



UiT The Arctic University of Norway

Faculty of Science and Technology, Department of Mathematics and Statistics

Statistical Curve Analysis: Developing Methods and Expanding Knowledge in Health

Stig Uteng

A Dissertation for the Degree of Philosophiae Doctor

April 2022

Abstract

The analysis of curves can be claimed to be the core of most scientific ventures. In this dissertation, we focus on the statistical aspect of this type of analysis. Here, the curves originate from health and food-related areas and include improvements in blood glucose measurements, classification of moles, measurements of parameters during liver transplants in pigs, and data from the monitoring of the quality of fish. More specifically, the statistical curve analysis consists of several perspectives were all have some kind of intrinsic comparison effort. However, the main approaches in these studies are related to regression and the problem of finding suitable critical regions. The regression part consists of robust nonlinear regression and linear mixed models while the critical regions are found through classification and hypothesis testing in scale-space. By improving the critical decision boundaries through e.g. the Bonferroni correction of scale-space maps in Paper I, and developing features to improve decisions regarding the classification of moles in Paper II, we were able to obtain high sensitivity and specificity in the developed systems. Regression was an integral part of the classification effort in Paper II, the improvement of blood glucose measurements in Paper III, and the statistical analysis of parameters measured during liver transplantation in pigs in Paper IV.

Paper I is focused on maximizing sensitivity and specificity when detecting a significant change in the data. Here as in Paper II hyperspectral images are the source of data. The developed method produces a scale-space, where significant changes can be detected.

Paper II aims to maximize sensitivity, specificity, and precision in the classification of moles. This is accomplished through curves from subimages obtained from each channel of the hyperspectral images. These curves show characteristic features from three important classes of moles. By using these features through the regression of these curves, we accomplish high sensitivity, specificity, and precision in the classification pursuit.

In Paper III, we introduce a novel method for improving blood glucose estimation from continuous glucose measurements by using deconvolution. First, regression is used to estimate the parameters in the convolution kernel. Thereafter this response function was deconvolved through regression. In this way, we can estimate blood glucose from subcutaneous measurements. This gives a new method for controlling blood glucose levels which is of great importance for type 1 diabetes patients during and after exercise to avoid hypoglycemia.

Testing two different methods in liver transplantation of pigs, where the statistical analysis of curves was done through the application of linear mixed models, is the focus of Paper IV. An important output of this work is that the two treatments can be statistically distinguished through the use of linear mixed models.

List of Publications

- Paper I** Early detection of change by applying scale-space methodology to hyperspectral images, **Stig Uteng, Thomas Haugland Johansen, Jose Ignacio Zaballos, Samuel Ortega, Lasse Holmström, Gustavo M. Callico, Himar Fabelo and Fred Godtlielsen** *Applied Sciences* 2020; Vol. 10 (7)
- Paper II** Curve-based Classification Approach for Hyperspectral Dermatologic Data Processing, **Stig Uteng, Eduardo Quevedo, Gustavo M. Callico, Irene Castaño, Gregorio Carretero, Pablo Almeida, Aday Garcia, Javier A. Hernandez and Fred Godtlielsen** *Sensors* 2021; Vol. 21 (3)
- Paper III** Glucose level during Physical activity, **Giovanni Sebastiani, Stig Uteng, Fred Godtlielsen, Jan Polák and Jan Brož** *International Journal of Biology and Biomedical Engineering* 2020 Vol. 14
- Paper IV** Machine Perfusion of the Liver to Resuscitate and Reverse Ischemic Liver Injuries, **Stig Uteng, Anmar Numan, Geir Nedredal and Fred Godtlielsen** *Unpublished draft*

Acknowledgements

First and foremost I will thank my wife Gudrun for being the best and especially for being so patient. Second, I will thank Maje and Frida for being the great girls that they are. Third, I will thank Professor Fred Godtliebsen for making this dissertation possible. Also, I will thank my colleagues at Rødbygget for being very supportive.

Contents

Abstract	i
List of Publications	iii
Acknowledgements	v
Contents	vi
I Introduction	1
1 Background	3
1.1 Health-related background	3
1.2 Technical background	5
2 Methodology	7
2.1 Statistical Curve Analysis by Comparison and Classification	8
2.2 Statistical Curve Analysis by Regression	10
3 Summary and discussion	13
3.1 Paper I - Change detection in Hyperspectral Images by Statistical Curve analysis by Scale-space techniques	13
3.2 Paper II - Classification of Hyperspectral Images by Statistical Curve analysis	17
3.3 Paper III - Improvement of Blood Glucose Estimation by Statistical Curve analysis	19
3.4 Paper IV - Quantifying group difference by Statistical Curve analysis	20
3.5 Concluding remarks	25
Bibliography	27
II The Papers	33
4 Paper I: Early Detection of Change by Applying Scale-space Methodology to Hyperspectral Images	35

5	Paper II: Curve-based Classification Approach for Hyperspectral Dermatologic Data Processing	49
6	Paper III: Estimation of Blood Glucose Concentration During Endurance Sports	63
7	Paper IV: Machine Perfusion of the Liver to Resuscitate and Reverse Ischemic Liver Injuries	69

To Gudrun

Part I

Introduction

Background

One of the main incentives for research is health. We have seen astonishing examples of health-driven research during the Covid 19 pandemic, with the incredibly fast development of vaccines. One of the main reasons for this development speed was the application of data science, where statistics play an integral part (Waltz, 2021). In this dissertation, we will also apply statistics in health-related areas ranging from food quality to liver transplantation.

The dissertation is divided into two main parts. In Chapter 1, necessary background information is presented to give an understanding of the context in which the research is placed. It begins by giving an overview of the health-related subjects this dissertation consists of before moving on to the more technical subject, hyperspectral imaging, utilized in this dissertation. In Chapter 2 we cover the various methodologies used in the four included papers. Further, in Chapter 3, summaries and discussion of the four research papers are presented, which includes detailed lists of the contributions by the author. The second and final part is chapters 4–7, which consists of the four papers included in the dissertation.

1.1 Health-related background

1.1.1 Food quality

Food quality is of high importance due to the severe consequences of contaminated or degraded food on human health. We have now developed a considerable amount of technology to prevent food spoilage. Still, this is an important issue and an active research area, where currently the research focus is microbiological processes (Bevilacqua et al., 2016). It will, however, always be interesting to monitor these processes. Due to this, the aim of Paper I is a method for monitoring and detecting changes through the utilization of hyperspectral images (HSIs). This has been analyzed through a novel scale-space method (Uteng et al., 2020). The developed methodology has been applied to frozen fish to potentially detect when it starts to deteriorate.

1.1.2 Skin cancer

Skin cancer is one of the most common forms of cancer, with more than 1.5 million new cases worldwide in 2020 according to [Sung et al. \(2021\)](#), where most of the cases occur among the white population in Australia, New Zealand, and South Africa ([Wild et al., 2020](#)). Pigmented skin lesions (PSLs) consist of a wide variety of types including cancerous and non-cancerous PSLs ([Elder et al., 2018](#)). The diagnosis of PSLs is performed by dermatologists through visual inspection or dermoscopic cameras to inspect and analyze the lesion according to the ABCDE (Asymmetry, Border irregularity, Color, Diameter, and Evolving size, shape or color) rule to establish a preliminary diagnosis ([Jensen and Elewski, 2015](#)). Malignant melanoma is by far the most dangerous type of skin cancer, due to its rather high mortality rate ([Sung et al., 2021](#)). For this cancer type, early detection is crucial in an effective treatment because the cancer can be eliminated before it becomes metastatic. In fact, non-metastatic melanoma has been reported to have a 99% 5-year survival rate, but for metastatic melanoma, the 5-year survival rate drops to 25% ([American Cancer Society, 2022](#)). The findings of Paper II indicate that the classification of PSLs can be improved. Hence, it may be a valuable contribution to the important early phase of the treatment of malignant PSLs ([Uteng et al., 2021](#)).

