  $n_k(t) = [\text{Free Neighbor } k](t)$  the concentration of free molecules in neighbor  $k$  at time  $t$ . Let  $K$  denote the number of neighboring pixels, and  $n(t) = \sum_{k=1}^K n_k(t)$  the sum of the concentrations of free molecules in the  $K$  neighboring pixels at time  $t$ . With the assumption that the exchange of molecules with all neighboring pixels occurs at

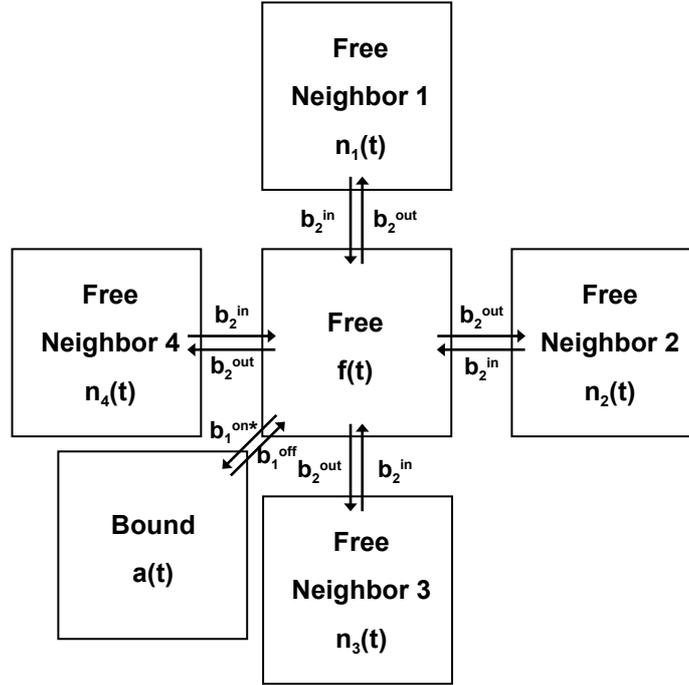


Figure 6.1: Interaction structure for a pixel with four neighbors and one interaction (e.g., binding reaction).

the same rate for each pixel, the change of the concentration of the free and bound unbleached molecules for one pixel can be described by the two differential equations

$$\frac{d}{dt}f(t) = -b_1^{\text{on}*}f(t) + b_1^{\text{off}}a(t) + b_2^{\text{in}}n(t) - Kb_2^{\text{out}}f(t), \quad (6.1)$$

and

$$\frac{d}{dt}a(t) = b_1^{\text{on}*}f(t) - b_1^{\text{off}}a(t). \quad (6.2)$$

Thus, a compartment model with three compartments is used to describe the change of the concentration of unbleached molecules in each pixel. A block diagram for this compartment model is given in Figure 6.2.

As the model introduced here is intended for the analysis of FRAP data sets of GFP-Dnmt1 expressed in mouse C2C12 myoblast cells and in S phase, Dnmt1 is involved in interactions with relatively small off-rates (see for example Schneider et al., 2013, Feilke, Schneider, and Schmid, 2015), i.e., the residence time of Dnmt1 at its binding sites is relatively long compared to the diffusion time,  $b_1^{\text{off}}$  is relatively small compared to  $b_2^{\text{in}}$  and  $b_2^{\text{out}}$ , and the term  $b_1^{\text{off}}a(t)$  in Equation (6.1) can be neglected. Therefore, Equation (6.1) reduces to

$$\frac{d}{dt}f(t) = b_2^{\text{in}}n(t) - (b_1^{\text{on}*} + Kb_2^{\text{out}})f(t). \quad (6.3)$$

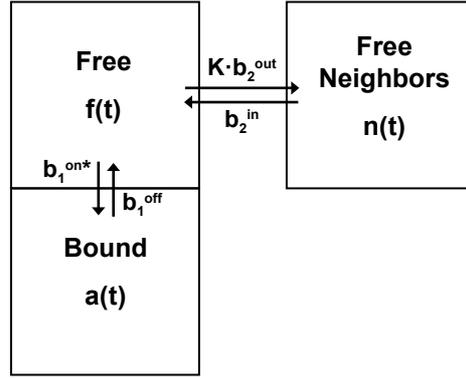


Figure 6.2: Compartment model with three compartments.

With boundary condition  $f(0) = f_0$ , which means that at time  $t = 0$  (the time of the bleaching) the concentration of unbleached free molecules equals  $f_0$ , the solution of this equation is

$$\begin{aligned}
 f(t) &= f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) + b_2^{\text{in}} n(t) * \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) \\
 &= f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) + b_2^{\text{in}} \int_0^t n(\tau) \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})(t - \tau)) d\tau \\
 &= f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) \\
 &\quad \cdot \int_0^t n(\tau) \exp((b_1^{\text{on}*} + Kb_2^{\text{out}})\tau) d\tau.
 \end{aligned} \tag{6.4}$$

We assume that  $n(\tau)$  is known and approximate the integral in (6.4) by a sum:

$$\begin{aligned}
 f(t) &= f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) \int_0^t n(\tau) \exp((b_1^{\text{on}*} + Kb_2^{\text{out}})\tau) d\tau \\
 &\approx f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}})t) \\
 &\quad \cdot \frac{t}{M} \sum_{m=1}^M n(\tau_m) \exp((b_1^{\text{on}*} + Kb_2^{\text{out}})\tau_m),
 \end{aligned} \tag{6.5}$$

with  $\tau_1, \dots, \tau_M$  equally spaced on  $[0, t]$ .

With boundary condition  $a(0) = a_0$ , the solution of Equation (6.2) is

$$\begin{aligned}
 a(t) &= a_0 \exp(-b_1^{\text{off}}t) + b_1^{\text{on}} f(t) * \exp(-b_1^{\text{off}}t) \\
 &= a_0 \exp(-b_1^{\text{off}}t) + b_1^{\text{on}} \int_0^t f(\tau) \exp(-b_1^{\text{off}}(t - \tau)) d\tau \\
 &= a_0 \exp(-b_1^{\text{off}}t) + b_1^{\text{on}} \exp(-b_1^{\text{off}}t) \int_0^t f(\tau) \exp(b_1^{\text{off}}\tau) d\tau.
 \end{aligned} \tag{6.6}$$

The solution for  $f(t)$  in Equation (6.5) is then inserted into the solution for  $a(t)$  in Equa-

tion (6.6). The integral in Equation (6.6) is then again approximated by a sum:

$$\begin{aligned}
a(t) &= a_0 \exp(-b_1^{\text{off}} t) + b_1^{\text{on}} \exp(-b_1^{\text{off}} t) \int_0^t f(\tau) \exp(b_1^{\text{off}} \tau) d\tau \\
&\approx a_0 \exp(-b_1^{\text{off}} t) + b_1^{\text{on}} \exp(-b_1^{\text{off}} t) \\
&\quad \cdot \frac{t}{M} \sum_{m=1}^M \{ [f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}}) \tau_m) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}}) \tau_m) \\
&\quad \cdot \int_0^{\tau_m} n(x) \exp((b_1^{\text{on}*} + Kb_2^{\text{out}}) x) dx] \exp(b_1^{\text{off}} \tau_m) \} \\
&\approx a_0 \exp(-b_1^{\text{off}} t) + b_1^{\text{on}} \exp(-b_1^{\text{off}} t) \\
&\quad \cdot \frac{t}{M} \sum_{m=1}^M \{ [f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}}) \tau_m) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{out}}) \tau_m) \\
&\quad \cdot \frac{\tau_m}{P} \sum_{p=1}^P n(x_p) \exp((b_1^{\text{on}*} + Kb_2^{\text{out}}) x_p)] \exp(b_1^{\text{off}} \tau_m) \},
\end{aligned}$$

with  $\tau_1, \dots, \tau_M$  equally spaced on  $[0, t]$  and  $x_1, \dots, x_P$  equally spaced on  $[0, \tau_m]$ .

For each pixel, the observed value during FRAP recovery is the total fluorescence intensity  $\text{total}(t)$ , which is the sum of the concentration of free unbleached molecules  $f(t)$  and the concentration of bound unbleached molecules  $a(t)$  in a pixel at time  $t$  plus an error.

For the diffusion rates, we assume that  $b_2^{\text{in}} = b_2^{\text{out}}$ , i.e., that the input of molecules from the neighboring pixels happens at the same rate as the excretion of molecules to the neighboring pixels. Therefore, only one diffusion rate has to be estimated and

$$\begin{aligned}
\text{total}(t) &= f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{in}}) t) \\
&\quad + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{in}}) t) \frac{t}{M} \sum_{m=1}^M n(\tau_m) \exp((b_1^{\text{on}*} + Kb_2^{\text{in}}) \tau_m) \\
&\quad + a_0 \exp(-b_1^{\text{off}} t) + b_1^{\text{on}} \exp(-b_1^{\text{off}} t) \\
&\quad \cdot \frac{t}{M} \sum_{m=1}^M \{ [f_0 \exp(-(b_1^{\text{on}*} + Kb_2^{\text{in}}) \tau_m) + b_2^{\text{in}} \exp(-(b_1^{\text{on}*} + Kb_2^{\text{in}}) \tau_m) \\
&\quad \cdot \frac{\tau_m}{P} \sum_{p=1}^P n(x_p) \exp((b_1^{\text{on}*} + Kb_2^{\text{in}}) x_p)] \exp(b_1^{\text{off}} \tau_m) \}.
\end{aligned}$$

### 6.3 Bayesian nonlinear model

We use a hierarchical Bayesian model consisting of the three levels described in Section 4.1 of Chapter 4.

### 6.3.1 Data model

The total observed concentration of unbleached molecules in pixel  $j$ ,  $j = 1, \dots, J$ , at time  $t_i$ ,  $i = 1, \dots, T$ , is denoted by  $C_j(t_i)$ . We assume the model

$$C_j(t_i) = \text{total}_j(t_i) + \varepsilon_{ij} \quad (6.7)$$

for each pixel, where  $\varepsilon_{ij} \sim N(0, \sigma^2)$  are independent Gaussian noise terms and

$$\begin{aligned} \text{total}_j(t_i) = & f_{0j} \exp(-(b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})t_i) \\ & + b_2^{\text{in}} \exp(-(b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})t_i) \frac{t_i}{M} \sum_{m=1}^M n_j(\tau_m) \exp((b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})\tau_m) \\ & + a_{0j} \exp(-b_{1j}^{\text{off}}t_i) + b_{1j}^{\text{on}} \exp(-b_{1j}^{\text{off}}t_i) \\ & \cdot \frac{t_i}{M} \sum_{m=1}^M \{ [f_{0j} \exp(-(b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})\tau_m) + b_2^{\text{in}} \exp(-(b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})\tau_m) \\ & \cdot \frac{\tau_m}{P} \sum_{p=1}^P n_j(x_p) \exp((b_{1j}^{\text{on}*} + K_j b_2^{\text{in}})x_p)] \exp(b_{1j}^{\text{off}}\tau_m) \}. \end{aligned}$$

In this equation,  $K_j$  is the number of neighboring pixels for pixel  $j$ , and  $n_j$  is the sum of the concentrations of free molecules in the  $K$  neighboring pixels of pixel  $j$ .

### 6.3.2 Prior model

In the model in Equation (6.7), there are some pixel-specific parameters as well as some parameters that are common for all pixels.

For the pixel-specific pseudo-on-rate  $b_{1j}^{\text{on}*}$  and the pixel-specific off-rate  $b_{1j}^{\text{off}}$ , we use log-normal distributions

$$b_{1j}^{\text{on}*} = \exp(\theta_{1j}) \sim LN(\mu_1, \tau_1^2), \quad b_{1j}^{\text{off}} = \exp(\theta_{3j}) \sim LN(\mu_3, \tau_3^2)$$

as priors to ensure positiveness, since we know that these rates must be nonnegative (Schmid, Whitcher, Padhani, Taylor, and Yang, 2009). The pixel-specific parameters  $f_{0j}$  and  $a_{0j}$ , which describe the concentration of free and bound unbleached molecules at the time of the bleaching, respectively, sum up to  $\text{total}_j(t_1)$ . As we only observe  $C_j(t_1)$  and not  $\text{total}_j(t_1)$ , we use this quantity as a proxy for  $\text{total}_j(t_1)$ , and use uniform distributions of the form

$$f_{0j} \sim U(0, C_j(t_1)), \quad a_{0j} \sim U(0, C_j(t_1))$$

as priors for the parameters  $f_{0j}$  and  $a_{0j}$ . As  $f_{0j}$  and  $a_{0j}$  are not independent of each other, we additionally introduce the constraint  $a_{0j} + f_{0j} = C_j(t_1)$  for these parameters.

For the common variance parameter  $\sigma^2$ , we use the Inverse Gamma distribution  $\sigma^2 \sim IG(a, b)$ , as this distribution is a conjugate prior for the Gaussian distribution with known mean. For the common diffusion rate  $b_2^{\text{in}}$ , we use the log-normal distribution

$$b_2^{\text{in}} = \exp(\theta_2) \sim LN(\mu_2, \tau_2^2)$$

as prior to incorporate the knowledge that this rate must be nonnegative.

### 6.3.3 Hyper prior model

After the definition of the data model and the prior model, we have to determine the unknown parameters in the prior model, namely the parameters  $\mu_1$ ,  $\mu_2$ ,  $\mu_3$ ,  $\tau_1^2$ ,  $\tau_2^2$ , and  $\tau_3^2$  of the log-normal distributions in Subsection 6.3.2. We use  $\mu_1 = -6$ ,  $\tau_1^2 = 6$ ,  $\mu_2 = -2$ ,  $\tau_2^2 = 2$ ,  $\mu_3 = -10$ , and  $\tau_3^2 = 6$ , to map our assumption that  $b_{1j}^{\text{off}}$  takes the smallest values in expectation, followed by  $b_{1j}^{\text{on}*}$  and  $b_2^{\text{in}}$ , since the residence time of Dnmt1 at its binding sites is relatively long compared to the diffusion time. These values apply to all pixels in an image.

### 6.3.4 Posterior distribution and MCMC inference

We assume *a priori* independence of all unknown parameters. The full conditional distributions of the parameters of the nonlinear model can be derived from the posterior distribution and are given in the Appendix (Part B). To obtain samples from these distributions, a Markov chain Monte Carlo algorithm with Gibbs- and MH-update steps is used. In each iteration of the MCMC algorithm, we update the parameter values for the common as well as for the pixel-specific parameters. The parameter  $\sigma^2$  is drawn in a Gamma Gibbs step, because the full conditional distribution of this parameter is an Inverse Gamma distribution. The remaining parameters are drawn in MH-steps with random walk proposals, because their full conditional distributions are no standard distributions. Similar to Sommer et al. (2014), we do a parallel update of the pixel-specific parameters following a checkerboard pattern, using conditional independence from the remaining pixels given all neighboring pixels. By the checkerboard pattern, the pixels in an image are divided into “even” and “odd” pixels. Therefore, the update of the pixel-specific parameters alternates between the “even” and the “odd” pixels. The update scheme used for the update of all – common and pixel-specific – parameters is given in Figure 6.3.

The sum of the concentrations of free molecules is latent, and hence can not be observed in a FRAP experiment. As it is, however, necessary for the recovery model in Equation (6.7), it has to be calculated in each iteration of the update scheme described above. Therefore, the concentration of free molecules is calculated based on the current values of all parameters for all pixels in an image in each iteration. For the first iteration, starting values for the concentration of free molecules in all pixels have to be calculated. The procedure for the calculation of the starting values is described in the following.

## 1. Initialization:

- Set  $q = 0$ .
- Set starting values for the unknown parameters:  
 $b_{1j}^{\text{on}(0)}, b_{1j}^{\text{off}(0)}, b_2^{\text{in}(0)}, f_{0j}^{(0)}, a_{0j}^{(0)}, \sigma^2(0)$ .
- Compute starting values for the concentration of free molecules in all pixels:  $n_j(t_i)^{(0)}$ .

2. Set  $q = q + 1$ .3. Update the common parameters  $b_2^{\text{in}}$  (MH-step) and  $\sigma^2$  (Gibbs-step)  $\Rightarrow b_2^{\text{in}(q)}, \sigma^2(q)$ .

## 4. Update of the “even” pixels:

- Update the pixel-specific parameters (MH-steps)  
 $\Rightarrow b_{1j}^{\text{on}(q)}, b_{1j}^{\text{off}(q)}, b_2^{\text{in}(q)}, f_{0j}^{(q)}, a_{0j}^{(q)}, \sigma^2(q)$  for all “even” pixels.
- Compute the concentration of free molecules with the help of the current values of all unknown parameters  
 $\Rightarrow n_j(t_i)^{(q)}$  for all “even” pixels.

## 5. Update of the “odd” pixels:

- Update the pixel-specific parameters (MH-steps)  
 $\Rightarrow b_{1j}^{\text{on}(q)}, b_{1j}^{\text{off}(q)}, b_2^{\text{in}(q)}, f_{0j}^{(q)}, a_{0j}^{(q)}, \sigma^2(q)$  for all “odd” pixels.
- Compute the concentration of free molecules with the help of the current values of all unknown parameters  
 $\Rightarrow n_j(t_i)^{(q)}$  for all “odd” pixels.

6. Iterate steps 2-5 until  $q = Q$ , the number of iterations.

Figure 6.3: Update scheme.

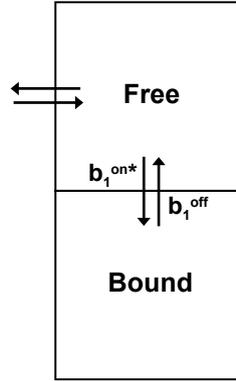


Figure 6.4: Compartment model with two compartments.

### Starting values for the concentration of free molecules in all pixels

In the model we use for the computation of starting values for the concentration of free molecules in all pixels, we assume one MC per pixel and disregard the concentrations of free molecules in the neighboring pixels. A compartment model with two compartments is used to describe the change of the concentration of unbleached molecules in each pixel in an image. A block diagram for this compartment model is given in Figure 6.4.

We use the notation introduced in Section 6.2. The change of the concentration of the free and bound molecules based on the compartmental model in Figure 6.4 can be described by the two differential equations

$$\frac{d}{dt}f(t) = -b^{\text{on}*}f(t) + b^{\text{off}}a(t) + D_f\nabla^2 f(t), \quad (6.8)$$

with  $\nabla^2$  the Laplacian operator and  $D_f$  the diffusion coefficient for free proteins, and

$$\frac{d}{dt}a(t) = b^{\text{on}*}f(t) - b^{\text{off}}a(t). \quad (6.9)$$

These differential equations are special cases of the differential Equations (5.1) and (5.2) in Chapter 5 for  $K = 0$ . As described in Chapter 5, we assume that  $f(t) = f_{\text{eq}}$ , a constant, and therefore, Equation (6.9) can be written as

$$\frac{d}{dt}a(t) = b^{\text{on}*}f_{\text{eq}} - b^{\text{off}}a(t) \quad (6.10)$$

and Equation (6.8) can be ignored. With boundary condition  $a(0) = a_0$ , the solution of Equation (6.10) is

$$a(t) = \frac{b^{\text{on}*}f_{\text{eq}}}{b^{\text{off}}} - \left( \frac{b^{\text{on}*}f_{\text{eq}}}{b^{\text{off}}} - a_0 \right) \exp(-b^{\text{off}}t), \quad (6.11)$$

similar to Equation (5.4) in Chapter 5. We assume again that the system is in equilibrium before the bleaching, i.e., before the bleaching we have  $\frac{d}{dt}f(t) = \frac{d}{dt}a(t) = 0$  and constant steady-state intensities  $f_{\text{eq}}$  and  $a_{\text{eq}}$ . Together with Equation (6.9) we then get

$$a_{\text{eq}} = \frac{b^{\text{on}*} f_{\text{eq}}}{b^{\text{off}}},$$

and can therefore reduce Equation (6.11) to

$$a(t) = a_{\text{eq}} - (a_{\text{eq}} - a_0) \exp(-b^{\text{off}}t).$$

For the sum of the concentration of free unbleached molecules and the concentration of bound unbleached molecules in a pixel at time  $t$  applies

$$\text{total}(t) = f(t) + a(t) = f_{\text{eq}} + a_{\text{eq}} - (a_{\text{eq}} - a_0) \exp(-b^{\text{off}}t).$$

With  $p_0 = f_{\text{eq}} + a_{\text{eq}}$  and  $p_1 = a_{\text{eq}} - a_0$ , this yields

$$\text{total}(t) = p_0 - p_1 \exp(-b^{\text{off}}t). \quad (6.12)$$

To assure that we get only nonnegative values for the fit of the recovery curve, we write  $p_0$  in Equation (6.12) as  $p_1 + \delta$  with  $\delta = a_0 + f_{\text{eq}}$  and fix  $\delta \geq 0$ . We use the following model, which can be derived from the nonlinear recovery model in Equation (6.12):

$$\begin{aligned} C_j(t_i) &= \text{total}_j(t_i) + \varepsilon_{ij} \\ &= p_{1j} + \delta_j - p_{1j} \exp(-b_j^{\text{off}}t_i) + \varepsilon_{ij}. \end{aligned}$$

We fitted this model with a nonlinear least squares method. A Levenberg-Marquardt algorithm implemented in the R function `nls.lm` (package `minpack.lm` (Elzhov et al., 2013)) was used. The fit of all pixels of an image was done in parallel. The R package `multicore` (Urbanek, 2014) was used for the parallelization.

## 6.4 Simulation

We simulated a FRAP image that is similar to a FRAP image showing the fluorescence recovery of Dnmt1. The simulated image consists of 18 pixels. We simulated half-nucleus FRAP (as opposed to circle FRAP or strip FRAP (Mueller et al., 2008; Sprague et al., 2004)), so that half of the pixels lie in the bleached area of the cell nucleus, and the other half lies in the area of the nucleus that has not been bleached. For the diffusion rate, which is common for all pixels, we assume  $b_2^{\text{in}} = 0.1$ . We assume that the pixel-specific off-rate  $b_{1j}^{\text{off}}$  is smaller than the pixel-specific pseudo-on-rate  $b_{1j}^{\text{on}*}$  for all pixels. As the off-rate and the pseudo-on-rate can vary between the pixels, we split the pixels in the image in three parts according to their off- and pseudo-on-rates. For one third of the pixels, we assume  $b_{1j}^{\text{on}*} = 6 \cdot 10^{-4}$  and  $b_{1j}^{\text{off}} = 6 \cdot 10^{-5}$ . For the second third of the pixels, we assume  $b_{1j}^{\text{on}*} = 6 \cdot 10^{-4}$  and  $b_{1j}^{\text{off}} = 9 \cdot 10^{-5}$ , and for the last third we

assume  $b_{1j}^{\text{on}^*} = 9 \cdot 10^{-4}$  and  $b_{1j}^{\text{off}} = 6 \cdot 10^{-5}$ . We assume  $f_{0j}$  and  $a_{0j}$  are zero for the pixels in the bleached area of the nucleus, and draw random numbers from the following normal distributions for the pixels in the unbleached area of the nucleus:  $f_{0j} \sim N(0.09, 10^{-6})$ ,  $a_{0j} \sim N(0.01, 10^{-6})$ . We assume that the concentration of the unbleached molecules is available for 781 time points. The first measurement has been done at the time of the bleaching of half the nucleus, and the following measurements have been done every 0.15 seconds after the bleaching. With the help of the differential Equations (6.2) and (6.3), the concentrations of the free and bound unbleached molecules can be calculated for all simulated time points. We discretize the differential quotients  $\frac{d}{dt}a(t)$  and  $\frac{d}{dt}f(t)$  in Equations (6.2) and (6.3), so that we have the difference quotients  $\frac{\Delta a}{\Delta t}$  and  $\frac{\Delta f}{\Delta t}$ , respectively:

$$\frac{\Delta a}{\Delta t} = b_1^{\text{on}^*} f(t) - b_1^{\text{off}} a(t),$$

$$\frac{\Delta f}{\Delta t} = -(b_1^{\text{on}^*} + K b_2^{\text{out}}) f(t) + b_2^{\text{in}} n(t).$$

Consequently, the change of the concentrations of free and bound unbleached molecules per  $\Delta t$  is

$$\Delta a = (b_1^{\text{on}^*} f(t) - b_1^{\text{off}} a(t)) \Delta t$$

and

$$\Delta f = (-(b_1^{\text{on}^*} + K b_2^{\text{out}}) f(t) + b_2^{\text{in}} n(t)) \Delta t.$$

For each simulated time point and for each pixel  $j$ , we computed

$$C_j(t_i) = f_j(t_i) + a_j(t_i) + \varepsilon_{ij} = \text{total}_j(t_i) + \varepsilon_{ij}$$

with  $\varepsilon_{ij} \sim N(0, 10^{-4})$ , that means, we added Gaussian noise with variance  $\sigma^2 = 10^{-4}$  to the simulated recovery curves of all pixels. The simulated recovery curves with Gaussian noise are given in Figure 6.5.

## 6.5 Conclusions and Discussion

The code for the fit of the simulated data was written in the programming languages R and C. The part of the code containing the Gibbs and MH-updates of the parameters of the nonlinear model was written in C, whereby the updates of the pixel-specific parameters was done in parallel according to the checkerboard pattern described in Subsection 6.3.4 for all pixels of the simulated image in order to reduce the runtime.

However, even with the parallelization in C, 1000 iterations of the MCMC algorithm took approximately five days using 9 cores on a linux server with 64 cores and 512 GB memory, what is a rather long time. Therefore, the MCMC algorithm was run with different settings for the

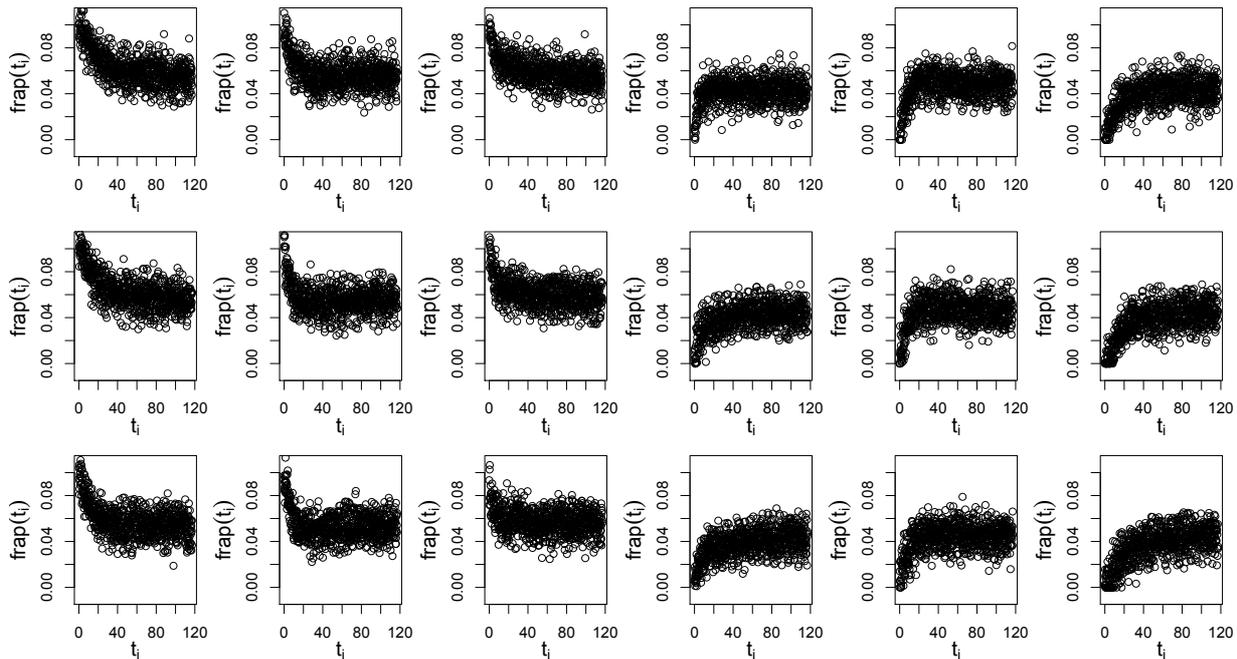


Figure 6.5: Simulated recovery curves with Gaussian noise.

proposal variances of the MH-updates for up to 15000 iterations without tuning the proposal variances. But the results were not satisfactory regarding the convergence of the MCMC algorithm. Manual tuning of the proposal variances, however, can not be realized within a reasonable time frame.

An alternative to the MCMC algorithm with manual tuning of the variances would be to use an adaptive MCMC algorithm as described in Haario et al. (2001) and applied for example in Weidemann et al. (2014). As for the simplest variant in Haario et al. (2001), however, 20000 iterations were needed, we found that an adaptive MCMC algorithm can as well not be realized within a reasonable time frame for our problem. Another possibility would be to use analytic approximations of the posterior distribution, such as Laplace's method (Tierney and Kadane, 1986). However, such methods are mostly not of sufficient accuracy for a moderate or high number of parameters, and it is difficult to numerically compute the associated Hessian matrices (Carlin and Louis, 2009). In structured additive regression models that include for example (generalized) linear models and (generalized) additive models, Integrated Nested Laplace Approximation can be used instead of MCMC sampling (Rue et al., 2009), and with that, the computational effort can be reduced. However, this approach can not be used with nonlinear regression models as the one used in this chapter.

Therefore, the implementation of an algorithm that tackles the problem described in this chapter in a reasonable time frame is left for future research.



# Chapter 7

---

## Boosting in nonlinear regression models with an application to DCE-MRI data

---

---

This chapter is mainly based on Feilke, Bischl, Schmid, and Gertheiss (2015). Compared to Feilke, Bischl, Schmid, and Gertheiss (2015), the abstract is slightly changed, and Subsection 7.2.1 is more detailed. Section 7.5 is in addition to the content in Feilke, Bischl, Schmid, and Gertheiss (2015), and Section 7.6 is accordingly changed. Some of the Figures of the Appendix in Feilke, Bischl, Schmid, and Gertheiss (2015) are additionally contained in this chapter and the content of this chapter is accordingly changed at some points compared to Feilke, Bischl, Schmid, and Gertheiss (2015).

### Abstract

For the statistical analysis of dynamic contrast-enhanced magnetic resonance imaging data, compartment models are a commonly used tool. By these models, the observed uptake of contrast agent in some tissue over time is linked to physiologic properties like capillary permeability and blood flow. Up to now, models of different complexity have been used, and it is still unclear which model should be used in which situation. In previous studies, it has been found that for DCE-MRI data, the number of compartments differs for different types of tissue, and that in cancerous tissue, it might actually differ over a region of voxels of one DCE-MR image.

The objective is to find the appropriate number of compartments and estimate the parameters of a regression model for each voxel in an DCE-MR image. With that, tumors in an DCE-MR image can be located, and for example therapy success can be assessed.

The observed uptake of contrast agent in a voxel of an image of some tissue is described by a concentration time curve. This curve can be modeled using a nonlinear regression model. We present a boosting approach with nonlinear regression as base procedure, which allows us to estimate the number of compartments and the related parameters for each voxel of an DCE-MR image. In addition, a spatially regularized version of this approach is proposed.

With the proposed approach, the number of compartments – and with that the complexity of the model – per voxel is not fixed but data-driven, which allows us to fit models of adequate complexity to the concentration time curves of all voxels. The parameters of the model remain nevertheless interpretable because of the underlying compartment model.

The proposed boosting approaches outperform all competing methods considered in this chapter regarding the correct localization of tumors in DCE-MR images as well as the spatial homogeneity of the estimated number of compartments across the image, and the definition of the tumor edge.

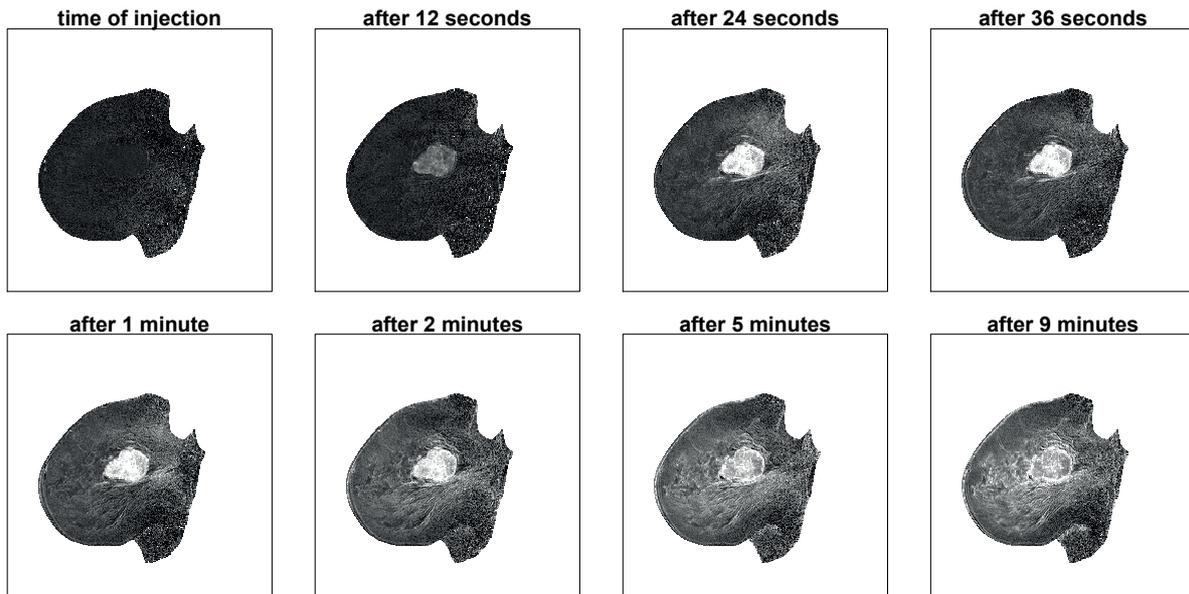


Figure 7.1: Excerpt of the DCE-MRI series of a breast cancer patient showing the concentration of contrast agent at eight different time points.

