SimTech

UNIVERSITY OF BERGEN
*Faculty of Mathematics and Natural Sciences*

Universität Stuttgart

INSTITUTE FOR MODELLING HYDRAULIC AND ENVIRONMENTAL SYSTEMS (IWS)

SIMULATION TECHNOLOGY DEGREE COURSE

Bachelor thesis

# Quantitative analysis for FluidFlower experiments using multichromatic tracer data

First Supervisor

Prof. Dr.-Ing. Dr.-Ing. h.c. Rainer HELMIG

Institute for Modelling Hydraulic and Environmental Systems
Universität Stuttgart

Second Supervisor

Dr. Jakub Wiktor BOTH

Department of Mathematics – Porous Media Group
Universitet i Bergen

Submitted by

| | |
|---|---|
| Author | Moritz MARQUARDT |
| Matriculation number | 3415391 |
| SimTech-Nr. | 168 |
| Submission date | July 2023 |

I hereby certify that I have prepared this thesis independently, and that only those sources, aids, and advisors that are duly noted herein have been used and/or consulted. I further agree that this thesis is made accessible for scientific purposes by the library of the Institute for Modelling Hydraulic and Environmental Systems of the University of Stuttgart (Publication according to §6 Abs. 1 UrhG). I agree that contents may be cited according to §52 UrhG.

Bergen, 26.06.2023

Moritz Marquardt

Hiermit erkläre ich, dass die vorliegende Arbeit von mir selbst und ohne fremde Hilfe, lediglich unter Benutzung der hier aufgeführten Literatur, angefertigt worden ist. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt. Ich stimme außerdem zu, dass die vorliegende Arbeit zu wissenschaftlichen Zwecken in den Bibliotheken des Instituts für Wasser- und Umweltsystemmodellierung der Universität Stuttgart aufgestellt und zugänglich gemacht wird (Veröffentlichung nach §6 Abs. 1 UrhG) und hieraus im Rahmen des §52 UrhG zitiert werden kann.

Bergen, 26.06.2023

Moritz Marquardt

# Kurzzusammenfassung

Im Rahmen der geologischen $CO_2$ Speicherung (GCS) stellt sich die Schlüsselfrage wie zuverlässig $CO_2$ im Boden verbleibt. Dazu werden Labor-Experimente an der Universität Bergen durchgeführt mit dem Ziel Simulationsmodelle zu entwickeln und verbessern und um gleichzeitig auf visuelle Art und Weise zu vermitteln, welche physikalischen Prozesse ein Teil dieser sind. Hierfür werden Tracer als Farbgeber genutzt, die im Kontakt durch Farbveränderungen die Entwicklung der $CO_2$-Konzentration im porösen Material anzeigen. Die zur Farbauswertung in Bergen entwickelte Bildanalyse-Software DarSIA wertete zunächst die Bilder des Experiments nur auf Basis von monochromatischen Informationen durch die Reduktion auf ein skalares Signal auf Pixelebene aus. Weil multichromatische Tracer jedoch ein besseres Abbild der dargestellten Konzentrationen und damit einen höheren Informationsgehalt ermöglichen, sollte die Erweiterung für eine multichromatische Auswertung der Bilder geprüft werden. Als Lösung zur Berechnung der dargestellten Konzentration wird daher in dieser Arbeit unter anderem eine Interpolation von definierten Farb-Konzentrationspunkten vorgeschlagen, die auf alle Farbpixel angewandt werden kann. Als Interpolationsmethode wird neben der Nearest-Neighbour Methode auch die Kernel Methode geprüft. Dabei hat sich die Kernel Methode im Ergebnis als vielversprechender Ansatz zur Berechnung der Konzentration von multichromatischen Tracern für die Integration in DarSIA erwiesen.

# Abstract

In the context of geological $CO_2$ storage (GCS), the key question is how reliably $CO_2$ remains in the underground. Laboratory experiments are being conducted at the University of Bergen with the aim of developing and improving simulation models and, at the same time, communicating in a visual way which physical processes are part of them. For this purpose, tracers are used as colour sensors, which allow to visualise the development of the $CO_2$ concentration in the porous material through colour changes on contact. Up until now, the image analysis software DarSIA, which has been developed in Bergen for colour-based analysis, evaluates the images of the experiment only based on monochromatic information by reducing them to a scalar signal at pixel level. However, because multichromatic tracers provide a better representation of the represented concentrations and thus a higher information content, the extension for an image analysis based on multichromatic information should be explored. As a solution for the determination of the represented concentration, this work therefore proposes, among other approaches, an interpolation of defined colour concentration pairs, which can be applied to all colour pixels. In addition to the nearest neighbour method, the kernel method is also tested as an interpolation method. As a result, the kernel method was identified to be a promising approach to determine the concentration of multichromatic tracers for an integration in DarSIA.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

Climate change is one of the most important challenges of our time. The global long-term shift of temperatures in the last two centuries is mainly caused by human activities [1] through the combustion of fossil fuels like coal, gas, and oil. These processes emit greenhouse gasses like $CO_2$ which, when released to the atmosphere, drive the climate change in a process called the greenhouse effect [2]. Currently the global $CO_2$ emissions and the concentration of $CO_2$ in the atmosphere are still increasing [2].

## 1.1 Motivation: Successful carbon capture and storage

To mitigate and turn around the climate change, the emissions and concentration of $CO_2$ in the atmosphere must be reduced. One of the major courses of actions to achieve this goal is Carbon Capture and Storage (CCS) according to most of the IPCC forecasting scenarios [2]. On the storage part, a promising and currently popular way is the injection and sequestration of $CO_2$ in geological subsurface reservoirs because of its maturity and large potential for scalability [3]. The first large scale $CO_2$ storage project is the Sleipner project which started in the 90s off the coast of Norway. A current CCS project in Norway is the Northern Lights project which aims to establish transport chains and the industrial sequestration of $CO_2$.

Subsurface reservoirs are porous and allow fluids like water and $CO_2$ to flow through the pores of permeable rocks, which makes it possible to inject and store $CO_2$. The modelling of $CO_2$ flow in the underground is important for the success of geological carbon storage (GCS). Simulations are especially important to estimate the long-term security (e.g., leak scenarios) of the sequestration and to optimise the storage site (e.g., well placement, injection strategies). Modelling and simulating those reservoirs are challenging tasks due to the large range of length and time scales. The reservoir can have a span of hundreds of kilometres in horizontal direction and the time horizon for $CO_2$ storage is in the order of thousands of years. At the same time, the porous media processes happen at the pore scale of the rocks. All things considered, in the end the success of $CO_2$ storage depends on the potential of long-term storage.

Because it is a relatively new field, long term data on the dynamics of $CO_2$ in the underground does not exist. This makes it difficult to calibrate, validate and improve the simulation models. To tackle this lack of information, research groups at the University of Bergen have introduced an experimental setup – the FluidFlower – to gather data on how $CO_2$ and other fluids behave in a lab setup [4]. The special feature of this experiment is the representability of real-world underground reservoirs at a smaller time and length scale and the repeatability of the experiments [5]. The concept of GCS and its representation in the FluidFlower is shown in Figure 1.1. High resolution photographs are used to capture the happenings and visual information in the FluidFlower. The gathered data can be used to test porous media models against it [6] and to gain insights into the happenings in the underground [4].

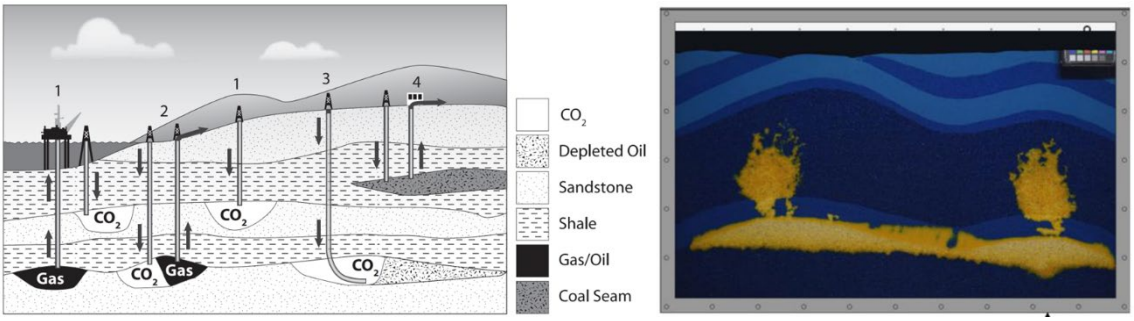


*Figure 1.1: CCS in the real world and in the lab; Left: Schematic of different kinds of GCS including pure $CO_2$ storage offshore and onshore and enhanced oil recovery (EOR) taken from [7]; Right: $CO_2$ injection in the medium sized FluidFlower geometry. The $CO_2$ is visible in yellow.*

The use of pH-sensitive dyes in the FluidFlower allows colour-based visualisation of the dynamics and flow of injected $CO_2$. This is based on the mechanism of $CO_2$ dissolving and – depending on its concentration – changing the pH-value of its solution to a non-neutral value which in turn changes the colour of the pH-sensitive dye in the solution. Colour and intensity of the pH-sensitive dye change depending on the concentration of dissolved $CO_2$. In more general terms, the pH-sensitive dye acts as a $CO_2$-concentration tracer. A tracer that changes its intensity based on one colour for different concentrations is called monochromatic throughout this paper. A tracer that additionally changes its colour depending on the concentration is called multichromatic. Examples of a monochromatic and two multichromatic tracers are shown in Figure 1.2.

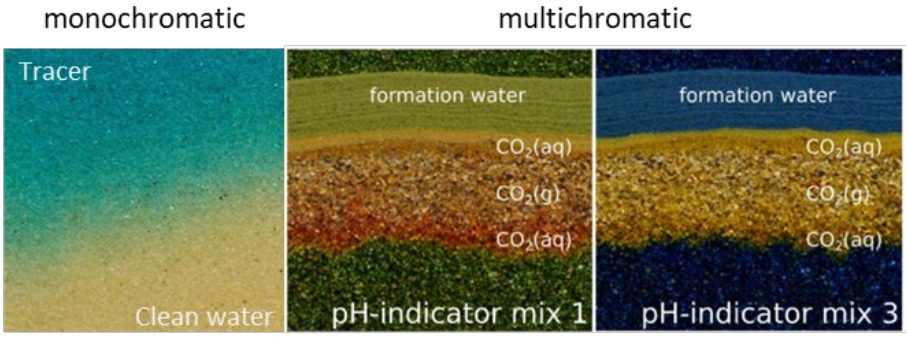


*Figure 1.2: Cropped region of the FluidFlower photographs with two examples of a multichromatic $CO_2$ concentration tracer and a monochromatic tracer. Left: Blue monochromatic tracer at the interface with the clean water; Right: Two multichromatic $CO_2$ tracers with two different pH-sensitive dyes. Formation water has a $CO_2$ concentration of 0. $CO_2$ (g) is $CO_2$ in its gaseous form and $CO_2$ (aq) is dissolved $CO_2$ indicated by the pH-sensitive dye. Taken from [8]*

Multichromatic tracers are preferably used since they potentially allow to visualise more information (concentration). In comparison to monochromatic tracers, multichromatic tracers have the potential to better represent and map the concentrations in different colours. This delivers more accurate results in a colour-based analysis.

The information visible in the FluidFlower photographs needs to be analysed quantitatively to make use of the gathered data. To analyse multichromatic tracer FluidFlower experiments quantitatively, the quantitative information as physical meaningful field variables (concentration) needs to be extracted from the information visible as tracer colours in the images.

## 1.2 Goal: Quantitative analysis of FluidFlower experiments

The task of obtaining information from an image is called image analysis. The porous media group at the Department of Mathematics at the University of Bergen has developed an image analysis toolbox called DarSIA (Darcy scale image analysis) [9]. The GitHub repository for DarSIA is publicly available [10]. DarSIA is a toolbox tailored to analysing and processing images with the goal of retrieving physical data from porous media experiments. Currently, DarSIA already contains a concentration analysis workflow which is capable of analysing experiments containing monochromatic information and extracting continuous concentration data. However, the analysis of multichromatic tracers is not yet supported.

Hence, the goal of this thesis is to develop and test a method for extracting pixel-wise concentration data from a two-dimensional photograph of a multichromatic tracer FluidFlower experiment. This general method can then be used to determine a concentration profile for the image.

## 1.3 Structure

Following this introduction, Chapter 2 contains a short summary of the Propaedeutic, which is attached as an appendix. The Propaedeutic covers relevant background information for this thesis, such as literature and theories. These fundamental studies are the starting point for this work.

Chapter 3 provides a description of a multichromatic tracer model experiment in the FluidFlower. This so-called model experiment will serve as the basis for the development and testing of the methods for a concentration analysis.

In Chapter 4, the image analysis toolbox DarSIA and its existing methods for a quantitative concentration analysis based on monochromatic pixel information will be discussed and adapted for the concentration analysis of multichromatic tracers.

In Chapter 5, a new approach is proposed to determine the tracer concentrations on pixel level of multichromatic tracer images using interpolation methods. Here, the kernel method will be suggested as an appropriate interpolation.

This thesis ends with a promising outlook in Chapter 6.

All code examples used in this thesis are publicly available in a GitHub repository [11].

# 2 Fundamental studies: Propaedeutic and Literature

The literature and theoretical background for this bachelor's thesis can be found in the Propaedeutic that is appended at the end of this document. The Propaedeutic precedes and lays the base for the studies in this thesis by covering and reviewing the current and relevant literature. This work is interdisciplinary, and the propaedeutic therefore covers a wide range of topics, which are divided into chapters 2-6. The following overview of the covered topics is taken directly from the introduction of the Propaedeutic [12]. It is recommended to read the Propaedeutic before reading this thesis. It will be cited and referenced throughout the thesis as [12]

The second chapter is about the carbon capture and storage (CCS) concept and will cover some motivational literature. This includes a short introduction in the storage aspect of CCS and specifically in the geological storage of carbons. This is the real-world application for the FluidFlower experiments and the studying of $CO_2$ dynamics in porous underground reservoirs.

The third part focuses on the FluidFlower concept. This is the experimental setup that is used as an example application for the DarSIA toolbox in this thesis. A description of the experimental setup with its limitations and properties and a motivation on how the experiments are linked to the real world and geological carbon sequestration is given.

The fourth part will cover the basic concepts of describing and modelling flow in porous media. This is important in understanding the physics in the porous structure of the FluidFlower and thus understanding the phenomena that are present in the experiment. This includes a discussion on scales of porous media and the upscaling of physics which will also lay the base and motivate concepts used in the DarSIA toolbox which will be presented in the fifth chapter.

The fifth chapter is about DarSIA (Darcy Scale Image Analysis), a toolbox for analysing and processing images with the goal of retrieving physical data from experimental images. Besides a brief introduction to DarSIA, central questions to be discussed in this chapter are: What information is visible in the pictures, and how to get from pore scale picture information to Darcy scale information by image processing.

In chapter 6, an introduction to kernel methods is given with the application of using it for multivariate interpolation. This will cover a mathematical definition of kernel methods and a reasoning on why it is useful and suitable for the application of sparse and multivariate interpolation.

# 3 FluidFlower: Multichromatic tracer model experiment

The methods discussed in this thesis are developed and tested using a model experiment designed to simplify the experimental setup while including all the important features necessary for a multichromatic analysis. In this work, this model experiment will be referred to as the model experiment. The model experiment is not a $CO_2$ injection experiment but a tracer injection experiment using a multichromatic tracer. This makes it possible to ignore the difficulties of $CO_2$ injection in the experimental set-up and the image analysis and concentrate on developing a method of analysing multichromatic tracers.

The model experiment for this work is conducted in a table-size FluidFlower rig (0.9m x 0.6m) in the lab of the Department of Physics and Technology at the University of Bergen [4]. The FluidFlower is a laboratory sized experiment consisting of a quasi-two-dimensional transparent container that can be filled with unconsolidated sands and fluids. The FluidFlower has several injection points in the back that make it possible to inject fluids to investigate multiphase flow in those porous media structures. The detailed setup of the FluidFlower and its properties can be found in chapter 3 of the Propaedeutic [12].

## 3.1 Simplifications

The first simplification of the model experiment is the sand geometry in the FluidFlower. The geometry consists of a single homogeneous sand as shown in Figure 3.1. This sand has the same porosity, permeability, pore size distribution and, most importantly, colour throughout the geometry. This greatly simplifies image analysis, allowing to concentrate on developing a methodology for the analysis of multichromatic tracers.

The second simplification concerns the fluids and their injection into the FluidFlower. Instead of conducting a $CO_2$ injection that is traced by multichromatic $CO_2$ tracers as described in the Introduction, a simpler multichromatic tracer injection is performed. The geometry is initially saturated with a neutral transparent fluid. The main idea of the experiment is to inject a multichromatic tracer at a constant rate into the porous geometry of the FluidFlower, creating a smooth front due to dispersion effects. This multichromatic tracer changes colour depending on its concentration in the initially neutral fluid. The temporal evolution of the injection and propagation of the injected plume is shown in Figure 3.1.

## 3.2 Data acquisition

A photo camera takes high resolution RGB photographs of the experiment (more can be found in chapter 5 of the Propaedeutic [12]) at discrete time-steps $t_i$ with $i = 0,1,2, \dots T$ which are equally distant. In Figure 3.1 the photograph of the experiment before the injection starts at timestep $t_0$ is displayed. This photograph will be called the baseline photograph and is denoted as $u_0 \coloneqq u_{t_0}$. Photographs taken after the injection start at time steps $t_i$ with $i \geq 1$ are called tracer photographs and will be denoted as $u_{t_i}$.

## 3.3 Conduction

The model experiment used in this thesis is shown in Figure 3.1. The tracer is green near the injection point, which is located in the centre of the green area. In this area the tracer has a high concentration close to 1. Closer to the interface with the neutral water, the tracer has a lower concentration and shows a blue colour.

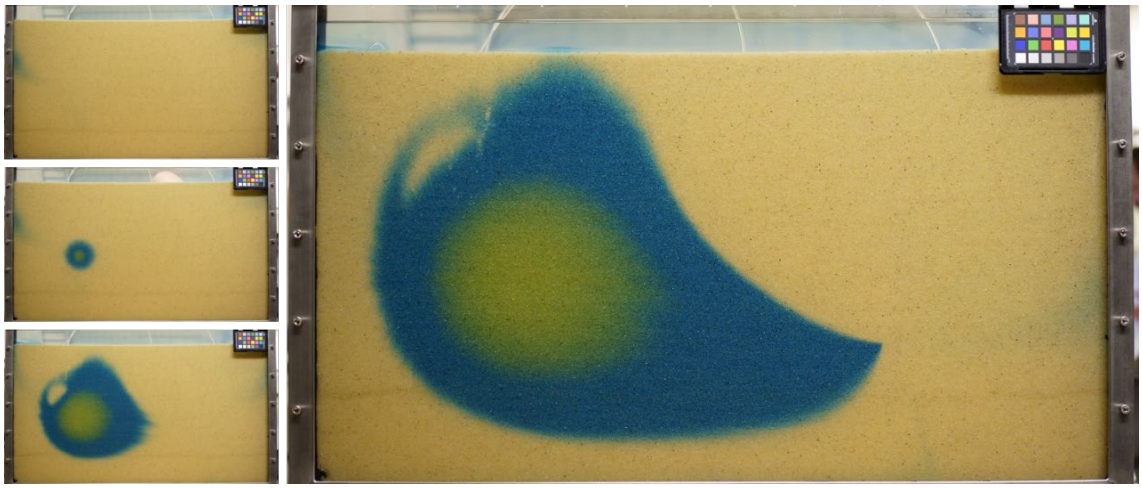

*Figure 3.1: 4 photographs in the time series of an example tracer model experiment; Top-left: Baseline image before the injection start; The other three photographs show the development of the experiment over time.*

The photographs of the experiment (Figure 3.1) show a slight contamination on the left side of the geometry with tracer residue from a previous experiment. It is also important to note that not only is fluid injected in the experiment, but fluid is also produced at the back of the FluidFlower through another of the many valves, which explains the complex shape of the tracer distribution.

## 3.4 Photograph interpretation

The advantage of the simplification of the model experiment is that the physics behind it is also simpler, which makes it possible to anticipate some qualitative features of the expected concentration results as a physical interpretation of the happenings in the FluidFlower.