1.1.3 Diabetes

One of the diseases with the highest global incidence is diabetes, counting 537 million adults (20-79 years) as of 2021, i.e. approximately 10% of the world population according to the [International Diabetes Federation \(2021\)](#). About 90% of these are diagnosed with type 2 diabetes. We will focus on type 1 diabetes mellitus (T1D) which results from failure of the pancreas to produce enough insulin due to loss of beta cells, caused by an autoimmune response ([Norman and Henry, 2015](#)). Due to the lack of insulin production, it has to be injected, and the amount has to be regulated according to measured glucose levels in the blood. The current direction is to automate these measurements and equipment to increase the measurement frequency is available today. This low invasive technology is denoted continuous glucose monitoring (CGM). Utilizing CGM, glucose concentration in subcutaneous interstitial tissue can therefore be reported e.g every five minutes ([Mian et al., 2019](#)). In physical activity (PA) a drop in blood glucose (BG) can result in hypoglycemia for T1D-patients and this poses a major fear for these patients. To this end, Paper III describes an improved method for estimating BG so that T1D-patients can perform PA more safely ([Sebastiani et al., 2020](#)).

1.1.4 Liver transplantation

Liver transplantation, also called hepatic transplantation, is the replacement of a non-functioning liver with a healthy liver from another person. This is a treatment option for end-stage liver disease and acute liver failure, although the availability of donor organs is a major limitation. To increase the donor pool, patients with irreversible brain damage in the intensive care unit is

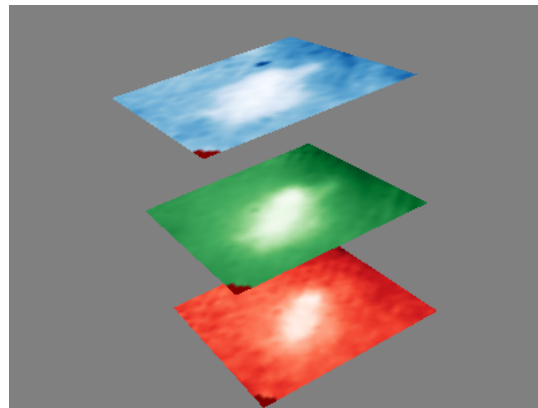
now becoming a new source of organs. These patients do not fulfill the brain death criteria. Withdrawal of life support will however cause the fulfillment of the death criteria. The condition when this occurs is called warm ischemia (WI), which means that oxygenated blood is being restricted to tissues, muscle groups, or organs of the body that is needed for cellular metabolism (i.e. to keep tissue alive). Machine perfusion techniques are currently being introduced into the clinic, with the aim of optimizing injured livers before implantation. There are two major schools in machine perfusion: normothermic machine perfusion (NMP) and hypothermic machine perfusion (HMP) (Marecki et al., 2017, Schlegel et al., 2019). Neither has yet proven its superiority. Machine perfusion may reverse the negative responses caused by the warm ischemia, and in addition, allow a performance assessment of the liver before transplantation. Two different methods of machine perfusion are investigated in Paper IV (Uteng et al., 2022).

1.2 Technical background

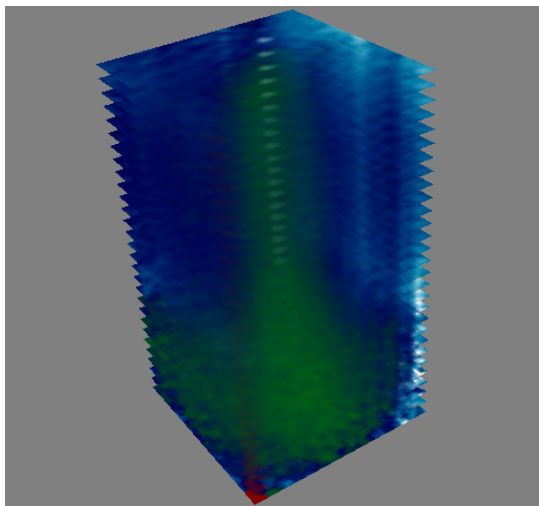
1.2.1 Hyperspectral images

HSI (Hyperspectral Imaging) is a non-destructive and non-invasive technology that combines conventional imaging and spectroscopy in one single mode (Li et al., 2013). In recent years, HSI, which was originally developed for remote sensing applications (Goetz et al., 1985), has become one of the most felicitous techniques for fast, non-destructive detection of food quality or PSL characteristics. The human visual system is, on average, considered to only perceive light with wavelengths in the approximate range from 380 to 750 nanometers, with long wavelengths - perceived as red, medium wavelengths - perceived as green, and short wavelengths - perceived as blue (Slinney, 2016). Most digital photography systems capture data across distinct bands of the electromagnetic (EM) spectrum, where each band corresponds to the primary colours: red, green, and blue (RGB), which corresponds to the human visual system. Each pixel in an RGB digital image is assigned the recorded luminance (amount of light) values for the respective spectral bands of each of the primary colors. The main objective of HSI is to assign each pixel with a very large amount of wavelength measurements, sampled uniformly across the entire bandwidth of the sensor, which then increases the opportunity to detect more than by only using information from an RGB image. HSI allows interactive analysis between spectral and surface properties of samples based on each pixel point of images acquired with a wide wavelength range of spectral information such as ultraviolet (UV), visible, near-infrared (NIR), or infrared (IR). The difference between an RGB image and an HSI can be observed in Fig. (1.1). Each pixel in the RGB image seen in Fig. (1.1(a)) will only store three values, but each pixel of an HSI will store the spectral signature of the material located in the corresponding coordinate, consisting of 125 values for the chosen camera. This creates a HSI-cube, with the ordinary two spatial image axes and an additional channel or frequency axis as seen in Fig. (1.1(b)) with its corresponding spectral signature curve in Fig. (1.1(c)). To get a better view of different signatures, Fig. (1.1(b)) shows

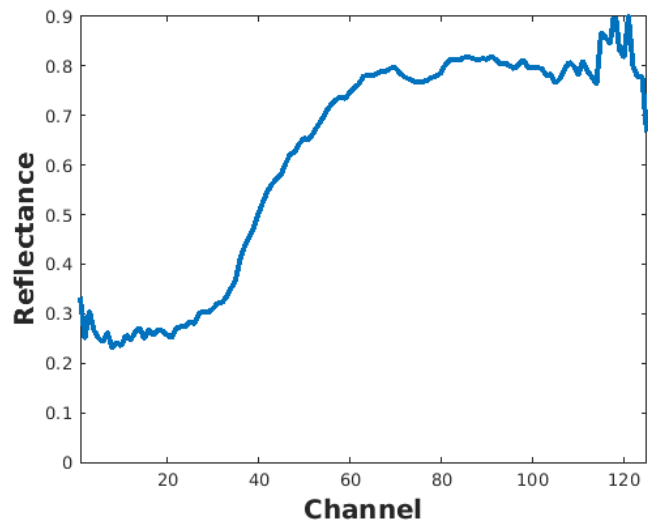
only 30 of the observed frequencies.



(a)



(b)



(c)

Figure 1.1: (a): RGB slices of a PSL analyzed in Paper II. (b): Part of the HSI cube of this PSL. The PSL can be seen as the green area through the cube. Typically, more texture is visible at the bottom of the cube. (c): The corresponding spectral signature curve of the same PSL at the coordinate $(x, y) = (17, 17)$. Reflectance is the surface's effectiveness in reflecting radiant energy.

In the applications in papers I and II, the channel axis consisted, as mentioned above, of 125 channels. This gives a rather high amount of information of each image coordinate, i.e. pixel, which is exploited in these papers.