## 7.1 Introduction

DCE-MRI is an imaging technique by which the blood supply of a tissue of interest can be recorded *in vivo*. A series of images is acquired by an MRI scanner that captures images of the tissue at several time points after the injection of a contrast agent (CA). An excerpt of the DCE-MRI series of a breast cancer patient showing the concentration of contrast agent at eight different time points can be found in Figure 7.1. The first image has been taken at the time of the injection of the CA, and the following images show the subsequent perfusion of the tissue with CA after 12, 24, and 36 seconds and after 1, 2, 5, and 9 minutes.

For each voxel of an image, the concentration of CA at each time point can be computed from the MR signal (Buckley and Parker, 2005). From the dynamic behavior of CA uptake, tumors can be located, malignancy and types of tumors can be inferred, tumors can be graded and therapy success can be assessed (Padhani et al., 2005; Liu et al., 2005).

For the analysis of the CA uptake behavior, model-based or data-driven methods can be used. An advantage of model-based methods is that they result in quantitative physiological parameters which characterize the amount and rate of capillary leakage (Padhani et al., 2005), as they are based on pharmacokinetic models describing the exchange of CA between different, well-mixed compartments (Tofts and Kermode, 1991). For data-driven methods such as nonparametric regression, usually no *a priori* compartment-structure is assumed (Schmid, Whitcher, Padhani, and Yang, 2009).

So far, several compartment models with a varying, *a priori* fixed, number of compartments have been proposed and it is not clear which model should be used in which situation. The two-

compartment exchange models proposed by Brix et al. (2009) and Sourbron and Buckley (2011), for example, consists of two different compartments for arterial plasma and interstitial plasma. Furthermore, a hierarchical Bayesian two-compartment model for the analysis of DCE-MRI data on voxel level has been proposed by Sommer and Schmid (2014), and a multi-tissue compartment model with a fixed number of compartments has been proposed by Port et al. (1999).

In different types of tissue, however, different numbers of compartments might be needed. Moreover, as tumor tissue is often heterogeneous, the adequate number of compartments might even vary over a field of voxels. Therefore, we propose to estimate the number of compartments for each voxel from the data.

Sommer et al. (2014) recently proposed a spatially regularized estimation method based on a multi-tissue compartment model for the estimation of the number of compartments and the related parameters per voxel. Thereby, they combined the advantages of model-based and data-driven methods, as the number of compartments is chosen data-dependently, but biologically interpretable parameter estimates are obtained.

Similar to Sommer et al. (2014), we base our approach on a multi-tissue compartment model. We, however, propose a more straightforward method to estimate the number of compartments and the parameters of this model: a boosting approach with nonlinear regression as base procedure. The original boosting algorithm (Freund and Shapire, 1996) arose in the machine learning community and was mainly used for classification Mayr et al. (2014a). Later, the concept of boosting was adapted to the field of regression modeling (Friedman et al., 2000; Friedman, 2001; Bühlmann and Yu, 2003; Bühlmann, 2006; Bühlmann and Hothorn, 2007), where it can be used in various settings to select predictors and estimate their effects on a univariate continuous response (Mayr et al., 2014b). Gradient boosting algorithms are currently gaining attention, as they can be very useful to address important research questions in modern biomedicine (Mayr et al., 2014b).

The boosting algorithm that will be introduced in Subsection 7.2.4 together with the voxelwise or the spatially regularized estimation procedure presented in Subsections 7.2.5 and 7.2.6 is a completely novel approach for a regression setting where the univariate response variable is described by a nonlinear parametric function. Specifically, it tackles the problem of the estimation of the number of compartments in a multi-tissue compartment model and the related parameter estimation. To the best of our knowledge, in contrast to the extensive literature on boosting in additive regression, boosting used in nonlinear parametric regression is described for the first time in Feilke, Bischl, Schmid, and Gertheiss (2015) and in this chapter.

The remainder of this chapter is organized as follows. In Section 7.2, the compartment model and the nonlinear regression model built upon it are introduced. Moreover, gradient boosting in general, the novel boosting algorithm, and the associated voxelwise and spatially regularized estimation procedures are described. In Section 7.3, the simulation setup and the results of the simulation studies are presented. The evaluation of the proposed method for *in vivo* data from a breast cancer study is found in Section 7.4. Section 7.5 deals with the assessment of therapy success in the framework of this study, and Section 7.6 gives a discussion on this chapter and concludes.

With the R package `dcmriboost`, which is available on GitHub (<https://github.com/feilke/dcmriboost>), voxelwise and spatially regularized boosting for

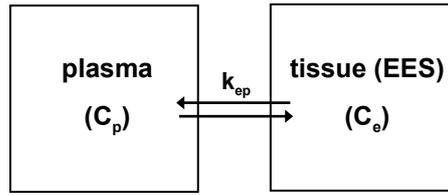


Figure 7.2: Tofts model.

DCE-MRI data as described in this chapter can be performed. This package contains the simulated data of Simulation 1 introduced in Section 7.3 of this chapter. Moreover, we uploaded two accompanying R files to GitHub. These files provide code for performing voxelwise boosting and spatially regularized boosting ( $\lambda = 10^{-10}$ ) for the data of Simulation 1.

## 7.2 Methods

### 7.2.1 Compartment model

For the analysis of DCE-MRI data, several compartment models with a varying number of compartments have been used. A fundamental compartment model is the so-called Tofts model (Tofts, 1997). A block diagram for the Tofts model is given in Figure 7.2.

In this model, the perfusion of CA is described with the help of an arterial plasma compartment (APC) and an interstitial space compartment, which is also called extravascular extracellular space (EES) (Tofts, 1997; Tofts et al., 1999). It is assumed that the contrast agent enters only the interstitial space and not the cells. This assumption holds typically when using low-molecular-weight tracers such as gadolinium diethylenetriaminepentaacetic acid (Gd-DTPA). CA is exchanged between the APC and the EES at constant rate  $k_{ep}$ .

The change of CA concentration in the EES can be described by the differential equation

$$\frac{d}{dt}C_e(t) = k_{ep}C_p(t) - k_{ep}C_e(t), \quad (7.1)$$

where  $C_e(t)$  is the CA concentration in the EES at time  $t$ , and  $C_p(t)$  the CA concentration in plasma at time  $t$ .  $C_p(t)$  is also called arterial input function (AIF). Given the initial condition  $C_e(0) = 0$ , which holds because prior to the CA injection, the concentration of CA in the EES is zero, the solution of this equation is

$$C_e(t) = k_{ep}C_p(t) * \exp(-k_{ep}t),$$

where  $*$  is the convolution operator such that

$$C_p(t) * \exp(-k_{ep}t) = \int_0^t C_p(t - \tau) \exp(-k_{ep}\tau) d\tau$$

as in Sommer et al. (2014). In the Tofts model, the concentration  $C(t)$  of CA at time  $t$  is the product of the concentration  $C_e(t)$  of CA in the EES at time  $t$  and the fractional volume of the

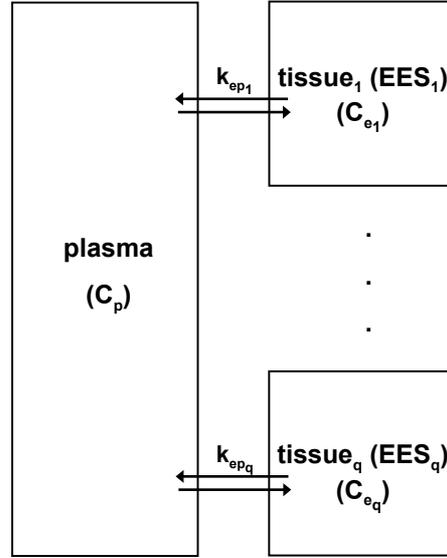


Figure 7.3: Multi-tissue compartment model with  $q$  tissue compartments.

EES  $v_e$ :

$$C(t) = v_e C_e(t).$$

With  $K^{\text{trans}} := v_e k_{ep}$  we have:

$$C(t) = C_p(t) * K^{\text{trans}} \exp(-k_{ep}t).$$

The parameters of interest, which have to be estimated, are  $k_{ep}$  and  $K^{\text{trans}}$ .  $k_{ep}$  is the rate constant at which the EES exchanges CA with the APC, and  $K^{\text{trans}} = v_e k_{ep}$  is the associated volume transfer constant. Observed tracer kinetics can often be well described by the Tofts model for healthy tissue despite the simplifying assumptions made (Sommer, 2013). However, as the microvasculatur in tumors often is highly heterogeneous (Yang and Kopp, 2011), the Tofts model fails to describe its observed uptake dynamics (Schmid, Whitcher, Padhani, Taylor, and Yang, 2009; Port et al., 1999), and therefore, more complex compartment models are needed (Sommer, 2013).

We use a multi-tissue compartment model with  $q$  tissue compartments. A block diagram for this model is found in Figure 7.3. This model is a generalization of the Tofts model:

$$C(t) = \sum_{k=1}^q v_{e_k} C_{e_k}(t) = \sum_{k=1}^q C_p(t) * K_k^{\text{trans}} \exp(-k_{ep_k}t). \quad (7.2)$$

It can also be obtained by solving a system of differential equations derived from the compartment model in Figure 7.3 similar to Equation 7.1 with some initial conditions (Seber and Wild, 1989).

With this model, it is possible to map tissue heterogeneity on the voxel level. For each voxel, we assume the compartment model in (7.2), and estimate the number of compartments  $q$ . Therefore, the number of compartments can vary over the voxels of an image. Moreover, each tissue compartment  $k$ ,  $k = 1, \dots, q$ , has a unique rate constant  $k_{\text{ep}_k}$  at which CA is exchanged between the APC and the EES, as well as an individual volume transfer constant  $K_k^{\text{trans}}$ . Of course, these parameters have to be estimated in addition to  $q$ .

### 7.2.2 Nonlinear regression model

Let the observed concentration of CA in voxel  $i$ ,  $i = 1, \dots, N$ , at time  $x_t$ ,  $t = 1, \dots, T$ , be denoted by  $Y_i(x_t)$ , and let  $C_i(x_t)$  be the expected concentration of CA in voxel  $i$  at time  $x_t$ . We assume that  $C_i(x_t)$  can be described by the nonlinear multi-tissue compartment model with  $q_i$  tissue compartments

$$C_i(x_t) = \sum_{k=1}^{q_i} C_p(x_t) * K_{ik}^{\text{trans}} \exp(-k_{\text{ep}_{ik}} x_t). \quad (7.3)$$

Furthermore, we assume that the observed concentration of CA in voxel  $i$  can be modeled by

$$Y_i(x_t) = C_i(x_t) + \varepsilon_{it} = \sum_{k=1}^{q_i} C_p(x_t) * K_{ik}^{\text{trans}} \exp(-k_{\text{ep}_{ik}} x_t) + \varepsilon_{it}, \quad (7.4)$$

where  $\varepsilon_{it}$  are independent Gaussian noise terms with mean zero and variance  $\sigma_i^2$  (Schmid et al., 2006).

For  $k_{\text{ep}_{ik}}$ , values between 0.05 and 20 are feasible, considering that the rate is positive and does typically not exceed 20 (Schmid et al., 2006), and for  $K_{ik}^{\text{trans}}$ , values in the interval  $[0.01, 20]$  are considered (Schmid et al., 2006; Parker et al., 1998), as these ranges of values are biologically realistic. Since in breast DCE-MRI there are no big vessels in the captured area from which  $C_p(x_t)$  could be measured, we use a bi-exponential population based arterial input function as proposed by Tofts and Kermode (1991):  $C_p(x_t) = D(a_1 \exp(-m_1 x_t) + a_2 \exp(-m_2 x_t))$ , with constant dose  $D$  of tracer (in mmol/kg) depending on the experimental conditions and fixed values  $a_1 = 3.99 \text{ kg/l}$ ,  $a_2 = 4.78 \text{ kg/l}$ ,  $m_1 = 0.144 \text{ min}^{-1}$ ,  $m_2 = 0.0111 \text{ min}^{-1}$  (Tofts and Kermode, 1991; Weinmann et al., 1984).

### 7.2.3 Gradient boosting

We use the boosting algorithm that is introduced in the following subsection to fit model (7.4), and thus to estimate the model complexity for each voxel of an MR image. We use a notation similar to the notation in Hofner et al. (2014), and adapt the general methodology presented there to the voxelwise DCE-MRI framework in this chapter.

We consider the response  $Y$  and the predictor function  $C =: f$ . The aim is to receive an optimal prediction of  $Y$ , which is achieved by minimizing the loss function  $\rho(Y, f) \in \mathbb{R}$  over all admissible prediction functions  $f$ . This means, the optimal prediction function

$$f^* := \arg \min_f \mathbb{E} [\rho(Y, f)] \quad (7.5)$$

has to be estimated, where  $\rho$  is assumed to be differentiable with respect to  $f$ . In our approach, we consider a nonlinear regression with response  $Y \in \mathbb{R}$  and use the  $L_2$  loss function  $\rho(Y, f) = \frac{1}{2}(Y - f)^2$ . As in practice, we deal with realizations  $Y(x_t)$ ,  $t = 1, \dots, T$ , of  $Y$ , the empirical risk  $\mathcal{R} := \sum_{t=1}^T \rho(Y(x_t), f(x_t))$  has to be minimized instead of the expected loss in equation (7.5).

### 7.2.4 Boosting algorithm

For each voxel  $i$ , we use the boosting algorithm in Figure 7.4 to minimize the empirical risk  $\mathcal{R}$  over  $f$ .

By step 4.c), model choice is performed, because compartments are iteratively accepted into the model only if the model is improved by the acceptance. A new compartment is accepted only if its  $k_{ep}$ -value differs substantially (i.e., by a factor of at least 5) from the  $k_{ep}$ -values of the already accepted compartments. We set a factor for the relative difference of the  $k_{ep}$ -values in the multi-tissue compartment model, because redundancy issues may occur if the exponential rates of two compartments are too similar. In Reich (1981), it has been shown that parameters in a sum of exponentials model are highly redundant if the exponential rates differ by less than a factor of 5 (Sommer, 2013). It is difficult to obtain a generalization of that for convolved exponentials, but it can be assumed that parameters are as well redundant in this case if the decay rates differ too little (Sommer, 2013). Therefore, we use the same factor also for the convolved exponentials. Moreover, in each iteration, we check if the RSS is decreased at least by a factor of  $10^{-8}$ .

The same base-learner can be selected several times during the execution of the boosting algorithm. Because of the additive update of compartments, the final boosting estimate of voxel  $i$  in iteration  $m_{stop_i}$  can be interpreted as an additive prediction function. The boosting algorithm iteratively fits the gradient of the loss function instead of fitting the original observations directly. In each update step, the current estimate  $\hat{u}_i^{[m]}$  is multiplied by the step length factor  $\nu$ . Therefore, the stepwise increments of the final estimator  $\hat{f}_i^{[m]}$  are small, and the overall minimum is only slowly approached. The choice of the step length factor  $\nu$  is of minor importance, as long as the chosen value is small (Schmid and Hothorn, 2008). It is essential that a small  $\nu$  is chosen, because otherwise the algorithm could overshoot the minimum of the empirical risk. We use  $\nu = 0.1$ , which is a typical value (Hofner et al., 2014). As  $u_i^{[m]}$  depends on  $x_t$ , it is from now on denoted by  $u_i(x_t)^{[m]}$ .

A substantial issue in the gradient boosting framework is the choice of the stopping iteration  $m_{stop}$ . Boosting algorithms should not be run until convergence, because otherwise, overfitting would be the consequence (Bühlmann and Hothorn, 2007; Mayr et al., 2012). Here, the optimal  $m_{stop}$  is chosen according to the Bayesian Information Criterion (BIC) (Hastie et al., 2009; Hansen and Yu, 2001). For the computation of the BIC, the degrees of freedom of the model have to be defined. As we fit a compartment model where each compartment is defined by two parameters that are biologically interpretable, we think that the number of parameters in the model is a somewhat natural definition of the degrees of freedom here.

1. Initialization: Set  $m = 0$ . Set  $\hat{f}_i^{[0]} = 0$ , because we fit a model without intercept. Generally,  $\hat{f}_i^{[m]}$  denotes the vector of function estimates in iteration  $m$  for voxel  $i$ .

2. Specify a set of base-learners. Here, the base-learners are single compartments

$$C_p(x_t) * K_{ik}^{\text{trans}} \exp(-k_{\text{ep}_{ik}} x_t) \quad (7.6)$$

from the multi-tissue compartment model in (7.3).

3. Set  $m = m + 1$ .

4. a) Compute the negative gradient of the loss function evaluated at the function estimates of the previous iteration:  $u_i^{[m]} = Y_i - \hat{f}_i^{[m-1]}$ , where  $Y_i$  is the original response vector.

b) Fit the nonlinear model in (7.6) to the negative gradient vector  $u_i^{[m]}$ . This is done by a nonlinear least squares method. As optimization procedure, a Levenberg-Marquardt algorithm implemented in the R function `nls.lm` (package `minpack.lm` (Elzhov et al., 2013)) is used (Ahearn et al., 2005). Hereby, the best base-learner, i.e., the base-learner that fits  $u_i^{[m]}$  best according to the residual sum of squares criterion is selected. The function estimate of the best base-learner, which minimizes the RSS, is denoted by  $\hat{u}_i^{[m]}$ .

As the output of the optimization procedure may depend on the starting values for the unknown parameters, ten initial guesses are drawn by latin hypercube sampling from the grid stretched by the biologically realistic ranges for the two parameters ( $k_{\text{ep}_{ik}} \in [0.05, 20]$ ,  $K_{ik}^{\text{trans}} \in [0.01, 20]$ ). The optimization procedure is performed with all ten initial guesses and the base-learner that leads to the smallest RSS is chosen. Details on the choice of the number of initial guesses are given in the final part of this section.

c) Decide if a new compartment is accepted into the model. We use the fact that a compartment  $k$  is characterized by its  $k_{\text{ep}}$ -value and

- accept the new  $k_{\text{ep}}$ -value, i.e., the new compartment, if the estimated  $k_{\text{ep}}$ -value resulting from the optimization procedure in b) differs from the  $k_{\text{ep}}$ -values of the compartments that have already entered the model in the previous iterations by a factor of at least 5.
- do not accept the new compartment otherwise. In this case, the model is fitted again with the previously accepted  $k_{\text{ep}}$ -value, i.e., only the parameter  $K^{\text{trans}}$  is estimated. The estimator  $\hat{K}^{\text{trans}}$  is determined analytically. Details can be found in the final part of this section.

d) Update  $\hat{f}_i^{[m]} = \hat{f}_i^{[m-1]} + v \hat{u}_i^{[m]}$ , with a step length factor  $v$ ,  $0 < v \leq 1$ .

5. Iterate steps 3 and 4 until  $m = m_{\text{stop}_i}$ , the stopping iteration.

Figure 7.4: Boosting algorithm.

|                                    | $\overline{MSE}$    | $\hat{q}_{av}$ | $q_c$ |
|------------------------------------|---------------------|----------------|-------|
| Boosting voxelwise (10 restarts)   | $1.7 \cdot 10^{-4}$ | 2.04           | 0.71  |
| Boosting voxelwise (1000 restarts) | $1.7 \cdot 10^{-4}$ | 2.03           | 0.71  |

Table 7.1: Average values for the 10 simulated images: Average MSE ( $\overline{MSE}$ ), estimated average number of tissue compartments ( $\hat{q}_{av}$ ), and percentage of voxels for which the number of tissue compartments is correctly estimated ( $q_c$ ).

### Details on the optimization in the boosting algorithm

The number of initial guesses for the optimization in step 4.b) of the boosting algorithm in Subsection 7.2.4 was determined in a benchmark experiment. In this experiment, we fitted a compartment model with one tissue compartment for the concentration time curves (CTCs) of the voxels contained in an image with four different algorithms:

- Levenberg Marquart (R function `nls.lm` in R package `minpack.lm` (Elzhov et al., 2013))
- Nelder-Mead (R function `constrOptim`)
- L-BFGS-B (R function `optim`)
- `rgenoud` (function `genoud` in R package `rgenoud` (Mebane and Jasjeet, 2011))

For each algorithm, we tried different numbers of restarts. The benchmark experiment was performed for all simulated images and two of the images from the breast cancer study. For each CTC and for each algorithm, we recorded the loss and the running time. We averaged the loss and the running time for each algorithm over all CTCs and determined the pareto front. According to the pareto front, we considered the Levenberg Marquart algorithm with 1-3 restarts a good solution. Conservatively, we chose the Levenberg Marquart algorithm with 10 restarts for the optimization in step 4.b).

To make sure that 10 restarts with the Levenberg Marquart algorithm are enough, we additionally ran the voxelwise boosting algorithm for the simulated data with 10 restarts and with 1000 restarts in the optimization in step 4.b). The results can be found in Table 7.1. The average mean squared error (MSE) for the 10 simulated images with 1000 restarts was  $1.7 \cdot 10^{-4}$ , the average number of compartments was 2.03, and for 71% of the voxels, the number of tissue compartments was estimated correctly. With 10 restarts, the MSE was  $1.7 \cdot 10^{-4}$ , the average number of compartments was 2.04, and for 71% of the voxels, the number of tissue compartments was estimated correctly. Hence, we concluded that 10 restarts are sufficient, because the average MSE and the percentage of voxels for which the correct number of tissue compartments was estimated were the same as with 1000 restarts, and the average number of compartments differed only marginally. Boxplots of the voxelwise MSE for all 10 simulated images for the comparison of the voxelwise boosting with 10 restarts and with 1000 restarts are plotted in Figure 7.5.

In step 4.c) of the boosting algorithm, in the case that we do not accept a new compartment,  $\hat{K}^{\text{trans}}$  can be determined analytically. By using a fixed  $k_{ep}$ -value, which is the previously accepted

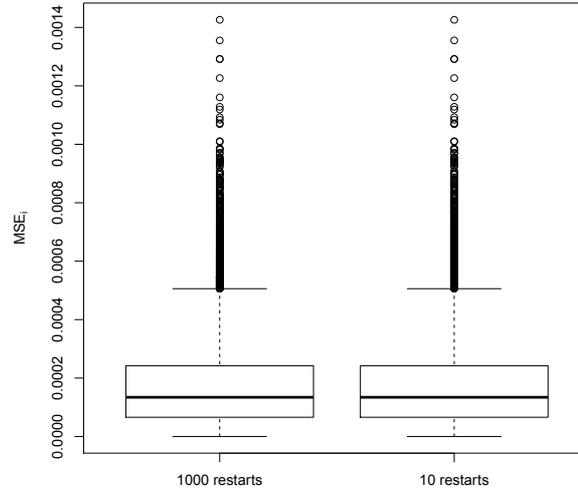


Figure 7.5: Voxelwise MSE for all 10 simulated images; voxelwise boosting with 10 restarts and with 1000 restarts.

$k_{ep}$ -value, the nonlinear regression model is linearized. Therefore, we have to solve a convex optimization problem and there is one local minimum which is also the global minimum. The solution of the optimization problem is given in what follows. We start with the nonlinear least squares estimate given in Equation 7.7 of Section 7.2.5:

$$\left( \hat{K}_{ik}^{\text{trans}[m]}, \hat{k}_{ep_{ik}}^{[m]} \right)' = \arg \min_{K,k} \left\{ \sum_t \left( u_i(x_t)^{[m]} - C_p(x_t) * K \exp(-kx_t) \right)^2 \right\},$$

with side conditions  $0.05 \leq \hat{k}_{ep_{ik}}^{[m]} \leq 20$  and  $0.01 \leq \hat{K}_{ik}^{\text{trans}[m]} \leq 20$ . As  $k$  is fixed in our case, with the convolution given in the Appendix of Sommer (2013) and

$$b(x_t) = \frac{a_1}{m_1 - k} (\exp(-kx_t) - \exp(-m_1x_t)) + \frac{a_2}{m_2 - k} (\exp(-kx_t) - \exp(-m_2x_t)),$$

this can be written as

$$\left( \hat{K}_{ik}^{\text{trans}[m]} \right)' = \arg \min_K \left\{ \sum_t \left( u_i(x_t)^{[m]} - DKb(x_t) \right)^2 \right\},$$

with side condition  $0.01 \leq \hat{K}_{ik}^{\text{trans}[m]} \leq 20$ . Minimizing

$$\sum_t \left( u_i(x_t)^{[m]} - DKb(x_t) \right)^2$$

means differentiating this function with respect to  $K$ , equating it to zero and solving the equation for  $K$ . This leads to

$$\hat{K}_{ik}^{\text{trans}[m]} = \begin{cases} 0.01, & K < 0.01 \\ \frac{\sum_t u_i(x_t)^{[m]} b(x_t)}{D \sum_t b(x_t)^2}, & 0.01 \leq K \leq 20 \\ 20, & K > 20 \end{cases}$$

The solution is a minimum as the second derivative with respect to  $K$ , which is  $2D^2 \sum_t b(x_t)^2$ , is greater than zero.

### 7.2.5 Voxelwise estimation

In a first approach, we estimate the parameters  $k_{\text{ep}_i} = (k_{\text{ep}_{i1}}, \dots, k_{\text{ep}_{iq_i}})'$  and  $K_i^{\text{trans}} = (K_{i1}^{\text{trans}}, \dots, K_{iq_i}^{\text{trans}})'$  independently for all voxels in an image, i.e., for each voxel  $i$ , we fully run the boosting algorithm described in Section 7.2.4.

In each iteration  $m$ , we have to minimize the residual sum of squares  $\sum_t (u_i(x_t)^{[m]} - \hat{u}_i(x_t)^{[m]})^2$ . Hence, the nonlinear least squares estimate is

$$\left( \hat{K}_{ik}^{\text{trans}[m]}, \hat{k}_{\text{ep}_{ik}}^{[m]} \right)' = \arg \min_{K, k} \left\{ \sum_t \left( u_i(x_t)^{[m]} - C_p(x_t) * K \exp(-kx_t) \right)^2 \right\}, \quad (7.7)$$

with side conditions  $0.05 \leq \hat{k}_{\text{ep}_{ik}}^{[m]} \leq 20$  and  $0.01 \leq \hat{K}_{ik}^{\text{trans}[m]} \leq 20$ .

### 7.2.6 Spatially regularized estimation

To take into account the spatial structure of the voxels in an image, we perform, as a second approach, a spatially penalized estimation. We use a two-dimensional neighborhood structure where adjacent voxels are neighbors, i.e., each voxel has up to four neighbors. Voxels at the edge of an image have less than four neighbors, whereas all other voxels have four neighbors.

For the penalization, we use again the fact that a compartment is characterized by its  $k_{\text{ep}}$ -value. While the contribution of a specific tissue compartment, i.e., the fractional volume  $v_e$  and consequently the volume transfer constant  $K^{\text{trans}}$  of this compartment, may vary considerably over a field of neighboring voxels, the rate constant  $k_{\text{ep}}$  at which the compartment exchanges CA with the APC should be similar for neighboring voxels (Tofts et al., 1999; Sommer, 2013). Thus, we penalize differences in the  $\hat{k}_{\text{ep}}$ -values of adjacent voxels. As a penalty, we use a ridge type penalty. Let  $\hat{k}_{\text{ep}_{j1}}^{[m]}, \dots, \hat{k}_{\text{ep}_{jq_j}}^{[m]}$  denote the  $q_j$   $\hat{k}_{\text{ep}}$ -values of voxel  $j$  in iteration  $m$ . Then, the nonlinear penalized least squares estimate for voxel  $i$  in iteration  $m$  is

$$\left( \hat{K}_{ik}^{\text{trans}[m]}, \hat{k}_{\text{ep}_{ik}}^{[m]} \right)' = \arg \min_{K, k} \left\{ \sum_t \left( u_i(x_t)^{[m]} - (C_p(x_t) * K \exp(-kx_t) + \lambda J(k)) \right)^2 \right\},$$

with side conditions  $0.05 \leq \hat{k}_{\text{ep}ik}^{[m]} \leq 20$  and  $0.01 \leq \hat{K}_{ik}^{\text{trans}[m]} \leq 20$ , penalty parameter  $\lambda$ , and penalization term

$$J(k) = \sum_{j \in \delta_i} \min_l ((\hat{k}_{\text{ep}ik}^{[m]} - \hat{k}_{\text{ep}jl}^{[m]})^2),$$

where  $\delta_i$  is the set of voxels in the neighborhood of pixel  $i$ .

This estimate corresponds to the estimate in (7.7), except for the penalization term  $\lambda J(k)$ , where  $J(k)$  corresponds to a Markov random field on the  $k_{\text{ep}}$ -values. By replacing the least squares base-learners by penalized variants, i.e., by the introduction of the penalization term  $\lambda J(k)$ , we ensure that neighboring voxels share similar  $k_{\text{ep}}$ -values, and therefore, spatial smoothness of these parameters is achieved.

Similar to Sommer et al. (2014), for efficiency reasons, we do a parallel update of the estimates of the voxels following a checkerboard pattern, using conditional independence from the remaining voxels given all neighboring voxels. By the checkerboard pattern, the voxels in an image are divided into “even” and “uneven” voxels. We do not fully run the boosting algorithm for each voxel as in 7.2.5. Rather, the updates of the parameter estimates are done alternately for the “even” and the “uneven” voxels. This means, in each iteration, the “even” and the “uneven” voxels are updated each once. In the first iteration, the  $\hat{k}_{\text{ep}}$ -values of the “even” and the “uneven” voxels are penalized to some starting  $\hat{k}_{\text{ep}}$ -values. To receive these starting values, we apply the voxelwise boosting algorithm (Subsections 7.2.4 and 7.2.5) to the mean CTC, which results from averaging the CTCs of all voxels in one image. From the second iteration on, the  $\hat{k}_{\text{ep}}$ -values of the “even” voxels are penalized to the current  $\hat{k}_{\text{ep}}$ -values of the “uneven” voxels and vice versa.

### 7.2.7 Refit

After the execution of the boosting algorithm with the voxelwise or the spatially regularized estimation procedure, for each voxel, the number of compartments and the  $\hat{k}_{\text{ep}}$ -values are considered as fixed and a refit of the model is done in order to get the final  $\hat{K}^{\text{trans}}$ -values (Candes and Tao, 2007; Gertheiss and Tutz, 2010; Tutz and Groll, 2013; Groll, 2013). By using fixed  $\hat{k}_{\text{ep}}$ -values, the nonlinear regression model is linearized, and therefore, the refit is done with the R function `solve.QP` (package `quadprog` (Turlach and Weingessel, 2013)).

### 7.2.8 Competing Methods

Within the framework of the simulation and for the analysis of real DCE-MRI data, we compare the results of our two boosting approaches (voxelwise and spatially regularized) with the results of fitting a Tofts model (Subsection 7.2.1), as well as the results of a voxelwise and a spatially regularized estimation approach as described in Sommer et al. (2014). The two approaches in Sommer et al. (2014) are based on basis functions  $\Psi_k(x_t) = C_p(x_t) * \exp(-k_{\text{ep}k} x_t)$ . Using those, Equation (7.3) becomes  $C_i(x_t) = \sum_{k=1}^{q_i} K_{ik}^{\text{trans}} \Psi_k(x_t)$ . A set of candidate values for  $k_{\text{ep}k}$  is considered such that  $\log(k_{\text{ep}k}) \in \{-3.0, -2.9, \dots, 3.0\}$ , and suitable values, i.e., compartments, have to be chosen. Therefore, the unknown parameters  $K_{ik}^{\text{trans}}$  are estimated, and the  $\hat{k}_{\text{ep}k}$ -values related

|              |               | $k_{ep_1}$ | $K_1^{\text{trans}}$ | $k_{ep_2}$ | $K_2^{\text{trans}}$ | $k_{ep_3}$ | $K_3^{\text{trans}}$ |
|--------------|---------------|------------|----------------------|------------|----------------------|------------|----------------------|
| Simulation 1 | outside tumor | 0.2        | 0.15                 | -          | -                    | -          | -                    |
|              | inside tumor  | 0.2        | 0.05                 | 1.4        | 1.3                  | -          | -                    |
|              | tumor edge    | 0.2        | 0.1                  | 1.4        | 0.6                  | 9.8        | 2.6                  |
| Simulation 2 | outside tumor | 0.15       | 0.1                  | -          | -                    | -          | -                    |
|              | inside tumor  | 1.5        | 1.3                  | 15         | 2.8                  | -          | -                    |
|              | tumor edge    | 0.15       | 0.08                 | 1.5        | 0.6                  | 15         | 2.6                  |
| Simulation 3 | outside tumor | 0.25       | 0.15                 | -          | -                    | -          | -                    |
|              | inside tumor  | 0.25       | 0.05                 | 1.25       | 1.2                  | -          | -                    |
|              | tumor edge    | 0.25       | 0.15                 | 1.25       | 0.4                  | 6.25       | 1.6                  |

Table 7.2: Simulation setup.

to positive  $\hat{K}_{ik}^{\text{trans}}$ -values are selected. As the usual ML-estimates are unstable or even not unique, two penalized approaches are proposed. In the first approach, the voxelwise regularized estimation, a penalized maximum likelihood estimator is used, with the (positive) elastic net (Zou and Hastie, 2005) being chosen as penalty. In the second approach, the spatially regularized estimation, the penalty term used in the first approach is modified in such a way that it enforces spatial smoothness of parameters of neighboring voxels. Differences in the  $\hat{K}^{\text{trans}}$ -values of adjacent voxels are penalized by a quadratic penalty term, which in our opinion, however, is not as straightforward as penalizing the  $\hat{k}_{ep}$ -values of adjacent voxels as done in the spatially regularized boosting approach proposed in this chapter.