The injection of a fluid into the FluidFlower is a porous transport process. The transport in a homogenous porous medium can be described using the porous transport equation which in 1 dimension looks like the following

$$\varphi \frac{\partial}{\partial t} c(x,t) + u\varphi \frac{\partial}{\partial x} c(x,t) - D \frac{\partial^2}{\partial x^2} c(x,t) = 0$$

with the porosity $\varphi$, the convective velocity $u$, the diffusion coefficient $D$ (in practice dispersion is incorporated) and the concentration $c(x,t)$. A more detailed description of the transport equation in multiple dimensions can be found in Chapter 4 of the Propaedeutic [12]. In 1 dimension, this equation has an analytical solution for $c(x,t)$ (from [13], [14])

$$c(x,t) = \frac{1}{2}\left(1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\varphi x - ut}{2\sqrt{D\varphi t}}} e^{-t^2} dt\right) + \frac{1}{2} e^{\frac{ux}{D}} \left(1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\varphi x + ut}{2\sqrt{D\varphi t}}} e^{-t^2} dt\right)$$

For prescribed boundary and initial conditions

$$c(0,t) = 1, t \geq 0$$

$$c(x,0) = 0, x \geq 0$$

$$c(\infty,t) = 0, t \geq 0$$

The term $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the so-called error function. The analytical solution using the following parameters can be seen in Figure 3.2.

$$t = 1, D = 0.01, u = 1, \varphi = 1$$

The chosen parameters are not related to the model experiment and are just chosen for illustrative purposes.

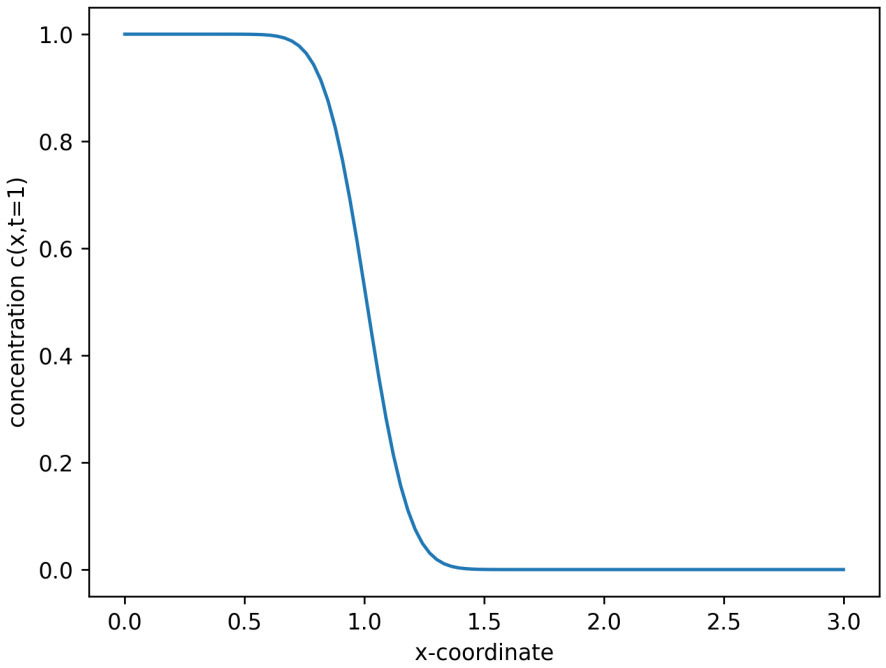


*Figure 3.2: Analytical solution to the transport equation in 1 dimension at $t = 1$ and D=0.01, u=1, φ=1*

Analogously, it is also possible to look at a 1 dimensional-like line in the FluidFlower experiment as drawn in Figure 3.3. The line covers all visible concentrations from close to the injection point on the left to the neutral fluid on the right. This choice of the position corresponds to the boundary conditions of the transport equation and corresponding analytical solution which makes it possible to qualitatively compare the concentration along that line to the analytical solution.

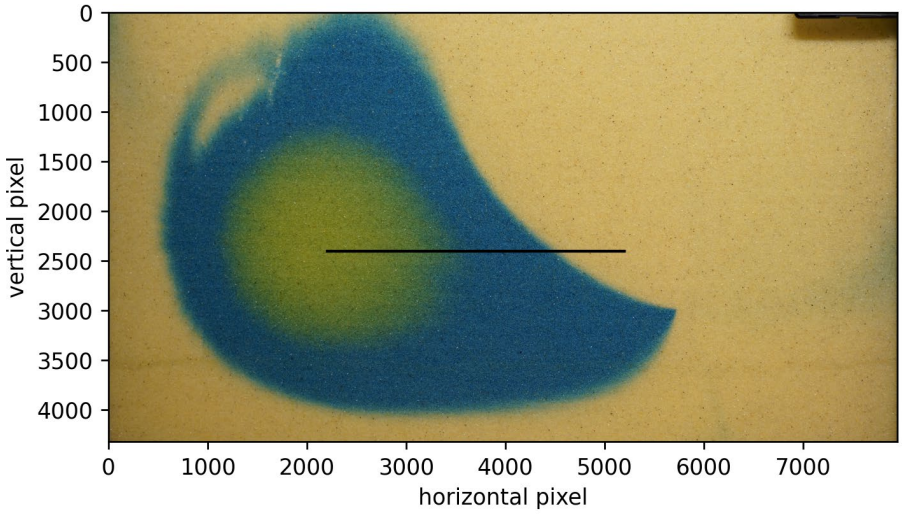


*Figure 3.3: A representative line in the multichromatic tracer model experiment from a region of concentration 1 to a region of concentration 0.*

According to the analytical solution shown in Figure 3.2, the qualitative expectations for the concentration profile are the following.

- Monotonically decreasing concentration profile from high concentrations to low concentrations.

- Inversed s-shape of the concentration curve
- Smooth continuous concentration profile.

These characteristics define a measurement for the comparison and evaluation of the in this thesis discussed methods. The resulting concentration profiles can be compared to the above described physical interpretation.

# 4 DarSIA: Concentration analysis and adaption for multichromatic tracers

DarSIA (Darcy Scale Image Analysis) is an open-source image analysis toolbox developed by the Mathematics Department at the University of Bergen [9], available publicly [10]. It offers capabilities to analyse images of porous media to extract physical data. DarSIA is especially designed to upscale information that is obtained at the pore scale via photographs or scans to a continuous Darcy scale information.

There is a particular workflow implemented in DarSIA called "ConcentrationAnalysis" that contains several image processing routines that can be used to extract concentration data from images. For this thesis important routines are the following

1) Preprocessing of the photograph
2) Subtracting images
3) Restoring and upscaling of information
4) Signal reduction
5) Signal modelling

Steps 1 to 3 are directly applicable to multichromatic tracer images without any modifications. Step 4, the signal reduction, is where the multichromatic information is reduced to monochromatic information by design, resulting in a loss of information. This means that in the existing workflow, tracers can only be processed based on a monochromatic analysis, which works for monochromatic tracers but means a loss of information for multichromatic tracers. This step will be adapted for multichromatic analysis as a first approach to multichromatic tracer concentration analysis.

The steps are applied to an example multichromatic tracer image to demonstrate the functionality, limitations, and adaptations of the existing analysis in the following sections of this chapter. The example tracer image $u_t$ will be an example timestep $t$ of the model experiment as seen in Figure 4.1 on the left.

## 4.1 Preprocessing of the photographs

The data acquisition of the experiments using a camera is often imperfect. For example, in Figure 4.1, the frame of the FluidFlower and the transparent top are visible in the photograph. Other common distortions are caused by automated corrections performed by the camera. The process of correcting these distortions is called preprocessing.

In this example, cropping is used to crop the FluidFlower geometry to where the tracer is visible. DarSIA has several other correction methods that are particularly important for multi-image analysis, such as colour and drift correction to correct for possible illumination fluctuations or unintended camera movements. Other corrections include bulge correction and curvature correction, which are particularly important when analysing images of the room-scale FluidFlower, which is curved and requires careful projection onto a planar rectangular domain to enable 2D image analysis. Furthermore, sophisticated correction routines in DarSIA make it possible to align grains in a image registration routine for multiple images. From now on, it is assumed for the images of the model experiment, that the sand grains in the images are aligned. These steps are essential for the accuracy of the quantitative results. However, in this methodological work the simple cropping method is sufficient. The results are shown in Figure 4.1.

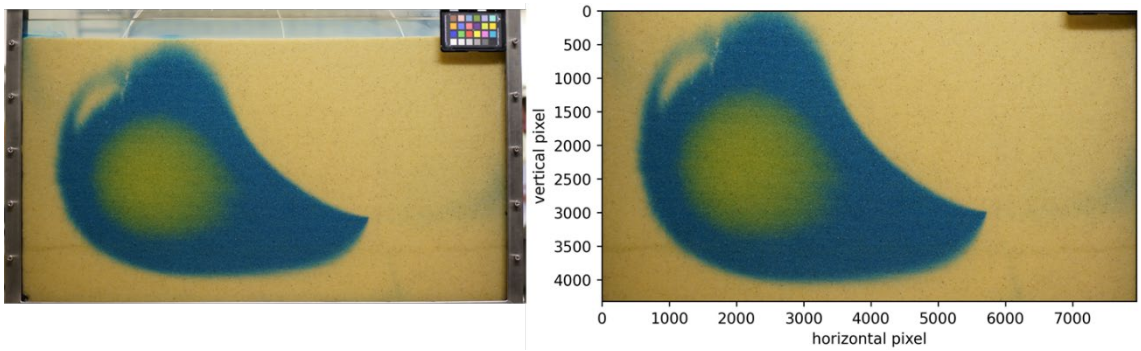


*Figure 4.1: A example timestep $t$ of a multichromatic tracer experiment; Left: Original photograph $u_t$; Right: The preprocessed corrected tracer image $\overline{u_t}$*

We denote the corrected baseline image before the injection start as $\overline{u}_0$ and the corrected tracer image at the timestep $t$ as $\overline{u}_t$.

## 4.2 Subtracting images

Taking the difference of images is used to subtract information that is present in both pictures. To track the dynamic changes (the tracer dynamics) in the FluidFlower, the static information such as the porous sand geometry is removed by subtracting images. Therefore, the baseline image $\overline{u_0}$ (Figure 4.2, left) where no tracer and only the sand geometry is visible is subtracted from the actual tracer image $\overline{u_t}$ (Figure 4.2, right). This results in the difference image $u_t'$.

$$u_t' = \overline{u_t} - \overline{u_0}$$

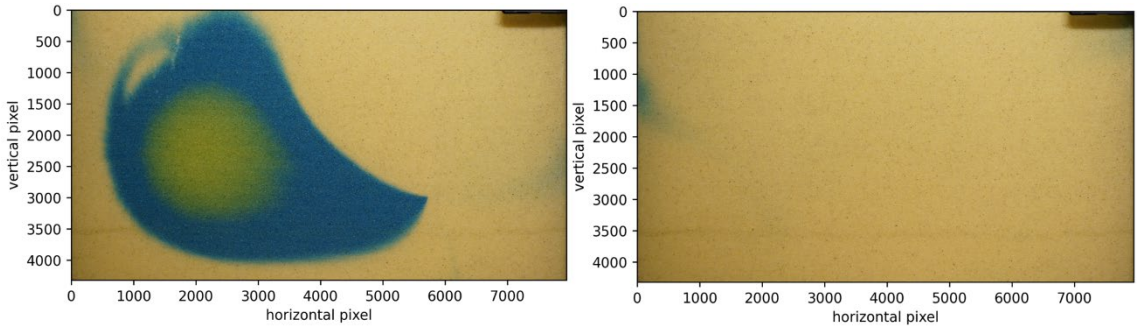


*Figure 4.2: Left; The corrected tracer image $\overline{u_t}$; Right: The corrected baseline image $\overline{u_0}$*

The pixel values resulting from that subtraction are not an RGB colour anymore since the subtraction of images produces negative values which is not part of the original RGB space anymore. As introduced in 3.2, the original tracer image $u_t$ is a RGB photograph. The RGB colours are assumed to be available as a float. This means that every pixel of the image $p \subset u_t$ is in $p \in [0,1]^3$. After subtracting the baseline from the tracer image, a pixel $p \subset u_t'$ now has a value $p \in [-1,1]^3$. Those pixel values are from now on called relative colours. The displaying of the difference image is not possible with conventional image displaying methods due to the negative relative colour values. However, one can still visualise the pixel-wise positive and the absolute value of negative relative colours separately using conventional image displaying methods to illustrate the subtraction effect. This can be seen in Figure 4.3.

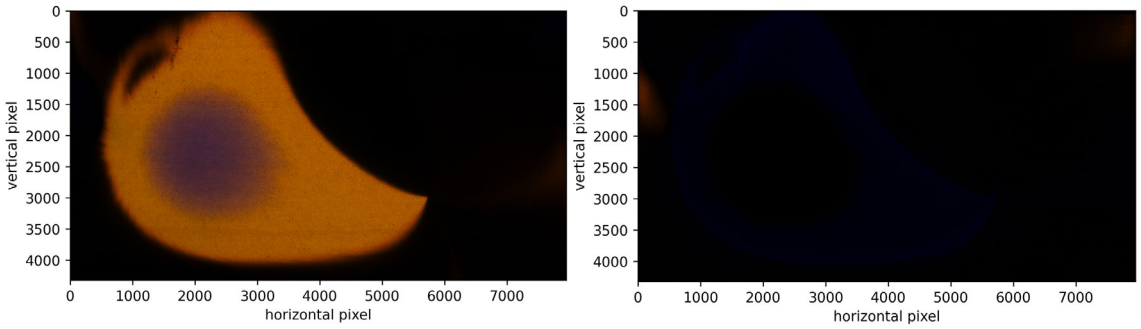


*Figure 4.3: Difference image; Left: Absolute value of negative values of the difference image visualised; Right: Positive values of the difference image visualised*

The procedure of subtracting images is not dependent on the colour space of the image and can also be done in other colour spaces.

## 4.3 Restoration and upscaling

Smoothing the image is the central part of the restoration and upscaling of the in the images contained physical information. A more detailed explanation of the upscaling of information and its relation to the upscaling of porous media can be found in the Propaedeutic in Chapter 5 [12].

In this work the isotropic Bregman algorithm [15] as implemented in the python library scikit-image [16] is used to perform a Total Variation Denoising (TVD) of the image. The parameters e.g. the denoising weight $\lambda$, the numerical stopping criterion $\epsilon$ and the maximum number of iterations for the algorithm $n_{max}$ are chosen as follows

$$\lambda = 0.025$$

$$\epsilon = 1 \times 10^{-4}$$

$$n_{\max} = 100$$

This choice aims on a smooth image so that the sand grains are not visible as sand grains anymore (compare to Figure 4.4). The resulting smooth image is denoted as

$$u_t^*$$

This smooth image can also be visualised using only the absolute value of the pixel wise negative values analogue to the section before which is shown in Figure 4.4.

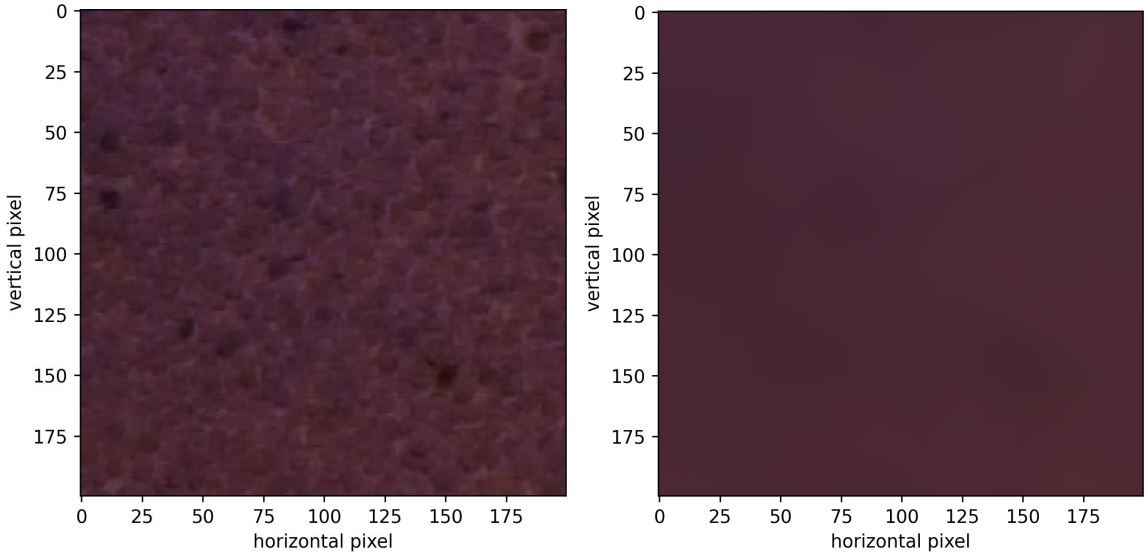

*Figure 4.4: Effect of the image smoothing; Left: The absolute value of the negative pixel-wise values of the difference image prior to the smoothing; Right: The absolute value of the negative pixel values of the smoothed image after applying the smoothing routine.*

The smoothing of the image with the Bregman algorithm is also applicable without modifications in a colour space that is not the RGB colour space.

## 4.4 Signal reduction

The signal reduction is used to reduce the multi-dimensional relative colour (resulting from the previous steps in the previous sections) in each pixel to a 1-dimensional scalar signal. This is the main routine (together with the signal modelling in 4.5) of extracting the concentration data from the tracer images in the existing DarSIA concentration analysis. In DarSIA a few signal reduction methods are implemented. Examples include the reduction to one channel of the relative colour or to one channel of the colour space transformed relative colour. Possible transformations are the greyscale and the HSV colour space. The transformation to the HSV space leads to the HSV signal reduction. It is important to note, that negative relative colours after transforming might result in negative HSV colours. Nevertheless, the routine can still be applied. By design, the HSV signal reduction as all the other implemented signal reductions only extracts a monochromatic information from the image. However, this reduction has been identified to be the most promising for an adaption to multichromatic information as a first approach to a multichromatic concentration analysis.

### 4.4.1 HSV signal reduction for monochromatic signals

Prior to the HSV signal reduction, the information (one tracer colour) needs to be characterised with a hue threshold interval by the user. This interval defines the range of for that tracer characteristic colour through its hue value. Each pixel of the image whose colour lies outside that hue threshold interval is set to 0. The image is then pixel-wise reduced to the value component of the remaining HSV colours. Here lies the limitation to monochromatic signals because only one hue range can be handled for characterisation.

This reduction approach is motivated by the idea, that the HSV value component represents the intensity of the characteristic colour which is assumed to correlate with the concentration of the tracer.

### 4.4.2 HSV signal reduction adapted to multichromatic signal analysis

Adapting the HSV signal reduction to multichromatic information is based on the idea of using multiple hue thresholds for multiple characteristic colours of the multichromatic tracer. A reduced signal is obtained for all characteristic colours using the existing HSV signal reduction with the corresponding threshold. In a next step, those signals are merged with weighted addition to one resulting signal.

The process will be demonstrated in a small region in the model experiment photograph (see Figure 4.5).

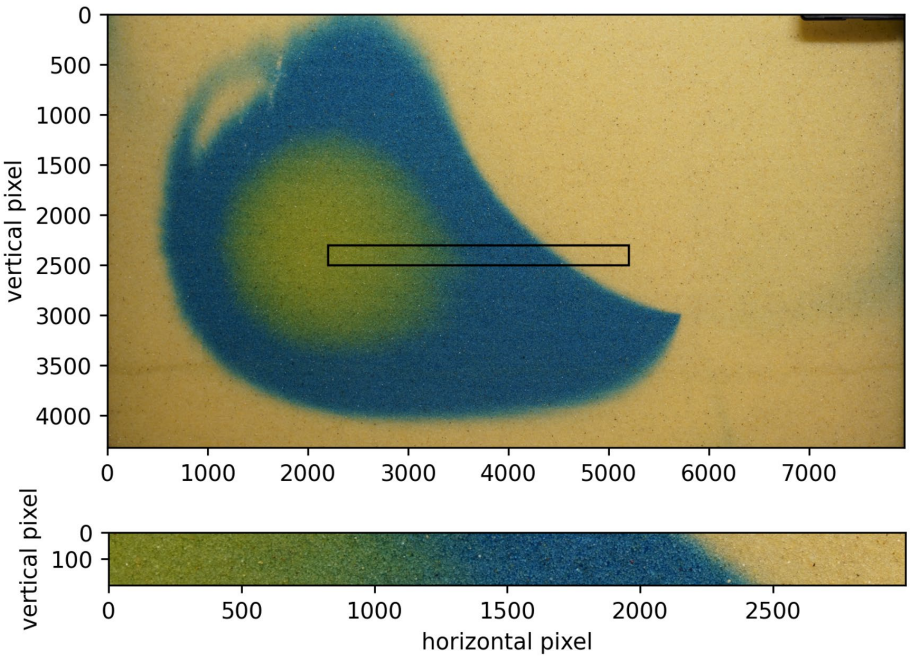


*Figure 4.5: Representative region of interest with all visible colours of the tracer*

The hue threshold values for each characteristic colour of the tracer are obtained by histogram analysis of the hue values in a small patch of the image where that colour is predominant. The histogram analysis shows which hue values are most frequent in that patch and limits containing those colours define the threshold. The two chosen patches for the two characteristic colours of the multichromatic tracer in the model experiment are shown in Figure 4.6.