Methodology

Curves have always been of fundamental importance in all scientific endeavors. In this dissertation curves ranging from pixel values of hyperspectral images (HSIs) to measured parameters from liver transplantation of pigs are analysed. To accomplish this curve analysis, mainly three fundamental statistical methods have been applied: comparison, classification, and regression. Comparison and classification have been the backbone of modern statistics since its dawn in the late nineteenth century. Comparison in its simplest form as statistical inference by the z-test and later the t-test is a very powerful and versatile method. These two approaches are similar through the goal of minimization of false positives, i.e. erroneously rejecting the null hypothesis, the type I error explained by [Larsen and Marx \(2018\)](#), which is the same as minimization of the false positive rate (FPR) ([Stehman, 1997](#)). Thus, both these approaches are connected to the problem of finding proper decision boundaries, i.e. the critical region, where the proper rejection of the null hypothesis is undertaken or where the correct classification is done. In this dissertation there are several axes of comparison; both the z-test and the t-test in the developed scale-space application and the effect of random effects in linear mixed models (LMMs). Further, comparison of different features of HSIs of pigmented skin lesions (PSLs), where these comparisons are related to the analysis of curves and classification. Regression is one of the oldest and most important statistical methods. The least-squares method was first published by Legendre, however, contributions by Gauss were possibly given earlier. A substantial part of statistics can be cast as regression problems, from analysis of variance (ANOVA) to structural equation modeling (SEM) ([Mair, 2018](#)). Also, there is a strong link between regression and classification as can be seen in the machine learning field, where almost all methods can be viewed as a regression approach, e.g. support vector machines (SVMs) and neural networks (NNs). NNs have in fact come on a par with humans regarding classification ability exemplified by the CIFAR classification problem. This is accomplished through the application of deep learning (DL), in particular convolutional NNs (CNNs) ([James et al., 2021](#)). So, regression has been developed in many different directions, whereas in our applications, we have chosen the rather classical types such as robust regression and LMMs. Thus, the focus of this dissertation is statistical curve analysis, where the main approaches are

1. The problem of finding proper decision boundaries in comparison and classification
2. Regression

All the papers have some intrinsic comparison content. However, in papers, I and II comparison and classification are the main methods. In papers II, III, and IV regression is the main or an important method. Thus, Paper II has both main approaches.

2.1 Statistical Curve Analysis by Comparison and Classification

As mentioned above, the two related methods, comparison and classification were employed in papers I and II, the first in the form of hypothesis tests and the second in the form of classification of PSLs as melanomas, malignant and benign. Another common feature of these two papers were the analysis of HSI curves, the first paper by the signature curves of the HSI-cube and the second by making a coordinate of the curve from a subimage of each channel. These coordinates consist of means and standard deviations of the subimages.

The predominant aim in these papers is to choose the most suitable decision boundaries. For Paper I this was done as a countermeasure for the low sensitivity and specificity in the method devised in [Hindberg et al. \(2019\)](#) when applied to HSIs. The designed method was a scale-space method for testing whether k multivariate data sets of the same dimension originate from the same distribution. On a more general level scale-space theory is a framework for representing signals on multiple scales, developed by the computer vision, image processing, and signal processing communities ([Lindeberg, 1994](#)). A recent comprehensive review of statistical scale-space methodology is constituted by [Holmström and Pasanen \(2017\)](#). There are quite a few directions of scale-space applications, where the first application in statistics was mode detection for uni- and bivariate density estimation ([Holmström and Pasanen, 2017](#)). The purpose of this methodology is to extract statistically significant features from noisy data at several scales, often corresponding to different levels of resolution in the underlying object of interest. Thus, at larger scales, small artifacts in images or curves may be seen as noise, but at smaller scales it does not make sense to consider comparisons with large-scale features. SiZer is one of the most important statistical applications of this methodology. This was developed by [Chaudhuri and Marron \(1999\)](#) and based on non-parametric regression with smoothing over different scales. The objective of the constructed scale-space map is mapping where the derivatives of the smoothed curves are significant. Hence, over a range of scales, is the underlying curve increasing significantly, is it significantly decreasing or is it flat? This gives a rather convenient instrument in curve analysis. In the application of Paper I, the aim was to find significant differences in HSIs by employing scale-space methodology. However, when testing for differences between spectral signature curves in different acquisitions it can be difficult to select the critical rejection thresholds due to

multiple testing. One possibility is to use the Bonferroni correction method designed to reduce false positives in testing multiple hypotheses, which was proposed by [Dunn \(1961\)](#) for confidence intervals ([Hochberg and Tamhane, 1987](#)). The same definition of position d and scale s as given in [Hindberg et al. \(2019\)](#) is used in Paper I. Thus, the location in the scale-space map is (d, s) . The critical quantile at the (d, s) location, when only using the Bonferroni correction, is then given by:

$$c(d, s)_B = t_{\frac{\alpha}{2p}, n-2}, \quad (2.1)$$

where $\alpha = 0.05$ is the significance-level and p the number of spectral channels of each spectral signature curve. As an alternative, we tried the statistical inference method described by [Hannig and Marron \(2006\)](#) to find suitable critical rejection thresholds for the scale-space map. This method is designed directly for this task. The global critical rejection threshold from ([Hannig and Marron, 2006](#)) is given by:

$$c(d, s)_G = \Phi^{-1} \left(\left(1 - \frac{\alpha}{2} \right)^{\frac{1}{p \sum_{k=1}^{n_s} \theta_k}} \right). \quad (2.2)$$

Here Φ is the normal cumulative distribution function and n_s denotes the number of rows in the scale-space map. Moreover, θ_k is given by

$$\theta_k = 2\Phi \left(\frac{\sqrt{3 \log p}}{2s_k} \right) - 1,$$

where s_k is the scale in row k . For comparison, the critical rejection threshold in ([Chaudhuri and Marron, 1999](#)) was given by:

$$c(d, s)_{SiZer} = \Phi^{-1} \left(\frac{1 + (1 - \alpha)^{\frac{1}{m}}}{2} \right),$$

where m is the number of data blocks. Although $c(d, s)_G$ and $c(d, s)_{SiZer}$ are somewhat different, by using the approximation: $(1 - \alpha)^{\frac{1}{b}} \approx 1 - \frac{\alpha}{b}$, we can get a better understanding of their similarity:

$$c(d, s)_{SiZer} \approx \Phi^{-1} \left(1 - \frac{\alpha}{2b} \right), \quad b = m \quad \text{and} \quad c(d, s)_G \approx \Phi^{-1} \left(1 - \frac{\alpha}{2b} \right), \quad b = p \sum_{k=1}^{n_s} \theta_k.$$

The critical values are further used to test if a new acquisition differs significantly from the existing acquisitions. Utilization of $c(d, s)_G$ gave a more conservative critical rejection threshold than $c(d, s)_B$. More details about this are given in [Section 3.1](#).

Paper II is a classification and regression paper, where we maximize sensitivity, specificity and precision through the, ideally, correct classification of PSLs in three classes as seen in ([Uteng et al., 2021](#)). These classes are melanoma, malignant and benign PSLs. However, correct classification is often not feasible. To maximize the correctness, we did a fair amount of feature

engineering. This also involved establishment of correct decision boundaries for these features, similar to the main aim of Paper I. As mentioned above, also here the curves originates from HSIs. However, here the curves used in the classification effort are constructed from the mean and standard deviation of subsets from the image of each channel in the HSIs. These subsets are coined regions of interest (ROIs). Here, the three phases: training, validation, and testing were conducted. In the training phase, the features which gave the best classification results were developed, which then were utilized in the validation and testing phases. Classification is a supervised learning task with a rather extensive application in ML, ranging from SVMs to NNs (Murphy, 2012). A useful method linking regression and classification is logistic regression which maps the data from the interval $(-\infty, \infty)$ to the interval $(0, 1)$. The logistic function or similar, e.g. the softmax function, is also the backbone of NNs, which have developed considerably in the last decade becoming the default data analysis method. As mentioned above we employ both supervised learning methods, regression and classification in Paper II. More details on this are in Section 3.2.