In both approaches from Sommer et al. (2014) as well as in the spatially regularized boosting approach, the tuning parameters are determined according to the BIC. The neighborhood structure in Sommer et al. (2014) and in this chapter coincide. In Sommer et al. (2014), the BIC is computed according to Hastie et al. (2009). For the DCE-MRI data, the variance  $\sigma_{\varepsilon}^2$  of the assumed normal distribution is estimated by the mean-squared error of a low-bias model (Hastie et al., 2009). We, however, compute the BIC for this data according to the formula  $n \cdot \log \left( \sum_{i=1}^N \sum_{t=0}^T (Y_i(x_t) - \hat{Y}_i(x_t))^2 \right) + p \cdot \log(n)$ , where  $p$  is the number of parameters, i.e., the sum of the estimated number of compartments over all voxels in an image, and  $n = N \cdot T$  is the number of voxels in an image multiplied by the number of time points at which an MR signal has been measured for this image. The approaches in Sommer et al. (2014) are from now on called ‘elastic net approaches’ (voxelwise and spatial elastic net approach, respectively).

## 7.3 Simulation studies

### 7.3.1 Simulation setup

For each of the simulations listed in Table 7.2, we simulated 10 DCE-MR images, each consisting of  $40 \times 50$  voxels. We simulated images which are similar to DCE-MR images revealing breast cancer, and therefore, three kinds of typical CTCs were simulated. For the tissue located outside the tumor, we simulated from a compartment model with one tissue compartment. Inside the tumor, we assumed a compartment model with two tissue compartments. For the voxels at the

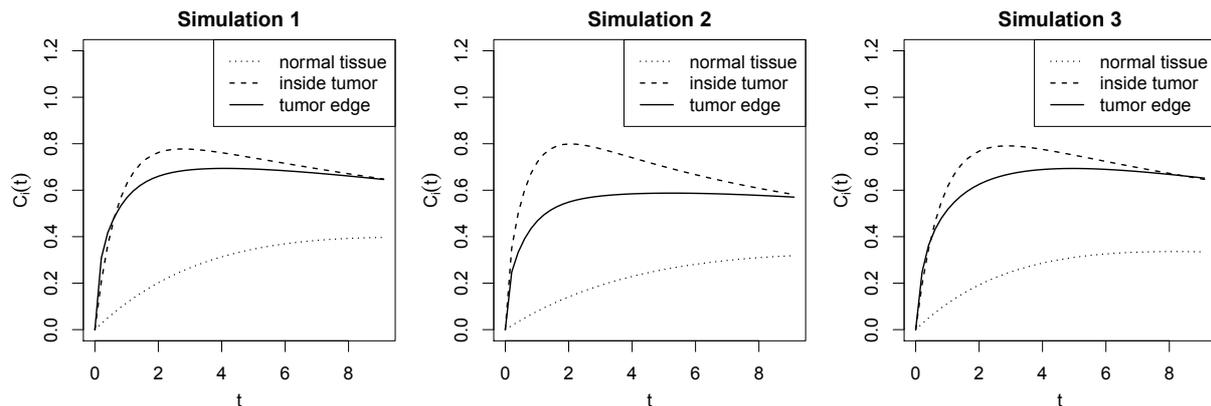


Figure 7.6: Simulated CTCs (without noise) for Simulations 1-3.

edge of the tumor, we simulated from a compartment model with three tissue compartments. We used different  $k_{ep}$ -values for the simulations, and the factor by which the  $k_{ep}$ -values differ in each of the three simulations is chosen differently as well. The  $k_{ep}$ -values differ by a factor of 7 in Simulation 1, by a factor of 10 in Simulation 2, and by a factor of 5 in Simulation 3. For all simulations, each simulated CTC consists of CA concentrations at 46 time points every 12 seconds, and we added Gaussian noise (standard deviation  $\sigma = 0.05$ ) to the CTCs of all voxels. A figure showing the simulated CTCs (without noise) for Simulations 1-3 can be found in Figure 7.6.

### 7.3.2 Results

In Figure 7.7, some results of Simulation 1 are shown for one of the ten simulated images. The results for the remaining simulated images are similar and can be found in the Appendix (Part C), as well as the results for the remaining simulations. The first image in the first row of Figure 7.7 shows the true number of compartments for all voxels in the simulated image. In the subsequent images in the first row, the estimated number of compartments  $\hat{q}$  is given for all methods under consideration. In general, when applying the Tofts model, for each voxel in an image, a compartment model with one tissue compartment is assumed. Therefore, the average number of compartments in an image is always 1 and the true average number of compartments is underestimated if the image contains voxels for which a compartment model with more than one tissue compartment is assumed, as it is the case for Simulation 1. It is obvious from the images in the first row of Figure 7.7 that for the boosting approaches, the estimated number of compartments across the image is spatially more homogeneous than for the elastic net approaches. This is most obvious for the tumor edge and the tissue located outside the tumor. Moreover, with the boosting approaches, the number of compartments is not as often overshooted as with the elastic net approaches, where in many cases a number of 4 or 5 compartments is estimated. In the second row of Figure 7.7, the voxelwise mean squared error (MSE)  $MSE_i = \frac{1}{T} \sum_{t=0}^T (C_i(x_t) - \hat{Y}_i(x_t))^2$  is plotted for all methods under consideration. For the Tofts model, it can be clearly seen that the

MSE is the highest at the tumor edge, followed by the MSE for the inner region of the tumor. The MSE is the lowest for the tissue outside the tumor. This shows that especially for the voxels at the tumor edge, the Tofts model is not sufficient and more complex compartment models are needed. For the boosting approaches, the MSE is comparatively large for the voxels inside the tumor, whereas for the elastic net approaches, voxels with a comparatively large MSE are a bit more scattered across the image.

In Table 7.3, for each simulation, the average MSE for the 10 simulated images is given for each considered method. The MSE for image  $k$  is calculated according to  $MSE_k = \frac{1}{N} \sum_{i=1}^N MSE_i$ , and the average MSE over all  $K$  images according to  $\overline{MSE} = \frac{1}{K} \sum_{k=1}^K MSE_k$ . Additionally, in Table 7.3, the true average number of compartments  $q_{av}^{true}$ , the estimated average number of compartments over all ten simulated images  $\hat{q}_{av}$ , and the percentage  $q_c$  of voxels for which the correct number of tissue compartments is estimated is given. It can be seen that the average MSE is approximately the same for all methods except for the Tofts model, where it is considerably larger than for the other four methods. For Simulation 1, boxplots of the voxelwise MSE for all considered methods for all 10 simulated images are available in Figure 7.8. When comparing the average MSE for the two voxelwise methods (boosting and elastic net) and the two spatial methods (boosting and elastic net) for Simulations 1-3, we can see that it is most of the times slightly larger for the elastic net approaches than for the boosting approaches. In all three simulation settings, the boosting approaches perform better than the elastic net approaches with regard to  $\overline{MSE}$  and  $q_c$ , as  $\overline{MSE}$  is at most as large as with the elastic net approaches, and  $q_c$  is always larger as with the elastic net approaches. For each simulation setting, the performance of the voxelwise and the spatially regularized boosting approach regarding  $q_c$  is similar, whereas for the elastic net approaches, the spatial approach always performs considerably better regarding  $q_c$  compared to the voxelwise approach.

## 7.4 Application to DCE-MRI data

### 7.4.1 Description of the data

For the clinical evaluation of our approach we used data of six breast cancer patients who have participated in a breast cancer study which has previously been reported on and which has already been analyzed (Ah-See et al., 2004; Schmid et al., 2005, 2006; Sommer et al., 2014). Per breast cancer patient, we used two scans recorded by a 1.5 T Siemens MAGNETOM Symphony scanner (Repetition time (TR) = 11 ms, Echo Time (TE) = 4.7 ms). One of the scans has been recorded at the beginning of the treatment (pre-treatment scan) and the second one after two weeks of chemotherapy (post-treatment scan). Each scan comprises three slices. For our analysis, we used only the central slice. Per patient and recording, the MR signal has been recorded every 12 seconds at 36 or 46 time points, respectively. At the start of the fifth MR signal recording, Gd-DTPA at a dose of 0.1 mmol per kg body weight has been injected with a power injector. The regions of interest used for the analyses cover the tumor as well as surrounding non-tumorous tissue. The smallest analyzed image consists of  $48 \times 81$  voxels, and the largest one of  $118 \times 115$  voxels.

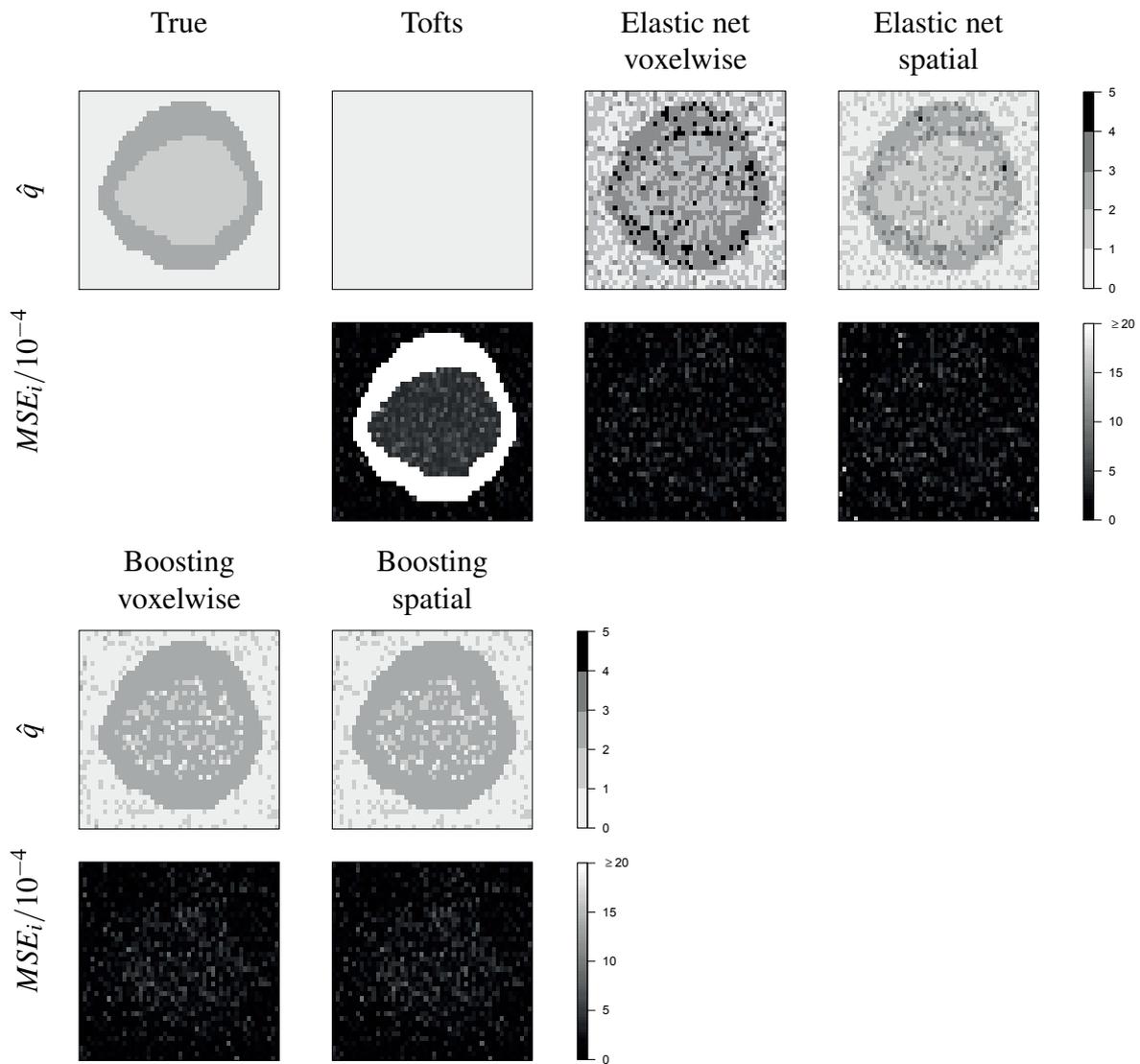


Figure 7.7: Results of Simulation 1 for the first of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

|   |  | $\overline{MSE}$    | $\hat{q}_{av}$ | $q_c$ |
|---|--|---------------------|----------------|-------|
| Simulation 1<br>( $q_{av}^{\text{true}}=1.78$ ) | Tofts  | $8.8 \cdot 10^{-4}$ | 1.00           | 0.48  |
|   | Elastic net voxelwise ( $\lambda = 1e-10, s=0.2$ ) | $1.8 \cdot 10^{-4}$ | 2.23           | 0.54  |
|   | Elastic net spatial ( $\lambda = 1e-07, s = 0.2$ ) | $1.9 \cdot 10^{-4}$ | 2.01           | 0.68  |
|   | Boosting voxelwise                                 | $1.7 \cdot 10^{-4}$ | 2.04           | 0.71  |
|   | Boosting spatial ( $\lambda = 1e-10$ )             | $1.7 \cdot 10^{-4}$ | 2.04           | 0.71  |
| Simulation 2<br>( $q_{av}^{\text{true}}=1.78$ ) | Tofts  | $7.2 \cdot 10^{-4}$ | 1.00           | 0.48  |
|   | Elastic net voxelwise ( $\lambda = 1e-10, s=0.2$ ) | $1.7 \cdot 10^{-4}$ | 2.20           | 0.54  |
|   | Elastic net spatial ( $\lambda = 1e-10, s = 0.2$ ) | $2.0 \cdot 10^{-4}$ | 1.95           | 0.69  |
|   | Boosting voxelwise                                 | $1.7 \cdot 10^{-4}$ | 1.98           | 0.74  |
|   | Boosting spatial ( $\lambda = 1e-10$ )             | $1.7 \cdot 10^{-4}$ | 1.98           | 0.74  |
| Simulation 3<br>( $q_{av}^{\text{true}}=1.78$ ) | Tofts  | $6.9 \cdot 10^{-4}$ | 1.00           | 0.48  |
|   | Elastic net voxelwise ( $\lambda = 1e-05, s=0.2$ ) | $1.8 \cdot 10^{-4}$ | 2.17           | 0.53  |
|   | Elastic net spatial ( $\lambda = 1e-10, s = 0.2$ ) | $1.7 \cdot 10^{-4}$ | 1.94           | 0.64  |
|   | Boosting voxelwise                                 | $1.6 \cdot 10^{-4}$ | 1.99           | 0.69  |
|   | Boosting spatial ( $\lambda = 1e-07$ )             | $1.6 \cdot 10^{-4}$ | 1.99           | 0.69  |

Table 7.3: Average values for the 10 simulated images: Average MSE ( $\overline{MSE}$ ), estimated average number of tissue compartments ( $\hat{q}_{av}$ ), and percentage of voxels for which the number of tissue compartments is estimated correctly ( $q_c$ ).

## 7.4.2 Results

In Figures 7.9-7.14, for both the boosting and the elastic net approaches the estimated number of compartments  $\hat{q}$  is plotted for the two images of each patient from the breast cancer study. In the last row of these figures, the concentration of CA one minute after the injection of the CA is plotted pre- and post-treatment as a reference. We will discuss only the results for one of the patients of the breast cancer study in detail (patient 2, Figure 7.10), as the difference in the results for the boosting and elastic net approaches is particularly evident for this patient. However, the results for the remaining patients are of similar quality. Patient 2 is a responder to the chemotherapy (Schmid et al., 2006). Therefore, the tumor is considerably smaller in the post-treatment image than in the pre-treatment image. We can see that with the boosting approaches, the tumor can be located much better than with the elastic net approaches. This is especially obvious for the post-treatment scan of patient 2, where the tumor can not be located correctly with the spatial elastic net approaches, but with the boosting approaches it can. Furthermore, with the boosting approaches, the estimated number of compartments across the image is spatially more homogeneous, but the tumor edge is still more clearly defined than with the elastic net approaches, which is especially obvious for the pre-treatment scan of patient 2.

The BIC and the estimated average number of compartments  $\hat{q}$  for the elastic net and the boosting approaches for all analyzed images can be found in Table 7.4 together with the corresponding tuning parameters. For all images except for the post-treatment scan of patient 3, the BIC for the elastic net approaches is smaller than the BIC for the boosting approaches, based on which one would usually favor the elastic net approaches. However, the BIC does not reflect the spatial structure of the image and the correct localization of the tumor. When comparing

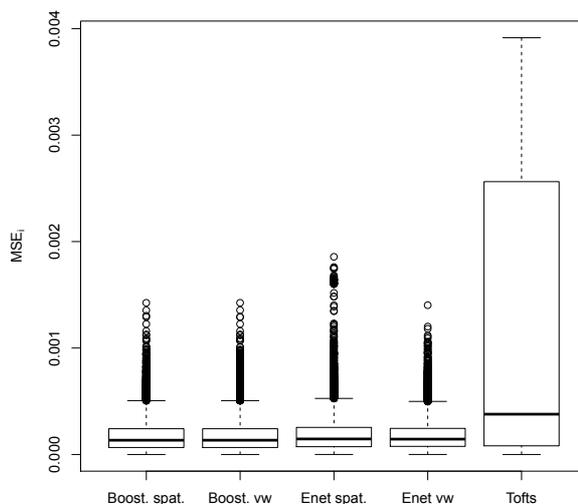


Figure 7.8: Voxelwise MSE for all considered methods for all 10 simulated images. The boxplots lie approximately in the same range for all considered methods except for the Tofts model, where the median and the 75%-quantile of the voxelwise MSE are much higher than for the other methods. For the spatial elastic net approach, there are slightly more outliers with marginally larger MSE-values than for the voxelwise elastic net and the boosting approaches.

the two boosting approaches presented, we see that the BIC for spatial boosting is smaller or approximately in the same range as the BIC for voxelwise boosting.

The tuning parameters  $\lambda$  and  $s$  for the elastic net approaches and  $\lambda$  for the spatially regularized boosting approach can be found in the last three rows of Table 7.4. Mostly, a small value of  $10^{-10}$  or  $10^{-7}$  is chosen for the penalization parameter  $\lambda$  in all of these approaches. An exception are the images of patients 5 and 6 and the post-treatment image of patient 4 (Figures 7.12-7.14), for which the value of the penalization parameter  $\lambda$  is  $10^{-3}$  for the spatially regularized boosting approach. Therefore, for these five images, the biggest influence of the spatial regularization can be found, and a clear difference between the spatial structure of the number of compartments for the voxelwise and the spatially regularized boosting approach is apparent. The estimated number of compartments is distributed more homogeneously across the image for the spatially regularized approach.

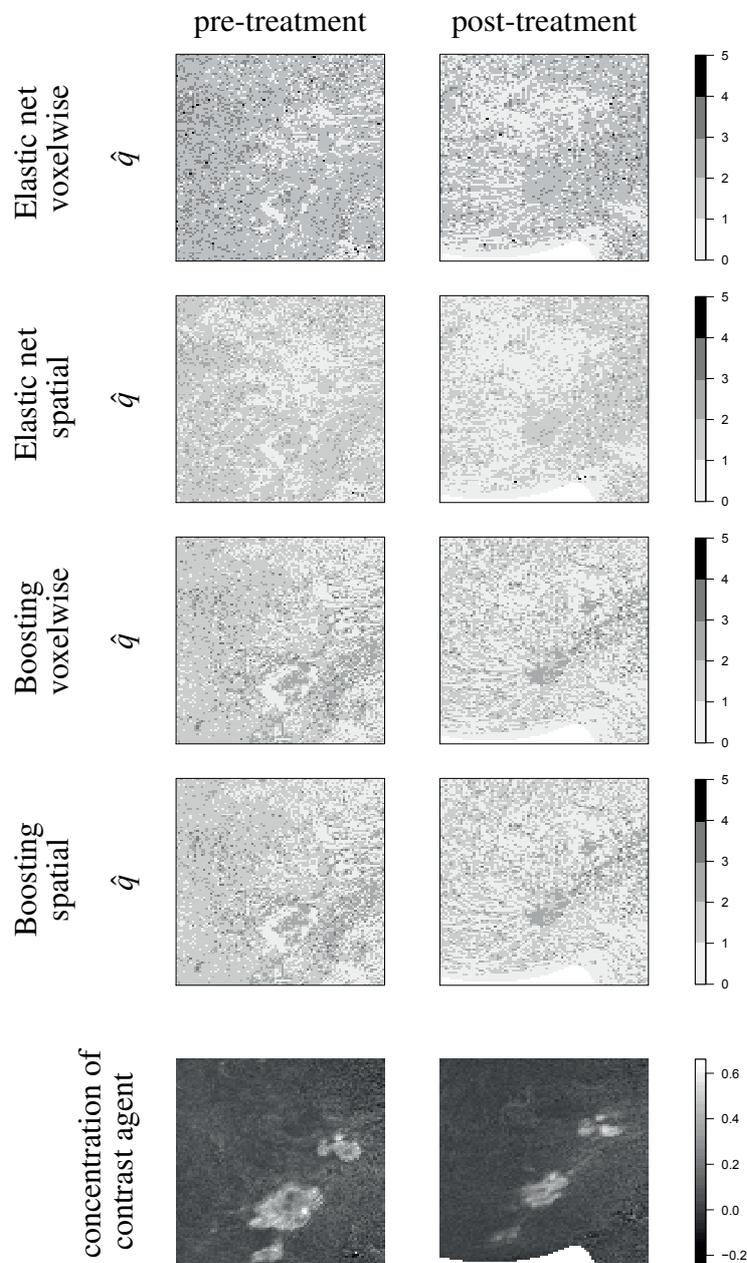


Figure 7.9: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 1. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

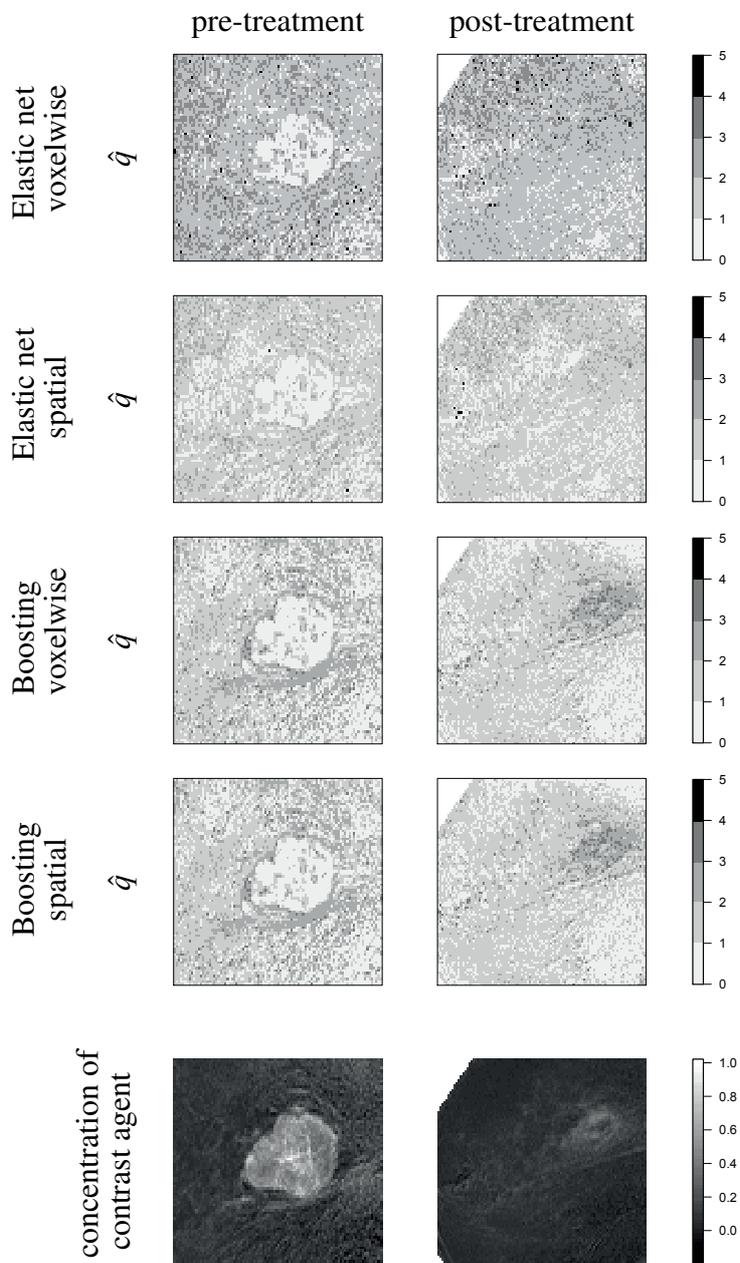


Figure 7.10: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 2. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

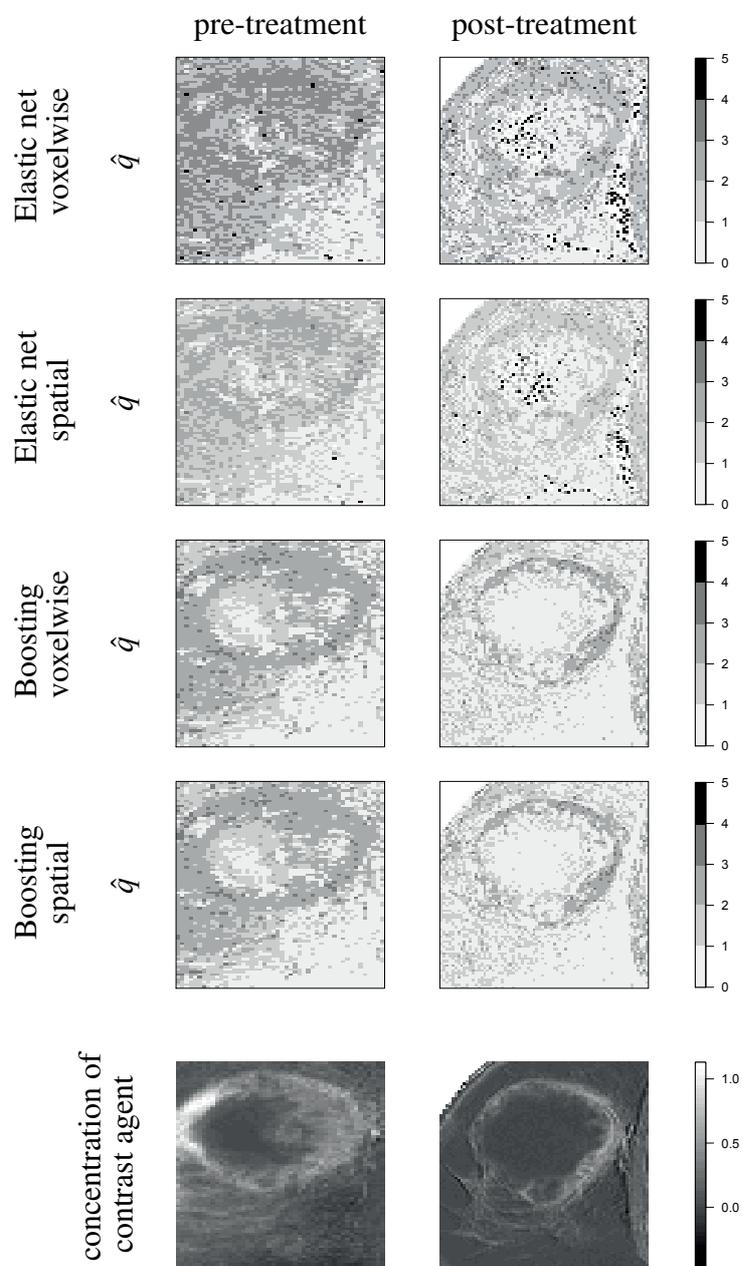


Figure 7.11: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 3. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

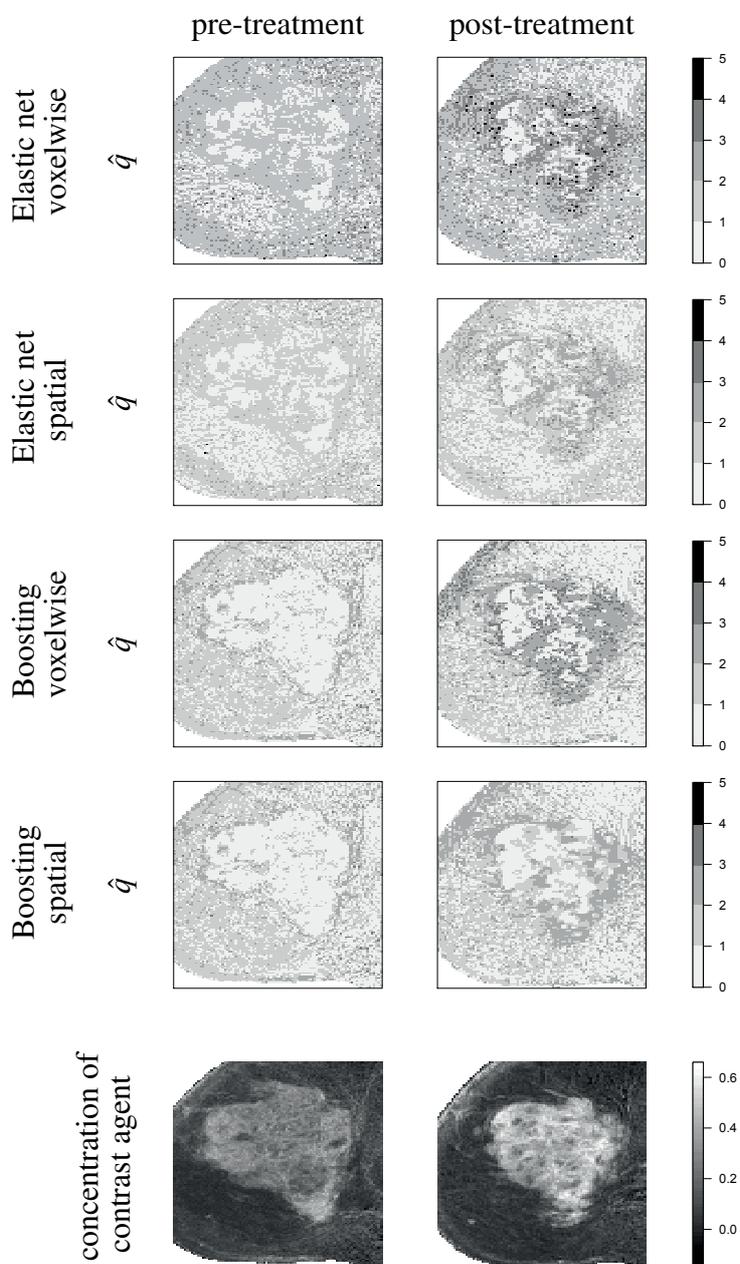


Figure 7.12: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 4. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

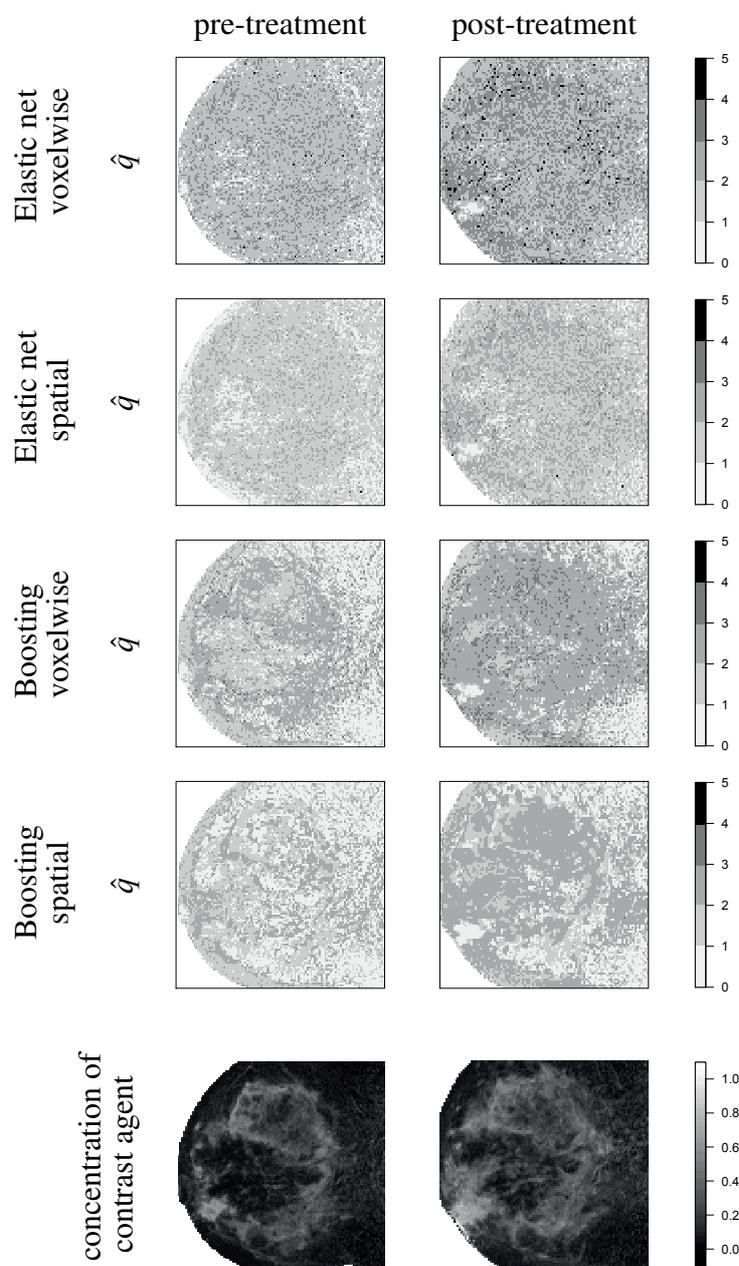


Figure 7.13: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 5. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