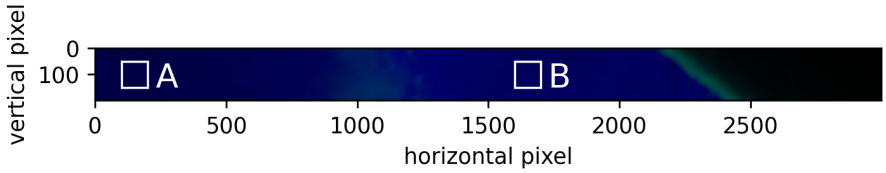


*Figure 4.6: HSV transformed image region of Figure 4.5 with two patches in each characteristic colour region*

The concentrations for the tracer colours are assumed to be 1 for patch A and 0.9 for patch B. These choices are arbitrary but comply with the physical interpretation of the tracer image (3.4) and lead to the hue interval concentration pairs as shown in Table 4.1.

*Table 4.1: Hue interval concentration pairs for the patches in the model experiment as shown in Figure 4.6*

| Patch *P* | Hue interval | concentration |
|---|---|---|
| A | [-0.5, -0.4] | 1 |
| B | [-0.08,-0.04] | 0.9 |

Using those hue intervals, the scalar signals for each interval can be extracted using the existing HSV reduction routine. For the example (Figure 4.5), the two retrieved scalar signals for each characteristic colour are shown in Figure 4.7.

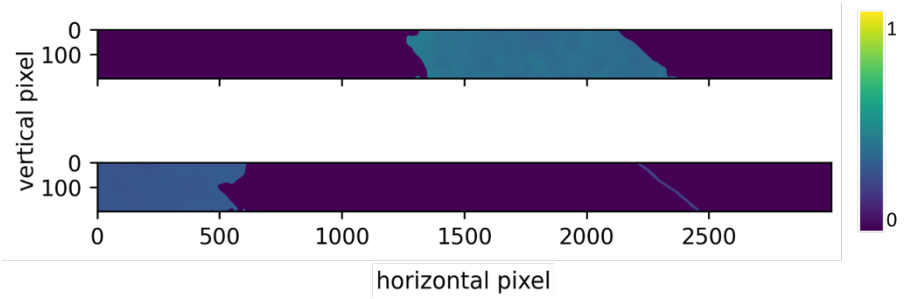


*Figure 4.7: The two signals retrieved with the HSV signal reduction for both characteristic colours.*

These signals are now modelled together to one signal. In its simplest form this could be plain addition of the signals which can be seen in the upper plot in Figure 4.8. A more sophisticated approach is to normalize both signals to the same intensity and then scale them with their corresponding concentration from the hue concentration pairs. This is done in the lower signal in Figure 4.8.

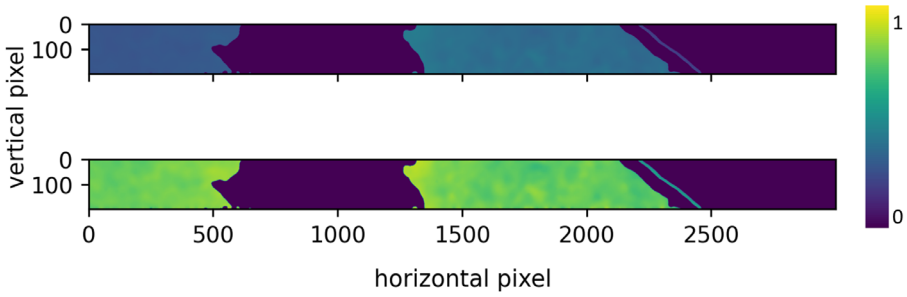


*Figure 4.8 Top: both signals added. Bottom: normalised and weighted addition of the two signals*

To compare this result to the physical interpretation of the image as described in 3.4, the concentration profile is reduced to 1 dimension. This is done by taking the average value of each vertical pixel column in the concentration profile. This reduction results in a 1-dimensional concentration profile which can be seen in Figure 4.9 for this case.

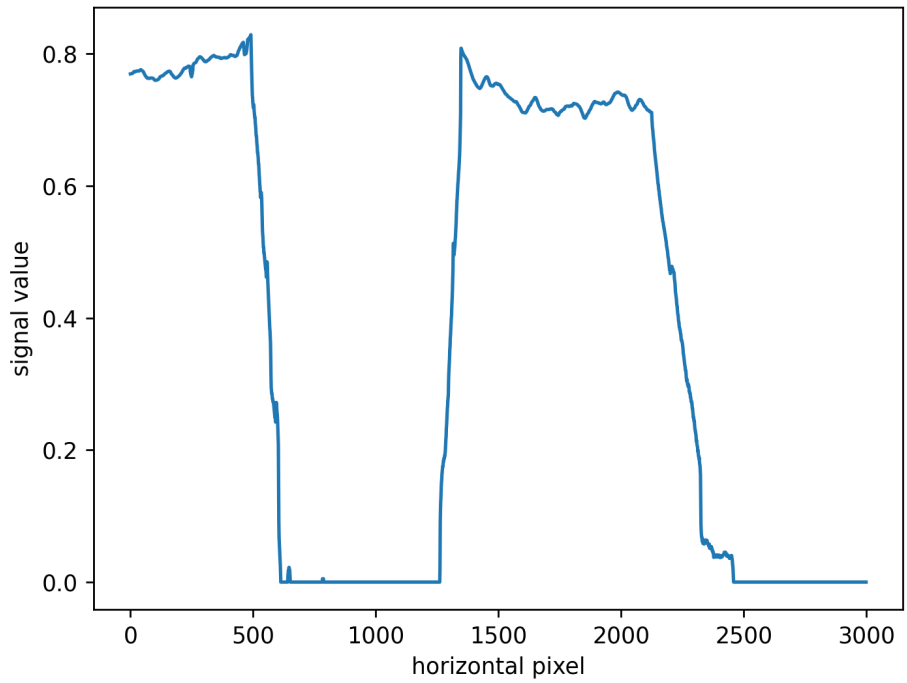


*Figure 4.9: 1d reduction of the weighted addition signals as shown in Figure 4.8 in the lower plot.*

The result of merging the two signals in a linear model shows the resulting concentration profile and its 1d reduction is not smooth as it would be expected according to 3.4. The transition zone between the tracer and the neutral fluid on the right is not properly described as it is more an edge than a gradient. Also, the region with zero concentration in the middle of the image is not physically reasonable as we would expect a smooth monotone behaviour.

The gab can be explained with the incomplete decomposition of the colour space since the hue thresholds do not cover the whole range of tracer colours. However, the use of additional hue range concentration pairs in the transition zone between the colours also did not achieve the desired results. Also, different weighting approaches or wider choices of the hue ranges were not satisfying. The ever-present problem is the non-smoothness of the resulting merged signal with hard edges as already seen in Figure 4.8.

The unsatisfying results could be resulting from the loss of information when the colours which characterise the information in each pixel are reduced to one colour channel while ignoring the information present in the other colour channels. This is particularly significant for multichromatic tracers as shown above.

In conclusion, the HSV signal reduction works for monochromatic tracers as demonstrated in other works [8] but the extension to multichromatic tracers is not satisfactory.

## 4.5 Signal modelling

The scalar signal that is obtained by the signal reduction must be converted to a physically meaningful variable, the concentration. In DarSIA a few simple models are implemented for that purpose which include a linear model and thresholding models.

After applying these models as the final step in the existing concentration analysis, the resulting concentration profile can be used for quantitative analysis of the experiment.

## 4.6 Interim valuation of DarSIA concentration analysis

In conclusion, the existing concentration analysis routines of preprocessing, subtraction of images and restoration, are usable for the concentration analysis of multichromatic tracer images. The step signal reduction has been adapted to multichromatic tracers but the loss of information in that approach produces poor resulting concentration profiles which do not correspond to the expected physics. This limiting part of the concentration analysis will be the objective of the next chapter.

In the following chapter a new approach for converting a pixel-colour to a tracer concentration is suggested and demonstrated that bases on the concept of interpolation. This method aims on replacing the signal reduction and modelling in the concentration analysis with a more general method.

# 5 Interpolation approach using kernel methods

As seen in the previous chapter, the conversion of colours to tracer concentrations at pixel level is not satisfactory for multichromatic tracers in the existing DarSIA concentration analysis due to the limitations of the signal reduction. In this chapter, interpolation is suggested as a new alternative to the signal reduction as means of converting the colours of a pixel to a concentration with the hope that it delivers better results for multichromatic tracers. The interpolation method builds upon the routines as described in 4.1-4.3 and complements them to a multichromatic tracer concentration analysis by converting the relative colours in the pixels of the smooth image (compare to 4.3) to a concentration value of the multichromatic tracer based on known data resulting in a concentration profile for the tracer image.

Chapter 5 is structured as follows. First, the problem of a colour to concentration conversion on pixel level is mathematically described as an interpolation problem which is the novel suggestion of this chapter. Second, the known data for the model experiment is described as the data basis for the application of the interpolation method. The rest of the chapter then discusses varieties in the design of the interpolation method with the goal of finding the best suiting out of a selected set of approaches. These include the kernel method and the nearest neighbour interpolation.

All code examples used in this thesis are publicly available in a GitHub repository [11].

## 5.1 Interpolate known data as problem solution

Interpolation estimates the value of unseen data based on known data. In the context of multichromatic tracers this seems like a promising approach, as shown below.

Assume a known number $n \in \mathbb{N}, n > 0$ of relative colours $x_i \in [-1,1]^3 \subset \mathbb{R}^3$ and their corresponding tracer concentrations $y_i \in [0,1] \subset \mathbb{R}$. These defined colour-concentration pairs are

$$(x_i, y_i) \in [-1,1]^3 \times [0,1] \ , i = 1, \dots, n$$

The here chosen range $[0,1]$ of concentration values is an arbitrary choice motivated by volumetric concentrations. It is easy to modify the concept for values between 0 and 14 which would be suitable for pH-values.

The interpolation problem is now to find the interpolant

$$s: [-1,1]^3 \to \mathbb{R}$$

that takes relative colours from $[-1,1]^3$ and assigns a concentration value under the constraint that the function satisfies

$$s(x_i) = y_i \qquad \forall i = 1, \dots, n$$

This is a multivariate interpolation problem since the photograph colours – and consequently, relative colours – live in a multi-dimensional space. A more general mathematical definition of the interpolation problem is given in the Propaedeutic in Chapter 6 [12].

The interpolant based on the known colour concentration pairs can now be applied to every pixel of the image (unseen data) to obtain an estimation of the concentration in each pixel resulting in a concentration profile for the whole image.

Interpolation approaches and strategies for finding the interpolant are various. In this work, the nearest neighbour interpolation and the kernel method with the linear and gaussian kernel will be discussed. The kernel method has been chosen because of its various advantages and suitability for multivariate interpolation.

Additionally, it is possible to translate the image to another colour space prior to the interpolation. Consequently, the known colours and the interpolation depend on the chosen colour space. In this work, the RGB (original) and the LAB space are discussed. The LAB space is chosen as an alternative because it is a uniform colour space, and the hope is that an interpolation in that space gives better results. More on colour spaces can be found in the Propaedeutic [12] in Chapter 5.

The colour concentration pairs are obtained by experiments or expert knowledge. It is not clear how different interpolation approaches behave for different amount of input expert knowledge.

To find the best suiting interpolation method variant of those variations, the following research questions must be investigated and answered in this chapter.

Q1: Which interpolation approach is best suited?
Q2: Which colour space is best suited?
Q3: What is the behaviour of the interpolation for different numbers of known colour concentration pairs?

In this chapter, the methods will be tested and demonstrated in a small region of the model experiment (Chapter 3), as described in 4.4.2, and the result will be qualitatively compared to the physical interpretation of the tracer images, as described in 3.4 as a metric for success.

## 5.2 Known data in the model experiment

Known data is the basis for every interpolation approach. For a multichromatic tracer, this known data is the knowledge about specific colours of that tracer and their corresponding concentrations. This knowledge is provided by the user as colour concentration pairs for the respective chosen colour space.

The multichromatic tracer in the model experiment displays three characteristic colours. For a high concentration almost 1, it displays a greenish colour; for lower concentrations, a blueish colour; and in the limit concentration 0, it is transparent, as seen in the upper plot in Figure 5.1. Those colours correspond to relative colours in the smooth RGB and smooth LAB image, as shown in the two lower plots in Figure 5.1. The known colours for those smooth images, on which the interpolation will be performed, are retrieved using characteristic patches for each characteristic colour as displayed in Figure 5.1. At the end of this section, it is described how the user can retrieve known colours for the colour concentration pairs.

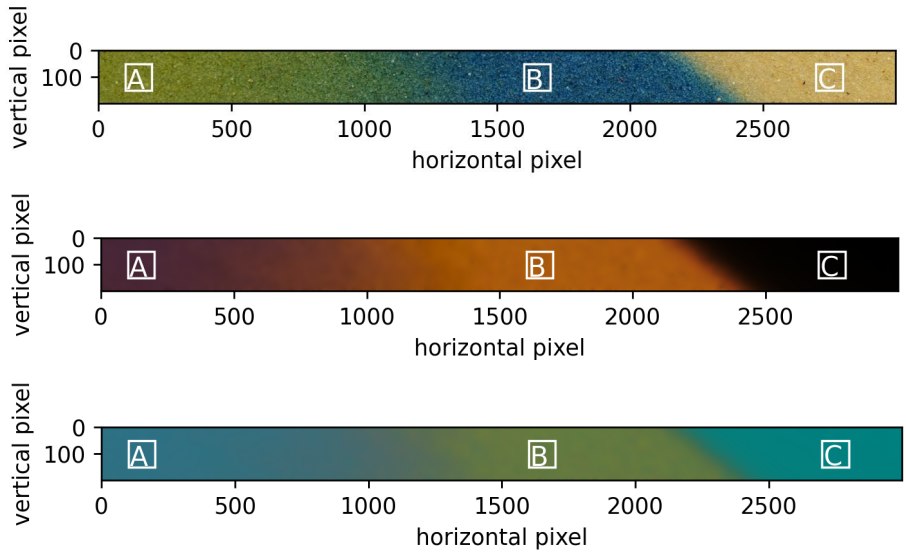


*Figure 5.1: Three patches for three characteristic colours; Top: Original photograph of the examined region of the tracer experiment; Middle: RGB smooth image; Bottom: LAB smooth image*

The known colour concentration pairs for the model experiment in the RGB space are given in Table 5.1 and the pairs for the LAB space are given in and Table 5.2.

*Table 5.1: Known colour concentration pairs in the RGB space*

| Patch $P$ | Characteristic colour $x_P$ | Concentration $y_P$ |
|---|---|---|
| A | (-0.29, -0.15, -0.21) | 1 |
| B | (-0.67, -0.35, 0.05) | 0.9 |
| C | (-0.01, -0.01, 0.01) | 0 |

*Table 5.2: Known colour concentration pairs in the LAB space*

| Patch $P$ | Characteristic colour $x_P$ | Concentration $y_P$ |
|---|---|---|
| A | (-0.17, 0.45, 0.51) | 1 |
| B | (-0.39, 0.47, 0.25) | 0.9 |
| C | (-0.01, 0.51, 0.49) | 0 |

The known concentration for each characteristic colour is obtained by expert knowledge of the tracer. However, in the methodical context of this thesis, the known concentrations are chosen arbitrarily, but in compliance with the basic physical understanding of the experiment (compare with section 3.4). The concentration in patch B is chosen to be low. According to the physical interpretation in 3.4, a concentration much closer to 1 would be expected and yet, for the development and discussion of the methodology, it is beneficial to choose a slightly lower concentration.

It is possible to choose more characteristic colours and concentrations as the known data points for the interpolation. This allows for more input of expert knowledge. This and how this could improve the interpolation is discussed in chapter 5.5.4.

*Histogram analysis of a representative patch to retrieve known colours from an image*

The characteristic colours of a tracer need to be extracted from the smooth image resulting from the previous image processing routines as described in chapter 4.1-4.3 in different colour spaces. This subsection describes the procedure of extracting known colours from an image using multivariate histograms. The method of retrieving characteristic colours is independent of the chosen colour space but will be demonstrated in the RGB space.

The characteristic colour is retrieved by analysing a small patch in a region, where that colour is mostly present. Such patches are displayed in Figure 5.2.

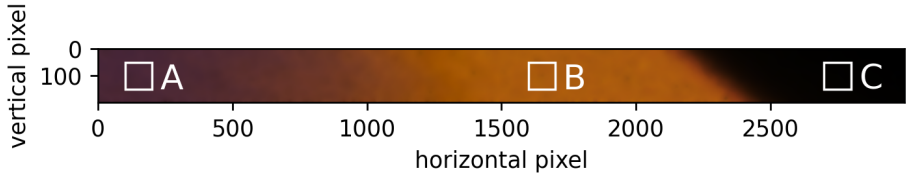


*Figure 5.2: The smooth RGB image with three different patches A, B and C, for three identified characteristic colours. For visualisation purposes, the absolute value of the relative colours is visualized.*

The characteristic colour is chosen to be the colour which is the most frequent in the patch. To retrieve that colour, a histogram analysis is used. A patch $P$ is a region of the image with relative colours $x \in [-1,1]^3$ in each pixel. These relative colours live in a three-dimensional space. Consequently, extracting the most frequent colour in a histogram style approach leads to multivariate histogram analysis. This procedure is based on [17]. The set $[-1,1]^3$ is divided into equally sized bins. In three dimensions, these have the shape of a cube. The patches have here the size of $100 \times 100$ pixels, which means a total of

$n = 10\,000$ pixels. When using the square root rule for histograms, this means to use $k = \sqrt{n} = \sqrt{10000} = 100$ bins in each dimension. The bins then are defined as

$$b_{i,j,l} = [e_{i-1}, e_i] \times [e_{j-1}, e_j] \times [e_{l-1}, e_l], \qquad i,j,l = 1,2,3,\dots,k$$

with the edges

$$e_l = -1 + 2\frac{l}{k}, \qquad l = 0,1,2,\dots,k$$

The absolute frequency $h$ in such a bin $h(b_{i,j,l})$ is calculated via

$$h(b_{i,j,l}) = \left|\{x \in P \mid x \in b_{i,j,l}\}\right|$$

with $|\cdot|$ being the number of elements in the set.

After computing the absolute frequency for all bins, the characteristic colour $x_P$ of the patch $P$ is defined as the value in the middle of the three-dimensional bin which has the highest absolute frequency $b_{\text{char}} = b_{\bar{i},\bar{j},\bar{l}}$.

$$x_P = \begin{pmatrix} \frac{e_{\bar{i}} - e_{\bar{i}-1}}{2} \\ \frac{e_{\bar{j}} - e_{\bar{j}-1}}{2} \\ \frac{e_{\bar{l}} - e_{\bar{l}-1}}{2} \end{pmatrix}$$

It is important to note, that the characteristic colour $c_P$ is not necessarily present in the image because it is a constructed colour resulting from the histogram analysis.

## 5.3 Colour space discussion RGB vs. LAB

As mentioned in 5.1, it is possible to transform the image to different colour spaces prior to applying the interpolation. This means that the interpolation is realised in different colour spaces, which raises the question of the colour space which gives the best results.

It will be of interest to investigate the structure of the colour spaces by analysing the arrangement of the known tracer colours in the colour space in order to explain the behaviour of the discussed methods in the following.