2.2 Statistical Curve Analysis by Regression

The papers with a regression-centered methodology are papers II, III and IV, where Paper II has robust regression as an essential method that plays an integral part in the classification scheme. Through this we established a link between regression and classification, which is often accomplished in machine learning. Paper II's use of robust non-linear regression is justified through the rather convoluted ROI curves, where the observed features were amplified using this type of regression. Robust regression is often used when data contains outliers and the regression method must be less sensitive to these outliers (Andersen, 2008). There are several types of robust regression, where we used the robust non-linear regression adapted from (Holland and Welsch, 1977). This was done through the MATLAB[®] function `nlinfit` which is based on the Levenberg-Marquardt (LM) nonlinear least-squares algorithm, explained in Seber and Wild (2003). This algorithm finds the minimum of the function:

$$\|\mathbf{y} - \mathbf{w}^T f(\mathbf{x}, \mathbf{b})\|^2 = \sum_{i=1}^n (y_i - w_i f(x_i, \mathbf{b}))^2, \quad (2.3)$$

where the chosen weight-function was

$$w(z) = \frac{1}{(0.011 + 0.011z)^2}, \quad (2.4)$$

(x_i, y_i) are the measured curve points, \mathbf{b} the estimated vector of coefficients, and $f(x_i, \mathbf{b})$ the target function, e.g. a quartic polynomial function. The LM nonlinear least-squares algorithm is an extension of the Gauss-Newton algorithm, where a regularization term is added, due to often poorly conditioned Jacobian matrices in non-linear regression problems. However, this algorithm

is rather slow when large residuals occur, as is the case when outliers are present. This is amended in the `nlinfit` function by combining the LM nonlinear least square algorithm with the iteratively re-weighted least square (IRLS) algorithm. The weights are changed according to Eq. (2.4), with an argument-value, z , of adjusted residuals. The adjustment factor is given by

$$\frac{1}{\sqrt{1-h}}, h_i = \min \left(0.9999, \sum_j (Q_{ij})^2 \right),$$

where Q is one part of the QR-decomposition of the Jacobian matrix (Sauer, 2018). This adjustment factor is multiplied with the residuals from the previous iteration:

$$z_{k+1} = \frac{z_k \odot \left[\frac{1}{\sqrt{1-h_0}}, \dots, \frac{1}{\sqrt{1-h_n}} \right]}{\hat{\sigma}_k},$$

at the k 'th iteration, where \odot is elementwise multiplication and $\hat{\sigma}_k = \text{MAD}(\text{residuals}_k)/0.6745$ a robust estimate of the standard deviation (Median absolute deviation (MAD) is described in (Wilcox, 2017)). Thus, the residuals and the Jacobian matrix are controlling the re-weighting and makes it a robust regression method. This type of re-weighting is used by Dumouchel and O'brien (1991), however, with different h_i 's. The `nlinfit`-function was also applied in Paper III, but without the IRLS-option.

In Paper IV we used LMM (Linear Mixed Model) to analyse the data and explore any significant differences between two groups of treatments. The LMM can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon}, \boldsymbol{\epsilon} \sim N(0, \mathbf{I}\sigma_\epsilon^2), \quad (2.5)$$

where \mathbf{y} is the response-variable, \mathbf{X} is a fixed effect model matrix, $\boldsymbol{\beta}$ the coefficient vector, \mathbf{Z} the random-effects design matrix, \mathbf{b} the random effect coefficient-vector and $\boldsymbol{\epsilon}$ the Gaussian-distributed error-vector with covariance-matrix $\mathbf{I}\sigma_\epsilon^2$. We include random effects to quantify the dependency structure between the measurements. The default method in comparing groups is analysis of variance (ANOVA). However, this model will then not include random effects. One could use the repeated measurements variant of ANOVA, but it will neither be sufficient in modeling the complex structures in the measurements. Thus, the more comprehensive model is given by Eq. (2.5), where the needed factors in the problem are accounted for. The omission of random effects in the model is related to the problem of pseudo-replication, especially pointed out by Hurlbert (1984). This problem is extensively studied by Pinheiro and Bates (2000), where many examples of mixed-effects modeling are given. In order to increase flexibility we used the `brms`-package (Bürkner, 2018), a Bayesian approach on estimating the parameters in Eq. (2.5) (Robert and Casella, 2010). This package is heavily dependent on Stan (Stan Development Team, 2016), an R-package (R Development Core Team, 2017) based among other things on the Hamiltonian Monte Carlo algorithm (Liu, 2001). However, due to increased computational complexity, the `lmer`-function from the `lme4`-package was utilized, when a random slope in the LMM-model was

included (Bates et al., 2015). On a more general level can LMM be seen as a kind of regularization similar to ridge regression, as shown by Christensen (2019), i.e. forcing both the estimated coefficients and the error terms to be small. This gives a rather versatile and robust method which according to Gelman and Hill (2007) should be the default method in multilevel- and repeated measurements-analysis, which also can be considered as a robust regression method.

Summary and discussion

3.1 Paper I - Change detection in Hyperspectral Images by Statistical Curve analysis by Scale-space techniques

In this paper the aim was to detect a significant change in HSI curves, through a scale-space map, with high sensitivity and specificity (Uteng et al., 2020).

The main method of this paper is to test simultaneously for many different scales and positions (frequency bands or channels). The scale s equals the number of different frequency bands being summed across. To be specific, this means that scale $s = 1$ corresponds to the situation where we test if the observed values at spectral frequency d are different between acquisitions of spectral signatures. At scale, $s = 3$ and position d , a smoothing in terms of a weighted average of the observed values for spectral frequencies $d-1$, d and $d+1$ are used to test whether the acquisitions are different. The weights are calculated from an Epanechnikov kernel function (i.e. a parabolic function) described in Wand and Jones (1994), the same as in Hindberg et al. (2019) (the kernel of least bias and variance). The smoothing process can also be seen as a convolution. For other scales, completely analogous smoothing over the frequency bands is made and used to perform the tests. Note that by applying this smoothing, we can test for differences in the acquisitions at all locations for a large number of scales. So, the tests are performed at all p spectral frequencies for a total number of n_s different scales. Instead of looking at a single location or a single scale, the described scale-space approach can help detect changes that appear at several levels of smoothing, i.e resolution.

The method in this paper is divided into two procedures: the training procedure and the testing procedure. The training procedure at a location (d, s) is accomplished by comparing one acquisition to the others. To simplify the description we illustrate the methodology by testing for change in the sample mean, \bar{X} , over the pixels in the image. In order to accomplish high sensitivity and specificity our method used the t -test either with a Bonferroni correction or the inference approach suggested in Hannig and Marron (2006), which is based on a z -test. Further,

we assume that

$$\bar{X}_1 \sim N(\mu_1, \sigma^2) \quad (3.1)$$

for acquisition one and

$$\bar{X} = \frac{1}{n-1} \sum_{k=2}^n \bar{X}_k \sim N\left(\mu, \frac{\sigma^2}{n-1}\right), \quad (3.2)$$

for the remaining $n-1$ acquisitions. Here n is the total number of acquisitions. Then, we estimate σ^2 by the standard estimator for variance using all acquisitions apart from the one left out. In the case acquisition 1 is left out, this means that σ^2 is estimated by

$$S^2 = \frac{1}{n-2} \sum_{i=2}^n (\bar{X}_i - \bar{X})^2.$$

As mentioned above, in addition to the Bonferroni-corrected quantile in Eq. (2.1), we also tried the so-called global quantile given in Eq. (2.2). In the testing procedure, we test

$$H_0 : \mu_1 = \mu \text{ against } H_1 : \mu_1 \neq \mu$$

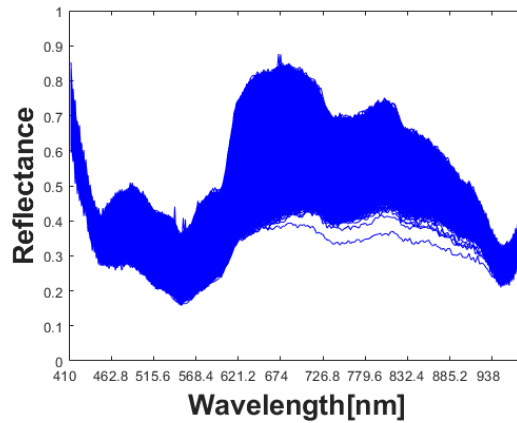
using the test statistic

$$T = \frac{\bar{X}_{\text{test}} - \bar{X}_{\text{train}}}{S \sqrt{1 + \frac{1}{n-1}}},$$