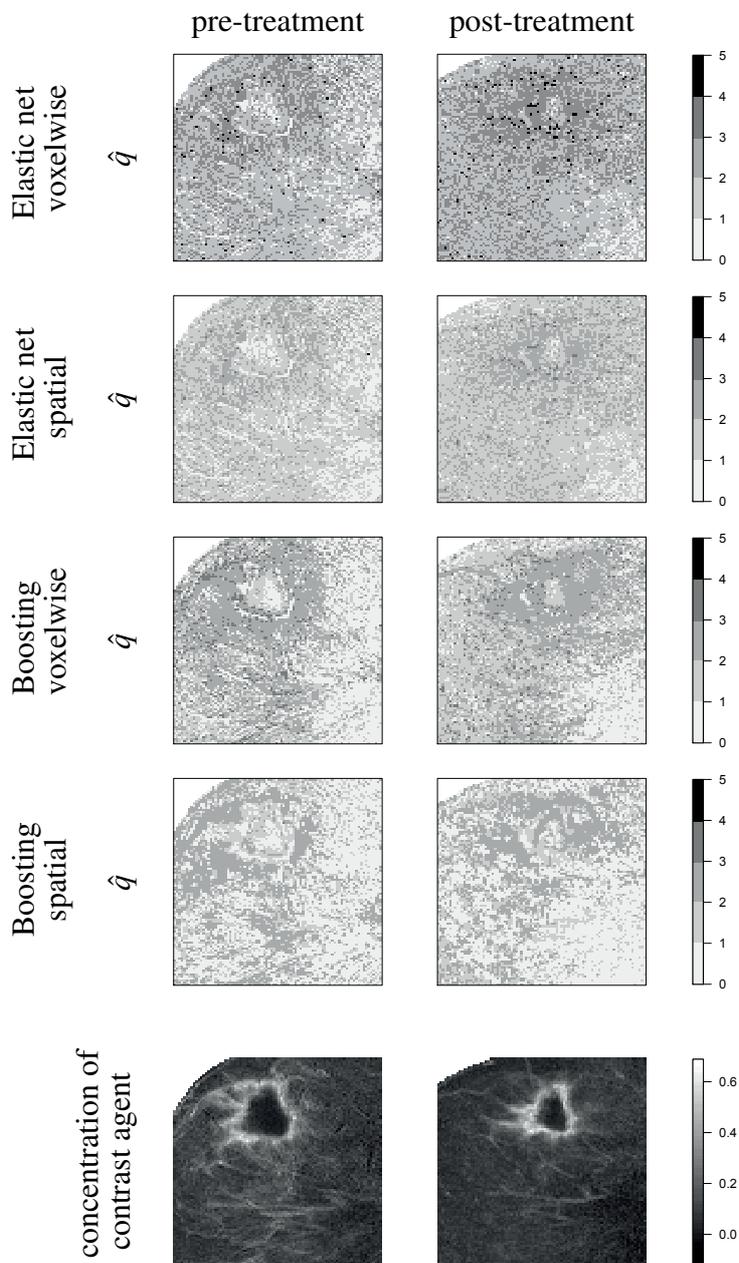


Figure 7.14: Results of the application to the DCE-MRI data: estimated number of tissue compartments  $\hat{q}$  for each voxel of the image of patient 6. Reference picture (last row): concentration of CA 1 minute after the injection of the CA.

|                       | patient 1 |           | patient 2 |           | patient 3 |           | patient 4 |           | patient 5 |           | patient 6 |           |
|-----------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                       | pre       | post      |
| BIC                   | 3904845   | 2971558   | 2258271   | 3314848   | 1099939   | 2103958   | 5256399   | 2821669   | 3640991   | 3577553   | 2121903   | 2111319   |
| Elastic net voxelwise | 3169567   | 2290706   | 2309692   | 2138783   | 1122943   | 1928639   | 3348782   | 2130934   | 2683162   | 2279462   | 1888564   | 1748292   |
| Elastic net spatial   | 3147522   | 2275503   | 2294001   | 2131346   | 1116738   | 1912825   | 3332692   | 2111086   | 2667538   | 2268167   | 1877224   | 1737165   |
| Boosting voxelwise    | 3185805   | 2304293   | 2310569   | 2143496   | 1124535   | 1879643   | 3355232   | 2148980   | 2731534   | 2334393   | 1906038   | 1753448   |
| Boosting spatial      | 3185398   | 2304577   | 2310615   | 2143351   | 1124490   | 1879275   | 3355391   | 2144395   | 2714865   | 2317842   | 1884041   | 1728813   |
| $\hat{q}$             | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      | 1.00      |
| Tofts                 | 2.01      | 1.74      | 2.06      | 2.07      | 2.24      | 3.30      | 2.82      | 3.02      | 3.09      | 3.37      | 2.15      | 2.37      |
| Elastic net voxelwise | 1.81      | 1.58      | 1.86      | 1.91      | 2.07      | 2.08      | 1.69      | 1.82      | 1.95      | 2.26      | 2.03      | 2.25      |
| Elastic net spatial   | 1.96      | 1.67      | 1.94      | 1.86      | 2.16      | 1.57      | 1.66      | 2.03      | 2.17      | 2.67      | 2.21      | 2.26      |
| Boosting voxelwise    | 1.96      | 1.67      | 1.95      | 1.85      | 2.16      | 1.57      | 1.66      | 1.85      | 1.90      | 2.31      | 1.90      | 1.85      |
| Boosting spatial      | 1e-10,0.4 | 1e-07,0.2 | 1e-10,0.3 | 1e-10,0.2 | 1e-07,0.3 | 1e-10,0.3 | 1e-10,0.2 | 1e-10,0.2 | 1e-10,0.2 | 1e-10,0.3 | 1e-10,0.3 | 1e-10,0.2 |
| $\lambda_{s}$         | 1e-10,0.4 | 1e-07,0.2 | 1e-10,0.3 | 1e-10,0.2 | 1e-10,0.3 | 1e-10,0.3 | 1e-10,0.2 | 1e-10,0.2 | 1e-10,0.2 | 1e-10,0.3 | 1e-10,0.3 | 1e-07,0.2 |
| $\lambda$             | 1e-10     | 1e-10     | 1e-07     | 1e-10     | 1e-07     | 1e-07     | 1e-10     | 1e-03     | 1e-03     | 1e-03     | 1e-03     | 1e-03     |

Table 7.4: Results of the application to the DCE-MRI data: BIC and estimated average number of tissue compartments  $\hat{q}_{TV}$  for all images of all six patients.

## 7.5 Assessment of therapy success

Whereas, in DCE-MRI, the rate constant  $k_{ep}$  describes the shape of a CTC, the fractional volume  $v_e$  and consequently the volume transfer constant  $K^{trans}$  relate to absolute values of a CTC (Sommer, 2013). The rate constant  $k_{ep}$  is always greater than the transfer constant  $K^{trans}$ . Several physiologic interpretations contingent upon the balance between capillary permeability and blood flow in the tissue of interest are possible for the transfer constant  $K^{trans}$  (Tofts et al., 1999). In a high-permeability scenario, i.e., if tracer flux is flow-limited, the volume transfer constant  $K^{trans}$  “equals the blood plasma flow per unit volume of tissue” (Tofts et al., 1999). In a permeability-limited scenario, i.e., if tracer flux is permeability-limited,  $K^{trans}$  “equals the permeability surface area product per unit volume of tissue” (Tofts et al., 1999). Thus,  $K^{trans}$  gives information about the balance between capillary permeability and blood flow in the tissue of interest, and is therefore used in order to assess therapy success for the patients in the breast cancer study described in Subsection 7.4.1.

A one-sided two-sample Wilcoxon rank sum test (also called ‘Mann-Whitney test’), was performed to compare the pre-treatment and post-treatment  $K^{trans}$ -values of the six patients from the breast cancer study in 7.4.1. We performed this test for each patient using the  $K^{trans}$ -values resulting from the application of the voxelwise boosting approach to the DCE-MRI data from the breast cancer study. The null hypothesis of the test was that the distributions of the pre-treatment and post-treatment  $K^{trans}$ -values of one patient differ by a location shift of  $\mu = 0$ . For patients 1, 2, 3, and 6, the null hypothesis was rejected at a significance level of 0.05 ( $P < 2.2 \cdot 10^{-16}$ ). This means that the true location shift is greater than zero, i.e., the distribution of the pre-treatment  $K^{trans}$ -values is shifted to the right of the distribution of the post-treatment  $K^{trans}$ -values. This suggests that patients 1, 2, 3, and 6 are responders to the therapy. For patients 4 and 5, the null hypothesis was not rejected at a significance level of 0.05 ( $P = 1$ ), which suggests that these patients are nonresponders to the therapy. This is in accordance with the true responder status of the patients Schmid et al. (2006). The distributions of the pre-treatment and post-treatment  $K^{trans}$ -values for each patient can be found in Figure 7.15. Therefore, therapy success can be assessed by using the  $K^{trans}$ -values resulting from the application of voxelwise boosting in order to identify responders and nonresponders of the therapy.

## 7.6 Conclusions and Discussion

We proposed two boosting approaches for data-driven model choice and parameter estimation for DCE-MRI data. Both approaches build on a multi-tissue compartment model and combine the advantages of data-driven and model-based approaches, as the number of compartments is estimated from the data for each voxel of an image, and the parameters nevertheless remain interpretable. Based on the results of the simulation studies and the results of the application to DCE-MRI data of six breast-cancer patients, we conclude that the boosting approaches outperform the Tofts and elastic net approaches regarding the correct localization of the tumor as well as the spatial homogeneity of the estimated number of compartments across the image, and the definition of the tumor edge. Therefore, we favor the boosting approaches over the elastic net

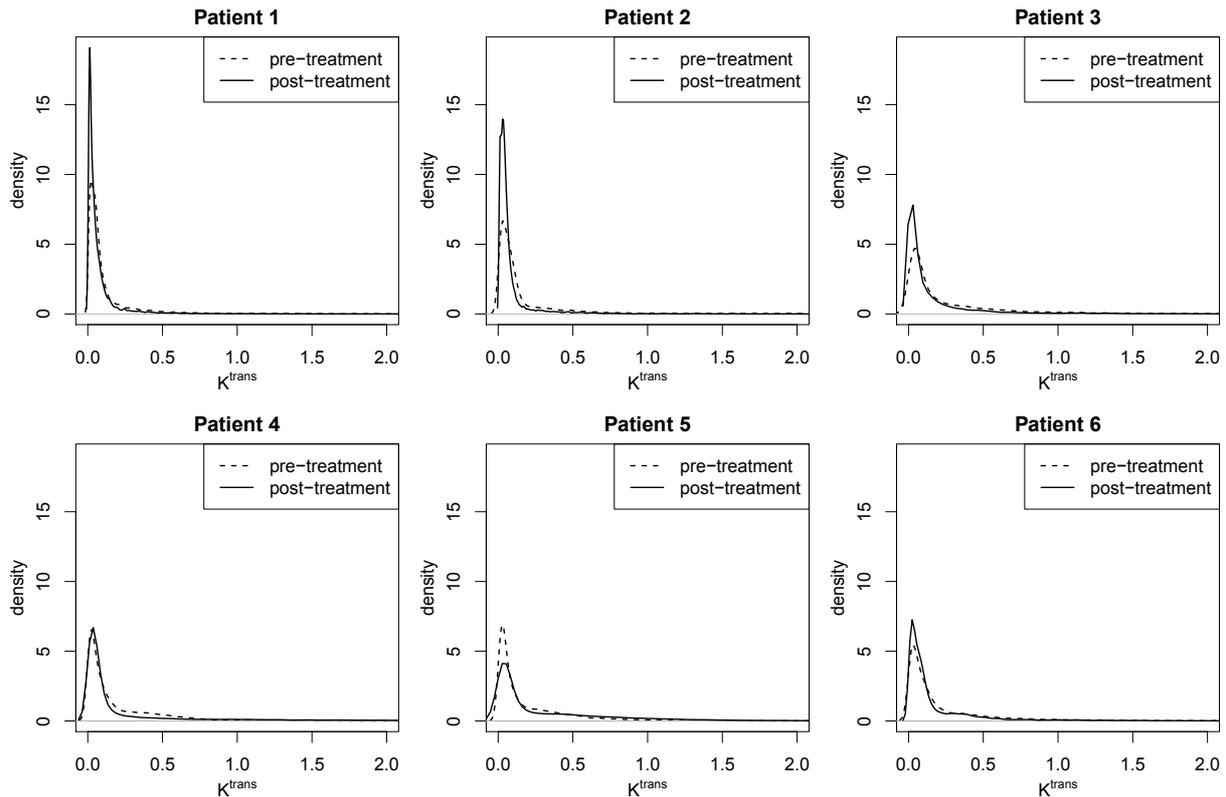


Figure 7.15: Distributions of the pre-treatment and post-treatment  $K^{\text{trans}}$ -values for each patient.

approaches, even though the BIC was smaller for the elastic net approaches than for the boosting approaches for most images of the breast-cancer patients.

When comparing the plots of the estimated number of compartments  $\hat{q}$  (Figures 7.9-7.14), the BIC, and the estimated average number of tissue compartments  $\hat{q}_{av}$  (Table 7.4) for the two boosting approaches applied to the DCE-MRI data, it appears that there is no major difference between the results of these two approaches, except for the images of patients 5 and 6 and the post-treatment image of patient 4. For these images, the number of compartments is distributed more homogeneously across the image for the spatially regularized approach, and for the images of patients 5 and 6, the BIC is considerably smaller for the spatially regularized approach compared to the voxelwise boosting approach. It was indeed a surprising result that the proposed penalization did not have a large impact on the nonlinear boosting. The voxelwise boosting approach already turned out to determine the number of compartments quite well and found numbers of compartments similar to the spatial elastic net approach in all of the simulations and most of the real situations. Therefore, we think that the spatial penalization could not strengthen the boosting approach in the same way it strengthened the elastic net approach. As there is also a large difference in computation time, in order to get results as quickly as possible, we would recommend to favor the voxelwise boosting approach over the spatially regularized version. In order to get results where the spatial structure is taken into account properly, however, we would

|   | $\overline{MSE}$    | $\hat{q}_{av}$ | $q_c$ |
|---|---------------------|----------------|-------|
| Boosting voxelwise (factor 5)                   | $1.7 \cdot 10^{-4}$ | 2.04           | 0.71  |
| Boosting voxelwise (factor 4)                   | $1.7 \cdot 10^{-4}$ | 2.31           | 0.57  |
| Boosting voxelwise (factor 3)                   | $1.8 \cdot 10^{-4}$ | 2.69           | 0.48  |
| Boosting spatial (factor 5, $\lambda = 1e-10$ ) | $1.7 \cdot 10^{-4}$ | 2.04           | 0.71  |
| Boosting spatial (factor 4, $\lambda = 1e-05$ ) | $1.8 \cdot 10^{-4}$ | 2.27           | 0.58  |
| Boosting spatial (factor 3, $\lambda = 1e-05$ ) | $1.9 \cdot 10^{-4}$ | 2.53           | 0.52  |

Table 7.5: Influence of the chosen factor for the relative difference of the  $k_{ep}$ -values in the boosting algorithm. Average values for the 10 simulated images of Simulation 1: Average MSE ( $\overline{MSE}$ ), estimated average number of tissue compartments ( $\hat{q}_{av}$ ), and percentage of voxels for which the number of tissue compartments is estimated correctly ( $q_c$ ).

recommend to use the spatially regularized version. The computation time for the spatially regularized boosting for all 12 images was approximately 19 days on a linux server with 64 cores and 512 GB memory using 40 cores in parallel, whereas the voxelwise boosting for all 12 images took approximately 2 hours on the same machine using 40 cores in parallel (For comparison: The computation times for the competing methods on the same machine using 40 cores in parallel were: spatial elastic net: approximately 2 hours, voxelwise elastic net: approximately 3 minutes.). For the spatially regularized boosting, the updates of the parameter estimates are done alternately for the voxels in an image, which increases the runtime considerably in comparison to the voxelwise boosting, where the parameter estimates are updated independently of each other for the voxels in an image. All code for the analyses was written in R. For the parallelization of R code, we used the R packages `parallel`, `BatchJobs` and `BatchExperiments` (Bischi et al., 2014). For the independent parameter updates and the refit in the voxelwise boosting approach, the parallelization was done with the packages `BatchJobs` and `BatchExperiments`. For the alternating updates in the estimation process in the spatially regularized boosting approach, we used the package `parallel`. For the refit, we used the packages `BatchJobs` and `BatchExperiments`. The fact that we had to use a different package for the estimation process in the spatially regularized boosting approach might also contribute to the increased computation time for this approach.

In order to avoid redundancy issues in the boosting algorithm, we set a factor by which the  $k_{ep}$ -values in the multi-tissue compartment model must differ at least. We chose a factor of 5, as we know from Reich (1981) that parameters in a sum of exponentials model are highly redundant if the exponential rates differ by less than a factor of 5, and a generalization of that for convolved exponentials is difficult to obtain. Table 7.5 shows the influence of the choice of this factor within the boosting algorithm on the average MSE ( $\overline{MSE}$ ), the estimated average number of compartments ( $\hat{q}_{av}$ ), and the percentage of voxels for which the correct number of tissue compartments is estimated ( $q_c$ ) for alternative values of 3 and 4 for Simulation 1. From this table, we see that the smaller the chosen factor, the larger the  $\overline{MSE}$ , the larger  $\hat{q}_{av}$ , and the smaller  $q_c$ . However, also with factors 3 or 4, the  $\overline{MSE}$  is at most as large as the  $\overline{MSE}$  for the spatial elastic net approach (cf. Table 7.3). But with factors 3 or 4,  $q_c$  is considerably reduced compared to factor 5. The influence of the chosen factor on  $q_c$  with the spatial boosting approach

is not as strong as with the voxelwise boosting approach, which is probably due to the spatial penalization.

In order to avoid overfitting of the boosting algorithm, we used early stopping. If we run the voxelwise boosting algorithm for example on the 10 simulated images of Simulation 1 without early stopping, the estimated average number of tissue compartments  $\hat{q}_{av}$  is 2.68. With early stopping,  $\hat{q}_{av}$  is 2.04 (cf. Table 7.3), which makes clear that without early stopping, the number of compartments is clearly overestimated.

The results of the simulation studies and the application to real world DCE-MRI data indicate that additional complexity is needed especially at tumor edges, and the Tofts model is not capable of reflecting this complexity. Using the approaches presented, the number of compartments is estimated per voxel. Thus, important information about the tissue heterogeneity is gained. This can not be done with *a priori* fixed model architectures.

Therapy success can be assessed by performing a Wilcoxon rank sum test on the pre-treatment and post-treatment  $K^{\text{trans}}$ -values of each patient resulting from the application of voxelwise boosting to the DCE-MRI data of the patient.

# **Chapter 8**

---

## **Conclusion**

---

## 8.1 Summary

*In vivo* imaging techniques can shed light on dynamic and structural aspects of organisms, cells or cellular sub-compartments in biology as well as of tissues of interest in medicine.

The two imaging techniques in the focus of this thesis are FRAP, a fluorescence microscopy technique in biology, and DCE-MRI, an imaging technique used for example in oncology. With FRAP, the binding behavior of molecules inside organisms, cells or cellular sub-compartments can be investigated *in vivo*. With DCE-MRI, the blood supply of a tissue of interest can be recorded *in vivo*.

After the introduction, this thesis starts with giving some theoretical background on compartment models and differential equations, nonlinear regression, and Bayesian data analysis. In the body of this thesis, novel methods for the estimation and model selection for dynamic biomedical images that promote image analysis in the fields of fluorescence microscopy and oncology are proposed.

For the first imaging application in this thesis, the quantitative analysis of FRAP data, two different hierarchical Bayesian models for parameter estimation in a compartment model have been newly developed: a nonlinear mixed-effects model for the analysis of a series of FRAP images (Chapter 5), and a nonlinear model for the pixelwise analysis of FRAP data (Chapter 6). The proposed nonlinear mixed-effects model is useful for the joint analysis of data from FRAP experiments on various similar cell nuclei, since estimates of the off-rates of the interactions the molecules of interest are involved in are attained and the model moreover provides insight into the variation of the off-rates in the population of cell nuclei. This is a novel approach in the field of FRAP analysis. Although a simplified kinetic model is used, the approach can easily be adapted to other FRAP experiments and thus promotes image analysis in fluorescence microscopy. The evaluation of the method was done on half-nucleus FRAP data, also in comparison with different kinds of fixed-effects models. The DIC, which served as a measure of the model fit for the comparison of the mixed-effects model to the fixed-effects models, is lower or in approximately the same range for the mixed-effects model in comparison to the considered models without random effects for all considered cell cycle phases.

The proposed nonlinear model for the pixelwise analysis of FRAP data, where information from the neighboring pixels is included into the nonlinear model for each pixel, is advantageous in cases where the spatial structure of a cell nucleus is of interest and could thus allow a deeper insight into the dynamic and structural aspects of cells of interest. As the evaluation of the proposed model in a simulation study could not be realized within a reasonable time frame, this task is left for future research.

For the second imaging application in this thesis, the quantitative analysis of DCE-MRI of the breast, an estimation and model selection approach based on boosting, with which the number of compartments in the compartment model that is applied at the voxel level is flexible, has been newly developed (Chapter 7). With this approach, data-driven model choice and parameter estimation for DCE-MRI data can be done. The advantages of data-driven and model-based approaches are unified in this approach, as the number of compartments is estimated from the data for each voxel of an image, but the parameters still remain interpretable. In a simulation

study and in an application to DCE-MRI data, the proposed boosting approach outperforms all considered competing methods with respect to the correct localization of the tumor, the spatial homogeneity of the estimated number of compartments across the image, and the definition of the tumor edge. Moreover, with a Wilcoxon rank sum test, therapy success could be assessed in the framework of a breast cancer study. For these reasons, the approach promotes image analysis in oncology.

## 8.2 Discussion

As each chapter in the body of this thesis contains an independent discussion, we now focus on some general topics concerning the applications in this thesis.

### Discussion on compartment models

All newly developed methods presented in this thesis are based on compartment models. In the FRAP applications in Chapters 5 and 6, a compartment model with a fixed number of compartments is assumed. In Chapter 7, the number of compartments in the compartment model is chosen data-driven.

In compartmental modeling, several general assumptions can be made (Tofts, 1997; Sommer, 2013), and some of them will be discussed in the following. One of the assumptions that applies for all compartment models considered in this thesis is time invariance, which means that the parameters in a compartment model are assumed to be constant during the time of data acquisition (Tofts, 1997; Sommer, 2013). For the compartment models considered in this thesis, this assumption concerns the rates of exchange between the compartments. Though we can imagine that the exchange properties in a compartment model may change over a longer period, for the relatively short period of image acquisition, it seems reasonable to make this assumption.

Another assumption generally made in a compartmental model is that the compartments are well-mixed and homogeneous (Tofts, 1997; Sommer, 2013). In the case that this condition is not fulfilled for a simple compartment model with a small number of compartments, it may be necessary to increase the number of compartments in the model, and therefore, to use a more complex compartment model. The results of the simulation study and the application of the boosting approach to real world DCE-MRI data in Chapter 7 for example indicated that in tumorous tissue, complex compartment models are needed especially at tumor edges, and less complex models like the Tofts model are not capable of reflecting this complexity. By choosing the number of compartments in a data-driven way as done in Chapter 7, strong deviation from the assumption of well-mixed and homogeneous compartments is avoided. Another way to prevent strong deviation from this assumption is to use a high image resolution, i.e., to define a compartment model at the pixel or the voxel level rather than for a region of interest. This was done in the applications in Chapters 6 and 7 of this thesis. For a high image resolution, the assumption of well-mixed and homogeneous compartments is more likely valid than for a low image resolution, in particular for compartment models with a small number of compartments, and therefore, models with a reasonable number of compartments can be used.

### Discussion on the used estimation methods

Another characteristic that the newly developed methods presented in this thesis have in common is that they all use nonlinear regression.

In Chapters 5 and 6, Bayesian approaches are proposed in order to estimate the parameters in nonlinear regression models. By using a Bayesian approach, consistency problems by specifying starting values are avoided. Moreover, convergence issues that can occur with algorithms for nonlinear model fitting in a frequentist framework do usually not arise with a Bayesian approach. Another advantage of Bayesian models is that they are very flexible. This advantage was for example exploited in Chapter 5, where mixed-effect priors on nonlinear parameters were incorporated into a nonlinear regression model. Bayesian approaches are, however, in general computationally more demanding than frequentist estimation approaches. For the Bayesian inference in Chapters 5 and 6, MCMC simulations are done. These are associated with a high computational effort. An alternative to MCMC inference would be to approximate the posterior distribution analytically. Analytical approximation methods are, however, mostly not of sufficient accuracy for a moderate or high number of parameters that have to be estimated, and it is difficult to numerically compute the associated Hessian matrices (Carlin and Louis, 2009). In structured additive regression models that include, among others, (generalized) linear models and (generalized) additive models, Integrated Nested Laplace Approximation (INLA) can be used instead of MCMC sampling (Rue et al., 2009). Thereby, the computational effort can be reduced. With the nonlinear regression models used in this thesis, however, INLA can not be used. Bayesian inference could be done without any problems for the proposed nonlinear mixed-effects model in Chapter 5. For the nonlinear model at the pixel level in Chapter 6, however, Bayesian inference could not be carried out satisfactorily, and the intended evaluation of the method in the framework of a simulation study could not be done successfully.

In Chapter 7, an estimation and model selection approach based on boosting is proposed, and we therefore are in a frequentist setting. In order to reduce the computational effort that usually results from fitting a model several times using a great number of starting values, a suitable algorithm and an adequate number of initial guesses for the optimization in the boosting algorithm was determined in a benchmark experiment. In accordance with the results of this experiment, a Levenberg Marquart algorithm with a small number of starting values was then used for this optimization. Convergence issues did not occur in our application.

Bayesian and frequentist approaches differ moreover substantially in the interpretation of estimation results. Whereas parameter estimation in a frequentist framework usually results in point estimates and corresponding confidence intervals, parameter estimation in a Bayesian framework is based on the posterior distribution of the parameters (Gilks et al., 1996), from which for example point estimates and credibility regions can be computed.

### Discussion on the software in this thesis

The majority of the program code used in the approaches in this thesis was newly developed in the programming languages R and C. Two R packages were built based on that code. Most of the code used in Chapters 5 and 6 was written in C, because especially for the parts of the code containing Gibbs and MH-updates of the parameters of the nonlinear model this resulted in a remarkable reduction in runtime. The analyses in Chapter 7 could be done completely in R. Parts of the codes used for the analyses in Chapters 6 and 7 of this thesis were moreover parallelized in R or C. In general, by a proper parallelization of program code, the runtime of a program is reduced, and it therefore usually allows the fast execution of a program – in this thesis a program written in R or C – even if complex parameter updates have to be done. This applies for the DCE-MRI application in Chapter 7. For the FRAP application in Chapter 6, however, we reached our limits regarding computation time despite the parallelization of parts of the program code in C. The proposed Bayesian hierarchical model could not be evaluated in the framework of a simulation study in a reasonable time frame. Therefore, this task is left for future research.

## 8.3 Outlook

In each chapter of the body of this thesis, a newly developed method for parameter estimation solely or parameter estimation and model choice in nonlinear regression is presented that has been tailor-made for an application in dynamic biomedical imaging. More specifically, this thesis provides two new methods that can be used in image analysis in FRAP and in DCE-MRI applications, respectively (Chapters 5 and 7). Therefore, the methods proposed in Chapters 5 and 7 of this thesis promote image analysis in oncology and fluorescence microscopy, as they open up new possibilities for image analysis in these fields.

In the future, it would be desirable that the implementation of an algorithm that tackles the problem described in the FRAP application in Chapter 6 within a reasonable time frame could be realized in order to be able to further study the binding behavior of molecules taking the spatial structure of cell nuclei into consideration. This could for example allow the location of binding sites of molecules of interest in a cell nucleus and would therefore further promote image analysis in FRAP applications.



# **Appendix A**

---

## **Appendix to Chapter 5**

---

## Full conditional distributions

For  $K = 1, 2, k = 1, 2$ :

$$a_0 | \cdot \sim N\left(\frac{m}{v}, \frac{1}{v}\right), \text{ with } v = \frac{\sum_{j=1}^J T_j}{\sigma^2},$$

$$m = \sum_{j=1}^J \sum_{i=1}^{T_j} \frac{-C_j(t_i) + 1 - \alpha_{0j} - \sum_{k=1}^K a_{kj} \exp(-b_{kj}^{\text{off}} t_i)}{\sigma^2}.$$

$$\alpha_{0j} | \cdot \sim N\left(\frac{m}{v + \frac{1}{\tau_{\alpha_0}^2}}, \frac{1}{v + \frac{1}{\tau_{\alpha_0}^2}}\right), \text{ with } v = \frac{T_j}{\sigma^2},$$

$$m = \sum_{i=1}^{T_j} \frac{1 - C_j(t_i) - a_0 - \sum_{k=1}^K a_{kj} \exp(-b_{kj}^{\text{off}} t_i)}{\sigma^2}.$$

$$\tau_{\alpha_0}^2 | \cdot \sim IG\left(c_0 + \frac{J}{2}, \frac{1}{2} \sum_{j=1}^J \alpha_{0j}^2 + d_0\right)$$

$$p(b_k^{\text{off}} | \cdot) = p(\exp(f_k) | \cdot) \propto$$

$$\prod_{j=1}^J \prod_{i=1}^{T_j} \exp\left(-\frac{1}{2\sigma^2} (C_j(t_i) - (1 - a_{0j} - \sum_{k=1}^K a_{kj} \exp(-\exp(f_k) \exp(\phi_{kj}) t_i)))^2\right)$$

$$p(\beta_{kj}^{\text{off}} | \cdot) = p(\exp(\phi_{kj}) | \cdot) \propto \prod_{i=1}^{T_j} \left( \exp\left(-\frac{1}{2\sigma^2} (C_j(t_i) - (1 - a_{0j} - \sum_{k=1}^K a_{kj} \exp(-\exp(f_k) \exp(\phi_{kj}) t_i)))^2\right) \right)$$

$$\cdot \frac{1}{\tau_{\beta_k^{\text{off}}} \sqrt{2\pi} \exp(\phi_{kj})} \exp\left(-\frac{1}{2\tau_{\beta_k^{\text{off}}}^2} \phi_{kj}^2\right)$$

$$\tau_{\alpha_k}^2 | \cdot \sim IG\left(c_k + \frac{J}{2}, \frac{1}{2} \sum_{j=1}^J \alpha_{kj}^2 + d_k\right)$$

$$\tau_{\beta_k^{\text{off}}}^2 | \cdot \sim IG\left(e_k + \frac{J}{2}, \frac{1}{2} \sum_{j=1}^J \phi_{kj}^2 + g_k\right)$$

$$\sigma^2 | \cdot \sim IG\left(a + \frac{1}{2} \sum_{j=1}^J T_j, b + \frac{1}{2} \sum_{j=1}^J \sum_{i=1}^{T_j} (C_j(t_i) - (1 - a_{0j} - \sum_{k=1}^K a_{kj} \exp(-b_{kj}^{\text{off}} t_i)))^2\right)$$

If  $K = 1$ :

$$a_1 | \cdot \sim N\left(\frac{m}{v}, \frac{1}{v}\right), \text{ with } v = \sum_{j=1}^J \sum_{i=1}^{T_j} \frac{\exp(-2b_{1j}^{\text{off}} t_i)}{\sigma^2},$$

$$m = \sum_{j=1}^J \sum_{i=1}^{T_j} \frac{(-C_j(t_i) - \alpha_{1j} \exp(-b_{1j}^{\text{off}} t_i) + 1 - a_{0j}) \exp(-b_{1j}^{\text{off}} t_i)}{\sigma^2}.$$

$$\alpha_{1j} | \cdot \sim N\left(\frac{m}{v + \frac{1}{\tau_{\alpha_1}^2}}, \frac{1}{v + \frac{1}{\tau_{\alpha_1}^2}}\right), \text{ with } v = \sum_{i=1}^{T_j} \frac{\exp(-2b_{1j}^{\text{off}} t_i)}{\sigma^2},$$

$$m = \sum_{i=1}^{T_j} \frac{(-C_j(t_i) + 1 - a_{0j} - a_1 \exp(-b_{1j}^{\text{off}} t_i)) \exp(-b_{1j}^{\text{off}} t_i)}{\sigma^2}.$$

If  $K = 2$ , for  $k = 1, 2$ :

$$a_k | \cdot \sim N\left(\frac{m}{v}, \frac{1}{v}\right), \text{ with } v = \sum_{j=1}^J \sum_{i=1}^{T_j} \frac{\exp(-2b_{kj}^{\text{off}} t_i)}{\sigma^2},$$

$$m = \sum_{j=1}^J \sum_{i=1}^{T_j} \frac{(-C_j(t_i) - \alpha_{kj} \exp(-b_{kj}^{\text{off}} t_i) - a_{-kj} \exp(-b_{-kj}^{\text{off}} t_i) + 1 - a_{0j}) \exp(-b_{kj}^{\text{off}} t_i)}{\sigma^2},$$

where  $-k = 2$ , if  $k = 1$ , and  $-k = 1$ , if  $k = 2$ .

$$\alpha_{kj} | \cdot \sim N\left(\frac{m}{v + \frac{1}{\tau_{\alpha_k}^2}}, \frac{1}{v + \frac{1}{\tau_{\alpha_k}^2}}\right), \text{ with } v = \sum_{i=1}^{T_j} \frac{\exp(-2b_{kj}^{\text{off}} t_i)}{\sigma^2},$$

$$m = \sum_{i=1}^{T_j} \frac{(-C_j(t_i) + 1 - a_{0j} - a_k \exp(-b_{kj}^{\text{off}} t_i) - a_{-kj} \exp(-b_{-kj}^{\text{off}} t_i)) \exp(-b_{kj}^{\text{off}} t_i)}{\sigma^2},$$

where  $-k = 2$ , if  $k = 1$ , and  $-k = 1$ , if  $k = 2$ .