The structure of colour spaces RGB and LAB is not intuitive, and neither are the addition and subtraction of colours in that colour space. To investigate the arrangement of the tracer colours in the colour space, 9 patches are chosen (shown in Figure 5.3, top) in every region of the tracer to retrieve 9 colours that cover the whole range of the tracer.

These colours can then be plotted in the relative colour space $[-1,1]^3$, as shown in Figure 5.3. The same analysis can be done for the LAB space, which is shown in Figure 5.3 bottom.

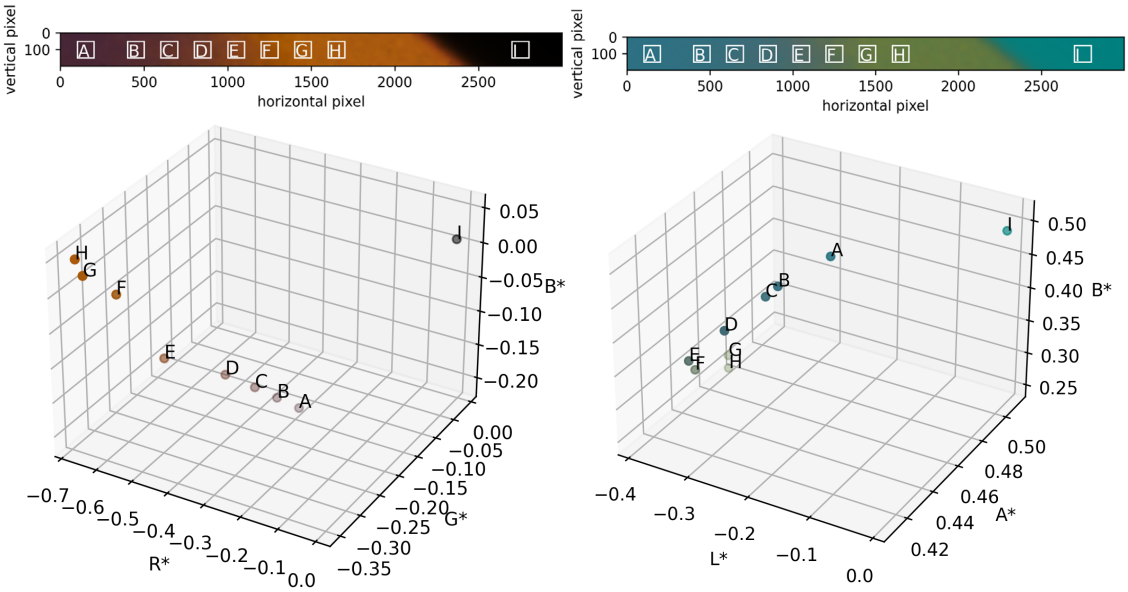


*Figure 5.3: Arrangement of the relative colours retrieved from patches in the colour space; Left: Arrangement of relative colours based on the RGB colour space. R*,G*,B* indicate the relative character of the colour channels; Right: Arrangement of relative colours based on the LAB colour space. L*,A*,B* indicate the relative character of the colour channels*

Comparing those two colour spaces is interesting, because it shows the different structure of the colours in the different colour spaces and it makes it clear once more that the whole image processing routine – including the kernel interpolation – depends on the choice of the colour space. It stresses the importance of research question Q2.

## 5.4 Nearest Neighbour interpolation

The simplest approach on interpolating multivariate data is the nearest neighbour interpolation (NN interpolation). The NN interpolation selects for each relative colour in the image $x$ the value of the closest known data point. This means that no other data points are considered, and the interpolation function is not smooth in terms of differentiating. The advantage is the simplicity of the algorithm.

The basis for that calculation is the definition of a metric $d: CS \times CS \to \mathbb{R}$ in the colour space $CS$ that measures the distance between two colours. The metric is colour space dependent. In the following, this procedure will be shown and discussed in the two colour spaces RGB and LAB, starting with the RGB space.

### 5.4.1 Nearest neighbour interpolation in RGB

In RGB, the distance $d_{RGB}(c_1, c_2)$ between two colours $c_1, c_2 \in RGB$ is calculated with the Euclidian norm

$$d_{RGB}(c_1, c_2) = \|c_1 - c_2\|_2 = \sqrt{\langle c_1 - c_2, c_1 - c_2 \rangle}$$

Other norms and metrics are possible to define on the RGB space, of which most aim to obtain a better fit for human perception.

Applying the nearest neighbour interpolation to the example region of interest gives the following resulting concentration profile.

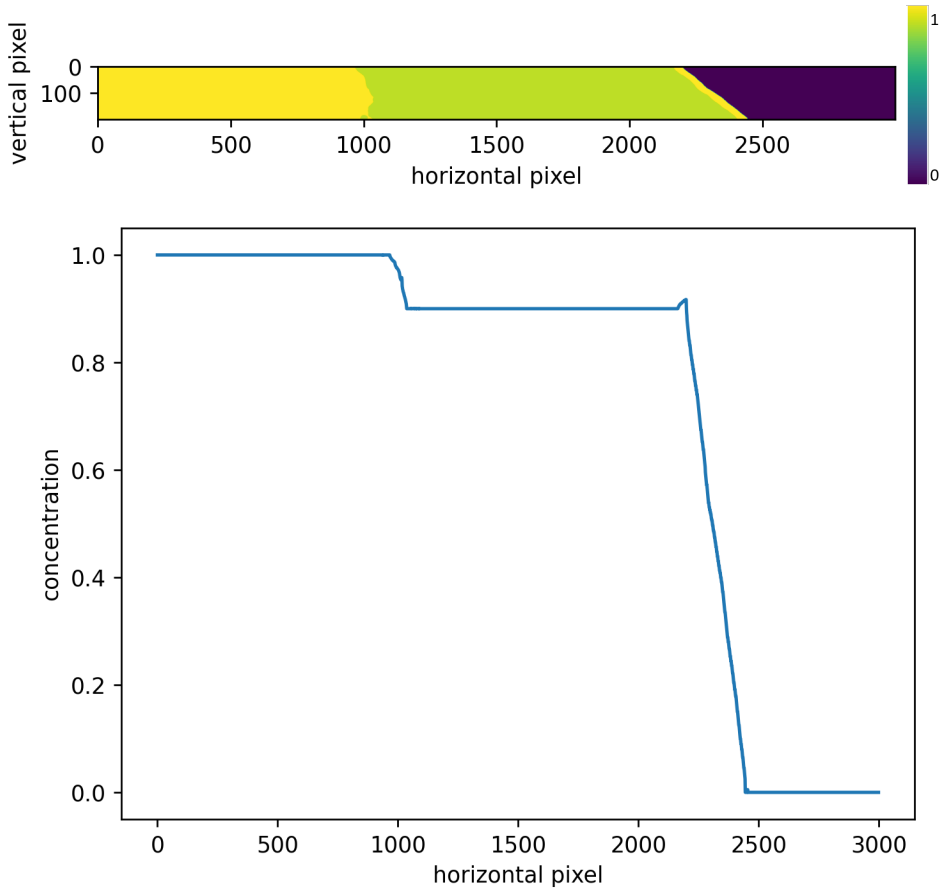


*Figure 5.4: Results of the nearest neighbour interpolation in RGB; Top: Concentration profile for region of interest as described in 5.2 with a concentration scale on the right; Bottom: Analogue to 4.4.2 1d reduced concentration profile*

The concentration profile is, by design of the nearest neighbour interpolation, not continuous and it shows a non-monotonic behaviour in the transition zone to the neutral fluid (pixels 2000-2500). Nevertheless, the resulting concentration profile has an overall good behaviour compared to the physical interpretation and expectation (3.4).

### 5.4.2 Nearest neighbour interpolation in LAB

The nearest neighbour interpolation can also be applied in the LAB space. The metric is again the Euclidian distance $d_{LAB}(c_1, c_2)$ between two colours $c_1, c_2 \in LAB$ and is calculated by

$$d_{LAB}(c_1, c_2) = \|c_1 - c_2\|_2 = \sqrt{\langle c_1 - c_2, c_1 - c_2 \rangle}$$

The resulting concentration profile can be seen in Figure 5.5.

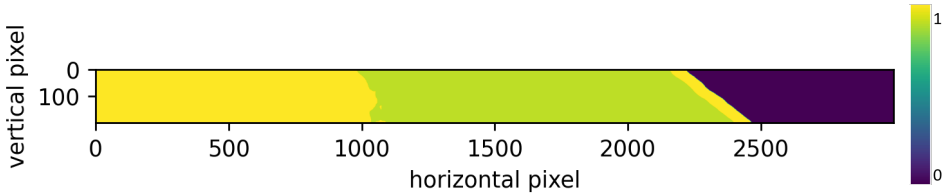


*Figure 5.5: Results of the nearest neighbour interpolation in RGB; Left: Concentration profile for region of interest as described in 5.2; Right: Concentration scale*

This concentration profile shows the same qualitative characteristics as the RGB variant of the nearest neighbour interpolation.

Summarising, the nearest neighbour interpolation delivers better results than the HSV-signal reduction result (compare to 4.4) and therefore strengthens the idea of an interpolation method. The bad behaviour of the interpolation in the transition zone to the neutral fluid (pixels 2000-3000) can be explained with the arrangement of the colours in the colour space as seen in 5.3. This behaviour can be improved by incorporating more colour to concentration pairs.

In conclusion, the nearest neighbour interpolation is a good simple interpolation approach that is suitable for a simple analysis using additional known data pairs. It is also possible that this is a satisfying approach for other tracers or colour spaces. However, in terms of robustness and smoothness of the interpolation, even for small known data points, an interpolation approach using kernel methods will be discussed in the following section.

## 5.5 Kernel interpolation for concentration analysis

Kernel interpolation is a new and advanced technique for addressing the multivariate interpolation problem described in 5.1. Its application is particularly well-suited for higher-dimensional and sparse data, which aligns with the characteristics of the problem. This method offers flexibility and proves to be a suitable choice.

### 5.5.1 Kernel interpolation method

Kernels will be used to find an interpolant in the three-dimensional space of the relative colours of the image. A formal definition of a kernel and its properties is given in the Propaedeutic, Chapter 6 [12]. The two kernels that are going to be used and tested in this chapter are the linear kernel $k_l$ and the gaussian kernel $k_g$ with the user-chosen parameter $\gamma$.

$$k_l(x_1, x_2) = \langle x_1, x_2 \rangle$$

$$k_g(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2), \qquad \gamma > 0$$

The interpolant for the known data is obtained by solving the following linear equation.

$$Ka = y$$

with the kernel matrix $K = \left(k(x_i, x_j)\right)_{i,j=1}^{N}$, the known relative colours $\{x_i, \dots, x_N\} \subset [-1,1]^3$, the known concentrations $y = (y_i)_{i=1}^{N}$ and the coefficient vector $a = (a_i)_{i=1}^{N}$. The known relative colours and corresponding concentrations as described in 5.2 are used.

The interpolant applied to each pixel x of the image is then defined via

$$s(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i)\,,\ x \in [-1,1]^3$$

This interpolant now estimates a concentration value $s(x) = c \in \mathbb{R}$ for a relative colour $x \in [-1,1]^3$. To retrieve the final concentration profile, the interpolant is evaluated in each pixel of the smooth image. It is important to note that there is no maximum principle and the value $s(x)$ can be outside of $[0,1]$.

After all, the kernel interpolation is a flexible method. The above-described procedure can be varied for different kernels, colour spaces and number of colour concentration pairs, as already pointed out in the research questions in 5.1. In the following subsections, we investigate the effect of these choices with the goal to deduct (through examples) a suitable combination.

### 5.5.2 Kernel interpolation with the linear kernel

The linear kernel $k_{lin}$ is a natural first choice since it is simple to compute and has no parameters. The kernel interpolation method using the linear kernel will be applied in the RGB space and in the LAB space with the known colour concentration pairs (Table 5.1, Table 5.2). For the RGB space, the resulting

concentration profile is shown in Figure 5.6 and the 1d reduction in Figure 5.7. For the LAB space, the resulting concentration profile and 1d reduction is shown in Figure 5.8.

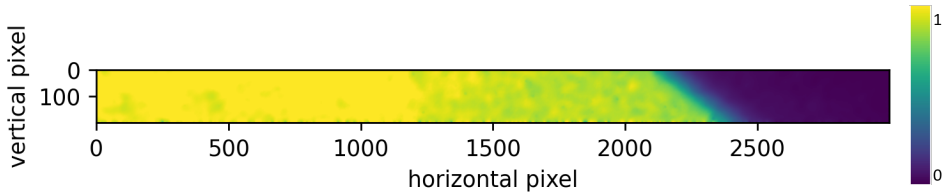


*Figure 5.6: Concentration profile using the linear kernel in the RGB space with three colour concentration pairs (Table 5.1). Concentration values over 1 are cut off to 1 and values below 0 are cut off to 0. Right: The concentration scale*

To compare this result to the physical interpretation of the image as described in 3.4, the concentration profile is reduced to 1 dimension analogue to 4.4.2. This reduction results in a 1-dimensional concentration profile which can be seen in Figure 5.7, for this case.

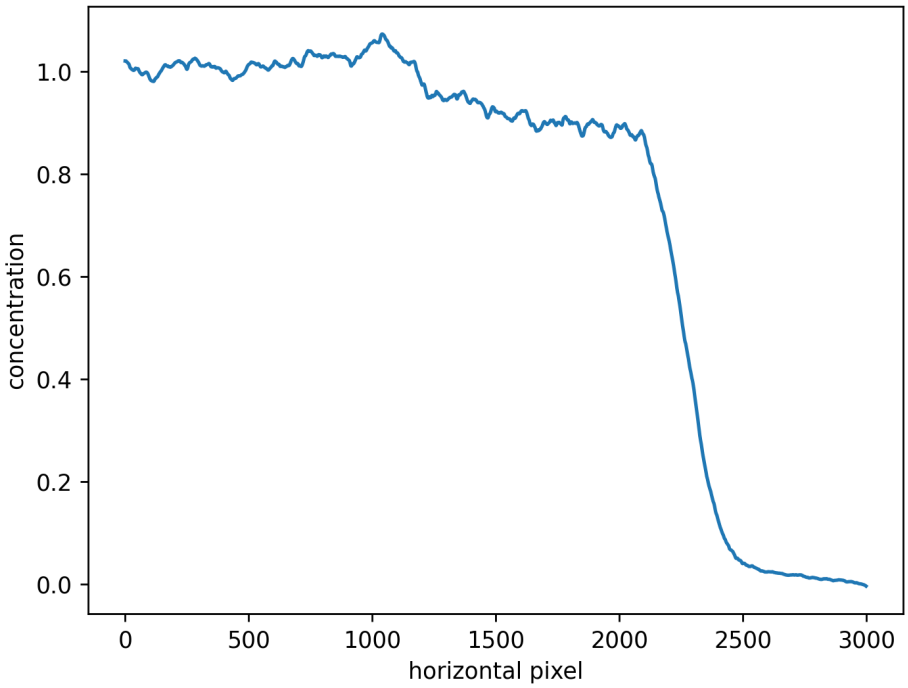


*Figure 5.7: 1d reduction of the concentration profile seen in Figure 5.6 (analogue to 4.4.2) using the linear kernel in the RGB space with three colour concentration pairs (Table 5.2)*

It can be seen that the transition zone on the right from tracer to neutral fluid (between pixels 2000 and 3000) corresponds to the physical interpretation of the tracer image as seen in Figure 3.2. The peak at pixel 1000 and the lack of monotony do not correspond to the physical understanding of the porous process (Figure 3.2), but in this mild form they still represent an acceptable result. Nevertheless, the question of better results and smoother profiles remains.

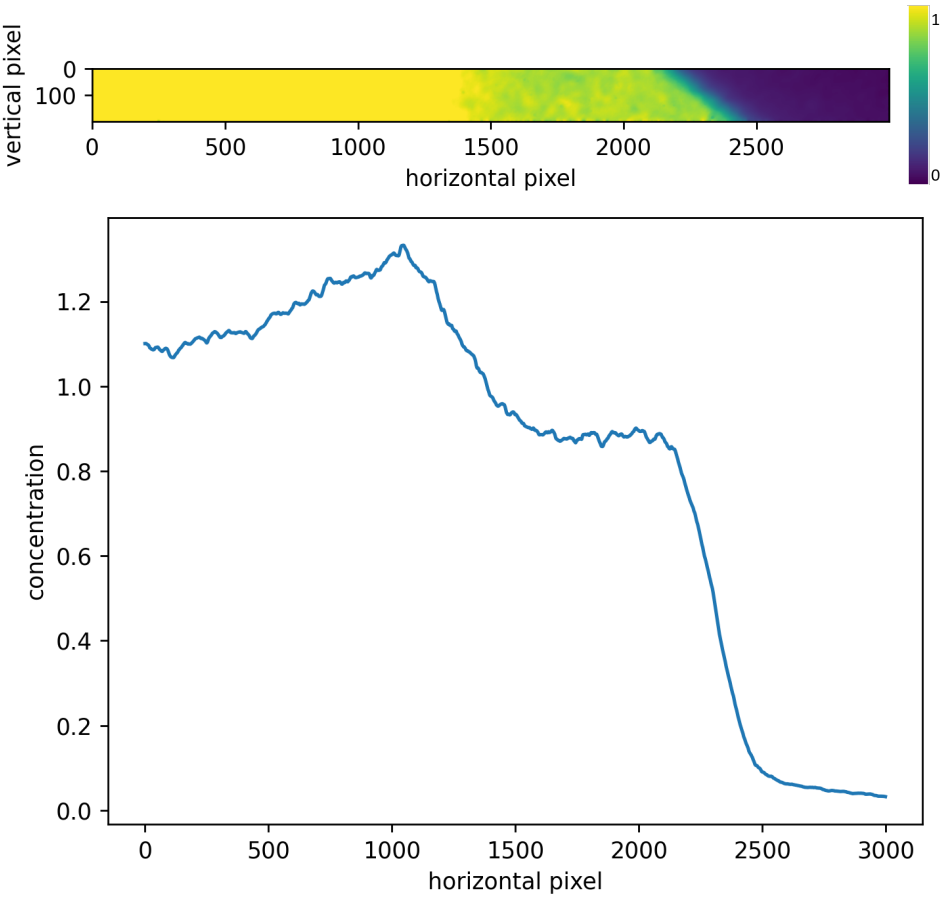


*Figure 5.8: Concentration profiles using the linear kernel in the LAB space with three colour concentration pairs (Table 5.2). Top: Concentration profile, Concentration values over 1 are cut off to 1 and values below 0 are cut off to 0. Bottom: 1d reduction of the concentration profile.*

The transition zone between pixels 2000 and 3000 is as in the RGB space correlating to the expectations in 3.4. However, the monotony is worse and the smoothness comparable to the RGB variant. The worse performance regarding monotony can be explained with the different structure of the LAB colour space in comparison to the RGB colour space, as visualised in 5.2 colour space analysis.

In conclusion, the linear kernel offers good results, especially in the RGB space, in comparison to the previously discussed nearest neighbour interpolation (5.4), with the advantage of no required kernel parameters.

### 5.5.3 Kernel interpolation with the gaussian kernel

In this subsection, the method of kernel interpolation will be applied with the gaussian kernel $k_g$. Properties and advantages of the gaussian kernel can be found in the Propaedeutic, Chapter 6 [12]. Mainly the hope is to get a smoother interpolation and thus concentration profile than with the linear kernel as seen in the section before 5.5.2.

The kernel parameter $\gamma > 0$ is chosen by the user from experience. As a technical note, values for the parameter $\gamma$ can be estimated by using a machine learning inspired approach, commonly referred to as the leave one out cross validation.

The interpolation method 5.5.1 is applied in the RGB and LAB space with the three known colour concentration pairs from Table 5.1 for RGB and Table 5.2 for LAB using the gaussian kernel $k_g$ with the parameter $\gamma = 9.73$ for RGB and $\gamma = 24.22$ for LAB. The resulting concentration profile is shown in Figure 5.9 for RGB and Figure 5.10 for LAB.