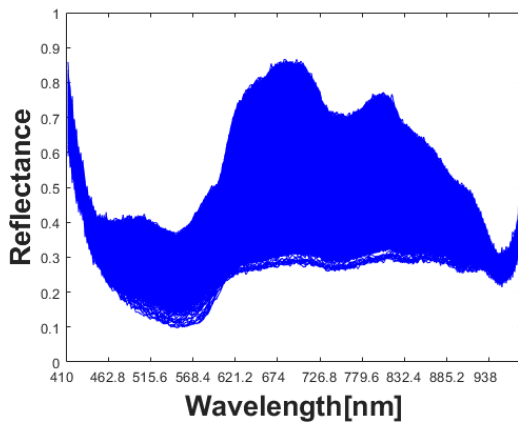
where H_0 is rejected if

$$|T| > c(d, s)_B \text{ or } |T| > c(d, s)_G.$$

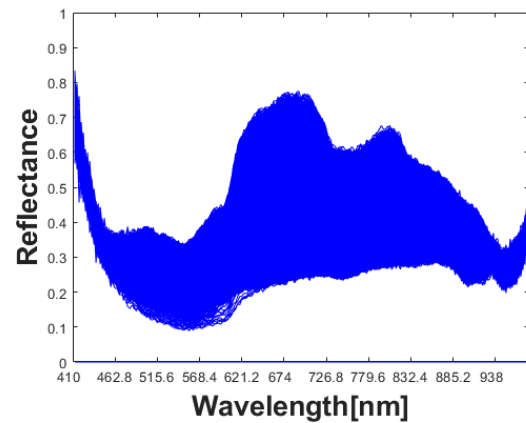
The results of the tests are summarized in a scale-space map, where the horizontal and vertical axes correspond to spectral frequency bands and scales, respectively. At each location (d, s) we perform a test, and the outcome is shown as a colored pixel, with red (blue) indicating a significant (not significant) difference at the position d for scale s . The main case study in this paper is the examination of the freshness of fish. An example taken from one fish at day 0 is given in Fig. (3.1(a)) and signatures for two different acquisitions for the same fish at day 4 are given in figures 3.1(b) and 3.1(c), respectively. Comparison of the signatures of Fig. (3.1(a)) with figures (3.1(b)) and (3.1(c)) results in the significance maps in Fig. (3.2). There the four panels show the results using two different parameters, the mean and the median. Statistically significant changes are detected with the mean and the median, but not with the standard deviation and the range. Typical FPR results are reported in Fig. (3.3), where we can see that the Hannig-Marron critical rejection point, $c(d, s)_G$, gives a slightly better performance.



(a)



(b)



(c)

Figure 3.1: (a): Plot of HSI-curves from acquisition number one from frozen fish at day 0. (b): Plot of HSI-curves from acquisition number two from frozen fish at day 4. (c): Plot of HSI-curves from acquisition number four from frozen fish at day 4. Reflectance is the surface's effectiveness in reflecting radiant energy.

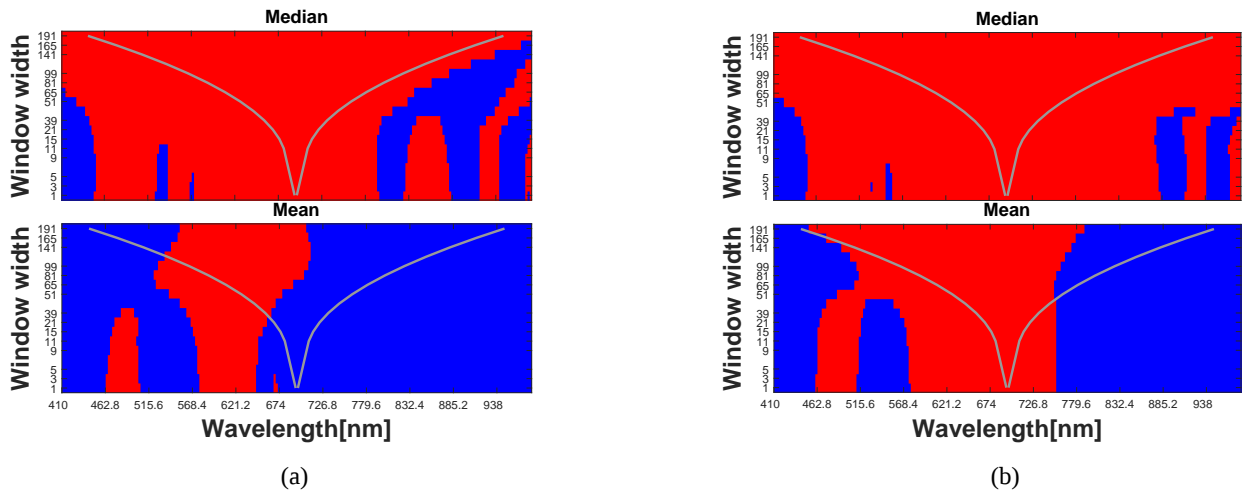


Figure 3.2: (a): Significance map for the comparison of day 0 with day 4 in acquisition number two using the median and the mean. (b): Significance map for the comparison of day 0 with day 4 in acquisition number four using the median and the mean. The Hannig-Marron rejection threshold was used in all maps.

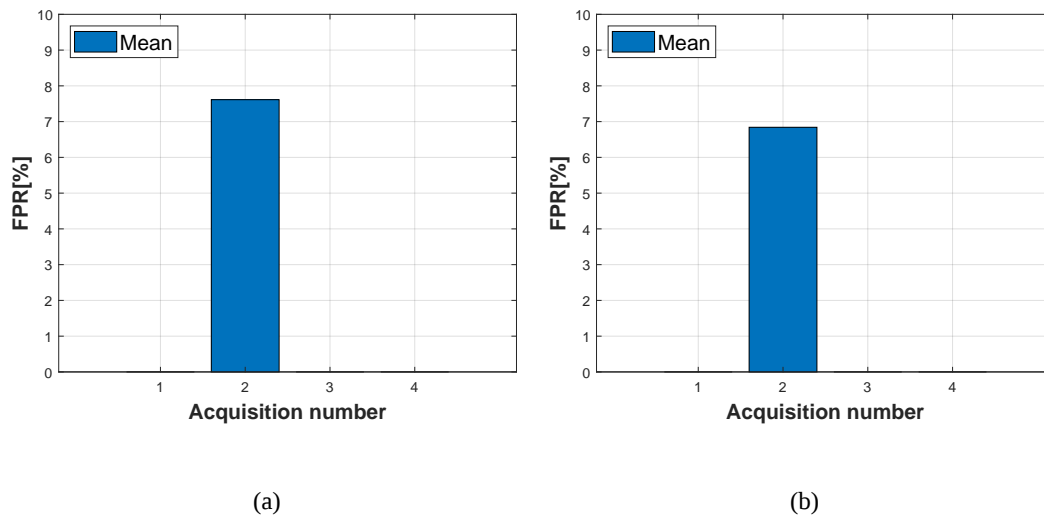


Figure 3.3: False positive rate (FPR) in the fish freshness example. (a): FPR from a leave-one-out test using Bonferroni correction to account for multiple testing. (b): FPR from a leave-one-out test and using the Hannig-Marron global rejection threshold.