# **Appendix B**

---

## **Appendix to Chapter 6**

---

## Full conditional distributions

### Pixel-specific parameters:

$$\begin{aligned}
 p(b_{1j}^{\text{on}^*} | \cdot) &= p(\exp(\theta_{1j}) | \cdot) \propto \prod_{i=1}^T \left( \frac{1}{\sigma} \exp\left(-\frac{1}{2\sigma^2}(C_j(t_i) - \text{total}_j(t_i))^2\right) \right) \\
 &\quad \cdot \frac{1}{\sqrt{\tau_1^2 \exp(\theta_{1j})}} \exp\left(-\frac{1}{2} \frac{(\theta_{1j} - \mu_1)^2}{\tau_1^2}\right) \\
 p(b_{1j}^{\text{off}} | \cdot) &= p(\exp(\theta_{3j}) | \cdot) \propto \prod_{i=1}^T \left( \frac{1}{\sigma} \exp\left(-\frac{1}{2\sigma^2}(C_j(t_i) - \text{total}_j(t_i))^2\right) \right) \\
 &\quad \cdot \frac{1}{\sqrt{\tau_1^2 \exp(\theta_{3j})}} \exp\left(-\frac{1}{2} \frac{(\theta_{3j} - \mu_3)^2}{\tau_3^2}\right)
 \end{aligned}$$

For  $f_{0j} \sim U(0, C_j(t_1))$ , we use the transformation  $g(f_{0j}) = \psi_{1j} = \log\left(\frac{f_{0j}-0}{C_j(t_1)-f_{0j}}\right)$ . It therefore follows that  $\psi_{1j} \in [-\inf, \inf]$ . With the transformation theorem for densities, we obtain the density of  $g(f_{0j}) = \psi_{1j}$ :

$$p(\psi_{1j} | \cdot) \propto \prod_{i=1}^T \left( \frac{1}{\sigma} \exp\left(-\frac{1}{2\sigma^2}(C_j(t_i) - \text{total}_j(t_i))^2\right) \right) \cdot \frac{C_j(t_1) \exp(\psi_{1j})}{(1 + \exp(\psi_{1j}))^2}$$

We use a similar transformation for  $a_{0j}$ . Additionally, we introduce the constraint  $a_{0j} + f_{0j} = C_j(t_1)$ . With this constraint, we get  $\psi_{2j} = -\psi_{1j}$ , and therefore only draw a proposal for  $\psi_{1j}$  in the MH-step in the MCMC algorithm, and calculate the proposal for  $\psi_{2j}$  from that.

### Common parameters:

$$\begin{aligned}
 \sigma^2 | \cdot &\sim IG\left(a + \frac{1}{2}TJ, b + \frac{1}{2} \sum_{j=1}^J \sum_{i=1}^T (C_j(t_i) - \text{total}_j(t_i))^2\right) \\
 p(b_2^{\text{in}} | \cdot) &= p(\exp(\theta_2) | \cdot) \propto \prod_{j=1}^J \prod_{i=1}^T \left( \frac{1}{\sigma} \exp\left(-\frac{1}{2\sigma^2}(C_j(t_i) - \text{total}_j(t_i))^2\right) \right) \\
 &\quad \cdot \frac{1}{\sqrt{\tau_2^2 \exp(\theta_2)}} \exp\left(-\frac{1}{2} \frac{(\theta_2 - \mu_2)^2}{\tau_2^2}\right)
 \end{aligned}$$

# **Appendix C**

---

## **Appendix to Chapter 7**

---

## Results of the simulation study - Additional figures

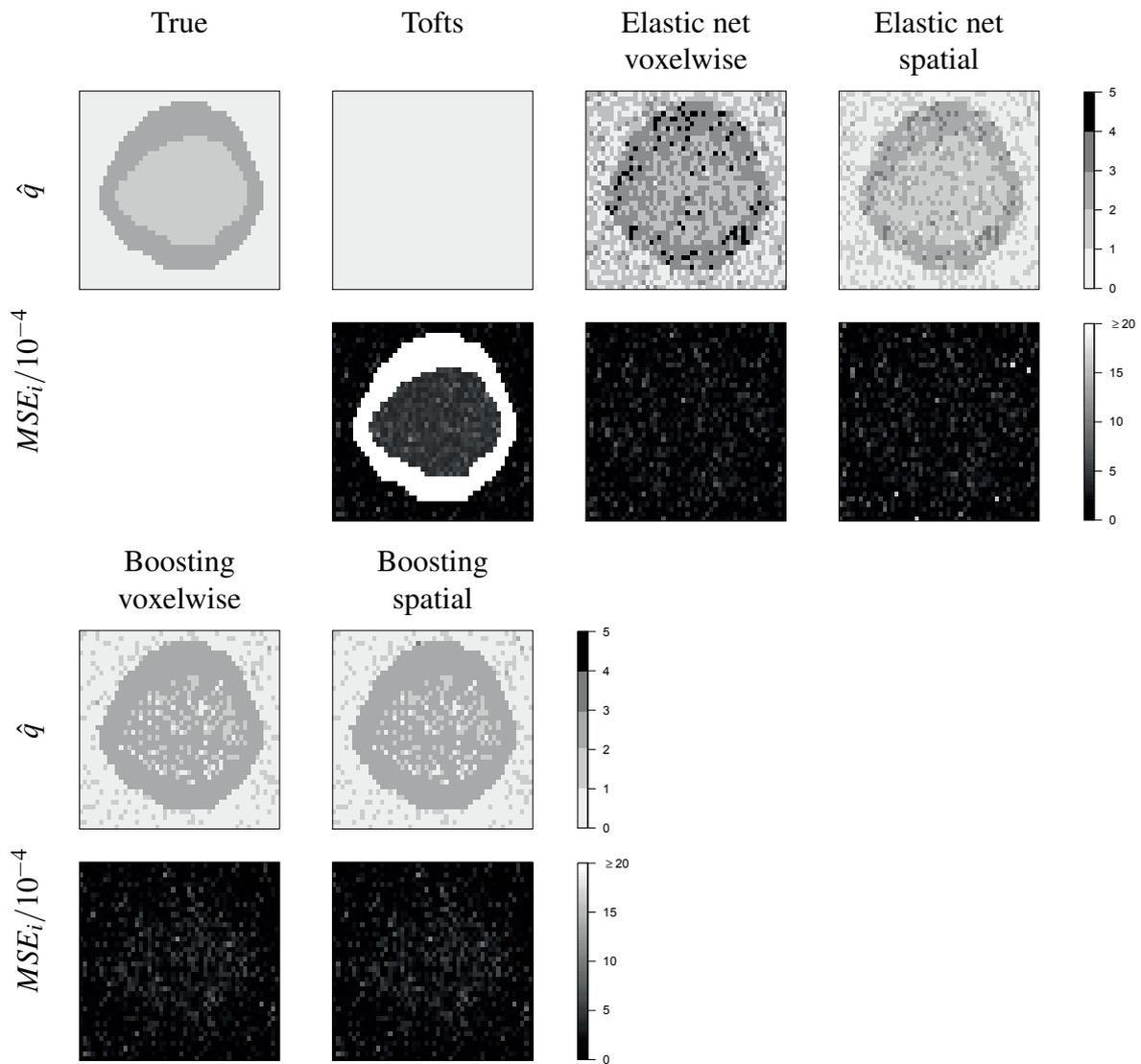


Figure C.1: Results of Simulation 1 for the second of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

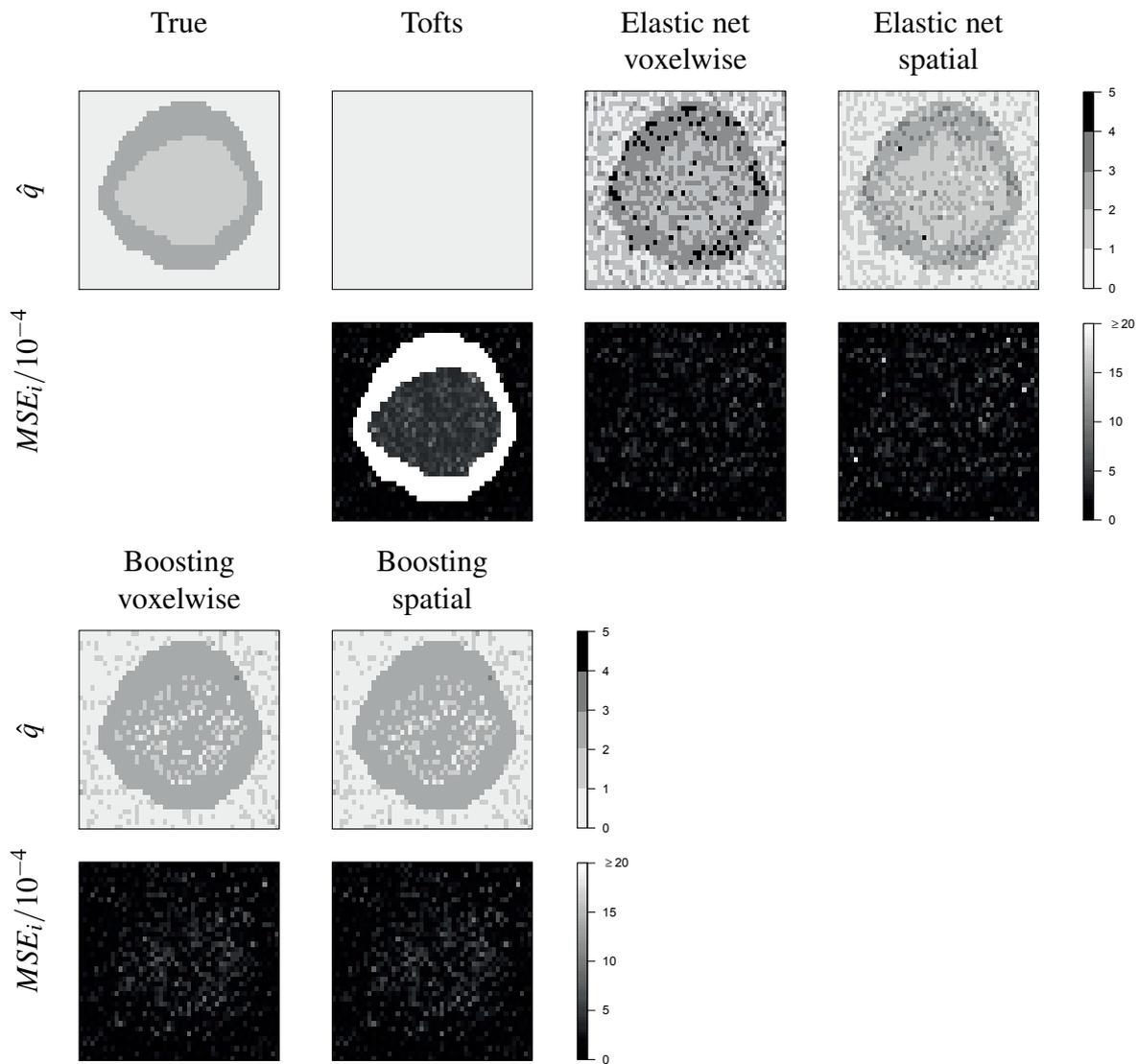


Figure C.2: Results of Simulation 1 for the third of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

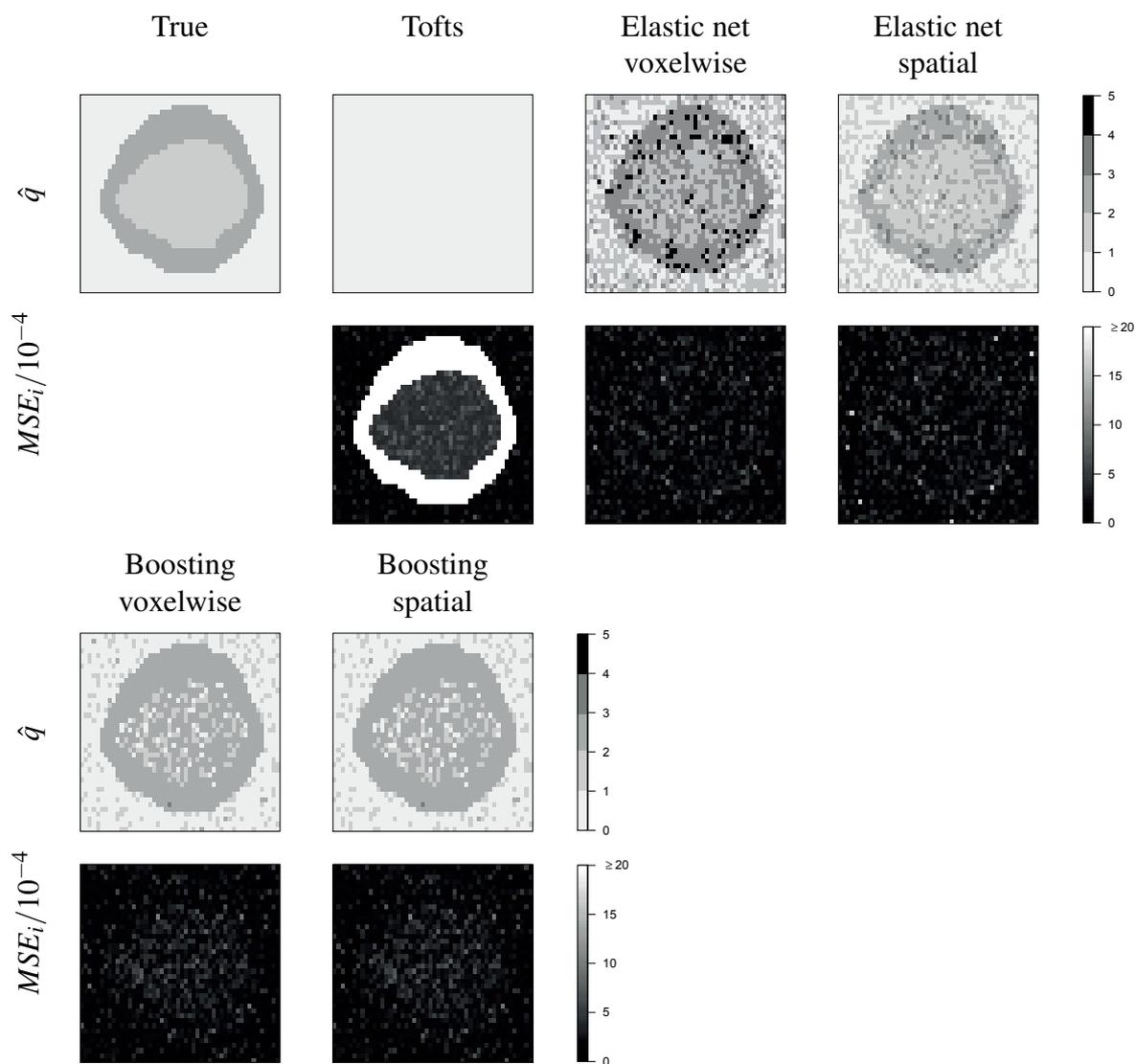


Figure C.3: Results of Simulation 1 for the fourth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

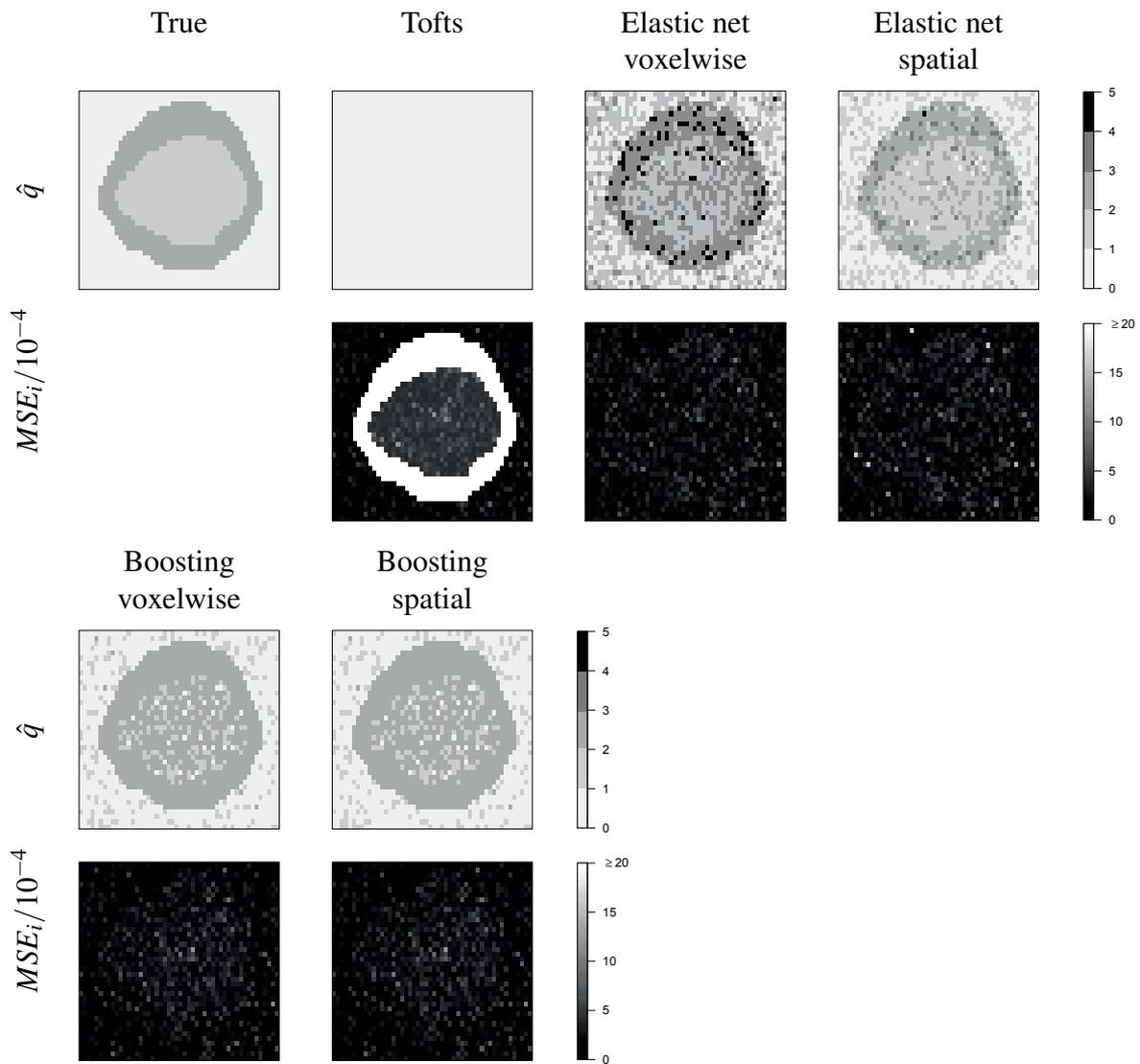


Figure C.4: Results of Simulation 1 for the fifth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

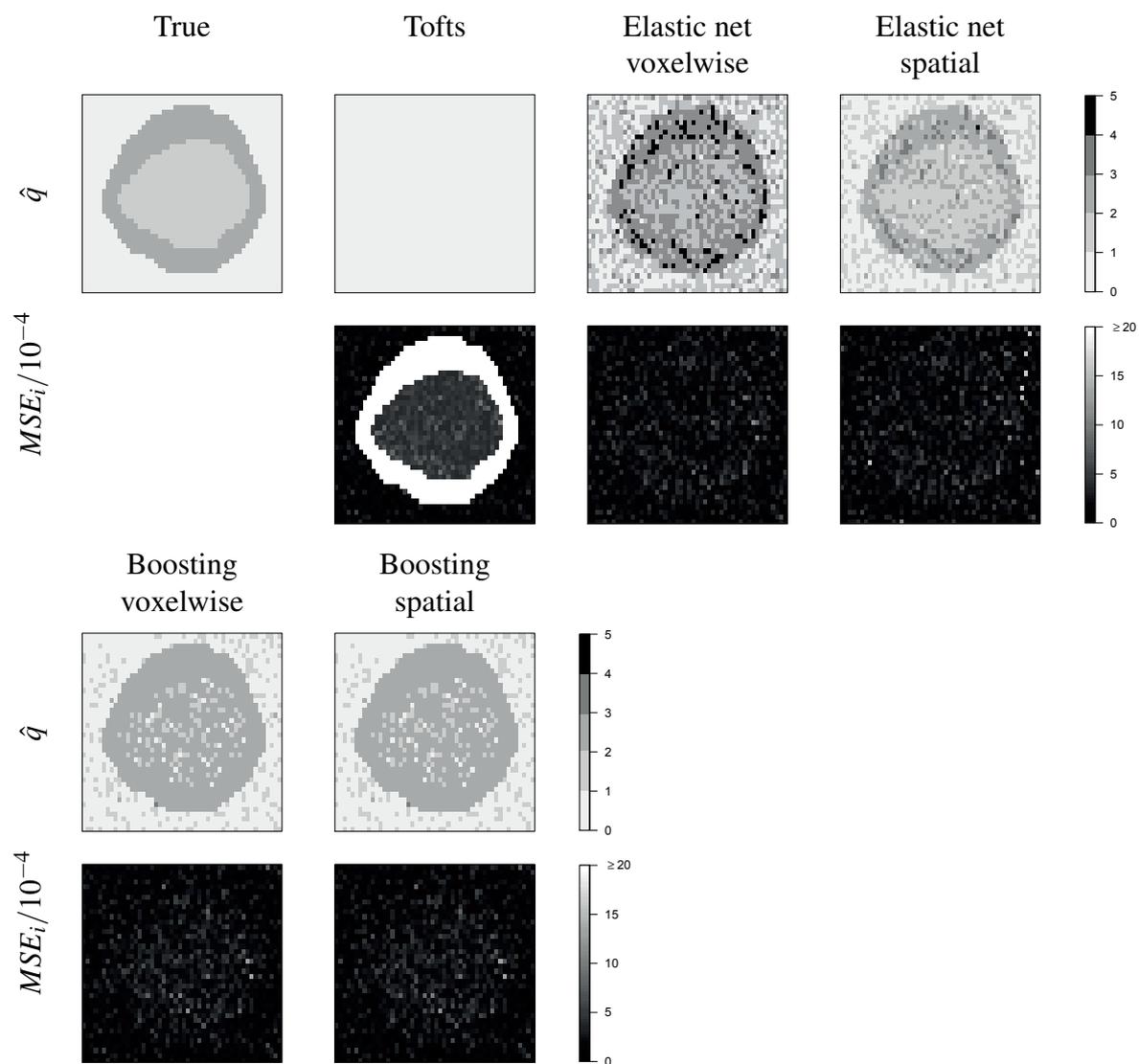


Figure C.5: Results of Simulation 1 for the sixth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

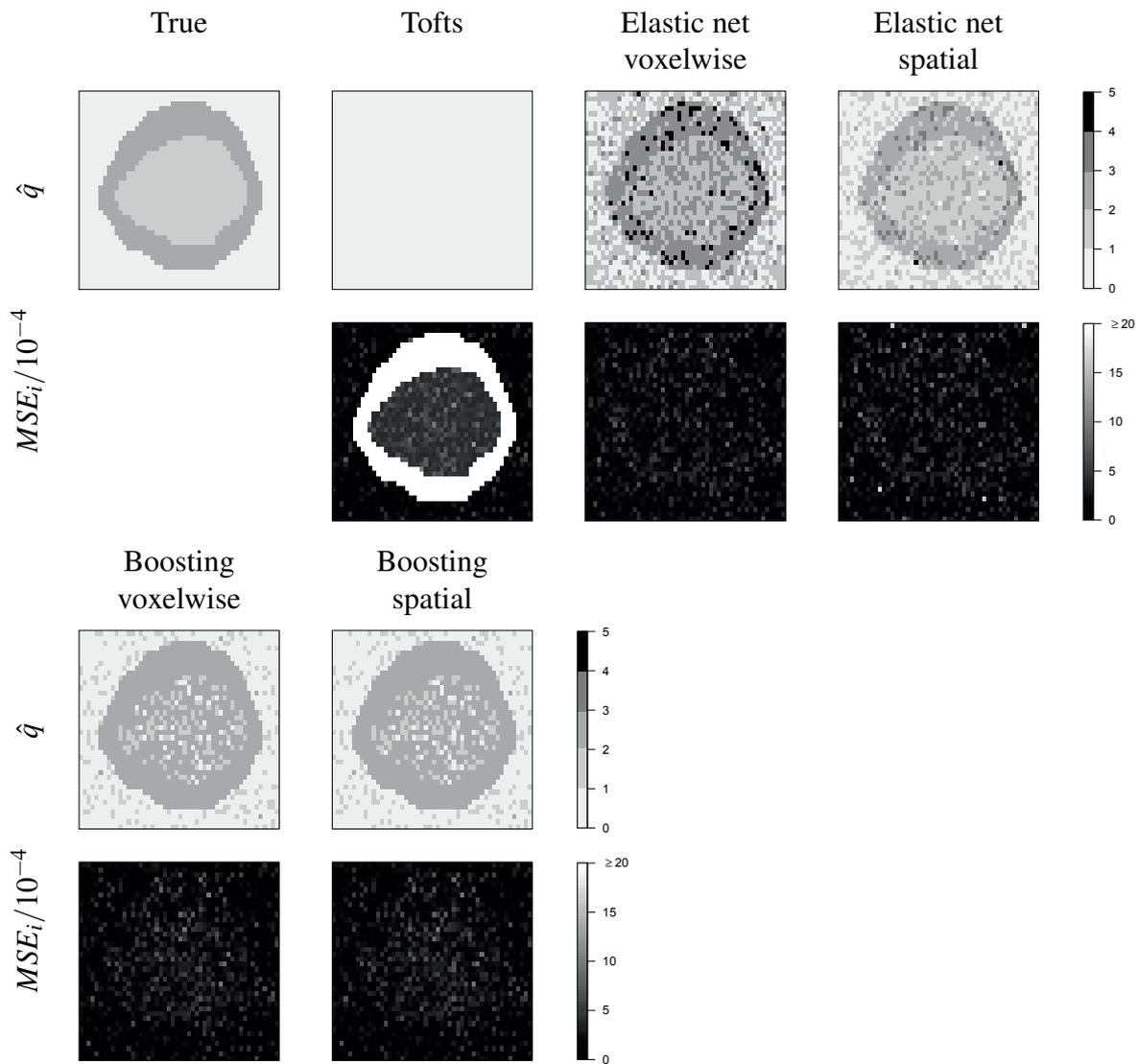


Figure C.6: Results of Simulation 1 for the seventh of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

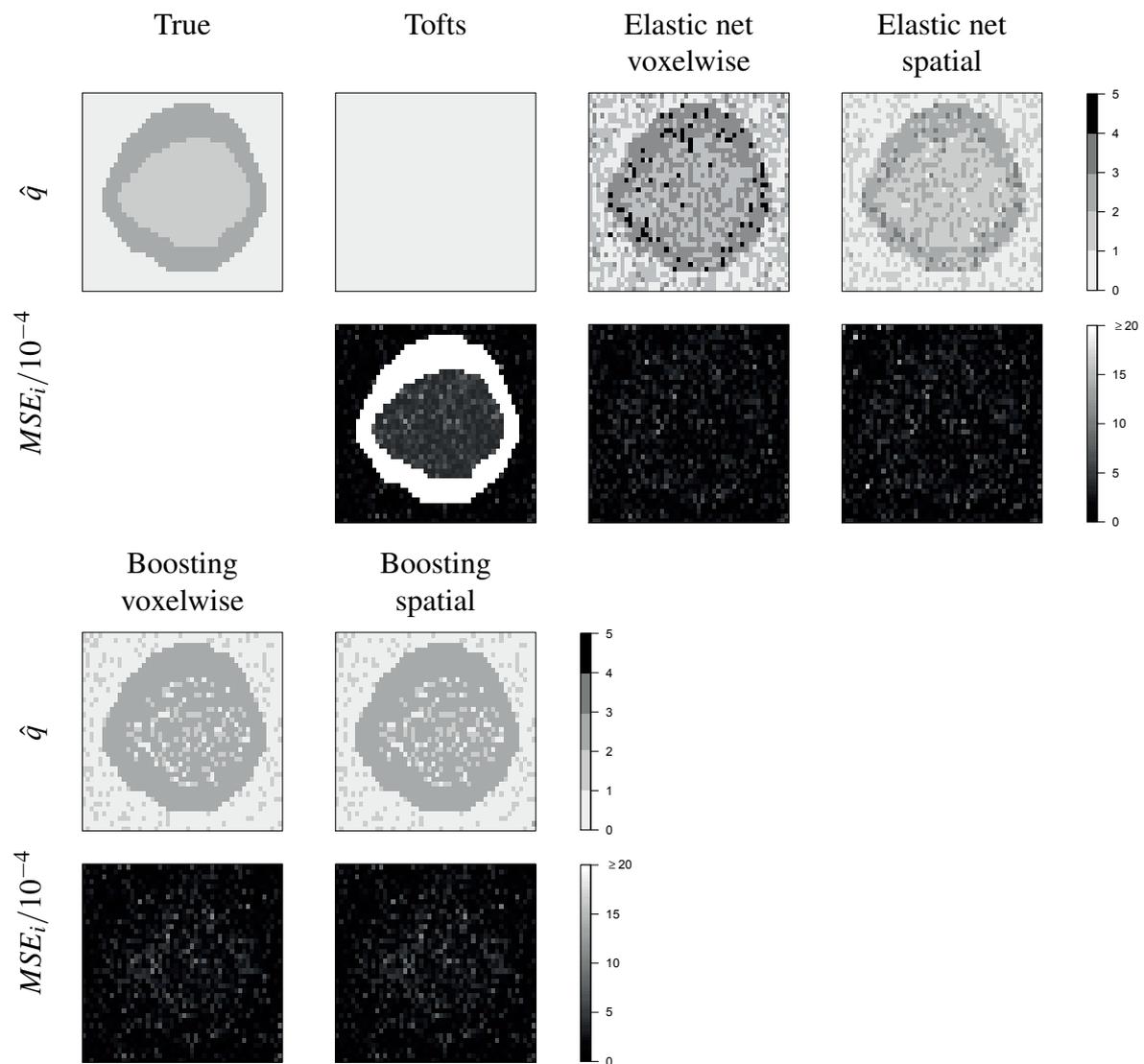


Figure C.7: Results of Simulation 1 for the eighth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

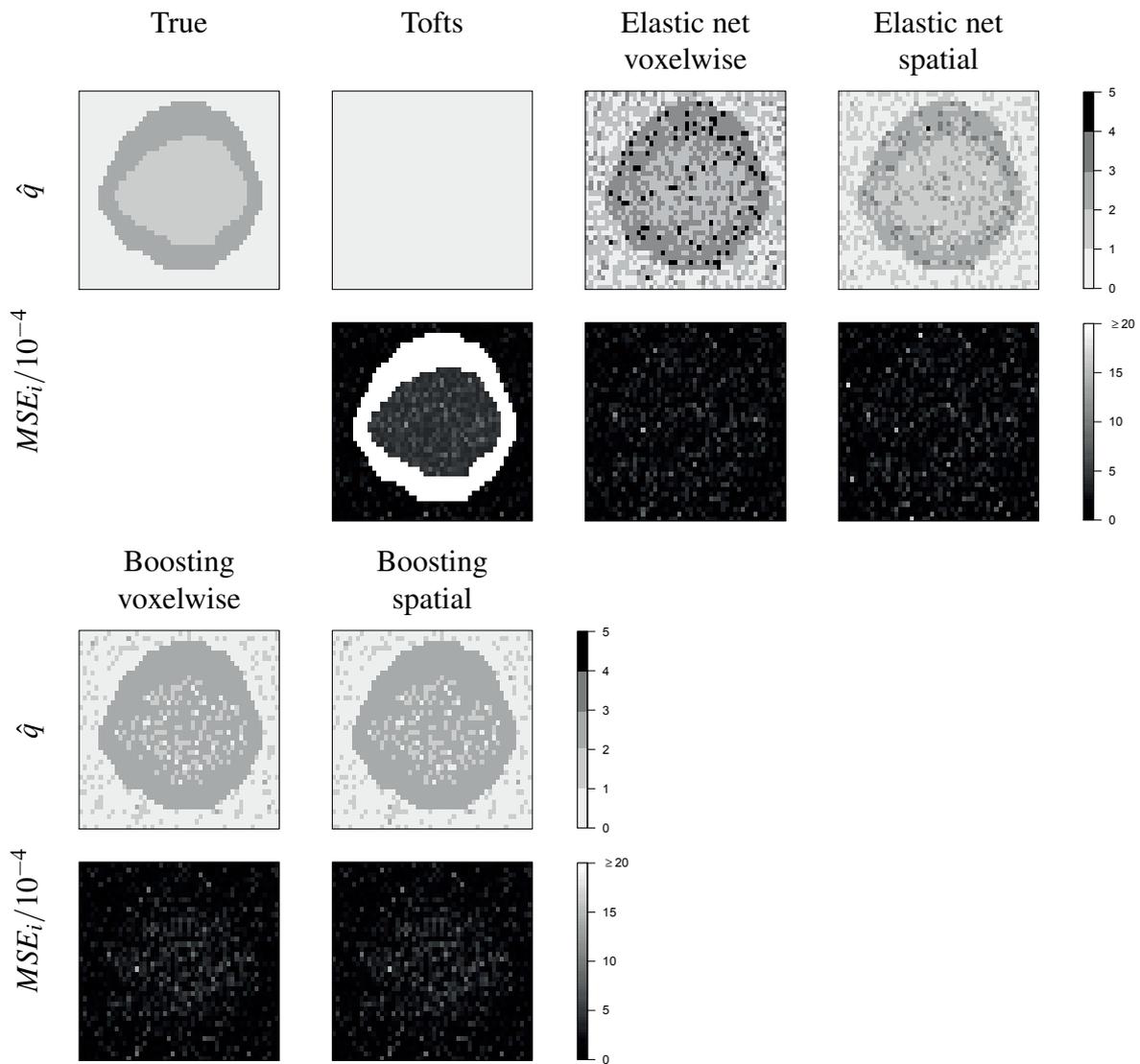


Figure C.8: Results of Simulation 1 for the ninth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