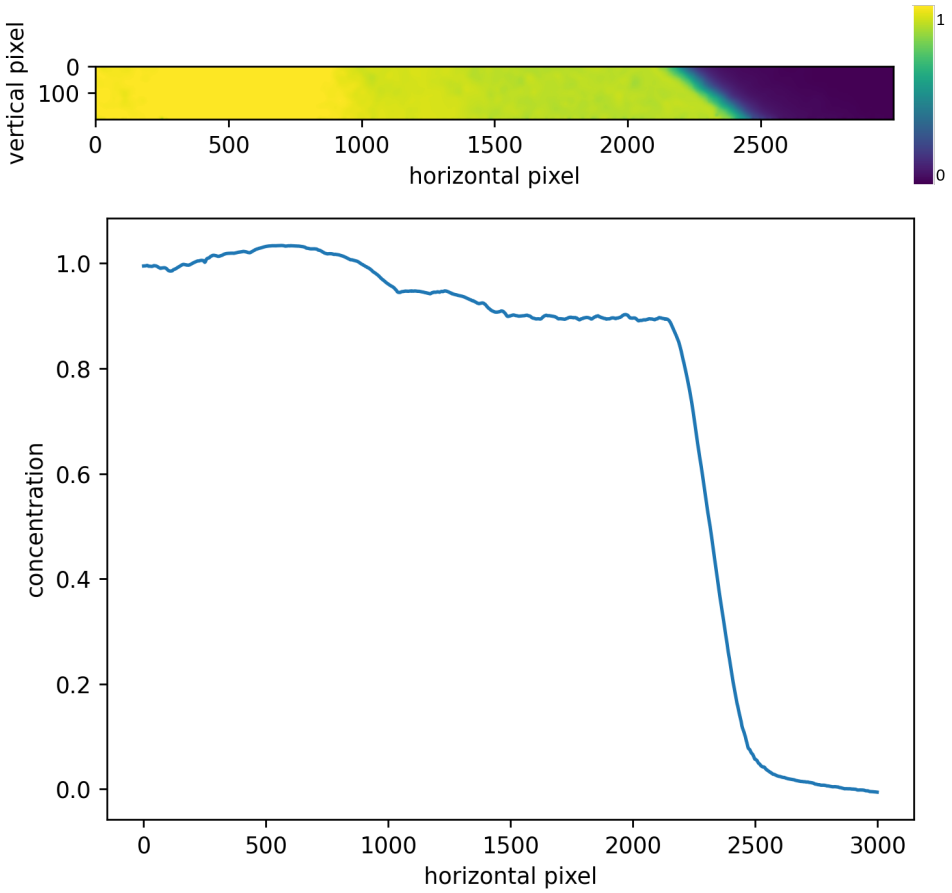


*Figure 5.9: Concentration profiles using the gaussian kernel in the RGB space with three colour concentration pairs. Top: Concentration profile, Concentration values over 1 are cut off to 1 and values below 0 are cut off to 0. Bottom: 1d reduction of the concentration profile.*

The concentration profile of the RGB variant is smoother compared to the linear kernel approach in RGB (Figure 5.7). The transition zone between 2000 and 3000 is mapped according to the physical interpretation of the image. The transition zone around 1000 is not corresponding to the physical expectations from 3.4 regarding monotony. Nevertheless, this concentration profile is already a very good approximation for the expected physical interpretation (3.4).

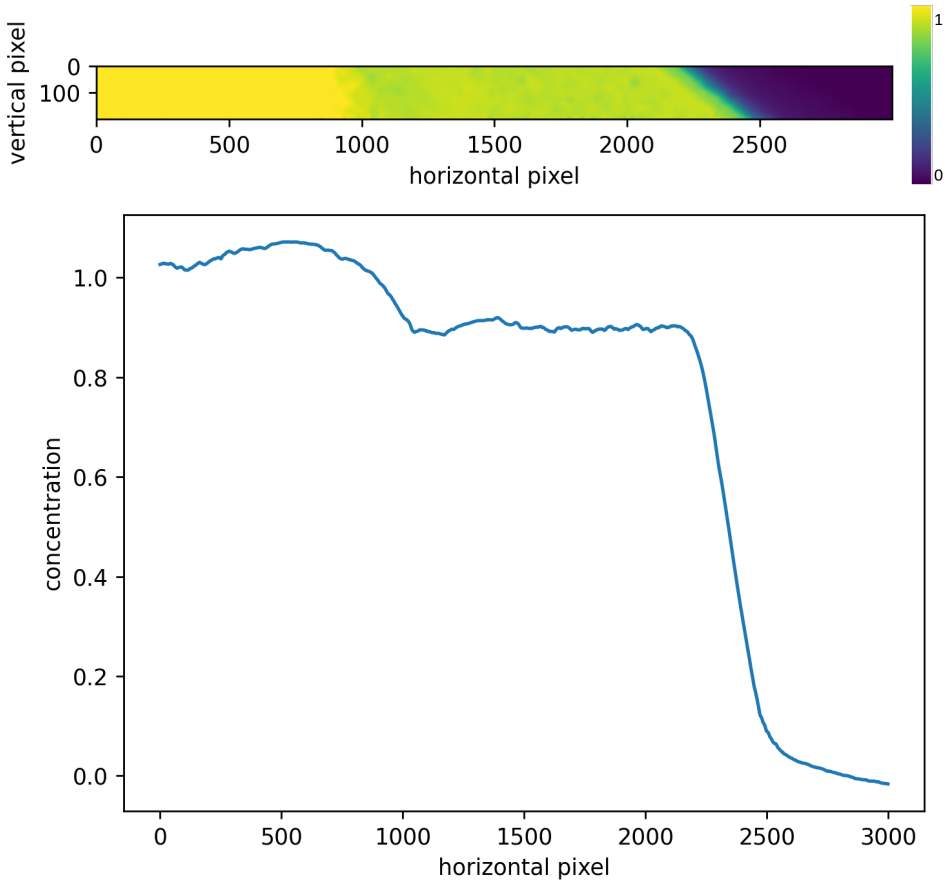


*Figure 5.10: Concentration profiles using the gaussian kernel in the LAB space with three colour concentration pairs. Top: Concentration profile, Concentration values over 1 are cut off to 1 and values below 0 are cut off to 0. Bottom: 1d reduction of the concentration profile.*

The resulting concentration profile for the LAB variant (Figure 5.10) is not monotone and the stair behaviour is not corresponding to the physical interpretation in 3.4. Overall, the profile is worse than the RGB variant (Figure 5.9).

Due to the worse performance of the LAB space both with the linear and the gaussian kernel, research question Q2 can be answered in combination with the considerations of 5.3. The RGB space is the preferred colour space in the context of the model experiment. From now on, only the RGB space will be considered.

Research question Q1 can also be answered after comparing the linear kernel and the gaussian kernel. The nearest neighbour interpolation is a simple approach, useful for fast calculations and when smoothness is not necessary. The linear kernel is useful because it does not require a parameter and it delivers good results for little amount of known data, with the drawback of non-smoothness and non-monotony. The gaussian kernel delivers smoother results than the linear kernel, but also lacks monotony.

To finally answer the research questions, in particular Q3, the behaviour and sensibility of the interpolation to additional known data pairs must be investigated.

### 5.5.4 Kernel interpolation with additional known data

The discontinuities and non-monotonous behaviour in the concentration profile can be explained because of the insufficient amount of input data (lack of known colour concentration pairs). This additional known data can be either retrieved from experiments or expert knowledge. To investigate the effect of additional known data points, the gaussian kernel is used in the RGB space using 9 instead of 3 known data pairs.

The data pairs are retrieved by choosing 9 patches that cover all tracer colours visible and using the method described in 5.2 to retrieve the characteristic colours. This is shown in Figure 5.11 for the RGB space.

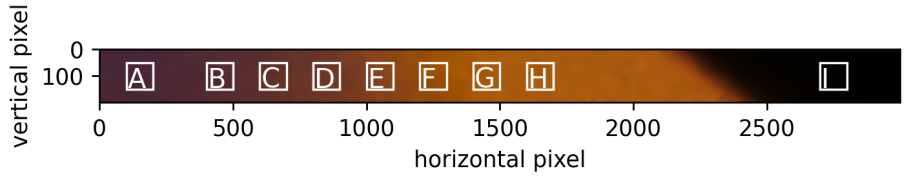


*Figure 5.11: Distribution of the patches used to retrieve additional colour concentration pairs in the smooth RGB image*

The concentration corresponding to the characteristic colour of the patches is chosen in a linear way by assigning them linear decreasing concentration values from A having 1 and H having 0.9. Patch I has concentration 0.

Applying the linear kernel interpolation method with those additional known colour concentration pairs, this leads to the concentration profiles as seen in Figure 5.12.

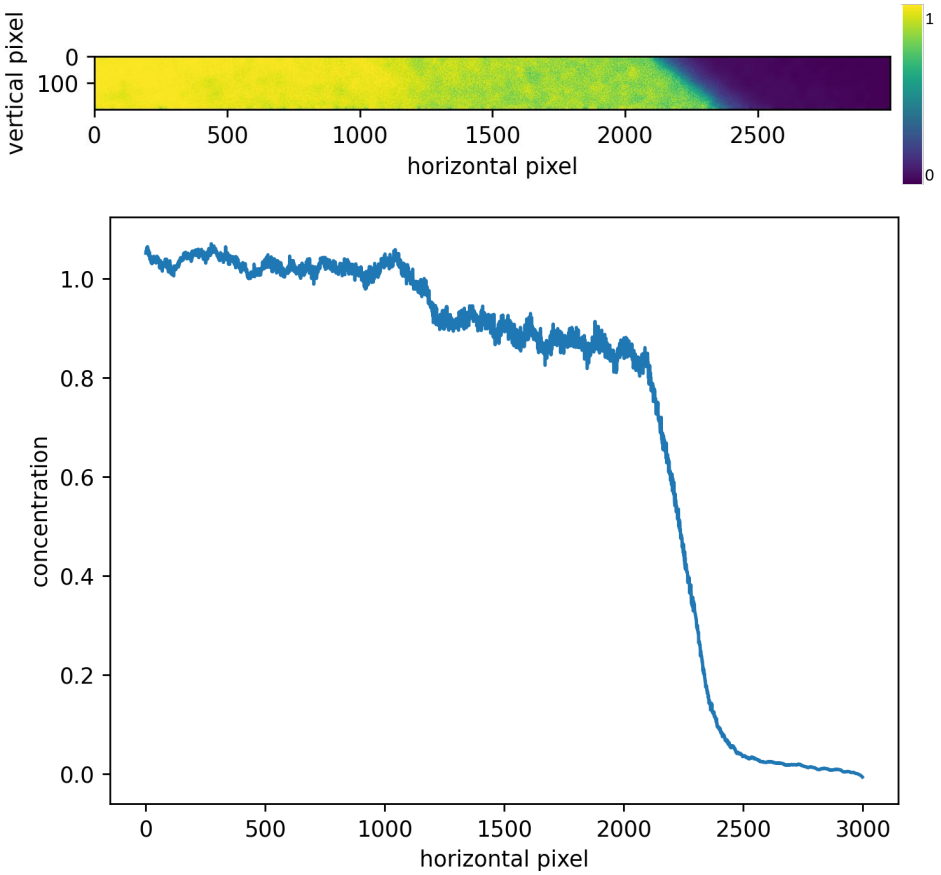


*Figure 5.12: Concentration profile (top) for the linear kernel using additional colour concentration pairs and its 1d reduction (bottom)*

The concentration profile for the linear kernel with additional known data points is less smooth than with only 3 characteristic colours (Figure 5.7) and the monotony is again missing. This result shows that the linear kernel is not beneficial for higher numbers of known colour concentration pairs.

Applying the gaussian kernel method with the parameter $\gamma = 9.73$ (same parameter as chosen in 5.5.3 for the RGB space), the resulting concentration profiles can be seen in Figure 5.13.

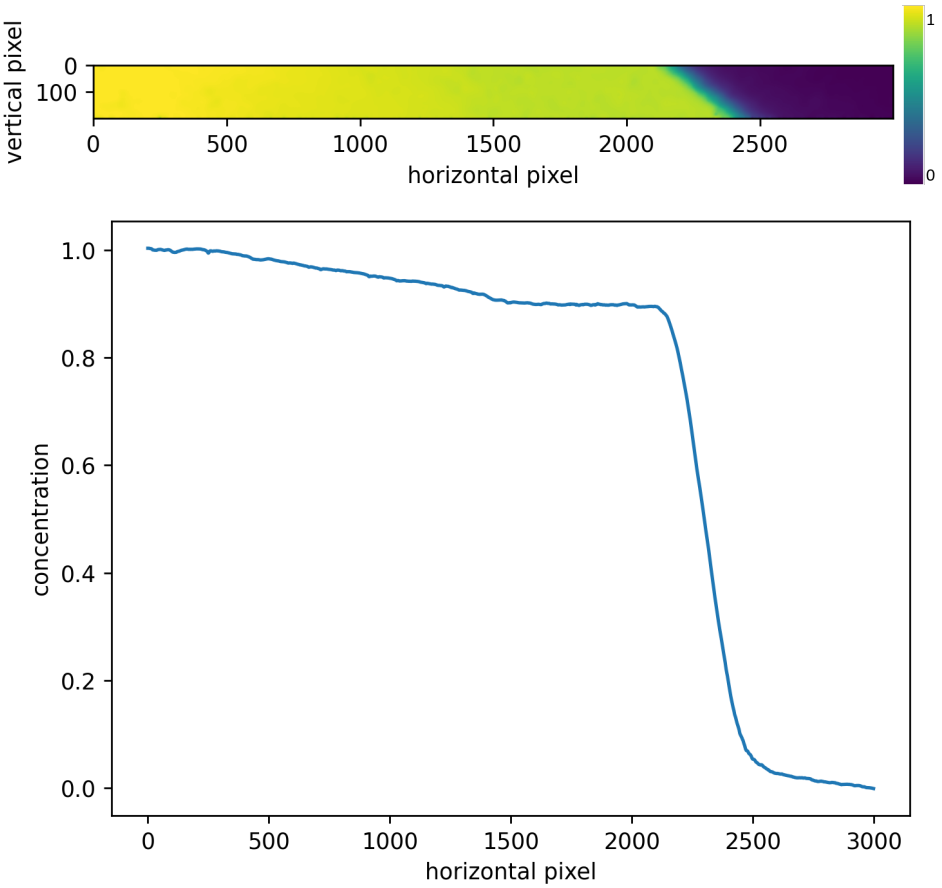


*Figure 5.13: Concentration profile (top) for the gaussian kernel using additional colour concentration pairs and its 1d reduction (bottom)*

The gaussian kernel interpolation with additional known data points delivers a smooth and monotone 1-dimensional concentration profile that corresponds to the expected physics (3.4) and the input expert knowledge of a linear slope in the concentration as defined in the additional known colour concentration pairs.

In conclusion, research question Q3 can be answered. The gaussian kernel delivers good results and maps the additional knowledge given by additional colour concentration pairs in the concentration profile. The linear kernel becomes unstable and oscillates for additional colour concentration pairs.

In conclusion, the gaussian kernel is the preferred method for an interpolation of colours to concentrations since it delivers smooth concentration profiles that correspond to the expected physics and input expert knowledge. However, for small amounts of known data, the linear kernel is an alternative option that has the advantage to not require a parameter.

The bottom line of this chapter is that the gaussian kernel is the preferred interpolation method when high numbers of known data should be considered. To test the possibility and demonstrate the ease of inputting expert knowledge to map colours to concentrations by interpolation, the following final test will be conducted.

## 5.6 Final demonstration

As a final test, to show the capabilities of introducing additional expert knowledge, the gaussian kernel interpolation method in the RGB space is applied to the image.

The known data pairs are the following.

- Known colours: Use the colours retrieved from the patches seen in 5.5.4

- Known concentrations: Use concentration 1 for patch A, 0.99 for patch H, 0 for patch I and linearly distributed values between 0.99 and 1 for patches B-G

This expert knowledge corresponds to the physics that are expected as seen in 3.4 and can be seen as a more realistic choice than in 5.2.

Using this input data, an interpolant is defined using the kernel method with the gaussian kernel, $\gamma = 9.73$ in the RGB space following the in 5.5.1 described method.

The resulting concentration profile and its reduction to 1d can be seen in Figure 5.14.

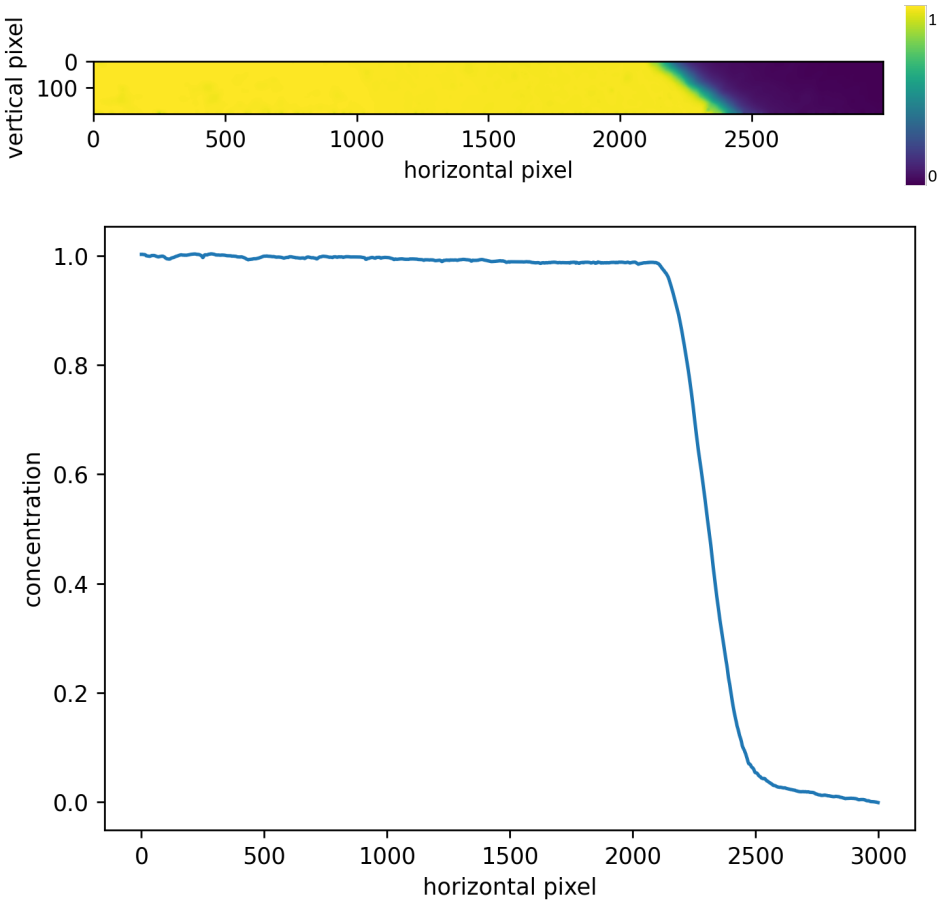


*Figure 5.14: Concentration profile (top) and its 1d reduction (bottom) for the described interpolation method*

The resulting concentration profile matches accurately the expert knowledge that has been input by the known colour concentration points as a linear slope can be seen from concentration 1 to 0.99. The transition zone to the concentration 0 is also accurately mapped according 3.4. The profile is smooth, monotone and s-shaped. For illustrative purposes, that interpolation method can also be applied to the whole tracer image as shown in Figure 5.15 to obtain a concentration map of the whole tracer image (Figure 4.1).