Our findings suggest that the proposed scale-space methodology can be successful in detecting small changes in an HSI. To be useful in practice, such a method must have both high sensitivity and high specificity and our results clearly show promise in this respect.

3.1.1 Future work

Freshness of food is an important area, where the method developed in this paper have potential to be integrated in monitoring devices. This method is also promising in the detection of change in PSLs. There is an ongoing process of accumulating HSIs related to PSLs in collaboration with several hospitals in the Canary Islands, Spain. We also want to design a successful PSL classifier for HSIs and this seems currently promising (Fabelo et al., 2019). However, a system capable of detecting changes in a mole will be even more important since detecting skin cancer early is a potential lifesaver. Our ultimate goal is to design a decision support tool based on HSIs which could be implemented on a smartphone and thus increase general availability.

3.1.2 Contributions by the author

- Developed and programmed the analysis method presented in the paper.
- Involved in the conceptualization of the scale-space change detection methodology.
- Did all the analysis presented in the paper.
- Was in charge of the review and editing process.
- Wrote the initial draft and was in charge of producing the final manuscript.

3.2 Paper II - Classification of Hyperspectral Images by Statistical Curve analysis

This paper Uteng et al. (2021) is mainly a classification paper, where we developed features through the regression of ROI curves. It also shows similarity to Paper I Uteng et al. (2020) since we want the classification scheme to have both high specificity and high sensitivity. The first issue was how to extract the ROIs. We found that the ROIs extracted on the edges or artifacts on the PSLs, using a subimage of 7x7, gave the best features. This can be interpreted as some kind of pixel gradient which gives some variation in the sub-image. In Fig. (3.4(a)) a typical ROI from one channel can be observed. After the ROI extraction, we formed curves with the mean of each channel's ROI as x-coordinate and the standard deviation as y coordinate. We then performed the regression as described in Section 2.2. The function utilized here was

$$f(x, \mathbf{b}) = b_0x^4 + b_1x^3 + b_2x^2 + b_3x + b_4,$$

with coefficients b_j , $j = 0, \dots, 4$, which were estimated by least squares. We were then able to classify the PSLs as melanoma, malignant and benign, with quite high sensitivity, specificity, precision, and accuracy. We used several features to classify the PSLs (i.e. these features were applied to the fitted curves):

3. Summary and discussion

- For the Melanoma PSLs: first derivative (df), second derivative (ddf), and combination of these.
- For the Malignant PSLs: Mean and product of max y and max x
- For the Benign PSLs: Mean and absolute value of the first derivative.

In Fig (3.4) we can view the typical regression curves together with the data.

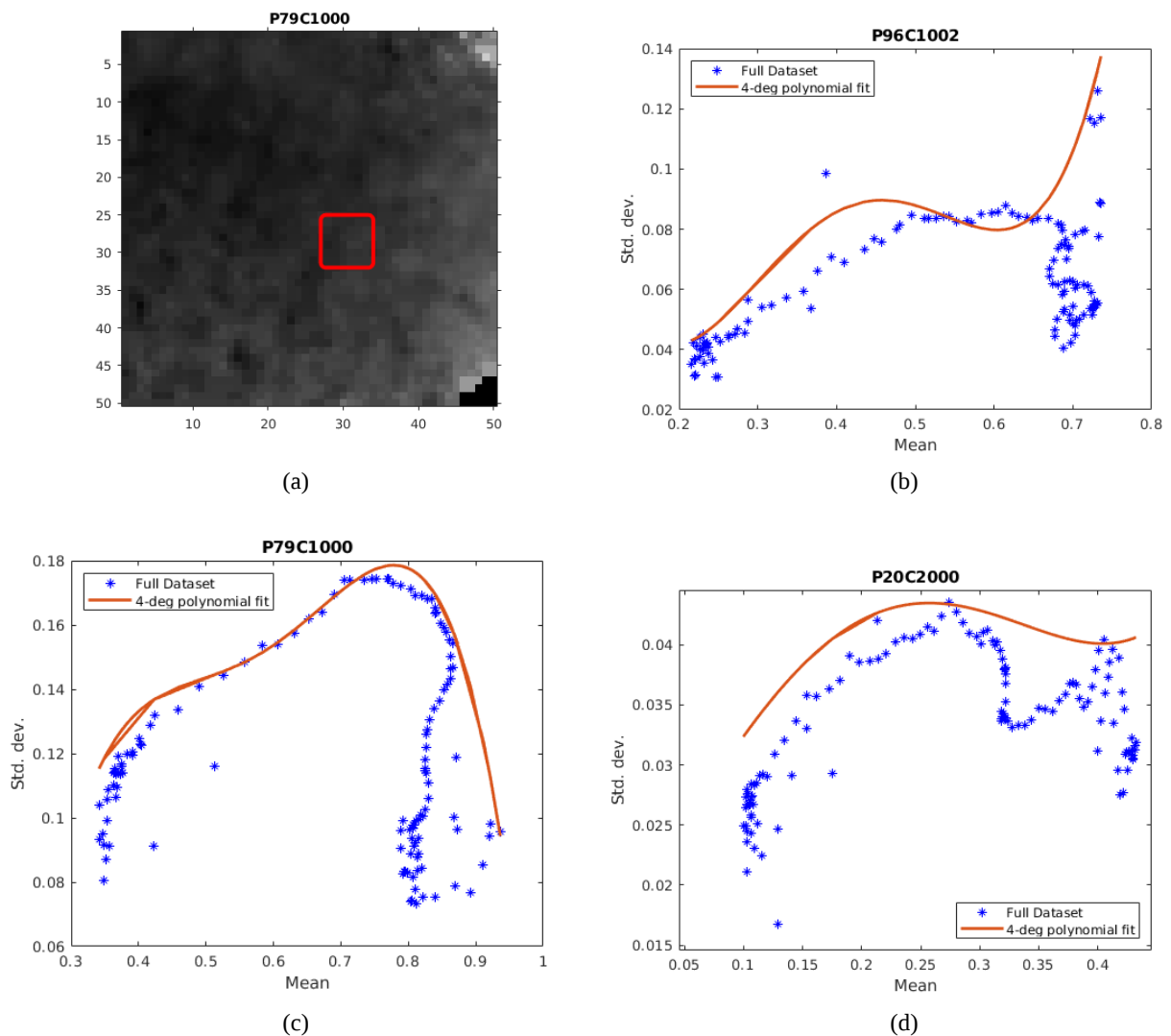


Figure 3.4: (a): A melanoma HSI with a typical example of a 7x7 ROI embedded (in red). (b): A typical example of a ROI curve from a melanoma HSI with a seemingly asymptotic y-value (i.e. it seems that for a mean (x-value) around 0.75 the standard deviation (y-value) of the fitted quartic polynomial function goes to infinity) originating from the ROI in Fig. (3.4(a)) (c): A typical example of a ROI curve from a malignant HSI with a rather high mean and maximum of the fitted quartic polynomial function. (d): A typical example of a ROI curve from a benign HSI with a relatively low mean of the fitted quartic polynomial function.

This work shows that it is possible to classify PSLs successfully in melanoma, malignant and benign classes.

3.2.1 Future work

A possible expansion of this work could be the establishment of statistical tests concerning the classification decisions based on the curve features.

3.2.2 Contributions by the author

- Developed and programmed the analysis method presented in the paper.
- Did most of the conceptualization of the curve classification methodology.
- Did all the analysis presented in the paper.
- Was in charge of the review and editing process.
- Wrote the initial draft and was in charge of producing the final manuscript.