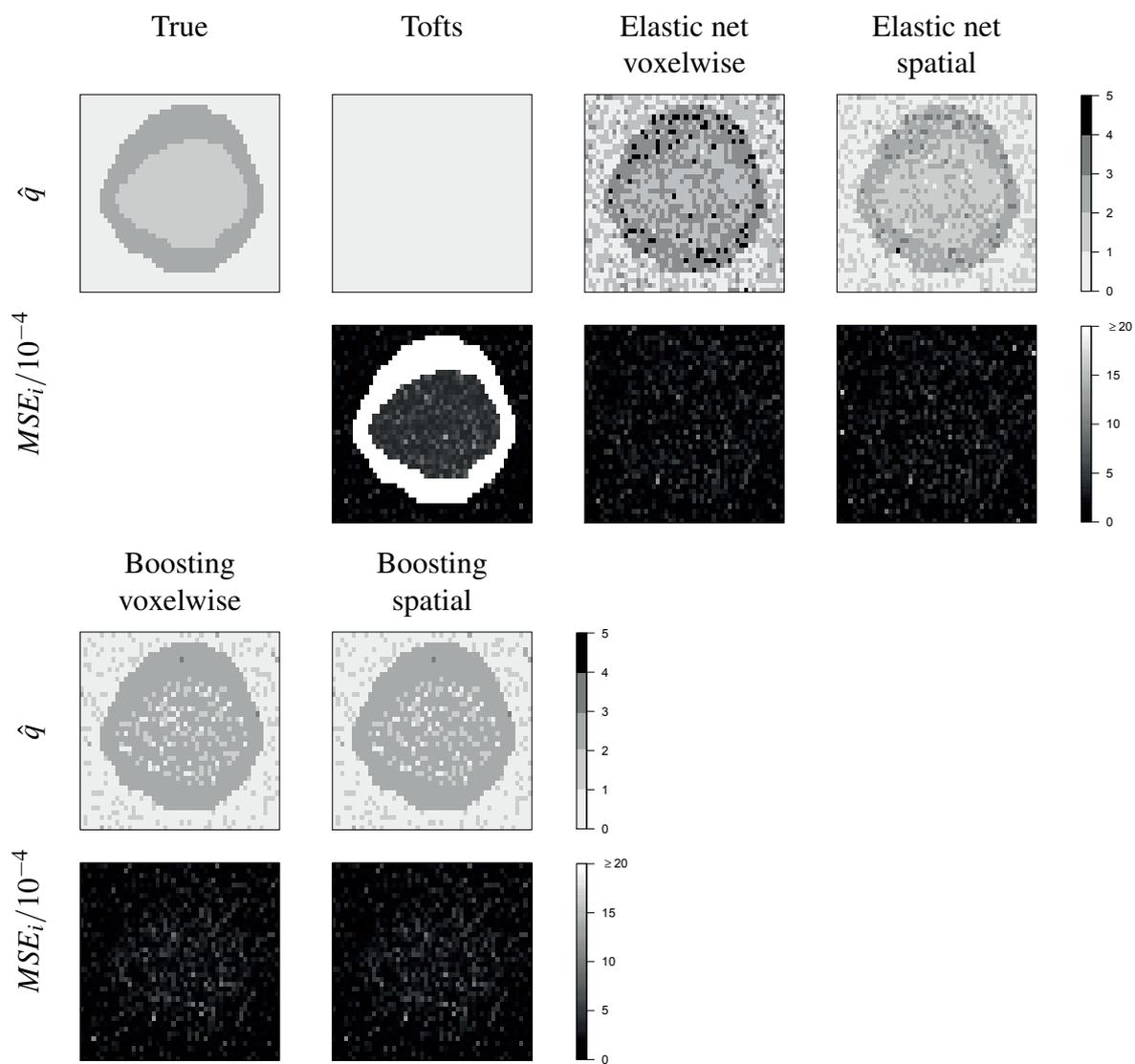


Figure C.9: Results of Simulation 1 for the tenth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

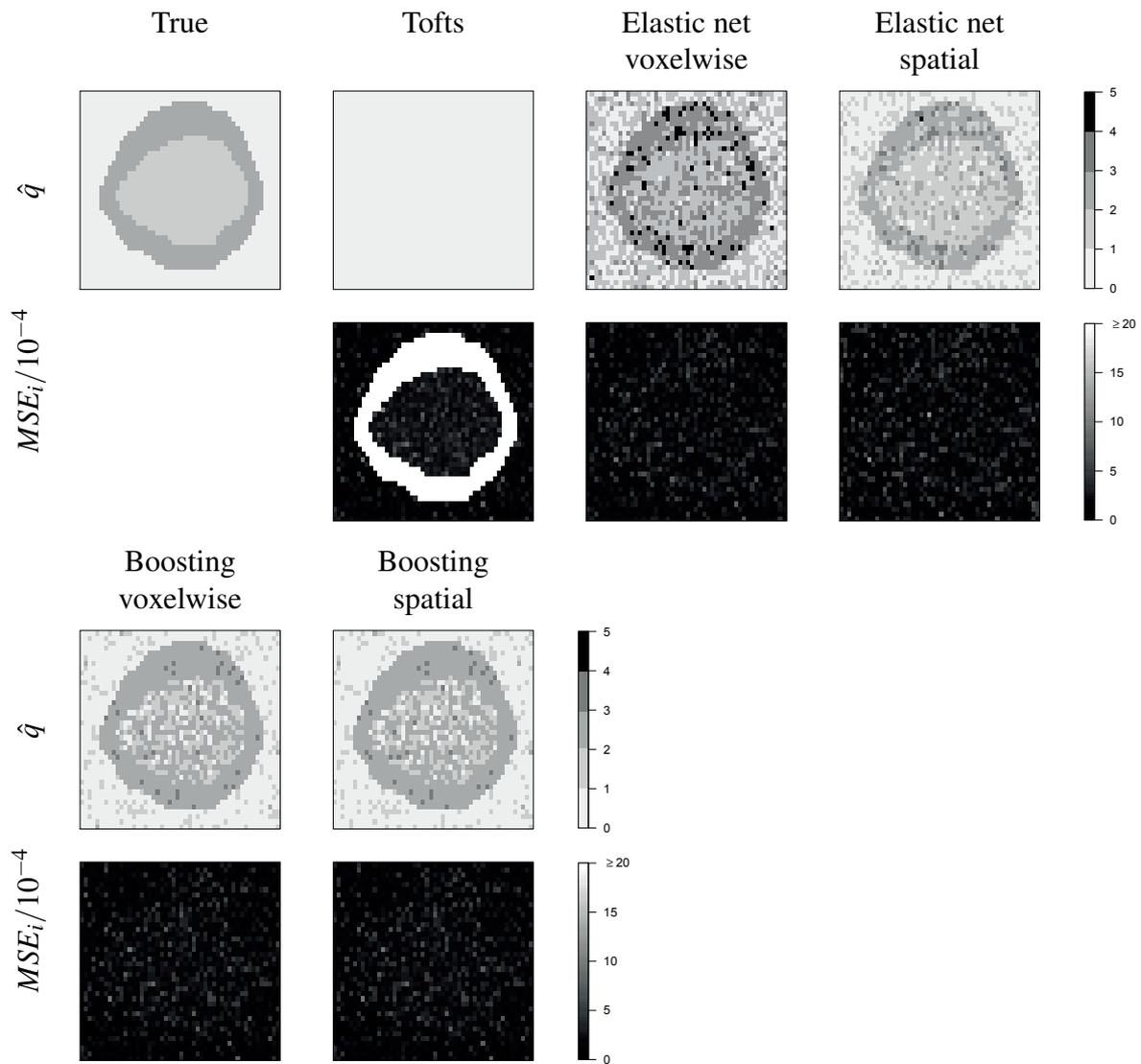


Figure C.10: Results of Simulation 2 for the first of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

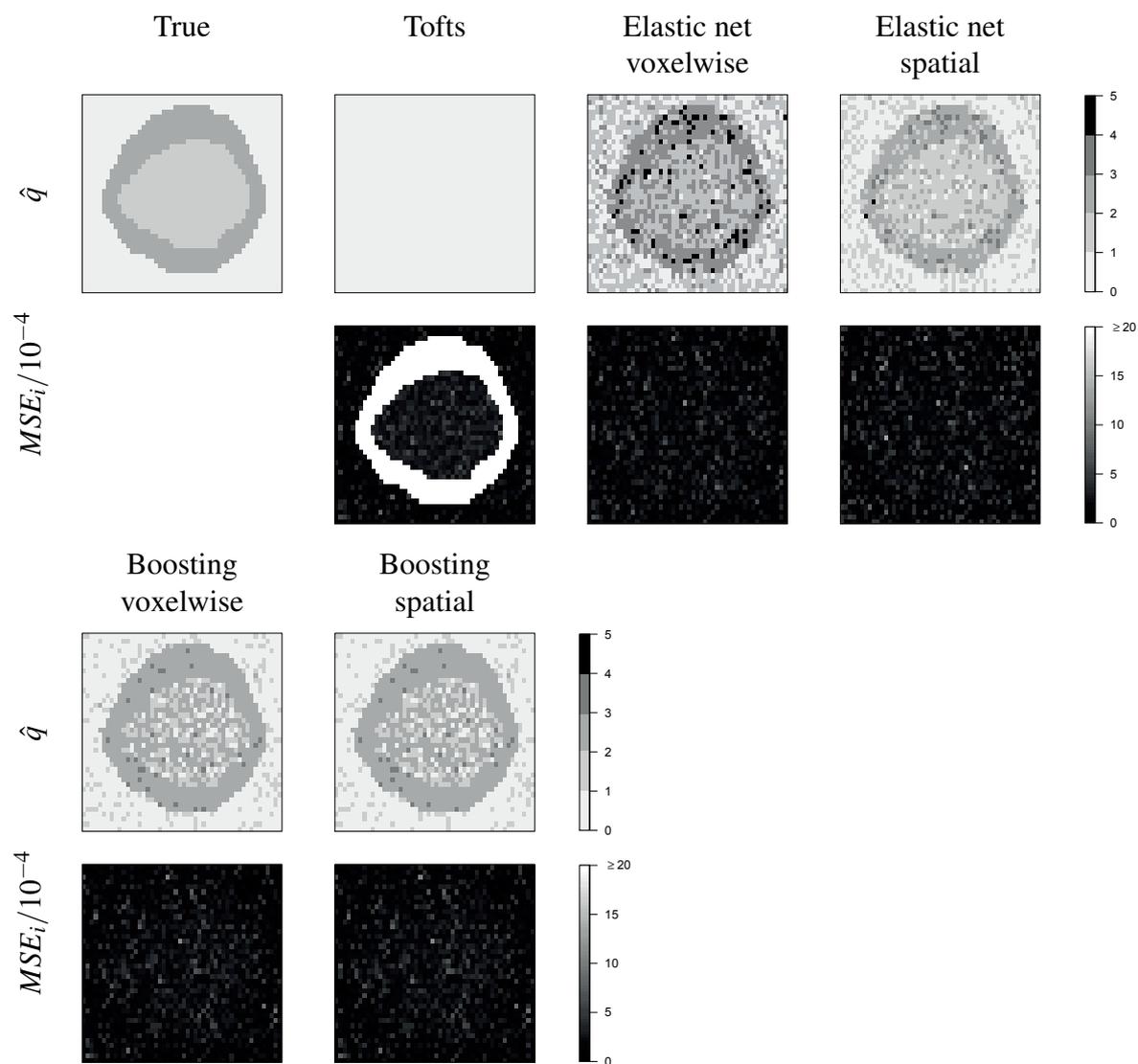


Figure C.11: Results of Simulation 2 for the second of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

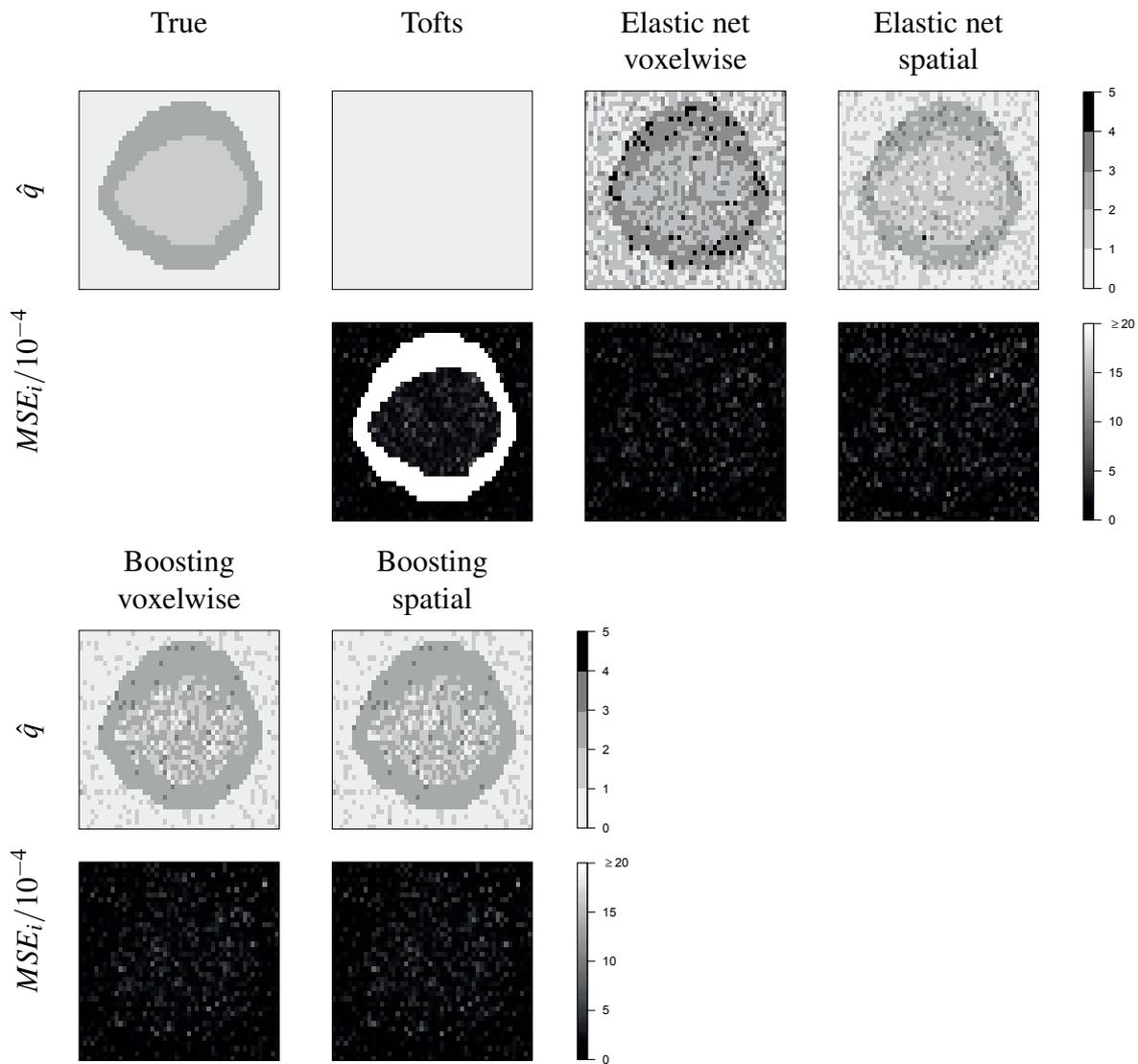


Figure C.12: Results of Simulation 2 for the third of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

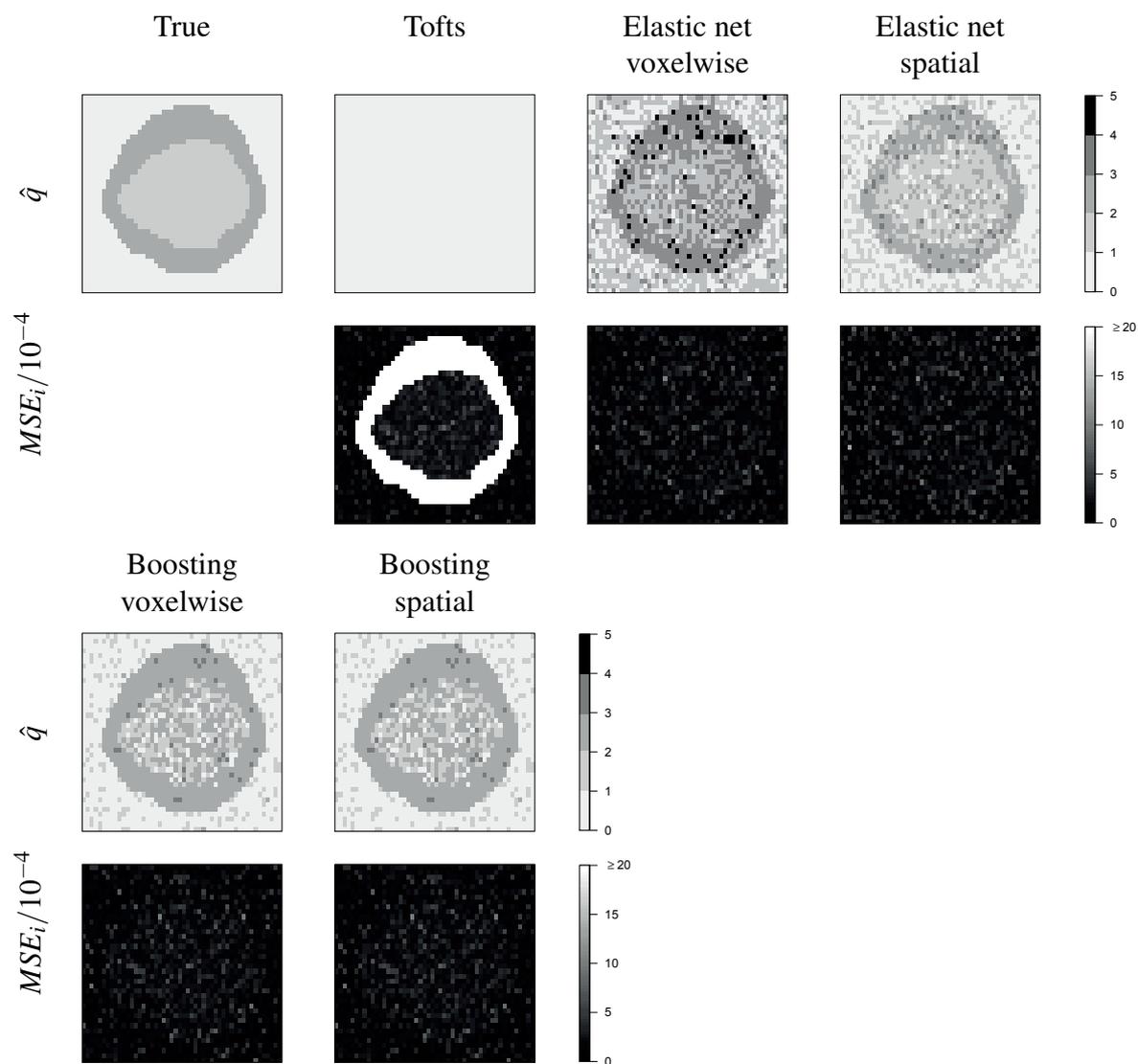


Figure C.13: Results of Simulation 2 for the fourth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

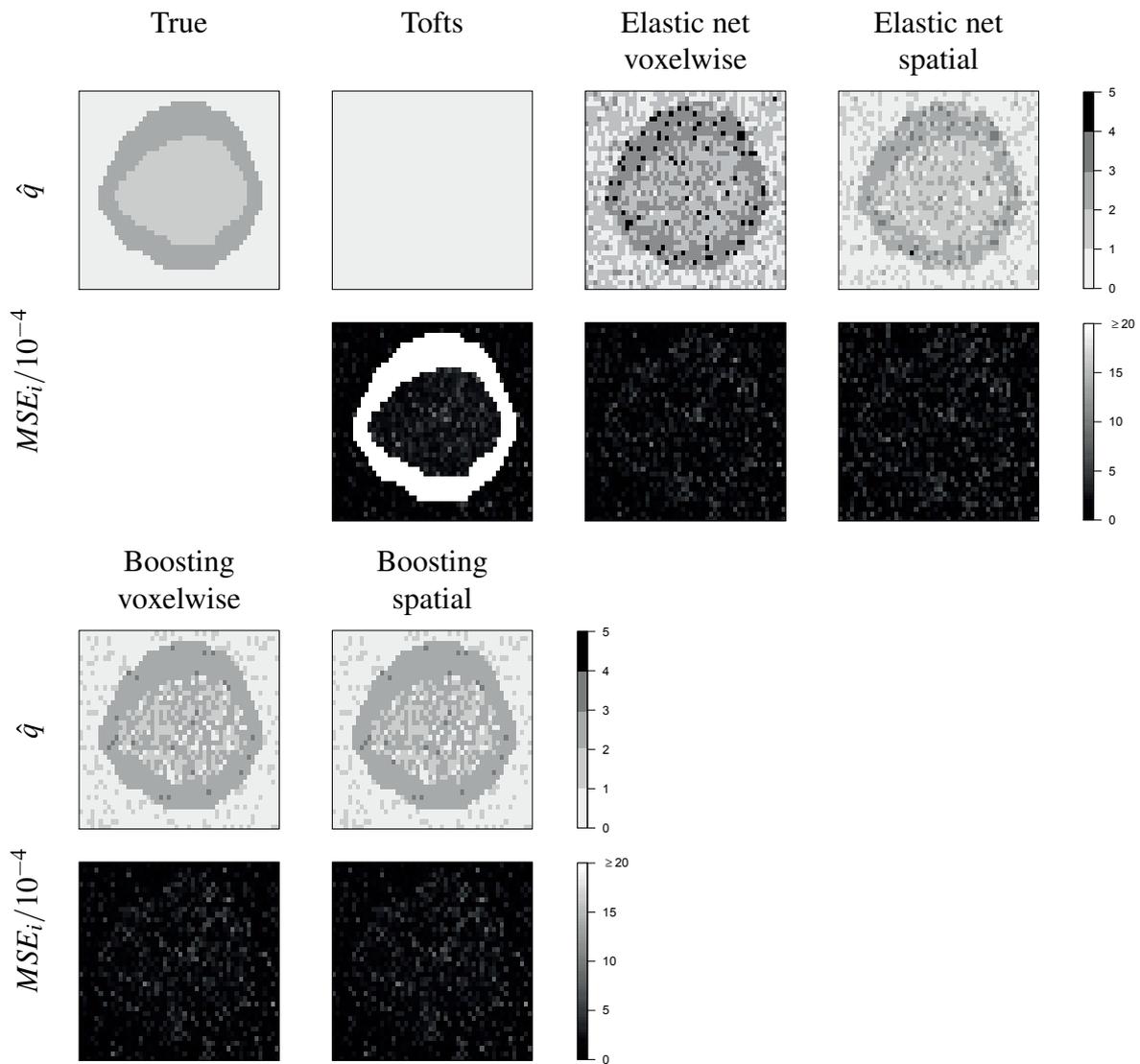


Figure C.14: Results of Simulation 2 for the fifth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

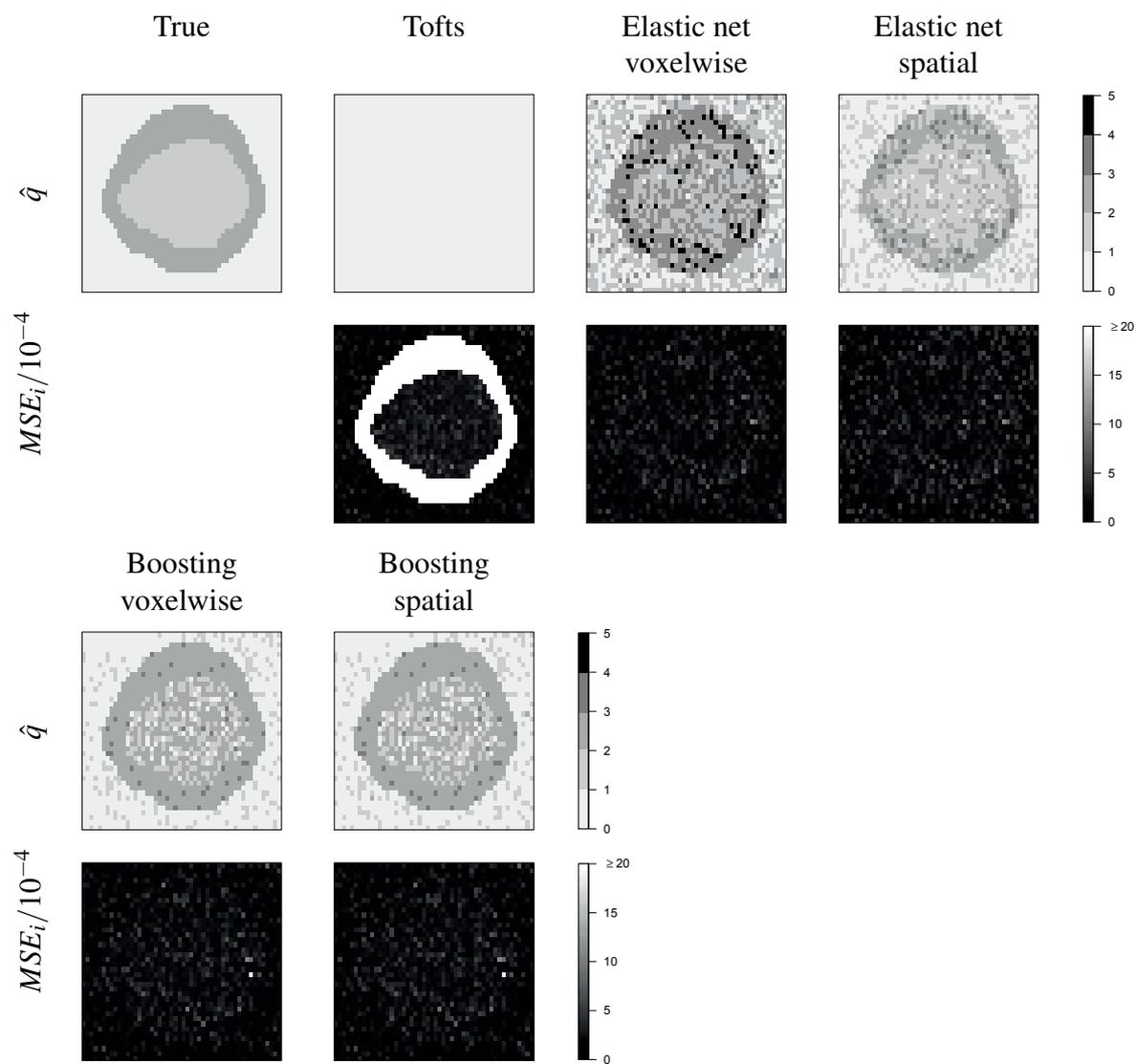


Figure C.15: Results of Simulation 2 for the sixth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

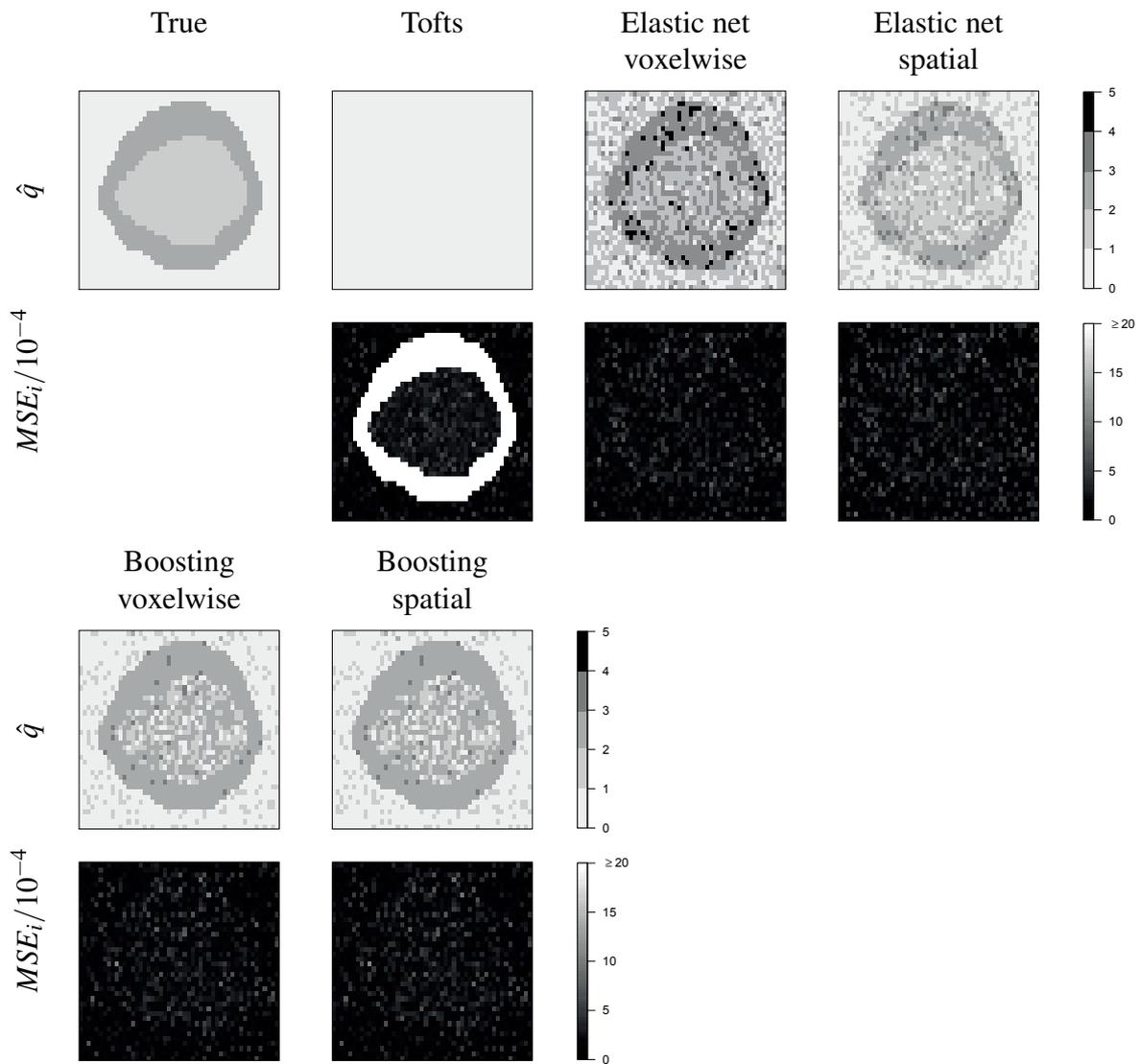


Figure C.16: Results of Simulation 2 for the seventh of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

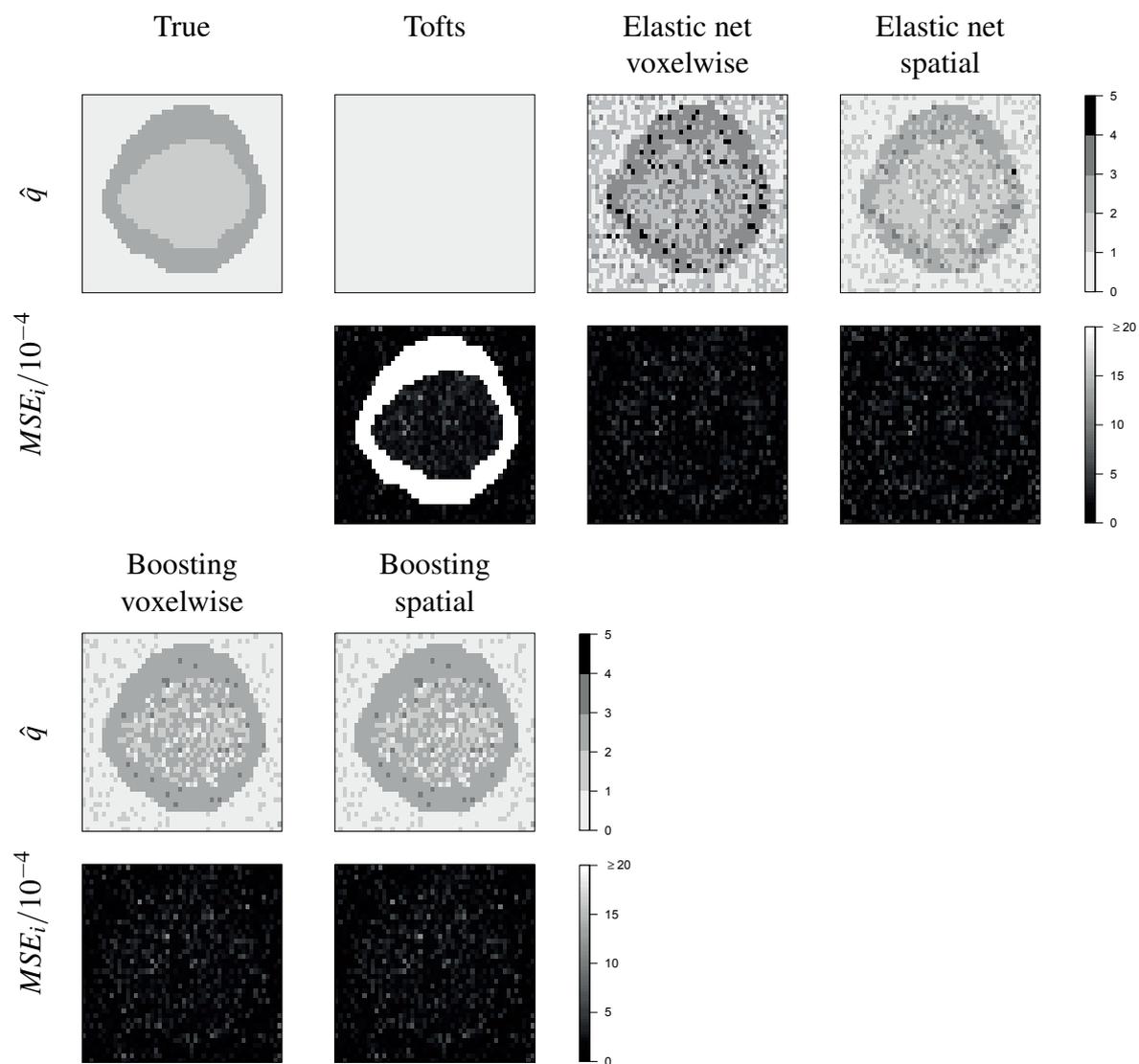


Figure C.17: Results of Simulation 2 for the eighth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