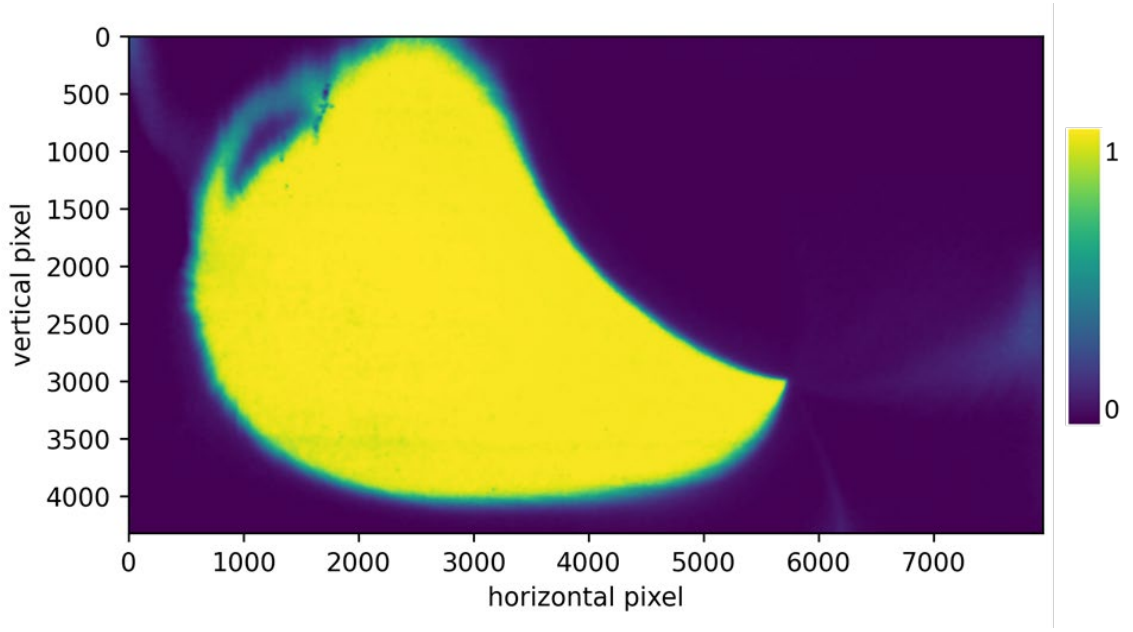


*Figure 5.15: Kernel interpolation as described in 5.5.1 and 5.6 applied to the tracer image (Figure 4.1) and a concentration colour scale (right)*

The shown results for the resulting concentration map are a success and look very promising. The physics of the porous media processes as described in 3.4 can be accurately mapped and represented in the concentration profile when using gaussian kernel interpolation based on known colour concentration pairs.

# 6 Conclusion and Outlook

The context of this work is to develop a method for analysing $CO_2$ concentrations in porous media of FluidFlower experiments. The objective was to find a method for analysing images of multichromatic tracers with the goal of extracting tracer concentrations on pixel level, which was not possible before. Novel methods have been presented in this work.

A first approach to adapt the existing DarSIA concentration analysis to multichromatic tracer information, in particular the adaptation of the signal reduction applied to images in the HSV colour space, proved unsatisfactory.

This work suggests replacing the signal reduction by an interpolation approach to determine concentration on pixel-level of tracer images using known data points. Additionally, the kernel method for interpolation is suggested, as it has shown to be promising for multichromatic tracer data. This method is suitable to convert complex colour information into physically meaningful data for quantitative analysis. The main advantage, especially when using the gaussian kernel, is a smooth and continuous interpolation that improves with the number of known data points measured in experiments or based on expert knowledge.

The kernel interpolation method has been applied to a model experiment in different variations with the following results.

(1) The linear kernel delivers good results for a small amount of known data whereas the gaussian kernel shows better results for higher numbers of known data.
(2) The RGB space has been shown as the preferred colour space in comparison to the LAB space.

Summarising these conclusions, the kernel interpolation method has been found promising to extend the functionality of DarSIA to analyse multichromatic tracer data. This extends the toolbox with better results for multichromatic tracer data while keeping the established signal reduction for monochromatic tracer data.

A possible next step should be to apply and test the extended method on real $CO_2$ FluidFlower experiments with multichromatic $CO_2$ tracers in order to verify the results. Depending on the qualitative performance, different kernels and their application in different colour spaces might be considered. The found method can be used to conduct quantitative research on FluidFlower experiments.

# References


[1] V. Eyring, N. Gillett, K. A. Rao, R. Barimalala, M. B. Parrillo, N. Bellouin, C. Cassou, P. Durack, Y. Kosaka, S. McGregor, S. Min, O. Morgenstern and Y. Sun, "Human Influence on the Climate System. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.

[2] P. Shukla, J. Skea, R. Slade, A. A. Khourdajie, R. v. Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz and (. J. Malley, "Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022.

[3] M. Bui, C. S. Adjiman, A. Bardow, E. J. Anthony and A. Boston, "Carbon capture and storage (CCS): the way forward," *Energy Environ. Sci.,* 2018.

[4] K. Eikehaug, M. Haugen, O. Folkvord, B. Benali, E. B. Larsen, A. Tinkova, A. Rotevatn, J. M. Nordbotten and M. A. Ferno, "Engineering meter-scale porous media flow experiments for quantitative studies of geological carbon sequestration," *arxiv,* 2023.

[5] M. A. Fernø, M. Haugen, K. Eikehaug, O. Folkvord, B. Benali, J. W. Both, E. Storvik, C. W. Nixon, R. L. Gawthrope and J. M. Nordbotten, "Room-scale CO2 injections in a physical reservoir model with faults," *arXiv,* 2023.

[6] B. Flemisch, "The FluidFlower International Benchmark Study: Process, Modeling Results, and Comparison to Experimental Data," *arXiv,* 2023.

[7] J. M. Nordbotten and M. A. Celia, Geological Storage of CO2: Modeling Approaches for Large-Scale Simulation, Wiley, 2011.

[8] M. Haugen, L. Salo-Salgado, K. Eikehaug, B. Benali, J. W. Both, E. Storvik, O. Folkvord, R. Juanes, J. M. Nordbotten and M. A. Ferno, "Physical variability in meter-scale laboratory CO2 injections in faulted geometries," *arxiv,* 2023.

[9] J. M. Nordbotten, B. Benali, J. W. Both, B. Brattekås, E. Storvik and M. A. Fernø, "DarSIA: An open-source Python toolbox for two-scale image processing of dynamics in porous media," 2023.

[10] J. W. Both, E. Storvik, B. Flemisch and J. M. Nordbotten, "DarSIA GitHub repository," 2023. [Online]. Available: https://github.com/pmgbergen/DarSIA.

[11] M. Marquardt, "GitHub repository for the Bachelor thesis Quantitative analysis for FluidFlower experiments using multichromatiic tracer data - scripts, data and results," 2023. [Online]. Available: https://github.com/moritzmarquardt/CA.

[12] M. Marquardt, "Propaedeutic for quantitative analysis for FluidFlower experiments using multichromatic tracer data," Bergen, 2023.

[13] A. Ogata and R. B. Banks, “A solution of the differentialequation of longitudinal dispersion in porous media: fluid movement in earth materials,” *US Government Printing Office,* 1961.

[14] L. Lapidus and N. R. Amundson, “Mathematics of adsorption in beds. vi. the effect of longitudinal diffusion in ionexchange and chromatographic columns,” *The Journal of Physical Chemistry 56,* p. 984, 1952.

[15] T. Goldstein and S. Osher, “The Split Bregman Method for L1-Regularized Problems,” *SIAM Journal on Imaging Sciences,* no. 2, pp. 323-343, 2009.

[16] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu and s.-i. contributors, “scikit-image: Image processing in Python,” *PeerJ,* p. e453, 2014.

[17] R. van den Boomgaard, “Lecture notes on Image processing and computer vision - Image histograms,” 2017. [Online]. Available: https://staff.fnwi.uva.nl/r.vandenboomgaard/IPCV20172018/LectureNotes/IP/Images/ImageHistograms.html. [Accessed June 2023].

[18] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett and J. Wilson, “SciPy 1.0: Fundamental Algorithms for Scientific,” *Nature Methods,* no. 17, pp. 261--272, 2020.

[19] D. Chen, M. Rojas, B. Samset, K. Cobb, A. Diongue Niang, P. Edwards, S. Emori, S. Faria, E. Hawkins, P. Hope, P. Huybrechts, M. Meinshausen, S. Mustafa, G.-K. Plattner and A.-M. Tréguier, “Framing, Context, and Methods. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change[,” Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021.

[20] H. Lee, J. Romero and (eds.), “Summary for Policymakers. In: Climate Change 2023: Synthesis Report.A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change,” IPCC, Geneva, Switzerland, (in press), 2023.

# Appendix

- Propaedeutic in full length

SimTech
Cluster of Excellence

University of Stuttgart
Germany

UNIVERSITY OF BERGEN
*Faculty of Mathematics and Natural Sciences*

B.SC SIMULATION TECHNOLOGY

# Propaedeutic

## Quantitative analysis for FluidFlower experiments using multichromatic tracer data


Supervisor

**Prof. Dr.-Ing. Dr.-Ing. h.c. Rainer Helmig**
Institute for Modelling Hydraulic and Environmental Systems (IWS)
Universität Stuttgart

Supervisor

**Dr. Jakub Wiktor Both**
Department of Mathematics – Porous Media Group
Universitet i Bergen

Submitted by

**Moritz Marquardt**
B.Sc. Simulation Technology
Universität Stuttgart
Matrikel Nr. 3415391

June 2023

# Contents

# Figures

# 1 Introduction to the Propaedeutic

The Propaedeutic precedes and lays the base for the studies in the bachelor thesis by covering and reviewing the current and relevant literature in that topic. The bachelor thesis with the title "Quantitative analysis for FluidFlower experiments using multichromatic tracer data" is a contribution to the DarSIA image processing toolbox with the goal of detecting and processing multichromatic tracers in images of porous media. In detail this is done at the example of multichromatic tracers in the porous structure of the FluidFlower experiments conducted at the University of Bergen with the goal of extracting the concentration of that tracer at each pixel of an image taken of the experiments. The FluidFlower experiments are carried out by the Department of Physics and the mathematical image processing part DarSIA is developed by the Department of Mathematics at the University of Bergen. This work is interdisciplinary and the propaedeutic will therefore cover a wide range of topics.

The second chapter following this introduction is about the carbon capture and storage (CCS) concept and will cover some motivational literature. This includes a short introduction in the storage aspect of CCS and specifically in the geological storage of carbons. This is the real world application for the FluidFlower experiments and the studying of $CO_2$ dynamics in porous underground reservoirs.

The third part focuses on the FluidFlower concept. This is the experimental setup that is used as an example application for the DarSIA toolbox in this thesis. A description of the experimental setup with its limitations and properties and a motivation on how the experiments are linked to the real world and geological carbon sequestration is given.

The fourth part will cover the basic concepts of describing and modelling flow in porous media. This is important in understanding the physics in the porous structure of the FluidFlower and thus understanding the phenomena that are present in the experiment. This includes a discussion on scales of porous media and the upscaling of physics which will also lay the base and motivate concepts used in the DarSIA toolbox which will be presented in the fifth chapter.

DarSIA (Darcy Scale Image Analysis) is a toolbox for analysing and processing images with the goal of retrieving physical data from experimental images. Besides a brief introduction to DarSIA, central questions to be discussed in this chapter are: What information is visible in the pictures, and how to get from pore scale picture information to Darcy scale information by image processing.

Determining concentrations of a tracer from images is done in the thesis using kernel methods for interpolation. In the last part an introduction to kernel methods is given with the application of using it for multivariate interpolation. This will cover a mathematical definition of kernel methods and a reasoning on why it is useful and suitable for the application of sparse and multivariate interpolation.

# 2 CCS and reservoir physics

Mitigating the climate change is one of the most important and relevant challenges of our time. Since $CO_2$ is one of the main drivers of the climate change, the $CO_2$ emissions must decrease drastically (IPCC, 2022). Historically the concentration of $CO_2$ in the atmosphere increases constantly as seen in the Keeling Curve (Figure 1). This development has to be turned around quickly to mitigate the climate change.

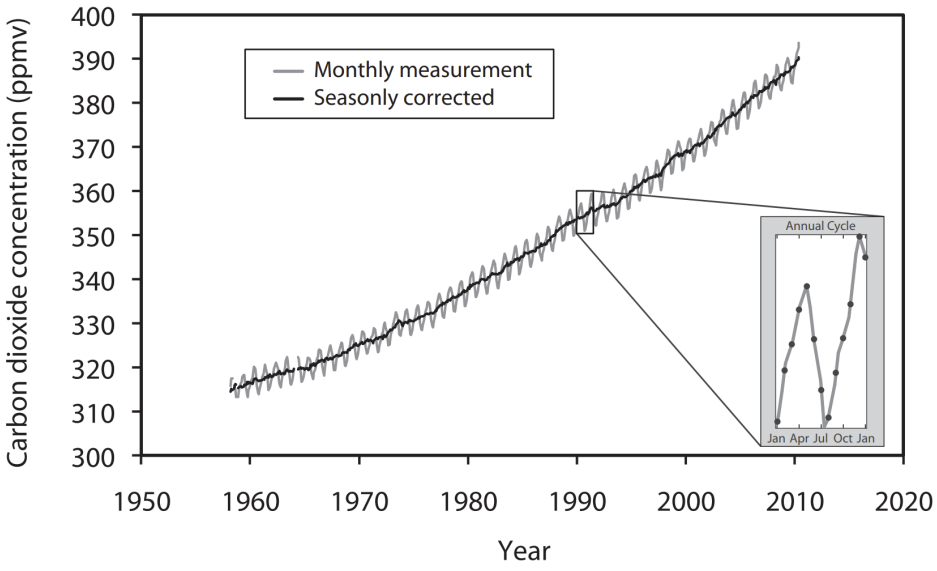


*Figure 1 $CO_2$ concentration measured at Mauna Loa (Keeling Curve) from (Nordbotten J. , 2012)*

The problem is that there exist drivers of carbon emissions that cannot be neutralized or take a long time to turn into carbon neutral processes. Because of that and the still increasing overall carbon emissions there is a need for carbon negative methods to compensate those emissions. One technique that can produce negative carbon emissions and can therefore help in decreasing the overall carbon emissions is carbon capture and storage (CCS) (IPCC, 2023) (a discussion of CCS itself is found in (Bui, Adjiman, Bardow, Anthony, & Boston, 2018)). One interesting and widely discussed approach is the geological storage of the $CO_2$ by sequestering it in underground reservoirs in so called saline aquifers (geological carbon sequestration (GCS)). Those are geological formations of porous rocks containing brine. Since the 70s $CO_2$ is injected into the subsurface, for example, in enhanced oil recovery (EOR); this means that $CO_2$ is injected into oil reservoirs to produce the oil to the surface. However, the first pure large scale industrial $CO_2$ storage project was the Sleipner project off the coast of Norway. Both, together with other kinds of geological $CO_2$ storage can be seen in Figure 2. Another current project in Norway in the field of CCS is the Northern Lights Project.

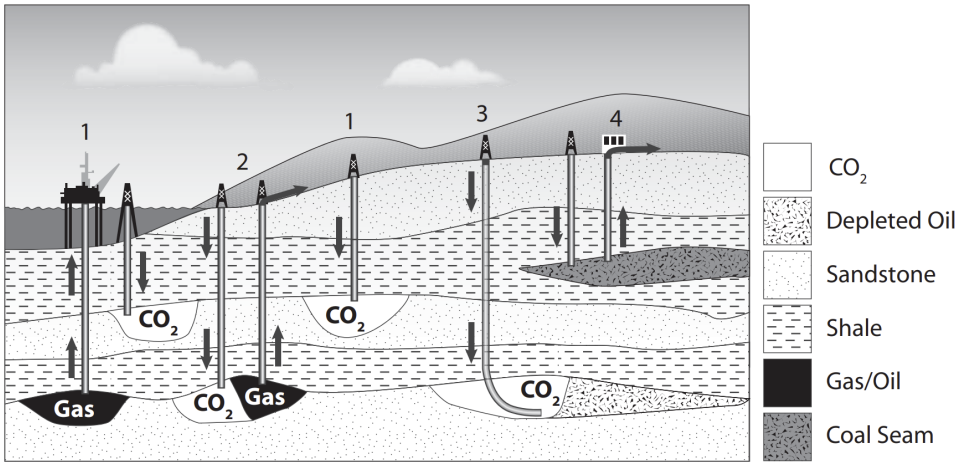


*Figure 2 Different kinds of geological $CO_2$ storage including pure $CO_2$ storage offshore and onshore from (Nordbotten J. , 2012)*

The Northern Lights Project is a key component of the Longship Project, which aims to establish large-scale carbon capture and storage (CCS) in Norway. The Northern Lights Project includes the transport of captured $CO_2$ from industrial sites across Europe and the storage in deep underground geological formations off the coast of Norway not far from Bergen.

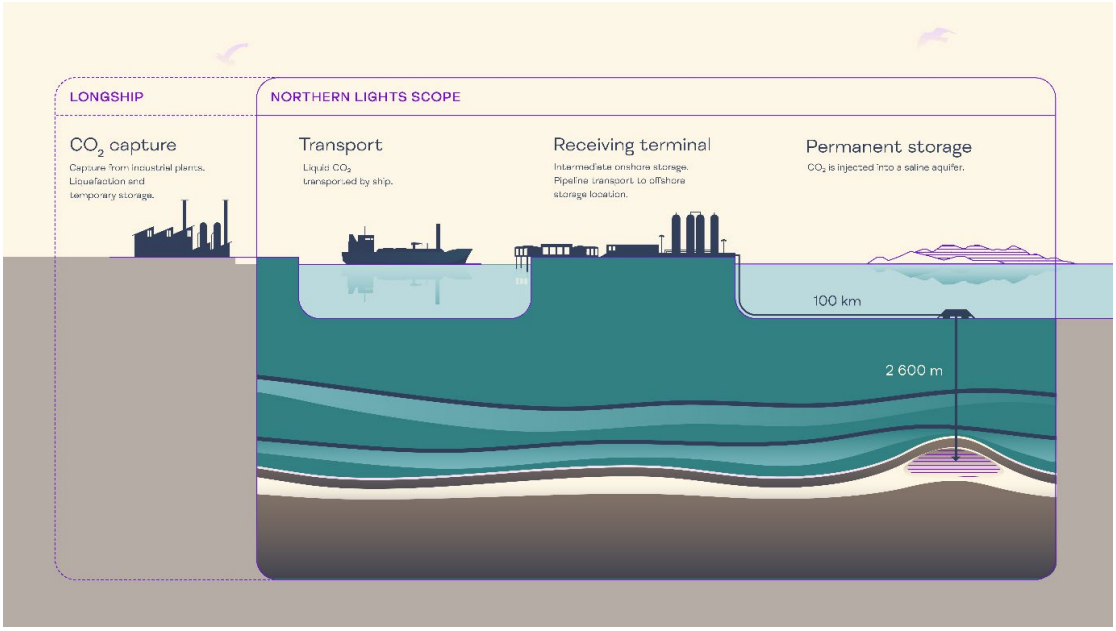


*Figure 3 Concept of the northern lights project inside the Longship project. (from https://norlights.com/ on 05/2023)*

There also exist other techniques of storing $CO_2$ like mineral storage and ocean storage which are not as mature and large scale by now as storing $CO_2$ in underground reservoirs (Bui, Adjiman, Bardow, Anthony, & Boston, 2018).

## 2.1 Geological carbon sequestration (GCS)

In this chapter the process of $CO_2$ injection and storage in a geological formation is described qualitatively following Chapter 3.10 in (Nordbotten J. , 2012).

The $CO_2$ is transported in the deep underground formations through a well and is then injected in the reservoir as also seen in Figure 2. When $CO_2$ is injected, it has a lower density than the brine and is therefore buoyant. Primarily it is trapped by structural trapping which in its simplest form means that the reservoir with a high permeability is overlain by a low permeable layer called the cap rock. This

trapping layer prevents the mobile $CO_2$ to move upwards and pass the trapping layer towards the surface. Over time, residual and solubility trapping become the main drivers for the $CO_2$ trapping. Solubility trapping is the process of $CO_2$ dissolving in the brine and increasing the density of the saturated brine which leads to the dissolved $CO_2$ propagating downwards, driven by gravity. Residual trapping describes the process of $CO_2$ being trapped inside the pore space and being held in place by capillary pressure. In the long run (over a time scale of 1000 to 10000 years), mineral trapping - the binding of $CO_2$ in the reservoir by geochemical reactions - becomes dominant as seen in Figure 4. Overall, this means that the security of the storage is increasing with time because solubility and mineral trapping are more stable and secure than pure structural trapping.