3.3 Paper III - Improvement of Blood Glucose Estimation by Statistical Curve analysis

In Paper III [Sebastiani et al. \(2020\)](#) the focus is on a problem with the estimation of blood glucose (BG) through continuous glucose measurements (CGM). This is related to the fact that we can only measure glucose concentration in subcutaneous interstitial tissue $G_s(t)$ at sampling times while we are interested more in temporal glucose levels in tissues and blood. Fast and large changes in glucose concentration in blood $G_b(t)$, can be observed during physical activity (PA). However, it is well known that related changes in $G_s(t)$ are seen after a time delay (lag time) ([Wei et al., 2010](#)). To describe both $G_b(t)$ and $G_s(t)$ we therefore adopted a model with two sine functions. The quantity $G_s(t)$ as a function of $G_b(t)$ is commonly modeled as in [Dicker et al. \(2013\)](#) by means of asymmetric convolution, due to causality, with exponential kernel $h(t)$ given by:

$$h(t) = \frac{1}{\rho} e^{-\frac{t}{\rho}}, \quad (3.3)$$

with ρ dependent on the duration of the time-lag. However, due to the two phases of glucose decrease and increase with different dynamics, we use two different versions of the convolution kernel in Eq. (3.3), one before the minimum location of $G_b(t)$ and the other one after it. The convolution becomes:

$$G_s(t) = \int_0^t G_b(\tau) h(t - \tau) d\tau + \epsilon, \quad \epsilon \sim N(0, \sigma^2), \quad (3.4)$$

which is discretized to:

$$\mathbf{G}_s = \mathbf{H}\mathbf{G}_b + \epsilon, \quad \epsilon \sim N(0, \sigma^2\mathbf{I}). \quad (3.5)$$

Given the functions $G_b(t)$ and $G_s(t)$ estimated for one or more subjects, the convolution kernel parameters are estimated by minimizing the squared error, as in Eq. (2.3), between $G_s(t)$ and its estimate by numerical convolution of $G_b(t)$. This minimization was here also, accomplished through the MATLAB® `nlinfit`-function, however, without the IRLS-option. In a new subject, model parameters of $G_s(t)$ are estimated from CGM measurements by minimizing mean squared error between data and model values at sampling times. Deconvolution of $G_s(t)$ is then performed to estimate model parameters for $G_b(t)$ by minimization of the squared error, also here similar to Eq. (2.3), between estimated $G_s(t)$ and the numerical convolution of $G_b(t)$ by Simulated Annealing explained in Kirkpatrick et al. (1983), i.e., not by the LM least squares algorithm. There are several alternative approaches to deconvolution, where the Wiener-filter first given by Wiener (1949) is rather popular. Bayesian deconvolution of time-series is discussed in some length in Kay (1993), also a popular approach in signal processing.

This gives a novel method for controlling BG levels also during and after exercise, which is of great importance for T1D-patients to avoid hypoglycemia. Also, the results showed that a statistical approach can be used to successfully estimate BG concentration along time during PA from CGM measurements.

3.3.1 Future work

This could be developed as a tool to monitor GB to maximize performance in endurance sports and also be used to reduce the risk of hypoglycemia in T1D subjects during PA.

3.3.2 Contributions by the author

- Did a substantial part of the analysis presented in the paper.
- Wrote the initial draft.
- Participated in the writing of the manuscript—reviewing and editing

3.4 Paper IV - Quantifying group difference by Statistical Curve analysis

In Paper IV the effects of liver transplantation in pigs were studied (Uteng et al., 2022). There are two major schools of in machine perfusion: normothermic machine perfusion (NMP) and hypothermic machine perfusion (HMP) (Marecki et al., 2017, Schlegel et al., 2019). Neither has yet proven its superiority. Here, there are two groups consisting of seven pigs each, where the total treatments for both these groups last in 720 min. In the first group, the livers from the

pigs are in warm ischemia (WI) and static cold storage (SCS) in 240 min before NMP and the second group WI and NMP in 720 min. Thus, we name the first group SCS and the second group NMP. Machine perfusion may reverse the negative responses caused by the warm ischemia, and in addition allowing a performance assessment of the liver before transplantation. Thus, it was investigated experimentally:

- What are the effects of ischemia/reperfusion injuries beyond the consensus of today (i.e. WI>30 min) in the liver parenchyma including the bile ducts and their potential reversal?
- Do the ischemic changes reflected in tissue metabolism improve once the organ is connected to the machine perfusion?
- Is controllable hemodynamics a prognostic marker to deem whether the livers can be transplanted or not?

The main goal of the statistical analysis of these data is to investigate the degree of difference between the SCS and NMP groups and how the methods compare as machine perfusion methods. The default method in comparing the differences of groups is the analysis of variance (ANOVA) when there are more than two groups and the t-test if there are only two groups. However, due to the problem of pseudoreplication, i.e. the observations on the response variable are not independent, we have to define dependence within and between the groups (Hurlbert, 1984). This is called a random effect, which is elaborated rather thoroughly in Pinheiro and Bates (2000), where they develop models and software packages to do inference in the resulting LMM. Further two of the three assumptions of repeated measurements ANOVA are violated: Normality of the error term and independence of measurements (sphericity is not relevant with only two groups) (Field et al., 2012). The current default software package doing this inference is lme4 presented in Bates et al. (2015), which is written for R (R Development Core Team, 2017). However, this package assumes Gaussian distributed error components, which was not fulfilled in our inference. Thus, we applied the R-package brms, presented in Bürkner (2018), which utilizes a Bayesian approach. Furthermore, the study presented in this paper has a multilevel two-way nested design and we wrote the model as:

$$y_{ijk} = Intercept + a_i + b_{ij} + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), \quad (3.6)$$

$$a_i \sim N(0, \sigma_{sn}^2), b_{ij} \sim N(0, \sigma_{Pig}^2),$$

where ϵ_{ijk} denotes the residual, and $N(0, \sigma_\epsilon^2)$ a Gaussian distribution with mean zero and standard deviation σ_ϵ , for the residuals. The parameters a_i and b_{ij} are the random effects with standard deviations σ_{sn} and σ_{Pig} , for Protocol and Pig, respectively, the *Intercept* is a constant, i , the groups, j , the individual pig and k , the measurement time points. The brms-package is a backend to RStan, the R interface to Stan (Stan Development Team, 2016), which applies the Hamiltonian Monte Carlo algorithm (Liu, 2001). This means that the intercept and the standard

deviations were given priors. The prior for the intercept was chosen to be a Gaussian distribution with mean zero and standard deviation 10. For all the standard deviations a half-Cauchy distribution with location zero and scale one truncated below zero, was chosen. To determine if there were any random effects we compared the fit of Eq. (3.6) with

$$y_{ijk} = \text{Intercept} + \epsilon_{ijk}, \quad \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), \quad (3.7)$$

by Watanabe's Widely Accepted Information Criterion (WAIC) (Vehtari et al., 2017).

We also observed some slope effects for some of the parameters and this was subsequently included in another model as *Time* for each Pig given by the model:

$$\begin{aligned} y_{ij} &= \text{Intercept} + \text{Time} + a_i + b_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_\epsilon^2), \\ a_i &\sim N(0, \sigma_{sn}^2), \quad b_{ij} \sim N(0, \sigma_{Pig}^2), \end{aligned} \quad (3.8)$$

Due to computational issues, we could not use the brms-package for the analysis. Thus, we had to use the lmer function from the lme4-package for this more complex model. Also, we compared this with the robust lmer package, developed by Koller (2016), which gave similar results. The results for all these computations showed clear evidence for random effects given by WAIC through the difference between the nullmodel in Eq. (3.7) and the random effects model in Eq. (3.6), the significance of the likelihood ratio test (LRT) and the difference between the null model (Eq.(3.7)) and the random-effects model (Eq.(3.8)) in the Akaike information criterion (AIC) (Akaike, 1992).