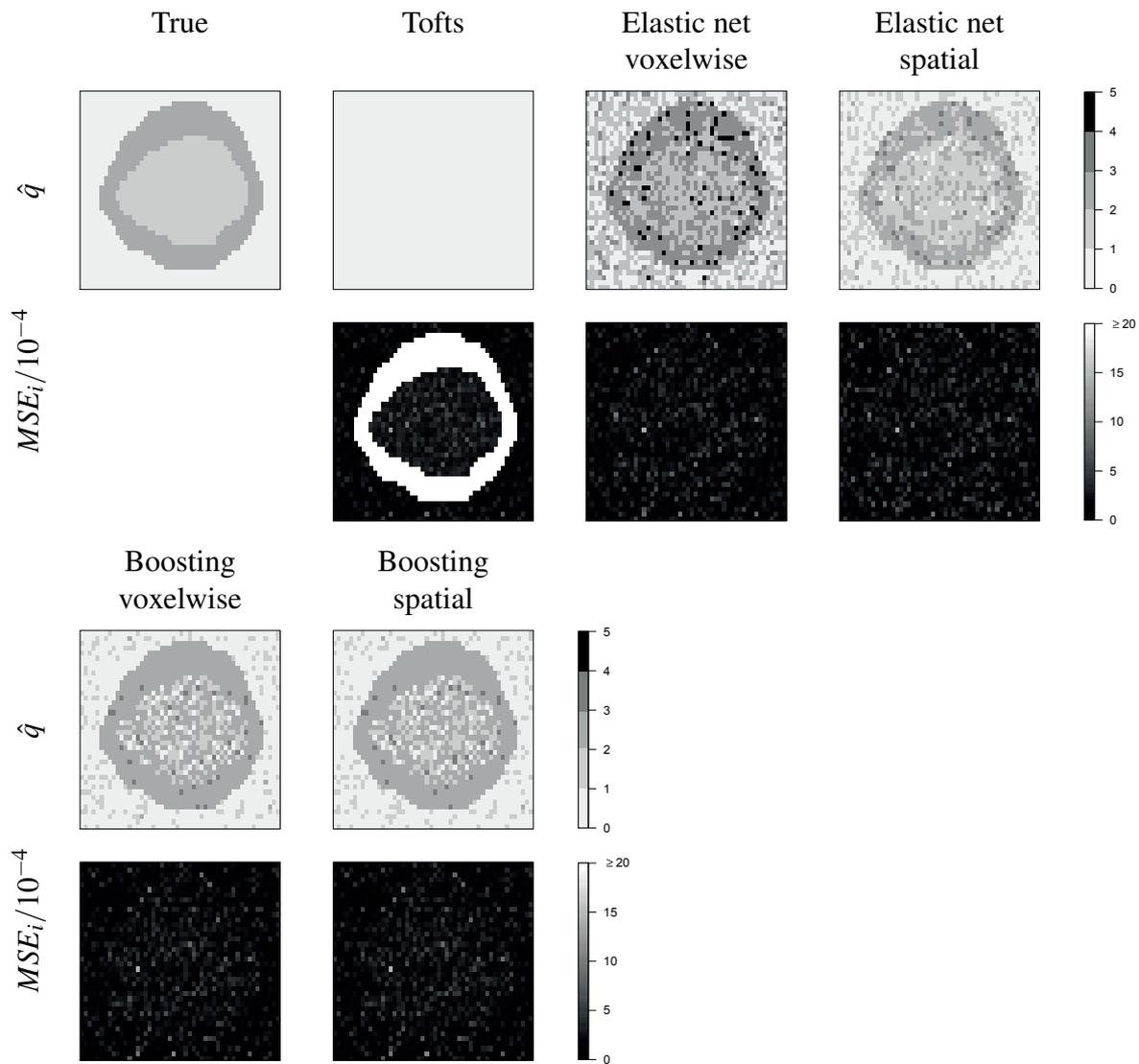


Figure C.18: Results of Simulation 2 for the ninth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

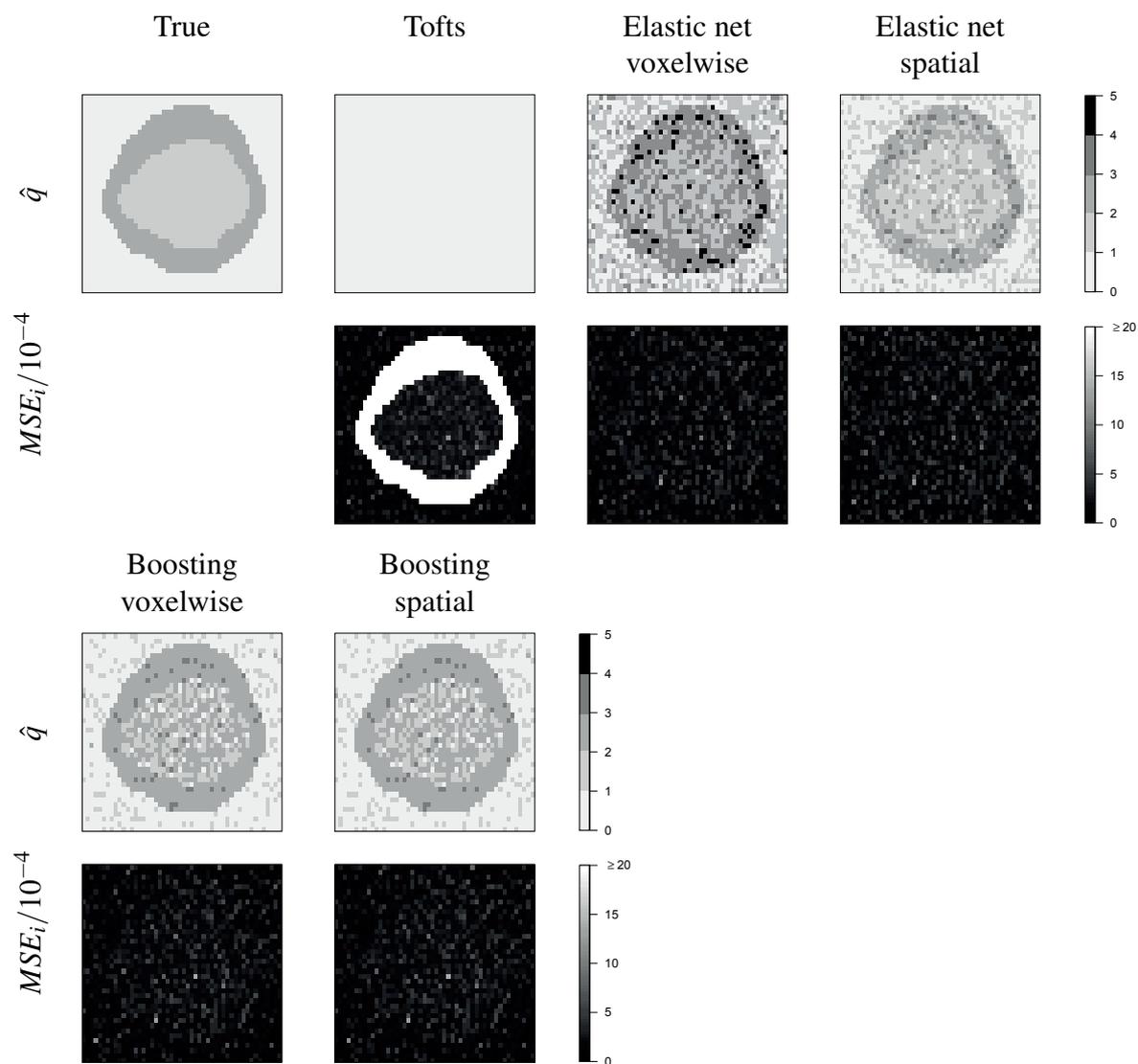


Figure C.19: Results of Simulation 2 for the tenth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

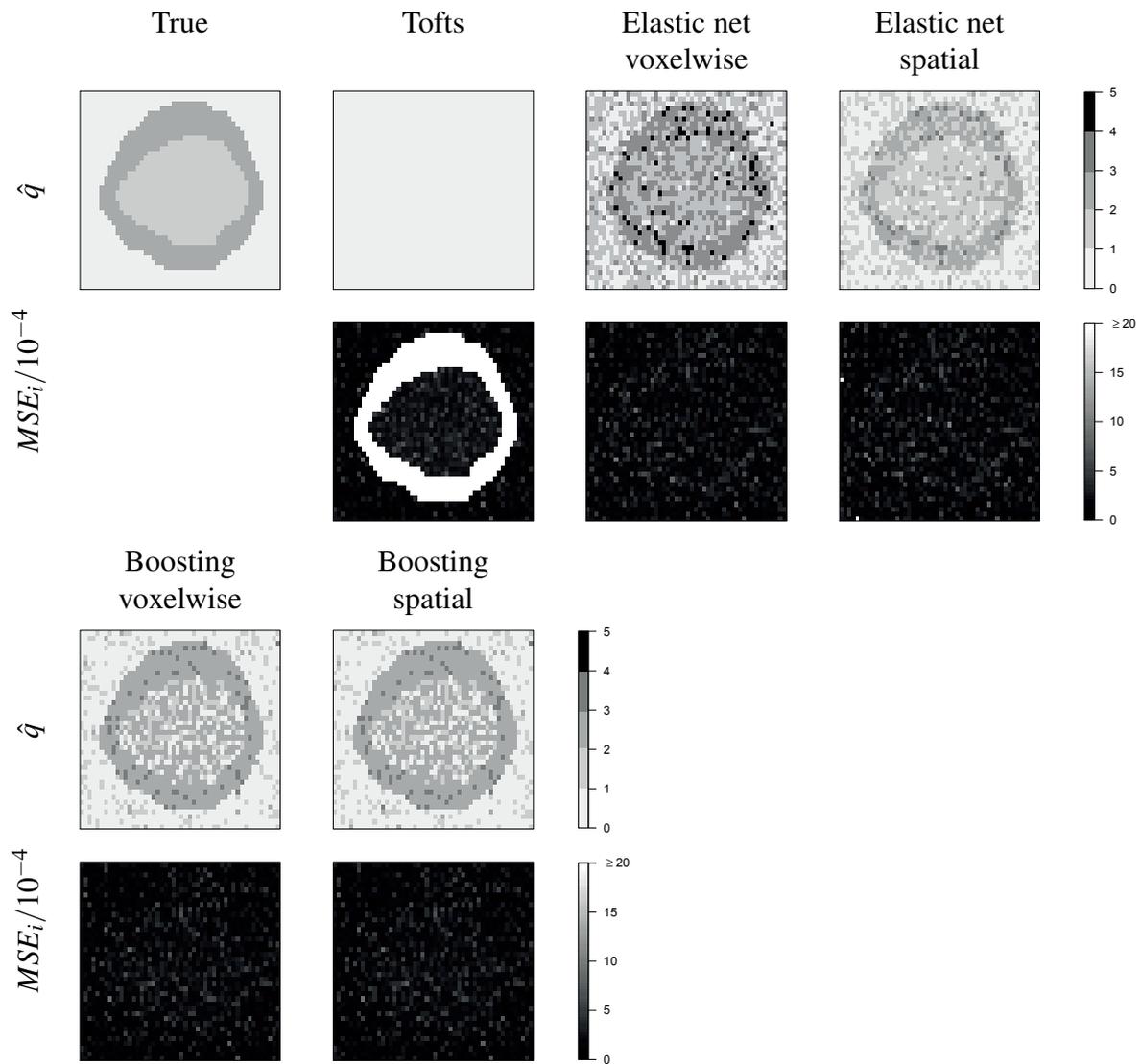


Figure C.20: Results of Simulation 3 for the first of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

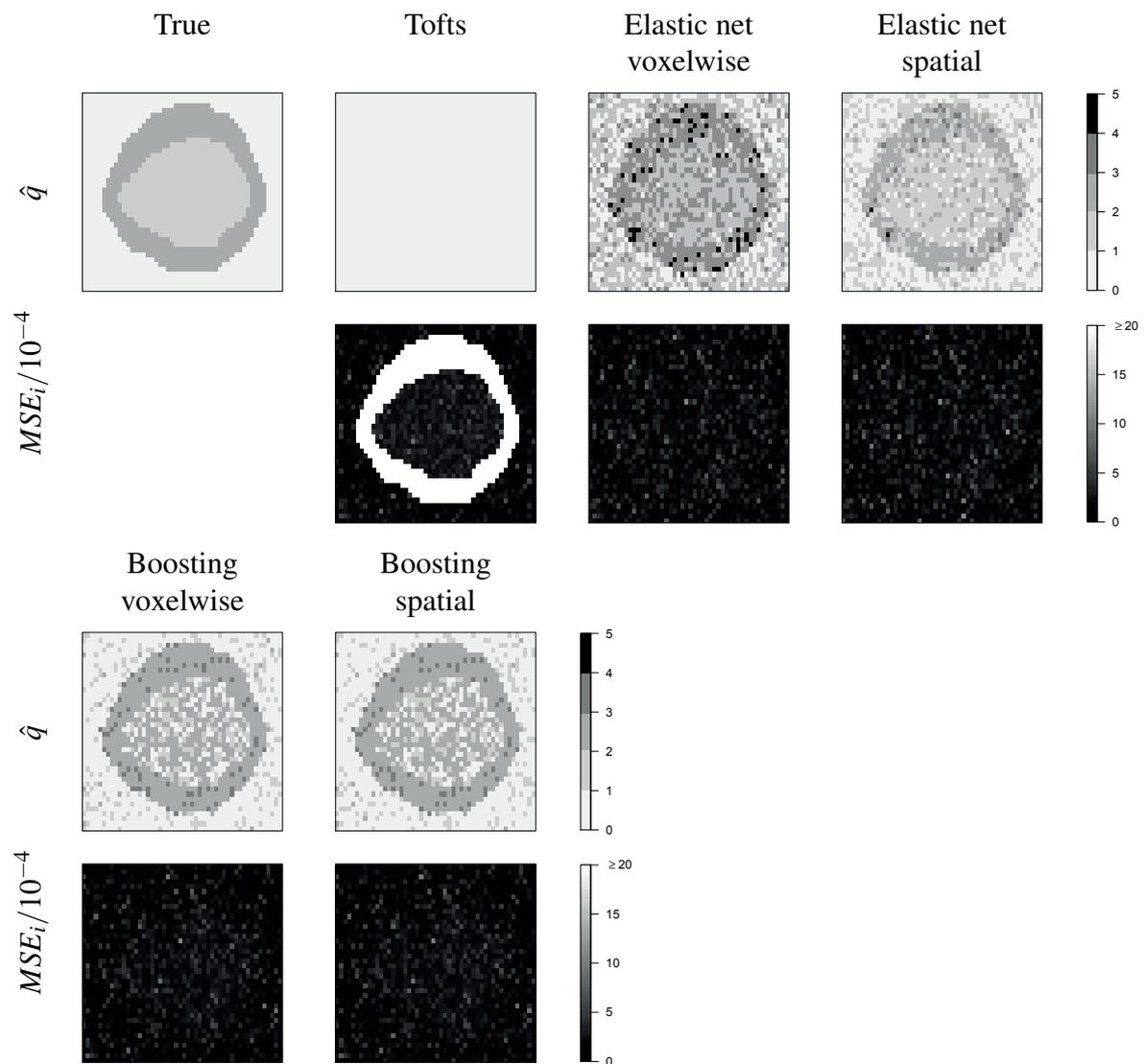


Figure C.21: Results of Simulation 3 for the second of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

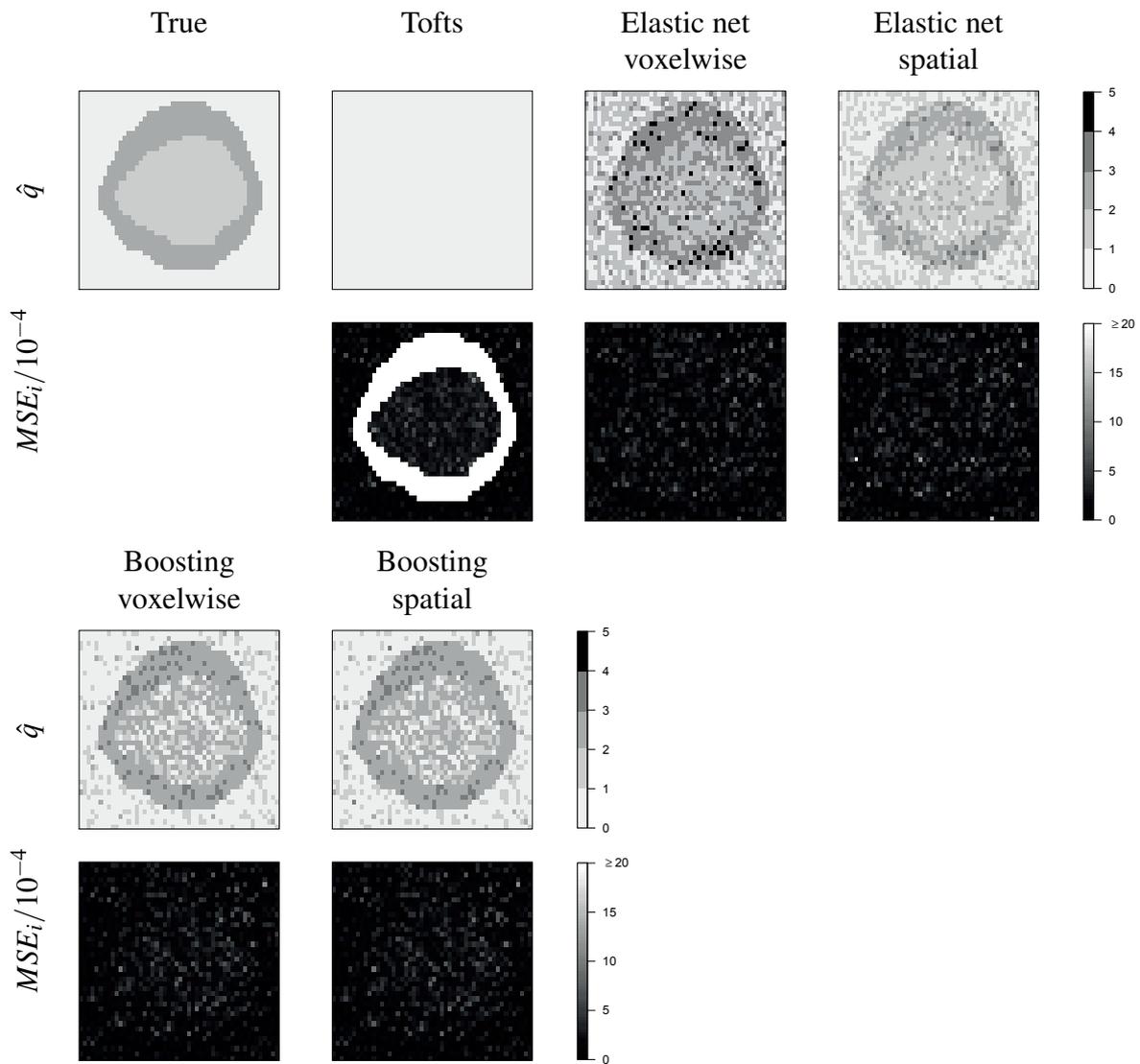


Figure C.22: Results of Simulation 3 for the third of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

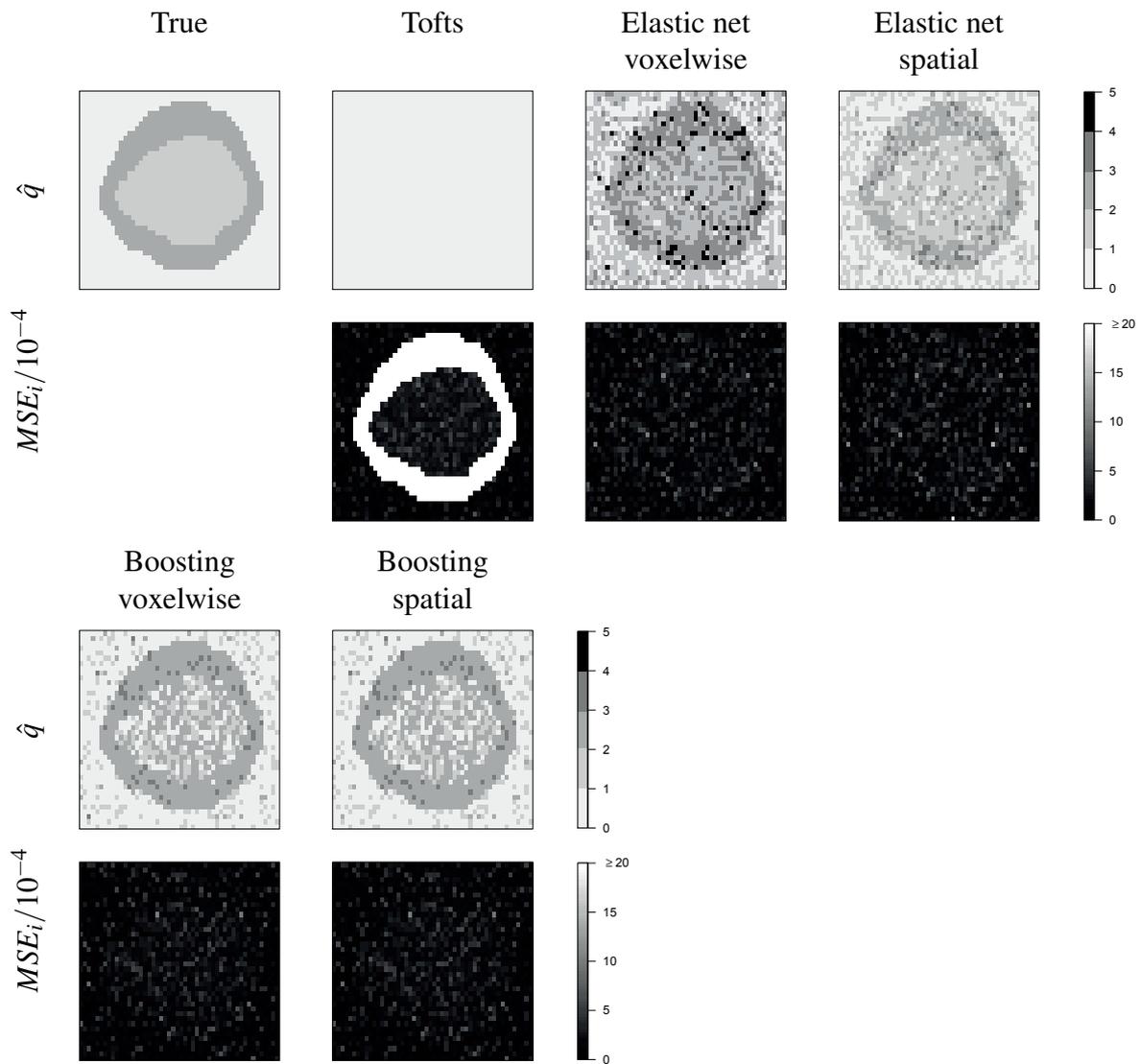


Figure C.23: Results of Simulation 3 for the fourth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

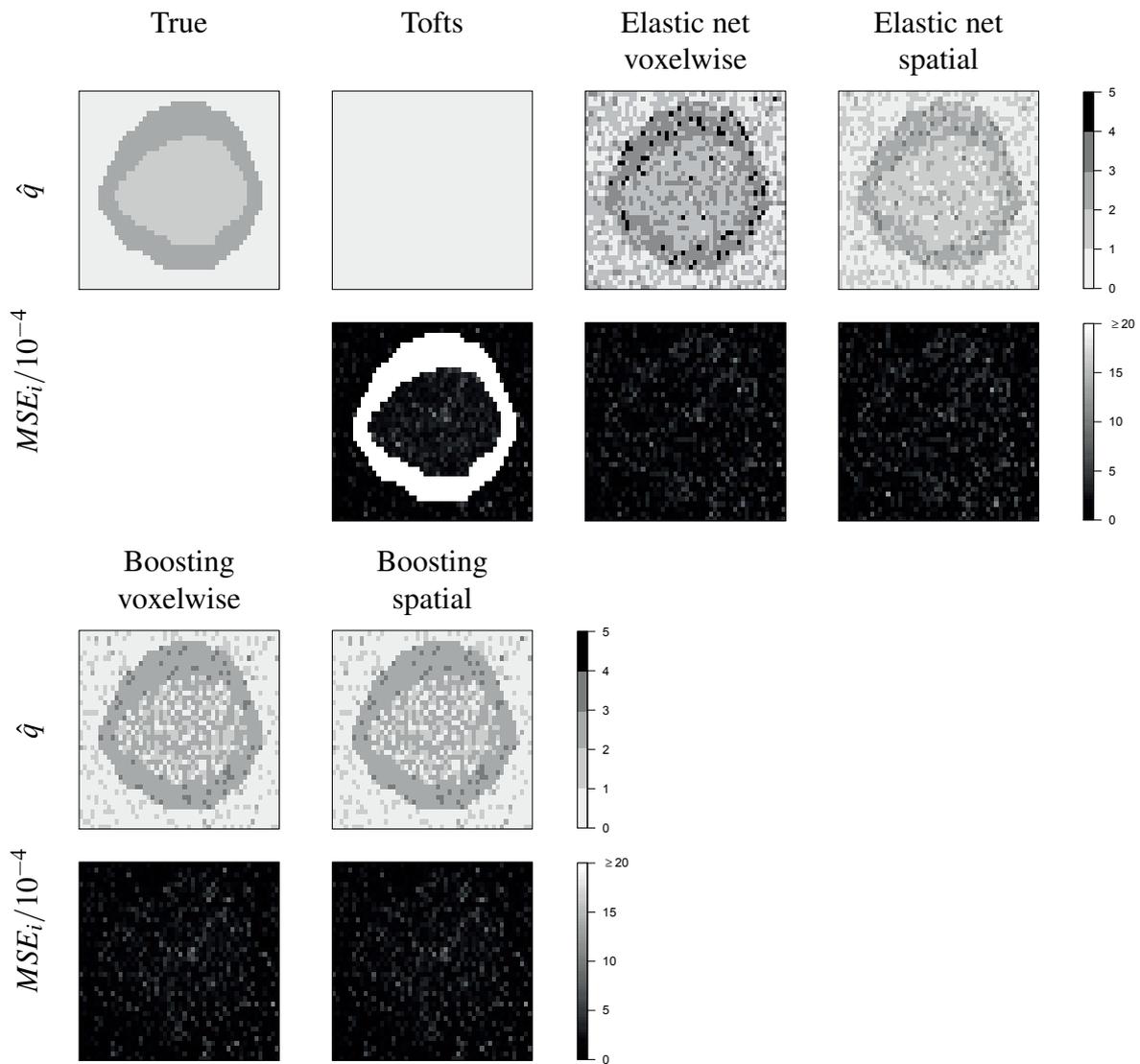


Figure C.24: Results of Simulation 3 for the fifth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

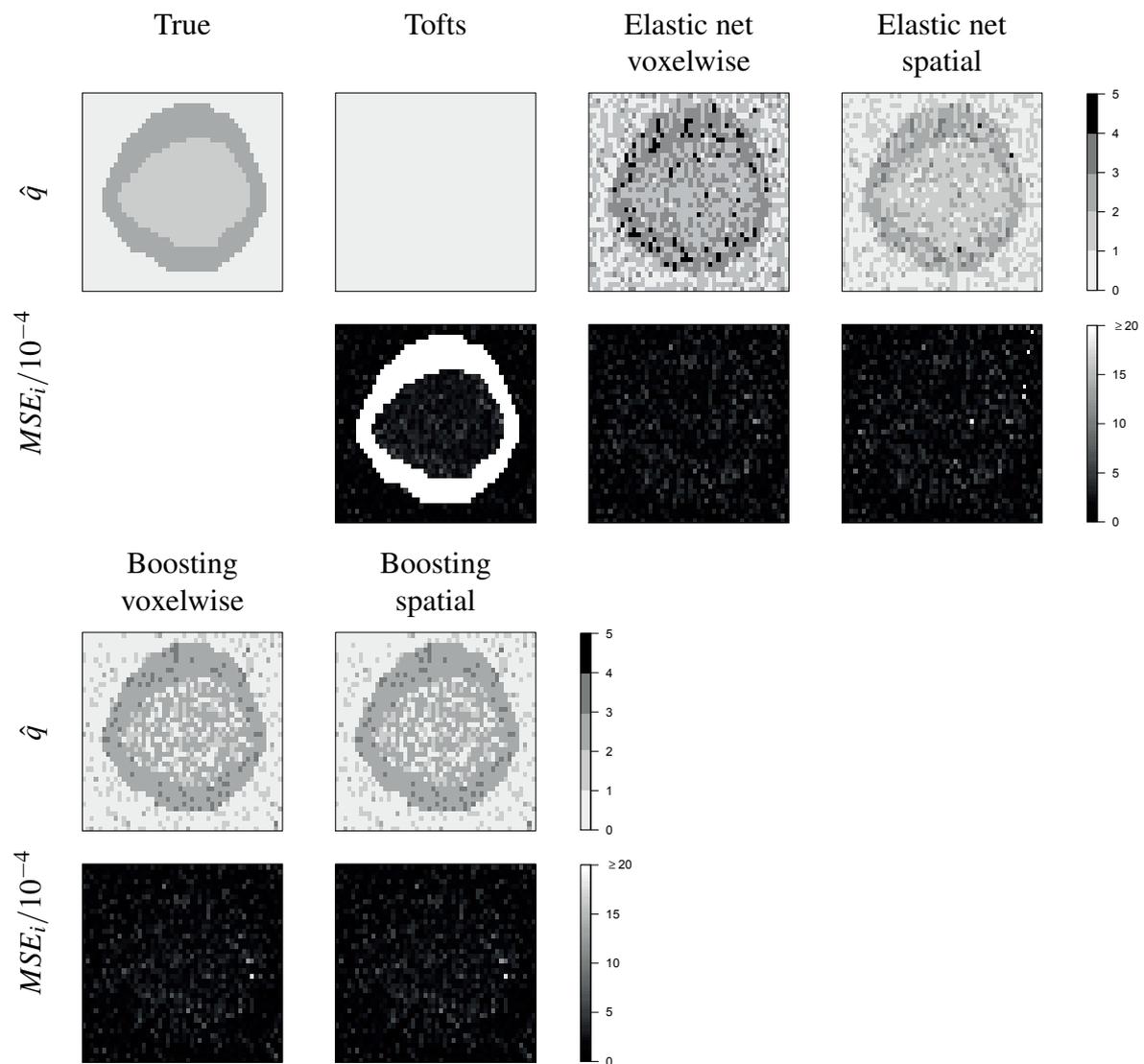


Figure C.25: Results of Simulation 3 for the sixth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