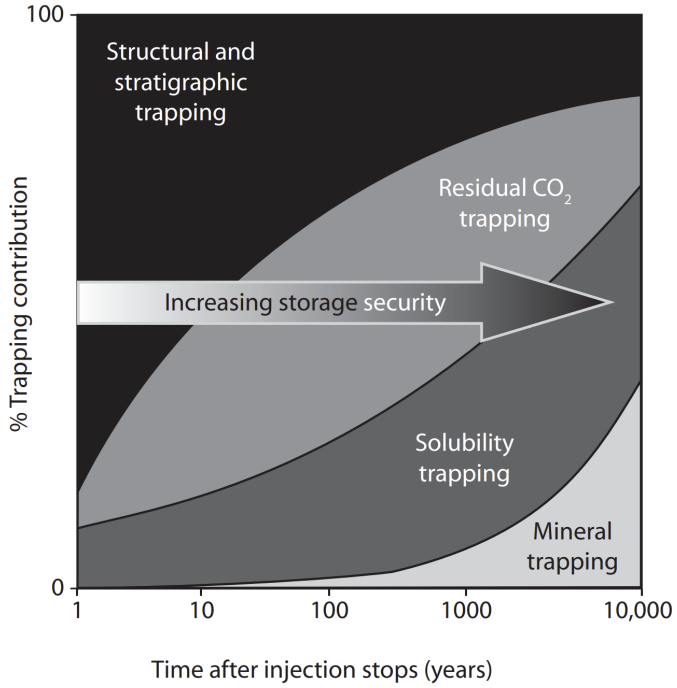


*Figure 4 Schematic of the evolution of $CO_2$ from free phase to less leakage - prone states (from IPCC, 2005) through (Nordbotten J. , 2012, p. 112)*

The main advantage of GCS is the large scalability and maturity since it is already being commercially used. However there are also disadvantages of GCS like storage security uncertainties, legal issues and the public perception.

To rank, optimize, screen and forecast those carbon underground storage sites and their storage security, it is necessary to use simulations and models of underground reservoirs at a large time scale that go far into the future. Since there exists no long-term data on carbon storage that could be used to validate and verify the simulations and models, there is a need for another solution. The FluidFlower experiments that are described in the following chapter are an approach to tackle this issue.

# 3 FluidFlower – The experimental set up

This chapter is based on the papers about the FluidFlower of Martin Fernø (Fernø, et al., 2023) and Kristoffer Eikehaug (Eikehaug, et al., 2023).

The FluidFlower is a laboratory sized experiment consisting of a quasi-two-dimensional transparent container that can be filled with unconsolidated sands and fluids. The FluidFlower exists in different scales, the room scale rig and the table size rig. The room scale rig with the dimensions of 3m x 2m has been used in the International Benchmark studies (Flemisch, 2023). The table sized rig with the

dimensions of 0.9m x 0.6m has been used in the experiments for this thesis. Quasi two dimensional means that the two dimensions, width and height, dominate the third dimension, depth, in orders of magnitude as seen in Figure 5. The sands can be arranged into geological structures (geometries) that represent real underground reservoirs. In Figure 5, in the room sized version of the FluidFlower, a geometry is visible that aims on reproducing the underground layers of the northern sea (Fernø, et al., 2023). The FluidFlower also has several injection points in the back that make it possible to inject fluids to investigate multiphase flow in those porous media structures. For example it is possible to inject $CO_2$ into the structure and trace the movements of it. This is for example done with the help of pH-indicators that are described in the chapter about fluids in the FluidFlower (3.2).

The goal of the FluidFlower concept is the possibility of making underground porous media flow accessible in the lab, to on the one hand, validate and compare porous media models to real experiments -this has been done in the International Benchmark Studies (Flemisch, 2023) where the porous media models of research groups all over the world were being tested against the FluidFlower experiment – and, on the other hand investigate $CO_2$ flow and trapping mechanisms. Hereby, the advantages of the FluidFlower are the realistic and simultaneous representation of trapping mechanisms and the high repeatability (more in 3.3). Another important aspect of the FluidFlower concept is the possibility of public outreach and to communicate the concept of CCS, as has been done in the Natural History Museum in Bergen. This setup can also be used in other contexts and applications.

The data acquisition of the experiment is done by camera with time lapse picture series. These pictures, taken in specified time steps, can be used to identify and track the dynamics and happenings in the FluidFlower by using image processing (more in chapter 5). Following standardized protocols, other data like the injection rates of fluids like $CO_2$ and the mass of displaced fluids in the overflow are measured and logged.

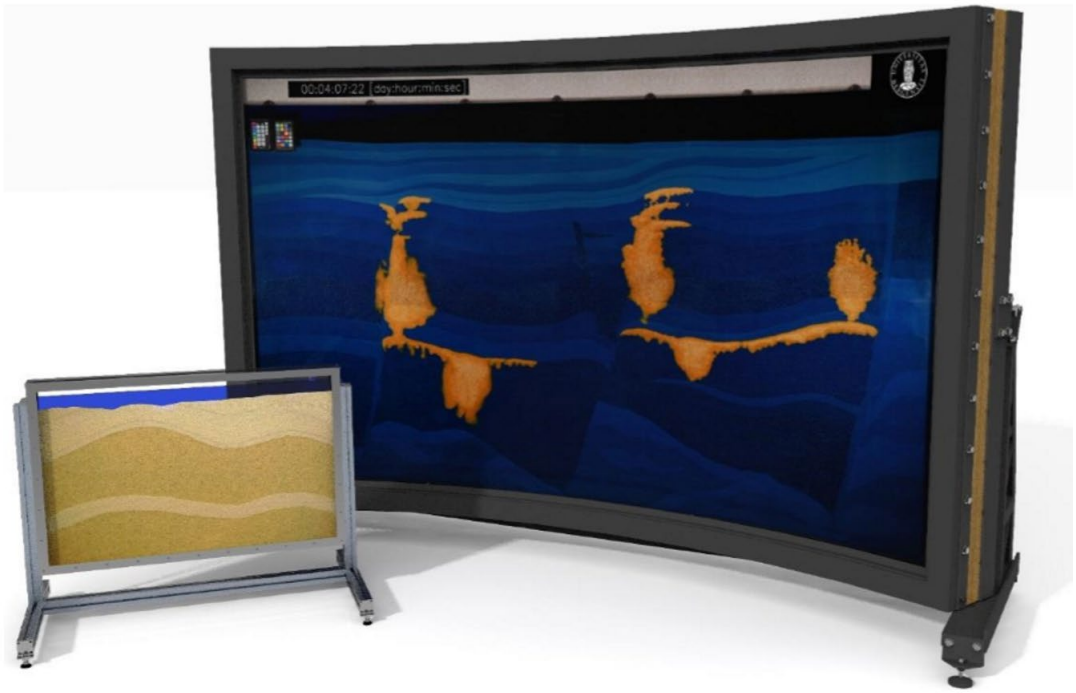


*Figure 5 The two available sizes of the FluidFlower. On the left the table sized rig and on the right the room sized rig. In the rig on the right a $CO_2$ injection in the geometry can be observed. From (Eikehaug, et al., 2023)*

The following two chapters about the setup of the FluidFlower for $CO_2$ injection experiments are based on the paper of Malin Haugen (Haugen, et al., 2023).

## 3.1 Porous media in the FluidFlower

The FluidFlower is filled with a variety of quartz sands with differing permeabilities and capillary entry pressures to replicate the characteristics of an underground geological reservoir. The sands are categorised based on their grain size, and for each grain size category, the permeability and capillary entry pressure are measured. All the sands used in the experiment have approximately the same porosity. Prior to its use in the FluidFlower experiment, the sands undergo an acid treatment to remove any carbonate impurities and are washed with tap water to prevent the occurrence of $CO_2$ mineralization. For a detailed description of the sand treatment process and property measurements, please refer to Haugen et al. (2023).

| Sand ID | Porosity [a] | K[D] [b] | Endpoint gas ($CO_2$) | | Endpoint water | | Capillary entry pressure |
|---|---|---|---|---|---|---|---|
| | | | $S_{wi}$ | $k_{rel.gas}$ | $1-S_g$ | $k_{rel.water}$ | Pc [mbar] |
| **ESF** | 0.44 | 44 | 0.32 | 0.09 | 0.86 | 0.71 | 15.0 |
| **C** | 0.44 | 473 | 0.14 | 0.05 | 0.90 | 0.93 | 3.3 |
| **D** | 0.44 | 1110 | 0.12 | 0.02 | 0.92 | 0.95 | 0.9 |
| **E** | 0.45 | 2005 | 0.12 | 0.10 | 0.94 | 0.93 | 0.26 [c] |
| **F** | 0.44 | 4259 | 0.12 | 0.11 | 0.87 | 0.72 | 0.10 [c] |
| **G** | 0.45 | 9580 | 0.10 | 0.16 | 0.94 | 0.75 | 0.01 [c] |

[a] average based on constructing each sand pack twice with the same sand sample
[b] average calculated from one ascending and descending cycle (9 datapoints) for each sand
[c] calculated from the power function of the trendline in the log-log plot of Pc against geometric mean grain width for ESF, C and D sand because gas accumulation was not detectable during the capillary entry pressure measurements.

*Figure 6 Porosity, absolute permeability K[D] and capillary entry pressure for a few of the different sands in the FluidFlower from (Haugen, et al., 2023, p. 5)*

## 3.2 Fluids in the FluidFlower

The porous sand geometry is initially saturated with a fluid consisting of an aqueous pH-indicator mixture referred to as the formation water. A pH-sensitive dye is used as the indicator that undergoes a colour change upon $CO_2$ dissolution that changes the pH-value. This colour change is essential for $CO_2$ detection using visual observation and image analysis. Different pH-indicators with varying colours and sensitivities exist, some of which not only change their colour intensity but also exhibit completely different colours depending on the $CO_2$ concentration. These are referred to as multichromatic tracers.

| Fluid (phase) | Composition | pH | Usage |
|---|---|---|---|
| Acid (aq) | Tap water with<br>- Above 0.1 M hydrochloric acid (HCl) | < 2 * | Remove carbonate impurities during sand cleaning |
| pH-indicator mix 1 (aq) | Deionized water with<br>- 0.14 mM bromothymol blue ($BTB^-$)<br>- 0.43 mM methyl red ($MRe^-$)<br>- 0.10 mM hydroxide ($OH^-$)<br>- 0.67 mM sodium ions ($Na^+$) | ~ 8.3 | Initial formation fluid to enable detecting $CO_2$ dissolution in the aqueous phase |
| pH-indicator mix 2 (aq) | Deionized water with<br>- 0.14 mM bromothymol blue ($BTB^-$)<br>- 0.10 mM hydroxide ($OH^-$)<br>- 0.24 mM sodium ions ($Na^+$) | ~ 8.3 | Initial formation fluid to enable detecting $CO_2$ dissolution in the aqueous phase |
| pH-indicator mix 3 (aq) | Deionized water with<br>- 0.75 mM bromothymol blue ($BTB^-$)<br>- 1.00 mM hydroxide ($OH^-$)<br>- 0.67 mM sodium ions ($Na^+$) | < 10.4** | Initial formation fluid to enable detecting $CO_2$ dissolution in the aqueous phase |
| $CO_2$ (g) | 99.999% - 5.0 purity | - | Injected as gaseous phase |
| Lye solution (aq) | Deionized water with<br>0.48 mM sodium hydroxide (NaOH) | < 10.7** | Reset/clean the system after experiments (flush out $CO_2$ saturated water and trapped free gas) |

* Maintained below 2 until stable. ** Depending on exposure to air

*Figure 7 Fluids used in the FluidFlower and their expected pH values from (Haugen, et al., 2023)*

Multichromatic tracers have the advantage that they can display and distinguish with higher sensitivity between more concentrations than a monochromatic tracer that only changes its colour intensity. It is therefore desirable to be able to handle these kind of tracers in the experimental design and image processing to get more accurate results (more on that in chapter 5.2). However, these multichromatic tracers are currently not compatible with DarSIA, which is only able to handle monochromatic data. The focus of this thesis is to optimize the image analysis and address the challenge of handling multichromatic tracers in DarSIA (more in 5.3).

## 3.3 Study $CO_2$ trapping mechanisms with the FluidFlower

The FluidFlower is a laboratory scale experiment that is conducted at laboratory conditions. This means the length scale of the investigated processes is somewhat smaller than in real reservoirs. Real underground reservoirs have a extend to tens or hundreds of kilometres in horizontal direction and hundreds of meters in vertical direction. The FluidFlower in contrast has dimensions of the order of a table. Also, the sands used in the FluidFlower have different properties than an actual reservoir rock. As seen when comparing the two tables Figure 6 and Figure 8, the permeability of the rocks in a reservoir is a few orders of magnitude smaller than the permeability of the sands in the FluidFlower. Also the temperature and pressure in the lab are different to the conditions in the reservoir.

| | Sandstone shallow (Viking 1240 m) | Sandstone deep (Basal 2732 m) | Carbonate shallow (Wabamun 1357 m) | Carbonate deep (Nisku 2049 m) |
|---|---|---|---|---|
| Residual saturation brine | 0.55 | 0.3 | 0.6 | 0.35 |
| Endpoint rel. perm. $CO_2$ | 0.35 | 0.55 | 0.55 | 0.2 |
| Permeability (mD) | 10 | 10 | 50 | 50 |
| Porosity (%) | 10 | 10 | 5 | 5 |

*Figure 8 Rock properties in an actual underground reservoir from (Nordbotten J. , 2012, p. 185)*

The fundamental processes of multiphase flow in porous media are still the same as at reservoir conditions and all the important trapping mechanisms are also present in the FluidFlower. A more detailed discussion on the similarity of the experiments and the real world can be found in (Fernø, et al., 2023).

Structural trapping is present with the different sand zones in the rig that have different permeabilities and capillary entry pressures. The structural trapping in the FluidFlower is caused by the contrast in the capillary entry pressure more than by the different permeabilities which would be the expected cause at field scale.

Dissolution trapping occurs when the $CO_2$ dissolves into the brine at the interface between the $CO_2$ and the originally present brine water phase. Residual trapping is also observed before the $CO_2$ dissolves into the water. Convective mixing occurs, when the water with dissolved $CO_2$ migrates downwards due to gravitation and the difference in density. This leads to fingering in the long run.

Mineral trapping is by design not part of the studies, to exclude pure chemical processes.

In conclusion, the flow and trapping mechanisms in the FluidFlower are representative for the mechanisms in the reservoir but at a different time scale. This results in an hour in the FluidFlower being equivalent to hundreds of years in the real world (See (Anthony R. Kovscek, 2023) for a detailed scale analysis).

# 4 Flow in porous media

This chapter gives a short overview over the dynamics and properties of flow in porous media and its modelling based on the book Multiphase Flow and Transport Processes in the Subsurface by Rainer Helmig (Helmig, 1997).

The FluidFlower is filled with different sands that can be described as a porous media since it is a material with pores (voids that allow fluids to flow). In section 4.2 the porosity which is one of the main characteristic of a porous medium will be presented. Injecting $CO_2$ and other fluids into the FluidFlower makes it also necessary to understand and model flow in porous media, to describe the happenings in the FluidFlower. In the FluidFlower, with two or more phases being present in the porous media, the processes are very complex. In this chapter, the basic concepts are presented to lay the basis for a more sophisticated understanding.

## 4.1 Porous Media and upscaling

A porous medium is a material containing void space called pores. This means that a lot of things around us are, to a different degree, porous. For example, our brain is a porous medium, but also rocks - including the sedimentary rocks in underground reservoirs - are porous. This stresses the importance of understanding and describing porous media.

Describing porous media is possible on different length scales. For example, on the molecular scale, all the properties of porous media and flow in porous media can be described by the interactions of the individual atoms and molecules with each other and boundary conditions. This can be done using, for example, molecular dynamics simulations. For large domains of porous media, this approach is not possible since the number of molecules is way too big and the calculations would exceed the available computational power.

At a slightly larger scale, the porous media can be described by explicitly describing the pores and their geometry. This is also not feasible when looking at larger domains of porous media because the pore structure might be unknown, impossible to find or too complicated to explicitly describe.

When looking at even larger scales, it is necessary to average the properties of porous media over a certain volume. This is a widely discussed and investigated topic in the porous media community. For different processes and different properties of the porous media it is necessary to consider different length scales. This is a very complex and interesting topic since it is necessary to describe huge systems of porous media (e.g. reservoirs) using averages while still taking into account that all the fundamental phenomena happen at the pore scale.

In this work especially the REV and Darcy scale will be of interest. The REV volume is the length scale on which porosity is defined (more in the following section) and the Darcy scale is the scale on which Darcy's law is defined (more in section 4.3).

## 4.2 Porosity and REV

Porosity is the main property to describe a porous medium. The porosity in a certain volume $V_0$ of the porous media is defined by

$$\varphi_{V_0} = \frac{V_{void}}{V_{total}}$$

with the void space (pore space) $V_{void}$ and $V_{total}$, the total volume of $V_0$.

It is easy to imagine that in an heterogenous medium the porosity changes with the choice of the size of $V_0$. For example when choosing $V_0$ very small so that it lays completely in a void space, the porosity is 1 but, when increasing $V_0$, the porosity decreases when also solid matter is included. Therefore the properties of the porous media are defined on a representative elementary volume (REV) where the parameters like porosity are ideally independent of the choice of $V_0$.

The definition of the REV is described by the example of the average mass density of a fluid in a two phase porous media system as seen in (Helmig, 1997, pp. 25, 26).

When choosing $V_0$ too small, the average mass density is not statistically representative of the whole medium. That is why, when increasing the size of $V_0$, the average mass density calculated in $V_0$ fluctuates depending on its size. These fluctuations disappear when the volume gets big enough which means that the average mass density is independent from the chosen averaging volume $V_0$. This length scale is the choice for the REV. When looking at larger length scales of $V_0$, fluctuations occur again because large heterogeneities, like fractures in the porous media, appear on that scale. This behaviour is shown in Figure 9.

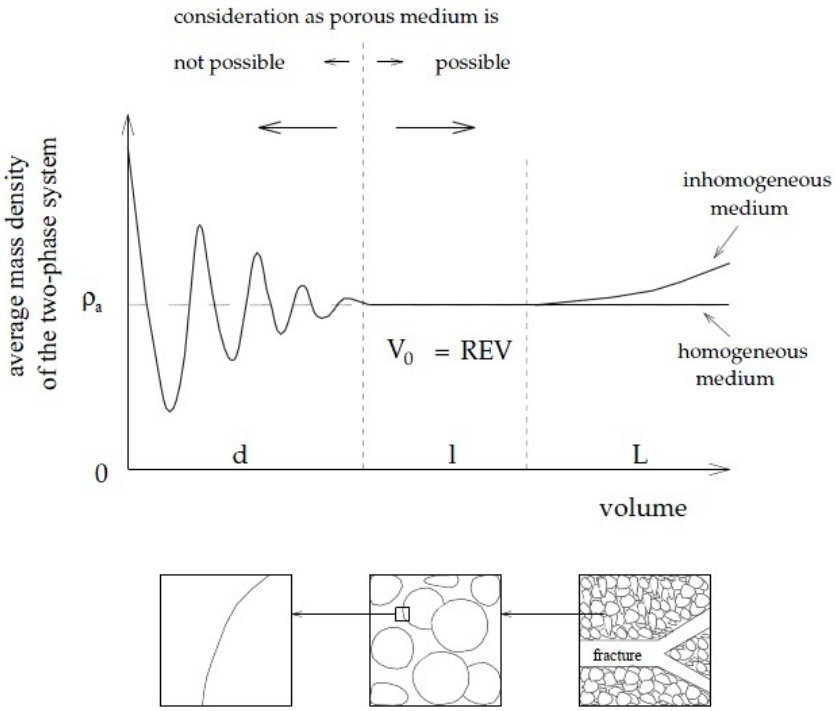


*Figure 9 Different length scales and the definition of the REV at the example of the average mass density of a fluid in a two phase porous media system. Taken from (Helmig, 1997)*

## 4.3 Darcy's law

Darcy's law is the fundamental relation to describe flow in porous media. It is based on the experiments of Henry Darcy on the flow of water through sand filters. In its simplest one dimensional form, it looks like the following

$$q_{Darcy} = -\frac{k}{\mu l}\Delta p$$

The flux $q_{Darcy} = \frac{V}{\Delta t \cdot A}$ is the flowing volume per time of a single fluid divided by the surface $A$ where it flows through. This flux is calculated using the viscosity $\mu$ of the fluid and the permeability $k$ of the medium. The length $l$ is the length of the medium and the pressure difference $\Delta p$ is the pressure potential that makes the fluid flow.

This hydraulic pressure can be calculated using the relation

$$\Delta p = \rho g \Delta h$$

With the fluid density $\rho$ and the so called hydraulic head $\Delta h$. These variables can be seen in Figure 10.

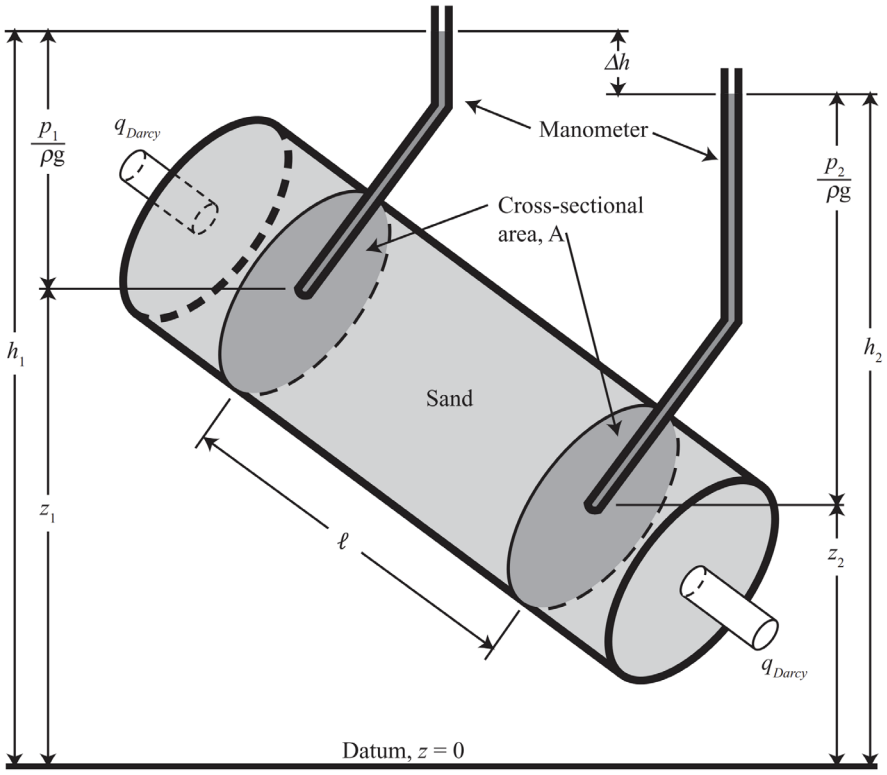


*Figure 10 Schematic of the Darcy experiment taken from (Nordbotten J. , 2012)*

The flow velocity $u$ of the fluid in the pores is different since the flow does not occur all over the surface due to the porous character of the medium. It is related with the flux via

$$u = \frac{q_{Darcy}}{\varphi}$$

It is also possible to write Darcy's law in a differential way when transferring to an infinitesimal length $l \rightarrow dl$. Also extensions for three dimensions and multiphase flow exist.

A derivation of Darcy's law from the momentum equation can be found in (Helmig, 1997).

As seen, Darcy's law exists on a scale level where the flow through porous media can be characterized by its permeability. This scale is the Darcy scale.

### 4.3.1 Mass conservation

The most fundamental term to describe a medium is the mass conservation. It states that the change of mass in a control volume (here the REV) is equal to the mass flux across its surface. It takes the following form

$$\frac{\partial}{\partial t}(\varphi\rho) + \nabla(\rho q) = h$$

With the flux $q$ and a sink or source term $h$.

This links the porosity with the mass $m = \varphi\rho$ and it links the porosity to Darcy's flow by including $q$. This is a model based on the mass conservation and it can be solved with given start and boundary values.

## 4.4 Transport in porous media

The pH-indicator that is present in the fluids of the FluidFlower can be seen as a component that flows with the fluid it is mixed into. As such, it can be described using the balance equation for a component which has the following form (Helmig, 1997, p. 115)

$$\frac{\partial(\varphi c)}{\partial t} + div(u\varphi c - \varphi D \nabla c) = r$$

with the components concentration $c = \varrho X$, the flux $q = u\varphi$, the molecular Diffusion $D$ and a sink/source term $r$.

The term $\varphi D \nabla c$ describes the non-convective mass flow, so the movement of the component that is caused by the Brownian motion and not by the flow of the fluid. The magnitude of order of this term is usually way smaller than the convection term and therefore vanishes.

The assumption is that the pH-indicator in the FluidFlower has small to no effect on the fluid flow. This is important because the indicator's purpose is to visualize physics without changing the physics being studied. In such passive tracer experiments, the concentration does not enter the fluid flow equations.

# 5 Image analysis and information upscaling

Image analysis describes the process of extracting information from an image using digital image processing. Digital image processing in general means applying algorithms or functions on an image.

In the case of the FluidFlower images, image analysis is for instance used with the goal of extracting concentration data from a photograph of the experiment. More accurately, the goal is to obtain the concentration of a chromatic tracer in the porous media in every pixel. This process also contains the upscaling of the information in the picture to Darcy scale which is equivalent to the upscaling of porous media to the REV scale (as discussed in chapter 4).

The particularity about the image processing of FluidFlower images is that it uses physical images that contain physical information, which we desire to extract, adding an extra requirement to the processing of the image in comparison to conventional techniques. The physical data must stay invariant in the images over the course of the processing.

DarSIA uses several image processing tools to extract information from the images. These include the smoothing or denoising of images, the cropping and pre-processing of images and the colour transformation of images.

A full description and all concepts are found in the paper for the DarSIA project (Nordbotten, et al., 2023).

## 5.1 Images and colour spaces

A digital two dimensional image is represented by a two dimensional grid consisting of pixels. The image can be seen as a function that assigns a value for each coordinate of those pixels and is defined on those pixels. The value it takes in each pixel can be different depending on the type of image. A coloured RGB image for example takes values from the three dimensional RGB colour space. In the following section, the RGB and the LAB colour space are introduced. However, there exist a variety of other colour spaces.

### 5.1.1 RGB colour space

In the RBG colour space, each colour is represented by a three dimensional vector (triplet) that takes the values of each of the three colour channels red, green and blue. Those values depend on how much the single colour channel contributes to the desired colour.

The range of values depends on the numerical representation of those colour channels. The most common is the 8 bit representation where the channels take integer numbers between 0 and 255. But also float numbers between 0 and 1 are helpful representations.

### 5.1.2 LAB colour space

The LAB or CIELAB colour space is a three dimensional colour space, meaning each colour is represented by a three dimensional vector. L stands for the lightness and a and b stand for the four colours magenta, green, blue and yellow. The speciality of the LAB colour space is that it is designed to be a so called uniform colour space, meaning that the perceptual difference between colours is directly related to its difference in the colour space.

A full description of the LAB colour space and the measurement of differences in that colour space is found in (Robertson, 1977).

## 5.2 DarSIA method of concentration-analysis

The concentration analysis in DarSIA is a framework for any kind of tracer concentration analysis. It has been used in the International Benchmark Studies (Flemisch, 2023) to separate the gas phase of $CO_2$ from the dissolved gas in the FluidFlower (this can be seen in Figure 11).

The concentration analysis is a collection and order of image processing routines to in the end get a map of concentrations of a tracer for each pixel of the image. The tracer is hereby a coloured tracer visible in the picture like, for example, a pH-indicator. The objective is to translate the colour of the tracer and intensity to a concentration.

In the following a few of those image processing routines that accomplish this are presented.

### 5.2.1 Pre-processing the image(s)

The first step is the pre-processing of the image. Here the image gets cropped to the region that is of interest. Also the colours of the image are corrected in order to have reproducible results. This is done by using a colour palette that is attached to the FluidFlower and visible in the image (the colour palette can be seen in Figure 5 in the top left corner of the big FluidFlower rig). Other corrections are bulk correction and curvature correction.

### 5.2.2 Smoothing and denoising the images

Smoothing the image is a trade-off between reducing noise in the image and not changing the image too much to preserve information. In the FluidFlower pictures, this means that we want to smooth over the single sand grains but still maintain edges and the distribution of the coloured tracer in the image. In DarSIA, this is done using the Rudin-Osher-Fatemi model (ROF-model). The presentation of that concept is based on the paper on DarSIA (Nordbotten, et al., 2023).

The above mentioned trade off to get the denoised image $\bar{u}$ takes the following form

$$\bar{u} = \arg \inf_{u} D(u_0, u) + N\left(u; \mu_{u_0}\right)$$

With $D(u_0, u)$ measuring the deviation between the image $u$ and the original noisy image $u_0$ and $N\left(u; \mu_{u_0}\right)$ measuring the noise of the image with the parameter $\mu_{u_0}$.

Following the ROF-model introduced in 1992 by Rudin, Osher and Fatemi (Rudin, Osher, & Fatemi, 1992) $D$ is set to be the weighted $L^2$ norm and $N$ is set to be the total variation (TV)

$$D(u_0, u; \omega_{u_0}) = \frac{1}{2} \| u_0 - u \|^2_{2,\omega_{u_0}} = \frac{1}{2} \int \omega_{u_0} (u_0 - u)^2 dV$$

and

$$N(u; \mu_{u_0}) = \mu_{u_0} |\nabla u|_1 = \mu_{u_0} \int |\nabla u| dV$$

Using the Euler Lagrange formalism, this leads to differential equations that are non-linear and expensive to solve. These can be solved numerically using, for example, the gradient descent method as proposed by the paper of Rudin Osher and Fatemi (Rudin, Osher, & Fatemi, 1992). Other more advanced methods for solving this problem are the Chambolle or the Bregman algorithm.

This is especially more expensive to solve than other denoising minimisation problems - like the Tikhonov regularisation, that uses the least squares error for the noise measurement. Nevertheless, the ROF regularisation is preferred because it preserves edges and therefore information.

This smoothing of the image over the sand grains in the FluidFlower is equivalent to the upscaling of the porous media properties from the pore scale to the REV scale. In the original image the pores of the sand grains are still visible. The smoothing then scales that up by making the concentration, which is represented by a colour in the picture, continuous over the pores. The parameter $\mu_{u_0}$ defines the scale of the smoothing. Higher values of $\mu_{u_0}$ means more smoothing and less details. A more detailed discussion on restoration and smoothing of DarSIA images can be found in (Nordbotten, et al., 2023).

### 5.2.3 Signal to concentration reducing and modelling

After smoothing the image, the final step is the translation of the colour to a concentration. Currently this is done in DarSIA in two steps. First, the colour is reduced to a monochromatic signal, meaning a scalar value of the signal for each pixel. Then, this value is modelled to a concentration using a simple linear model. Those models can also be calibrated when a whole time series of pictures exists.

## 5.3 Limitations

DarSIA has been successfully used in the international benchmark studies for the phase separation. Here, the concentration analysis only distinguishes between the tracer being present, which delivers concentration 1, and no tracer being present, which delivers the concentration 0 (see Figure 11 left).

It is possible to extract continuous concentration profiles for monochromatic tracers that change their colour intensity with their concentration using the above described methods of signal reduction and signal to concentration modelling.

However, currently it is not possible to analyse multichromatic tracers that take different colours depending on their concentration like for example pH-indicators. This is the main motivation for the thesis following up this propaedeutic since those indicators are able to deliver more accurate concentration data and results (an example multichromatic tracer experiment can be seen in Figure 11 on the right).

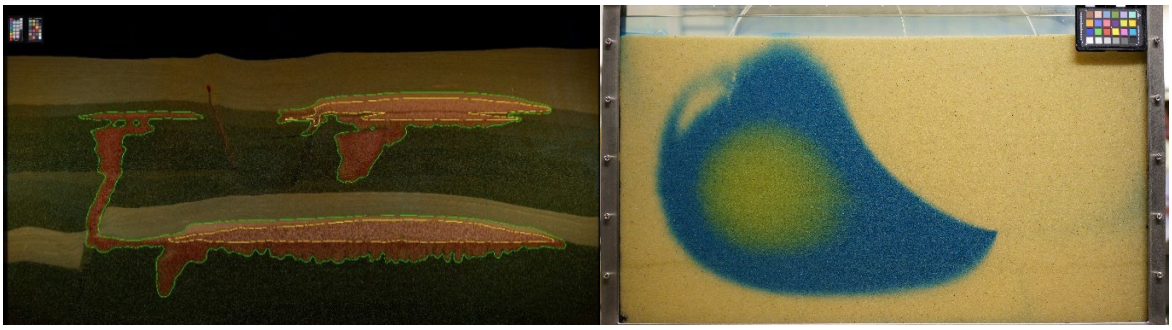

*Figure 11 Phase segmentation in the Benchmark studies (left) and a multicolour tracer experiment (right)*

# 6 Interpolation with kernel methods

Kernel methods are an approach for approximating and predicting data such as for example interpolation. This approach is especially popular in machine learning because it can easily handle higher dimensional data, as seen in the following sections. This chapter follows the lecture notes of Prof. Haasdonk on Approximation with Kernel Methods at University of Stuttgart in 2022 (Haasdonk, 2022).

## 6.1 Definition of kernel methods

The central element of the kernel methods is a kernel which is a symmetric function

$$k\colon \Omega \times \Omega \longrightarrow \mathbb{R}$$

on an arbitrary non-empty set $\Omega$. For example this could be the polynomial kernel

$$k_p(x_1, x_2) = (\langle x_1, x_2 \rangle + a)^b$$

with the parameters $a, b \in \mathbb{R}$. The simplest case of that polynomial kernel with $a = 0$ and $b = 1$ is the linear kernel

$$k_l(x_1, x_2) = \langle x_1, x_2 \rangle$$

Another example is the gaussian kernel

$$k_g(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2)$$

with the parameter $\gamma > 0$.

The model, that is then constructed is typically of the form

$$s(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i),\ x \in \Omega$$

with coefficients $\alpha_i$ and support vectors $\{x_i, \dots, x_N\} \subset \Omega$. The function $s(x)$ is called the approximant and determining the coefficients $\alpha_i$ is called training the model.

## 6.2 Interpolation using kernels

Using kernel methods is one of many possibilities for multivariate interpolation. Multivariate interpolation is the interpolation of a function of more than one variable which is the case if $\Omega$ has dimension $d > 1$. The points that are going to be interpolated are called scattered data points.

### *Interpolation problem definition*

Given a scattered data set which is a set of points $\{x_i, \dots, x_N\} \subset \Omega$ with corresponding target values $\{y_i, \dots, y_N\} \subset \mathbb{R}$, the interpolation problem is to find an interpolant $s\colon \Omega \to \mathbb{R}$ that satisfies $s(x_i) = y_i$.

*Kernel interpolation*

It is possible to use the aforementioned kernel methods to interpolate such scattered data. The interpolant is obtained by solving

$$Ka = y$$

with the kernel matrix $K = \left(k(x_i, x_j)\right)_{i,j=1}^{N}$, the support points $\{x_i, \dots, x_N\} \subset \Omega$, the target values $y = (y_i)_{i=1}^{N}$ and the coefficient vector $a = (a_i)_{i=1}^{N}$.

The interpolant is then the approximant from above and is defined via

$$s(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i),\ x \in \Omega$$

### 6.2.1 Existence and uniqueness of kernel interpolation

To show the advantages of kernel methods, a definition of definiteness and the Mairhuber-Curtis theorem will be presented in the following.

*Definition of kernel definiteness*

The kernel $k$ is denoted strictly positive definite (s.p.d.) if for all $n \in \mathbb{N}$ and all $X = \{x_i\}_{i=1}^{n} \subset \Omega$ with distinct points, i.e. $x_i \neq x_j, i \neq j$ holds

$$\alpha^T K \alpha > 0 \qquad \forall \alpha \in \mathbb{R}^n \setminus \{0\}$$

This condition is for example satisfied for the Gaussian kernel.

*Theorem: Existence and uniqueness of kernel interpolation*

Let $\{x_i, \dots, x_N\} \subset \Omega$ be distinct data points and the kernel $k$ being s.p.d.

Then there exists a unique solution of the form $s(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i),\ x \in \Omega$ that satisfies the interpolation condition $s(x_i) = \mathrm{y_i}$.

### 6.2.2 Advantages of kernel interpolation

To discuss the advantages of the kernel approach the Haar space is defined and the Mairhuber-Curtis theorem is presented.

*Definition of Haar space*

Let $\Omega \subset \mathbb{R}^d$ contain at least $N$ points and $V \subset \mathcal{C}(\Omega)$ be an $N$-dimensional linear space. Then $V$ is called a Haar space of dimension $N$ on $\Omega$ if for arbitrary distinct points $\{x_i\}_{i=1}^{N} \subset \Omega$ and arbitrary $\{y_i\}_{i=1}^{N} \subset \mathbb{R}$, there exists a unique function $s \in V$ with $s(x_i) = y_i$ for $1 \leq i \leq N$.

*Mairhuber-Curtis Theorem*

Let $\Omega \subset \mathbb{R}^d, d > 1$ be a set with nonempty interior. Then there exists no Haar space on $\Omega$ of dimension $N > 1$.

A proof can be found in (Haasdonk, 2022).

As a consequence of this theorem, the ansatz-space for the interpolation $V$ has to be chosen dependent on the data. This is, for example, fulfilled with the kernel approach which makes it appropriate for higher dimension interpolation.

In comparison, multivariate polynomial interpolation has the problem of dimensionality. For fixed p and growing dimension d, the dimension $\dim\left(\mathbb{P}_p\left(\mathbb{R}^d\right)\right) = \binom{d+p}{p}$ of the polynomial interpolation

ansatz space grows exponentially and therefore requires a lot of interpolation points to solve the interpolation problem.

In conclusion, the kernel methods are desirable to use for multivariate interpolation.

## 6.3 Example

In the case of this work, $\Omega = RGB \subset \mathbb{R}^3$ becomes the colour space. The scattered data points are colours $c_i \in \Omega$ in the colour space RGB and $v_i \in \mathbb{R}$ are their corresponding pH-values.

Interpolating these pH-values between the given colours kernel interpolation as described in the previous section can be used by applying the described procedure.

This is being used to develop and extend the functionality of DarSIA, the implementation and methodology is described in the thesis itself.

Without intending to complicate the limited presentation in this propaedeutic report, it is important to note that, in the realization of the interpolation for this work, the space $\Omega$ is actually not the RGB space anymore because the subtraction of images produces negative colours. Still, the above example provides and communicates the essence.

# 7 References


Anthony R. Kovscek, J. M. (2023). Scaling up FluidFlower results for carbon dioxide storage in geological media. *Arxiv*. Retrieved from https://arxiv.org/abs/2301.09853

Bui, M., Adjiman, C. S., Bardow, A., Anthony, E. J., & Boston, A. (2018). Carbon capture and storage (CCS): the way forward. *Energy Environ. Sci.* doi:10.1039/C7EE02342A

Eikehaug, K., Haugen, M., Folkvord, O., Benali, B., Larsen, E. B., Tinkova, A., . . . Ferno, M. A. (2023). Engineering meter-scale porous media flow experiments for quantitative studies of geological carbon sequestration. *arxiv*.

Fernø, M. A., Haugen, M., Eikehaug, K., Folkvord, O., Benali, B., Both, J. W., . . . Nordbotten, J. M. (2023). Room-scale CO2 injections in a physical reservoir model with faults. *arXiv*.

Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Association for Computing Machinery*. doi:10.1145/358669.358692

Flemisch, B. (2023). The FluidFlower International Benchmark Study: Process, Modeling Results, and Comparison to Experimental Data. *arXiv*. Retrieved from https://arxiv.org/abs/2302.10986

Haasdonk, P. B. (2022). *Lecture notes on Approximation with Kernel Methods.*

Haugen, M., Salo-Salgado, L., Eikehaug, K., Benali, B., Both, J. W., Storvik, E., . . . Ferno, M. A. (2023). Physical variability in meter-scale laboratory CO2 injections in faulted geometries. *arxiv*.

Helmig, R. (1997). *Multiphase Flow and Transport Processes in the Subsurface.* Springer Verlag.

IPCC. (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.). Cambridge, UK and New York, NY, USA: Cambridge University Press. doi:10.1017/9781009157926

IPCC. (2023). *Summary for Policymakers. In: Climate Change 2023: Synthesis Report.A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Core Writing Team, H. Lee and J. Romero (eds.). IPCC, Geneva, Switzerland, (in press).

Nordbotten, J. (2012). *Geological storage of CO2: Modeling Approaches for Large-Scale Simulation.* Wiley.

Nordbotten, J. M., Benali, B., Both, J. W., Brattekås, B., Storvik, E., & Fernø, M. A. (2023). DarSIA: An open-source Python toolbox for two-scale image processing of dynamics in porous media. doi:10.48550/ARXIV.2301.05455

Robertson, A. R. (1977). The CIE 1976 Color-Difference Formulae. *Color Research & Application*. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1520-6378.1977.tb00104.x

Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*. Retrieved from https://doi.org/10.1016/0167-2789(92)90242-F