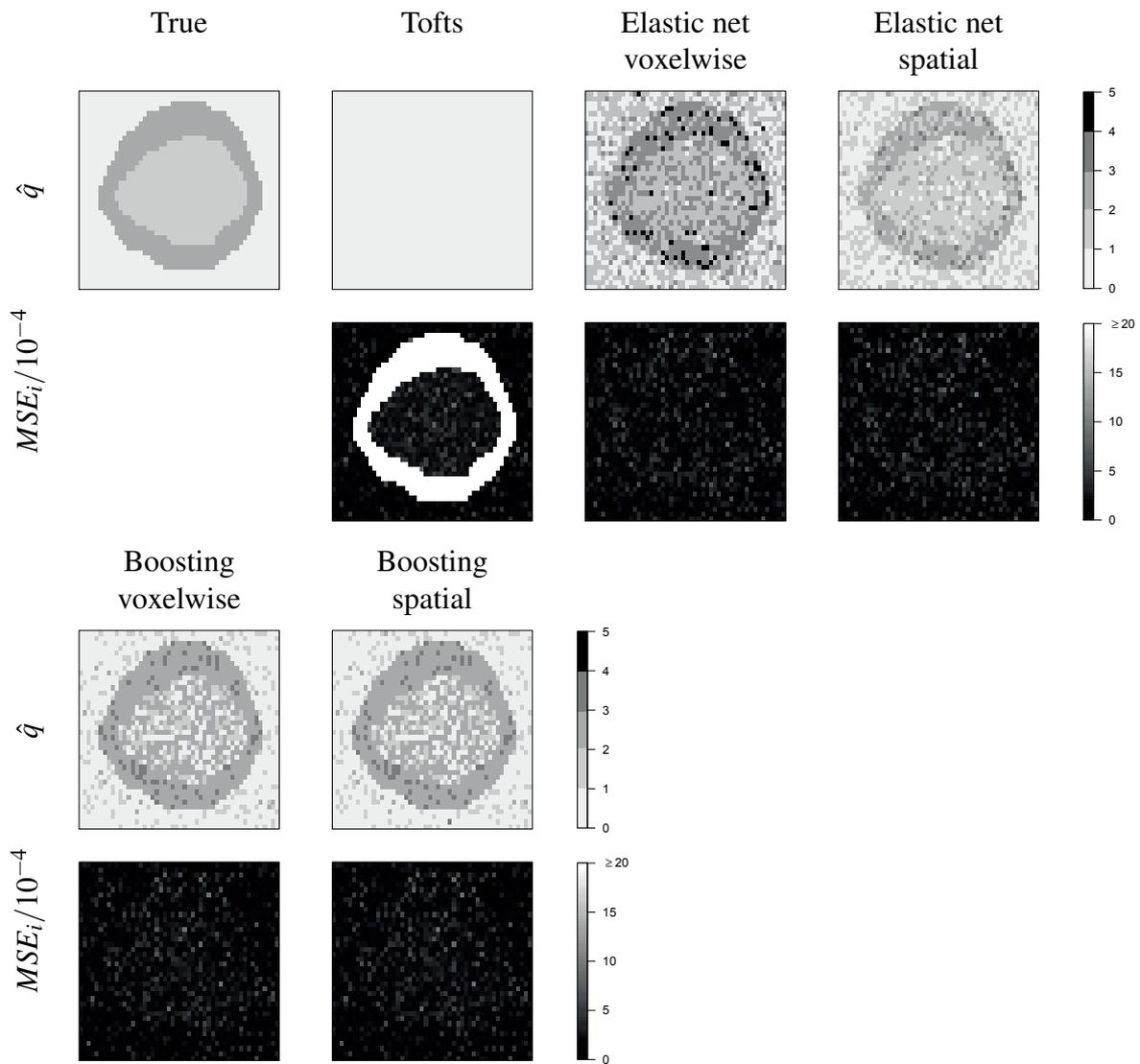


Figure C.26: Results of Simulation 3 for the seventh of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

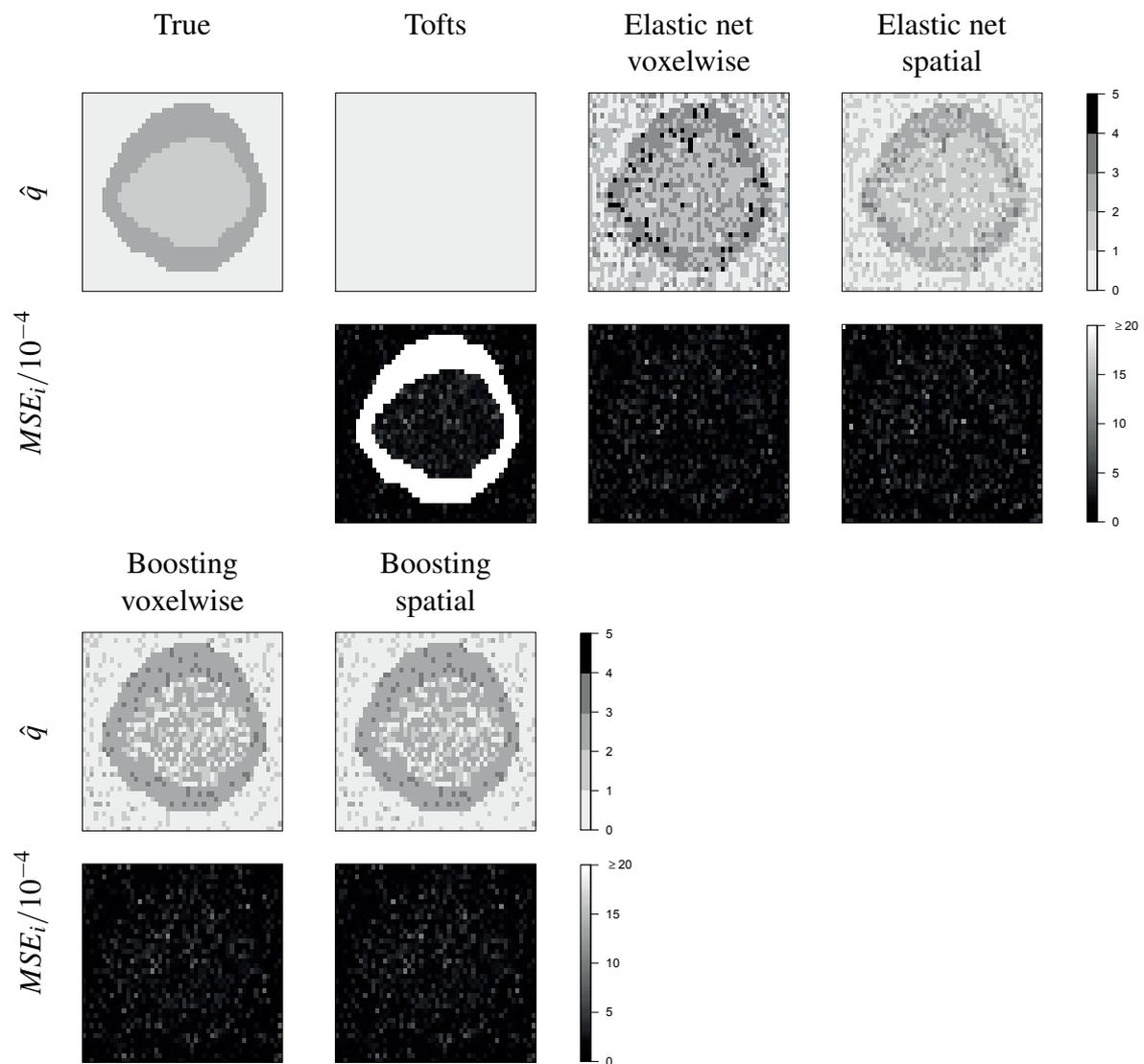


Figure C.27: Results of Simulation 3 for the eighth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

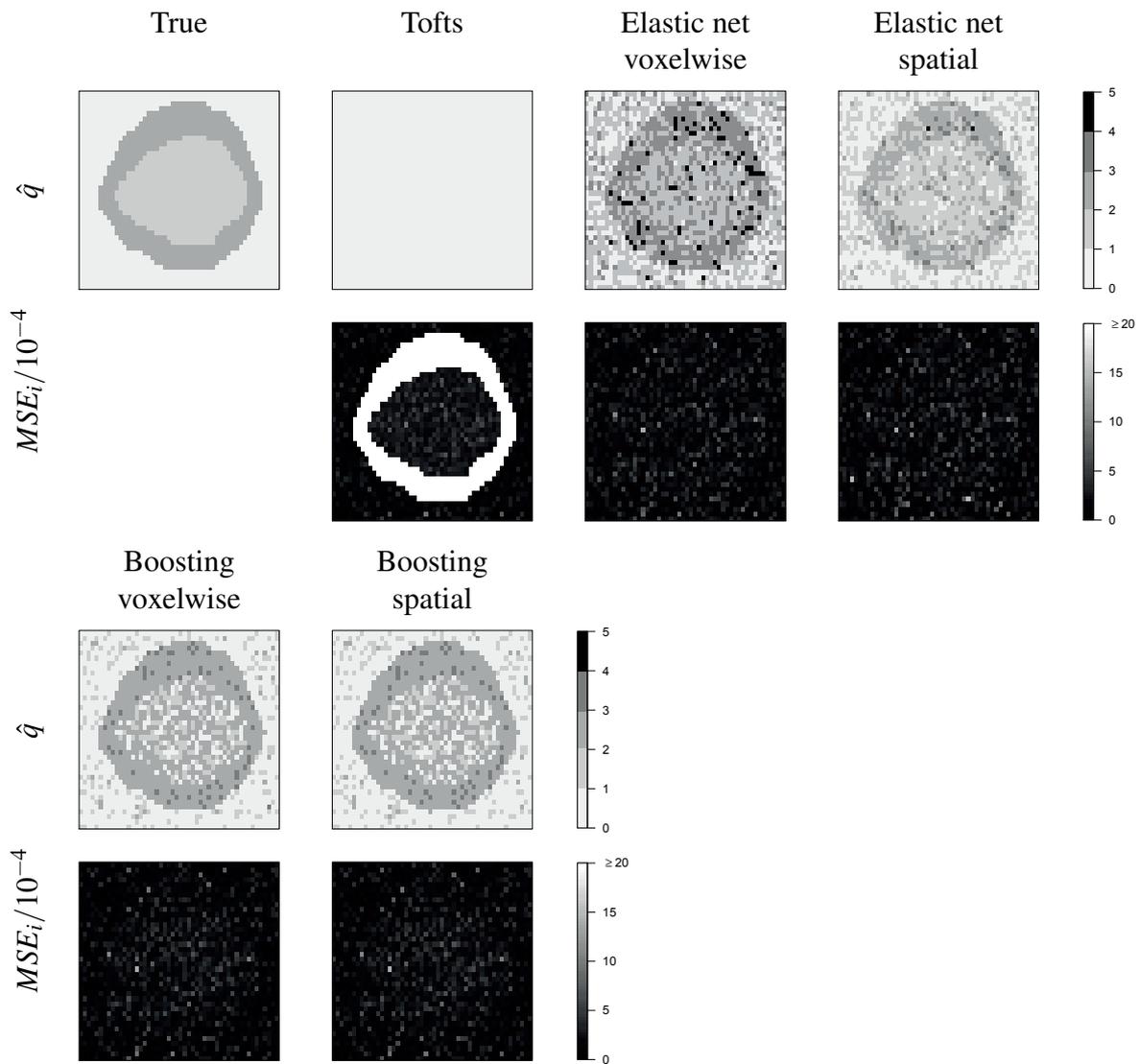


Figure C.28: Results of Simulation 3 for the ninth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

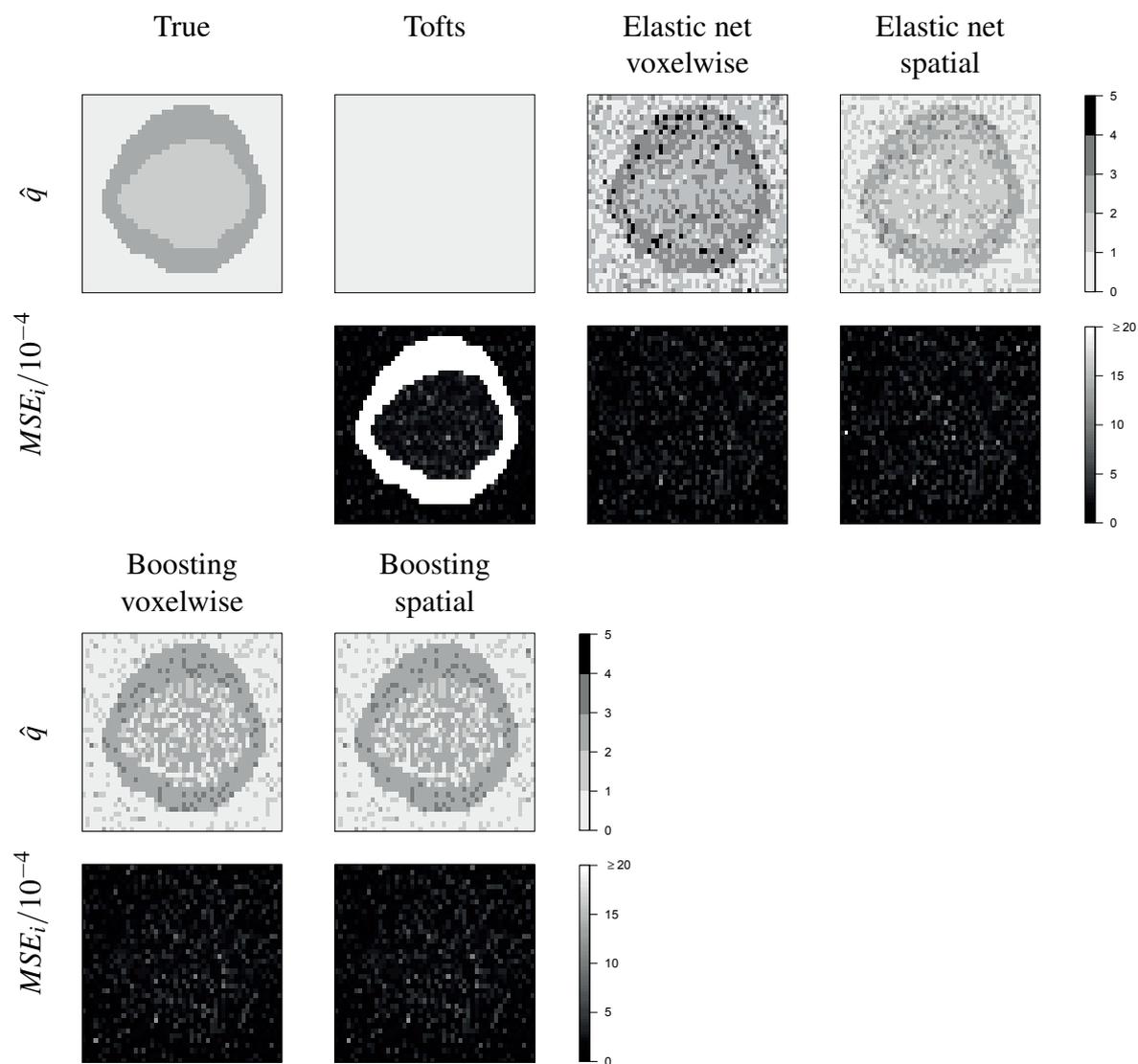


Figure C.29: Results of Simulation 3 for the tenth of the 10 simulated images: Estimated number of tissue compartments  $\hat{q}$  and MSE for each voxel of the image.

# Bibliography

- Ah-See, M.-L., A. Makris, N. Taylor, R. Burcombe, M. Harrison, J. Stirling, P. Richman, M. Leach, and A. Padhani (2004). Does vascular imaging with MRI predict response to neoadjuvant chemotherapy in primary breast cancer? *Journal of Clinical Oncology (Meeting Abstracts)* 22(14S), 582.
- Ahearn, T., R. Staff, T. Redpath, and S. Semple (2005). The use of the Levenberg-Marquardt curve-fitting algorithm in pharmacokinetic modelling of DCE-MRI data. *Physics in Medicine & Biology* 50(9), N85–N92.
- Anderson, D. H. (1983). *Compartmental Modeling and Tracer Kinetics*. Berlin, Heidelberg, Germany, New York, USA, Tokyo, Japan: Springer-Verlag.
- Bates, D. and D. G. Watts (1988). *Nonlinear regression analysis and its applications*. New York, USA: Wiley.
- Beaudouin, J., F. Mora-Bermúdez, T. Klee, N. Daigle, and J. Ellenberg (2006). Dissecting the Contribution of Diffusion and Interactions to the Mobility of Nuclear Proteins. *Biophysical Journal* 90, 1878–1894.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development* 16(1), 6–21.
- Bischl, B., M. Lang, O. Mersmann, J. Rahnenfuehrer, and C. Weihs (2014). BatchJobs and BatchExperiments: Abstraction mechanisms for using R in batch environments. *Journal of Statistical Software*. Accepted for publication.
- Brix, G., S. Zwick, F. Kiessling, and J. Griebel (2009). Pharmacokinetic analysis of tissue micro-circulation using nested models: multimodel inference and parameter identifiability. *Medical Physics* 36(7), 2923–2933.
- Brooks, S., A. Gelman, G. L. Jones, and X.-L. Meng (2011). *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC.
- Brooks, S. P. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7, 434–455.
- Brown, H. and R. Prescott (1999). *Applied Mixed Models in Medicine*. Chichester, UK: Wiley.

- Buckley, D. and G. Parker (2005). Measuring Contrast Agent Concentration in  $T_1$ -Weighted Dynamic Contrast-Enhanced MRI. In A. Jackson, D. Buckley, and G. Parker (Eds.), *Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology*, Chapter 5, pp. 69–79. Springer.
- Bühlmann, P. (2006). Boosting for high-dimensional linear models. *The Annals of Statistics* 34, 559–583.
- Bühlmann, P. and T. Hothorn (2007). Boosting Algorithms: Regularization, Prediction and Model Fitting. *Statistical Science* 22, 477–522.
- Bühlmann, P. and B. Yu (2003). Boosting with the  $L_2$ -Loss: Regression and Classification. *Journal of the American Statistical Association* 98, 324–338.
- Bulinski, J. C., D. J. Odde, B. J. Howell, T. D. Salmon, and C. M. Waterman-Storer (2001). Rapid dynamics of the microtubule binding of ensconsin in vivo. *Journal of Cell Science* 114, 3885–3897.
- Candes, E. and T. Tao (2007). The Dantzig selector: Statistical estimation when  $p$  is much larger than  $n$ . *The Annals of Statistics* 35(6), 2313–2351.
- Carlin, B. P. and T. A. Louis (2009). *Bayesian Methods for Data Analysis* (3rd ed.). Boca Raton, USA: Chapman & Hall/CRC.
- Carrero, G., E. Crawford, M. J. Hendzel, and G. de Vries (2004). Characterizing Fluorescence Recovery Curves for Nuclear Proteins Undergoing Binding Events. *Bulletin of Mathematical Biology* 66, 1515–1545.
- Coscoy, S., F. Waharte, A. Gautreau, M. Martin, D. Louvard, P. Mangeat, M. Arpin, and F. Amblard (2002). Molecular analysis of microscopic ezrin dynamics by two-photon FRAP. *Proceedings of the National Academy of Sciences of the United States of America* 99(20), 12813–12818.
- Dargatz, C. (2010). *Bayesian Inference for Diffusion Processes with Applications in Life Sciences*. Dissertation at the Faculty of Mathematics, Computer Science and Statistics of the LMU Munich. Munich, Germany.
- Dundr, M., U. Hoffmann-Rohrer, Q. Hu, I. Grummt, L. I. Rothblum, R. D. Phair, and T. Misteli (2002). A kinetic framework for a mammalian RNA polymerase in vivo. *Science* 298, 1623–1626.
- Elzhov, T., K. Mullen, A.-N. Spiess, and B. Bolker (2013). *minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds*. R package version 1.1-8.
- Eriksson, E. (1971). Compartment Models and Reservoir Theory. *Annual Review of Ecology and Systematics* 2, 67–84.

- Fahrmeir, L., T. Kneib, S. Lang, and B. Marx (2013). *Regression - Models, Methods and Applications*. Berlin, Heidelberg, Germany: Springer Berlin Heidelberg.
- Feilke, M., B. Bischl, V. J. Schmid, and J. Gertheiss (2015). Boosting in nonlinear regression models with an application to DCE-MRI data. *Methods of Information in Medicine*. Accepted for publication.
- Feilke, M., K. Schneider, and V. J. Schmid (2015). Bayesian mixed-effects model for the analysis of a series of FRAP images. *Statistical Applications in Genetics and Molecular Biology* 14(1), 35–51.
- Freund, Y. and R. Shapire (1996). Experiments with a New Boosting Algorithm. In *Proceedings of the Thirteenth International Conference on Machine Learning Theory*, Number 3, San Francisco, CA, pp. 148–156. Morgan Kaufmann Publishers Inc.
- Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics* 29, 1189–1232.
- Friedman, J., T. Hastie, and R. Tibshirani (2000). Additive logistic regression: a statistical view of boosting (with Discussion). *The Annals of Statistics* 28, 337–407.
- Fuchs, C. (2013). *Inference for Diffusion Processes with Applications in Life Sciences*. Heidelberg, Germany: Springer.
- Gelman, A., G. O. Roberts, and W. R. Gilks (1996). *Bayesian statistics 5*. Oxford, UK: Oxford University Press.
- Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–511.
- Gertheiss, J. and G. Tutz (2010). Sparse modeling of categorical explanatory variables. *The Annals of Applied Statistics* 4(4), 2150–2180.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996). *Markov chain Monte Carlo in practice* (1st ed.). London, UK: Chapman & Hall.
- Groll, A. (2013). *GMMBoost: Likelihood-based Boosting for Generalized mixed models*. R package version 1.1.2.
- Haario, H., E. Saksman, and J. Tamminen (2001). An adaptive Metropolis algorithm. *Bernoulli* 7, 223–242.
- Hansen, M. and B. Yu (2001). Model Selection and the Principle of Minimum Description Length. *Journal of the American Statistical Association* 96(454), 746–774.
- Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning* (2nd ed.). New York: Springer.

- Hemmerich, P., L. Schmiedeberg, and S. Diekmann (2011). Dynamic as well as stable protein interactions contribute to genome function and maintenance. *Chromosome Research* 19, 131–151.
- Herbst, P. G. (1963). Organizational Commitment: A Decision Process Model. *Acta Sociologica* 7(1), 34–46.
- Heuser, H. (1995). *Gewöhnliche Differentialgleichungen* (3rd ed.). Stuttgart, Germany: B.G. Teubner.
- Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid (2014). Model-based boosting in R: a hands-on tutorial using the R package mboost. *Computational Statistics* 29, 3–35.
- Jacquez, J. A. (1972). *Compartmental analysis in biology and medicine*. Amsterdam, Netherlands: Elsevier.
- Legrand, J., A. Sanchez, F. Le Pont, L. Camacho, and B. Larouze (2008). Modeling the Impact of Tuberculosis Control Strategies in Highly Endemic Overcrowded Prisons. *PLoS ONE* 3(5), e2100.
- Lipsitch, M., C. T. Bergstrom, and B. R. Levin (2000). The epidemiology of antibiotic resistance in hospitals: Paradoxes and prescriptions. *Proceedings of the National Academy of Sciences of the United States of America* 97(4), 1938–1943.
- Liu, G., H. S. Rugo, G. Wilding, T. M. McShane, J. L. Evelhoch, C. Ng, E. Jackson, F. Kelcz, B. M. Yeh, F. T. J. Lee, C. Charnsangavej, J. W. Park, E. A. Ashton, H. M. Steinfeldt, Y. K. Pithavala, S. D. Reich, and R. S. Herbst (2005). Dynamic contrast-enhanced magnetic resonance imaging as a pharmacodynamic measure of response after acute dosing of AG-013736, an oral angiogenesis inhibitor, in patients with advanced solid tumors: results from a phase I study. *J Clin Oncol* 23(24), 5464–5473.
- Mayr, A., H. Binder, O. Gefeller, and M. Schmid (2014a). Extending statistical boosting. An overview of recent methodological developments. *Methods of Information in Medicine* 53, 428–435.
- Mayr, A., H. Binder, O. Gefeller, and M. Schmid (2014b). The evolution of boosting algorithms. From machine learning to statistical modelling. *Methods of Information in Medicine* 53, 419–427.
- Mayr, A., B. Hofner, and M. Schmid (2012). The importance of knowing when to stop. A sequential stopping rule for component-wise gradient boosting. *Methods of Information in Medicine* 51(2), 178–186.
- Mazza, D., A. Abernathy, N. Golob, T. Morisaki, and J. G. McNally (2012). A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Research* 40(15), e119.

- McNally, J. G. (2008). Quantitative FRAP in Analysis of Molecular Binding Dynamics In Vivo. *Methods in cell biology* 85, 329–351.
- Mebane, Jr., W. R. and S. S. Jasjeet (2011). Genetic Optimization Using Derivatives: The rge-noud Package for R. *Journal of Statistical Software* 42(11), 1–26.
- Meyvis, T. K. L., S. C. De Smedt, P. van Oostveldt, and J. Demeester (1999). Fluorescence Recovery After Photobleaching: A Versatile Tool for Mobility and Interaction Measurements in Pharmaceutical Research. *Pharmaceutical Research* 16(8), 1153–1162.
- Mueller, F., D. Mazza, T. J. Stasevich, and J. G. McNally (2010). FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know? *Current Opinion in Cell Biology* 22, 403–411.
- Mueller, F., P. Wach, and J. G. McNally (2008). Evidence for a Common Mode of Transcription Factor Interaction with Chromatin as Revealed by Improved Quantitative Fluorescence Recovery after Photobleaching. *Biophysical Journal* 94(8), 3323–3339.
- Murase, K. (2004). Efficient Method for Calculating Kinetic Parameters Using  $T_1$ -Weighted Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* 51, 858–862.
- Padhani, A. R., M.-L. Ah-See, and A. Makris (2005). MRI in the detection and management of breast cancer. *Expert Reviews in Anticancer Therapy* 5(2), 239–252.
- Parker, G., J. Suckling, S. Tanner, A. Padhani, J. Husband, and M. Leach (1998). MRIW: Parametric analysis software for contrast-enhanced dynamic MR imaging in cancer. *Radiographics*, 18(2), 497–506.
- Phair, R. D., P. Scaffidi, C. Elbi, J. Vecerová, A. Dey, K. Ozato, D. T. Brown, G. Hager, M. Bustin, and T. Misteli (2004). Global Nature of Dynamic Protein-Chromatin Interactions In Vivo: Three-Dimensional Genome Scanning and Dynamic Interaction Networks of Chromatin Proteins. *Molecular and Cellular Biology* 24(14), 6393–6402.
- Pinheiro, J. and D. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. New York, USA: Springer Verlag.
- Plummer, M., N. Best, K. Cowles, and K. Vines (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* 6(1), 7–11.
- Port, R., M. Knopp, U. Hoffmann, S. Milker-Zabel, and G. Brix (1999). Multicompartment analysis of gadolinium chelate kinetics: blood-tissue exchange in mammary tumors as monitored by dynamic MR imaging. *Journal of Magnetic Resonance Imaging* 10, 233–241.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

- Ratkowsky, D. A. (1990). *Handbook of nonlinear regression models*. New York, USA: Dekker.
- Reich, J. G. (1981). *On parameter redundancy in curve fitting of kinetic data*. Kinetic data analysis: Design and analysis of enzyme and pharmacokinetic experiments. New York: Plenum Press.
- Reits, E. A. and J. J. Neefjes (2001). From fixed to FRAP: measuring protein mobility and activity in living cells. *Nature Cell Biology* 3(6), E145–E147.
- Roberts, G. O., A. Gelman, and W. R. Gilks (1994). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Research Report 94.16, Statistical Laboratory, University of Cambridge*.
- Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). *Journal of the Royal Statistical Society. Series B* 71, 319392.
- Savic, R. M., D. M. Jonker, T. Kerbusch, and M. O. Karlsson (2007). Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies. *Journal of Pharmacokinetics and Pharmacodynamics* 34(5), 711–726.
- Schermelleh, L., A. Haemmer, F. Spada, N. Rösing, D. Meilinger, U. Rothbauer, M. C. Cardoso, and H. Leonhardt (2007). Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Research* 35(13), 4301–4312.
- Schmid, M. and T. Hothorn (2008). Boosting additive models using component-wise P-splines. *Computational Statistics & Data Analysis* 53, 298–311.
- Schmid, V. (2004). *Bayesianische Raum-Zeit-Modellierung in der Epidemiologie*. Dissertation at the Faculty of Mathematics, Computer Science and Statistics of the LMU Munich. Munich, Germany: Verlag Dr. Hut.
- Schmid, V., B. Whitcher, A. Padhani, N. Taylor, and G.-Z. Yang (2009). A Bayesian hierarchical model for the analysis of a longitudinal dynamic contrast-enhanced MRI oncology study. *Magnetic Resonance in Medicine* 61(1), 163–174.
- Schmid, V., B. Whitcher, A. Padhani, and G. Yang (2009). Quantitative analysis of Dynamic contrast-enhanced MR images based on Bayesian P-Splines. *IEEE Transactions on Medical Imaging* 28(6), 789–798.
- Schmid, V., B. Whitcher, G. Yang, N. Taylor, and A. Padhani (2005). *Medical Imaging Computing and Computer-Assisted Intervention – MICCAI 2005*, Chapter Statistical analysis of pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging, pp. 886–893. New York: Springer. J. Duncan and G. Gerig, Eds.

- Schmid, V. J., B. Whitcher, A. R. Padhani, N. J. Taylor, and G.-Z. Yang (2006). Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging. *IEEE Transactions on Medical Imaging* 25(12), 1627–1636.
- Schneider, K. (2009). *Analysis of cell cycle dependent kinetics of Dnmt1 by FRAP and kinetic modeling*. Diploma Thesis at the Faculty of Biology of the LMU Munich. Munich, Germany.
- Schneider, K., C. Fuchs, A. Dobay, A. Rottach, W. Qin, P. Wolf, J. M. Álvarez Castro, M. M. Nalaskowski, E. Kremmer, V. Schmid, H. Leonhardt, and L. Schermelleh (2013). Dissection of cell cycle-dependent dynamics of Dnmt1 by FRAP and diffusion-coupled modeling. *Nucleic Acids Research* 41(9), 4860–4876.
- Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. New York: Wiley.
- Sommer, J. (2013). *Regularized estimation and model selection in compartment models*. Dissertation at the Faculty of Mathematics, Computer Science and Statistics of the LMU Munich. Munich, Germany: Verlag Dr. Hut.
- Sommer, J., J. Gertheiss, and V. Schmid (2014). Spatially regularized estimation for the analysis of dynamic contrast-enhanced magnetic resonance imaging data. *Statistics in Medicine* 33, 1029–1041.
- Sommer, J. and V. Schmid (2014). Spatial two-tissue compartment model for dynamic contrast-enhanced magnetic resonance imaging. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 63, 695–713.
- Sourbron, S. and D. Buckley (2011). On the scope and interpretation of the Tofts models for DCE-MRI. *Magnetic Resonance in Medicine* 66(3), 735–745.
- Spada, F., U. Rothbauer, K. Zolghadr, L. Schermelleh, and H. Leonhardt (2006). Regulation of DNA methyltransferase 1. *Advances in Enzyme Regulation* 46, 224–234.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 64(4), 583–639.
- Sporbert, A., P. Domaing, M. C. Cardoso, and H. Leonhardt (2005). PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. *Nucleic Acids Research* 33, 3521–3528.
- Sprague, B. L. and J. G. McNally (2005). FRAP analysis of binding: proper and fitting. *TRENDS in Cell Biology* 15(2), 84–91.
- Sprague, B. L., R. L. Pego, D. A. Stavreva, and J. G. McNally (2004). Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching. *Biophysical Journal* 86, 3473–3495.

- Tierney, L. and J. B. Kadane (1986). Accurate Approximations for Posterior Moments and Marginal Densities. *Journal of the American Statistical Association* 81(393), 82–86.
- Tofts, P. (1997). Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of Magnetic Resonance Imaging* 7(1), 91–101.
- Tofts, P., G. Brix, D. Buckley, J. Evelhoch, E. Henderson, M. Knopp, H. Larsson, T.-Y. Lee, N. Mayr, G. Parker, R. Port, J. Taylor, and R. Weiskoff (1999). Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: Standardized quantities and symbols. *Journal of Magnetic Resonance Imaging* 10, 223–232.
- Tofts, P. and A. Kermode (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging - 1. Fundamental concepts. *Magnetic Resonance in Medicine* 17, 357–367.
- Tsoukias, N. M. and S. C. George (1998). A two-compartment model of pulmonary nitric oxide exchange dynamics. *Journal of Applied Physiology* 85(2), 653–666.
- Turlach, B. A. and A. Weingessel (2013). *quadprog: Functions to solve Quadratic Programming Problems*. R package version 1.5-5.
- Tutz, G. and A. Groll (2013). Likelihood-based boosting in binary and ordinal random effects models. *Journal of Computational and Graphical Statistics* 22(2), 356–378.
- Urbanek, S. (2014). *multicore: A stub package to ease transition to 'parallel'*. R package version 0.2.
- van Royen, M. E., P. Farla, K. A. Mattern, B. Geverts, J. Trapman, and A. B. Houtsmuller (2009). Fluorescence Recovery After Photobleaching (FRAP) to Study Nuclear Protein Dynamics in Living Cells. *Methods in Molecular Biology* 464, 363–385.
- van Royen, M. E., A. Zotter, S. M. Ibrahim, B. Geverts, and A. B. Houtsmuller (2011). Nuclear proteins: finding and binding target sites in chromatin. *Chromosome Research* 19, 83–98.
- Weidemann, F., M. Dehnert, J. Koch, O. Wichmann, and M. Höhle (2014). Modelling the epidemiological impact of rotavirus vaccination in Germany-A Bayesian approach. *Vaccine* 32, 5250–5257.
- Weinmann, H., M. Laniado, and W. Mützel (1984). Pharmacokinetics of Gd-DTPA/dimeglumine after intravenous injection into healthy volunteers. *Physiological Chemistry and Physics and Medical NMR* 16, 167–172.
- Yang, X. and M. V. Kopp (2011). Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging: a review. *Journal of biomedicine & biotechnology* 2011, 1–12.

Yu, L. X. and G. L. Amidon (1999). A compartmental absorption and transit model for estimating oral drug absorption. *International Journal of Pharmaceutics* 186(2), 119–125.

Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society B* 67, 301–320.



# Eidesstattliche Versicherung

gemäß §8 Abs. 2 Pkt. 5 der Promotionsordnung vom 12.07.2011

Hiermit erkläre ich an Eides statt, dass die Dissertation von mir selbstständig,  
ohne unerlaubte Beihilfe angefertigt ist.

München, den 11. März 2015

---

Martina Feilke