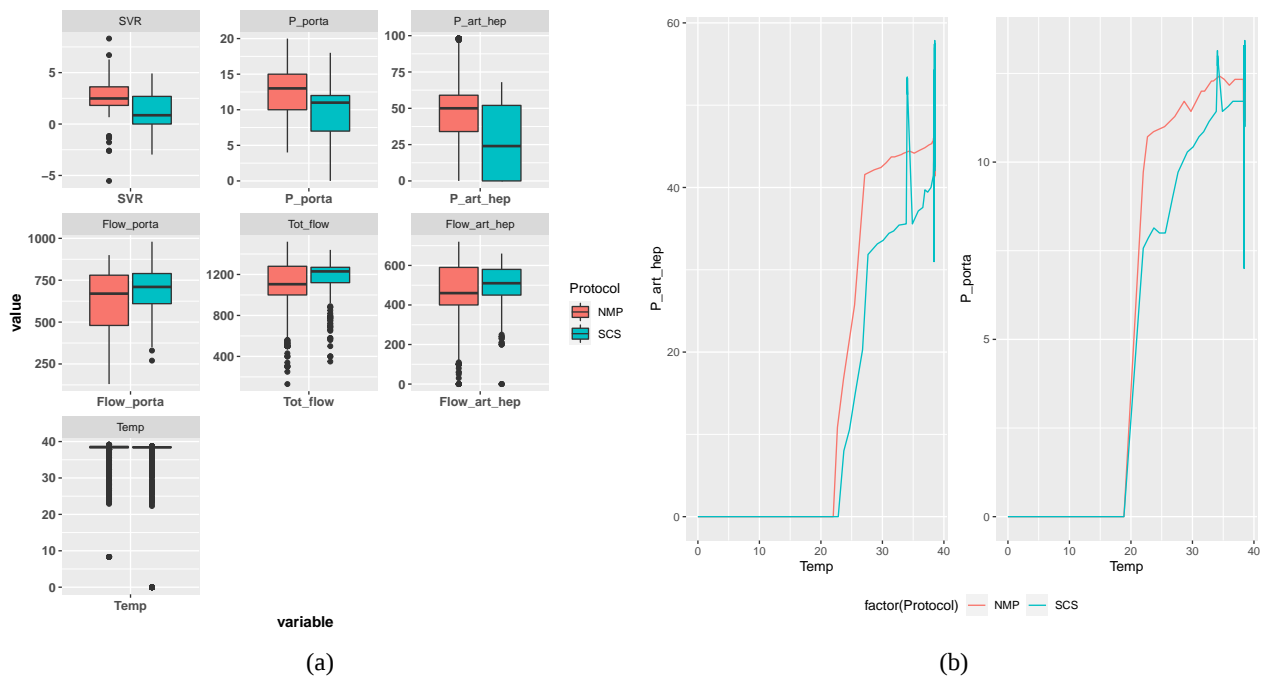


Figure 3.5: (a): Boxplots of the hemodynamic data for the NMP and the SCS groups.(b): Plot of P_art_hep and P_porta vs. Temperature for the NMP and SCS groups. All figures are made from the hemodynamic data.

Many of the variables in the SCS group were somewhat smaller than in the NMP group as seen in Fig. (3.5(a)) for the hemodynamic data, which was also the case for the microdialysis (MD) data. One such variable is systemic vascular resistance (SVR), indicating a lower probability of injuring the organ during machine perfusion. Also, according to the LMM analysis a significant difference between the two groups for most of the variables was detected.

Vasoconstriction is another important issue, which in the rewarming phase at the start of the machine perfusion period (from 4°C to 38°C) is typical. This vasoconstriction and the high pressures especially in the hepatic artery (P_art_hep) will commonly decline when the organ becomes normotherm (38°C). Beyond the rewarming phase, the pressures in the hepatic artery and the portal vein (P_porta) will deem whether an organ is transplantable or not. As seen in Fig. (3.5(b)), in the SCS group these parameters show a better prospect.

3. Summary and discussion

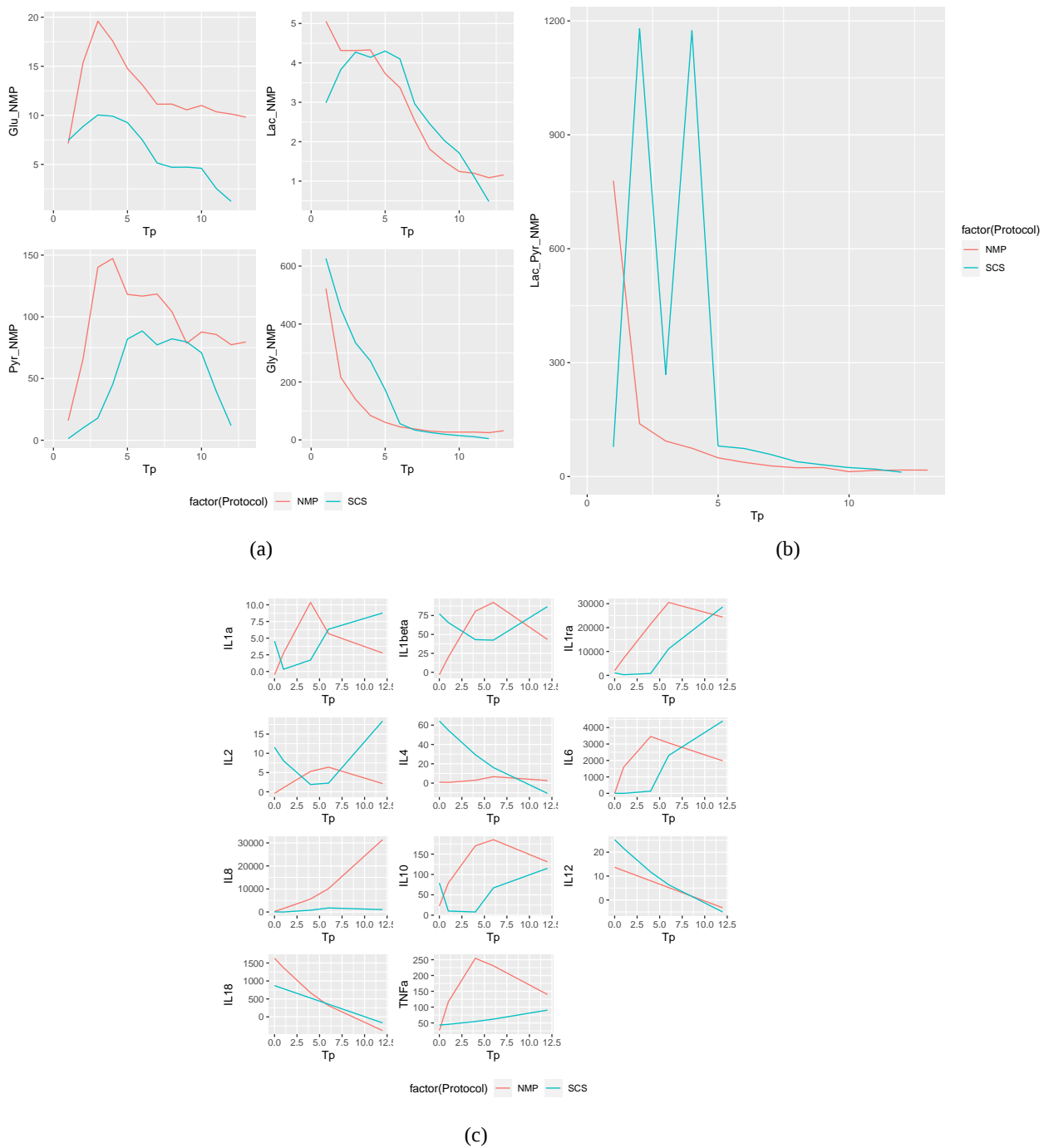


Figure 3.6: (a): Plot of glucose (Glu_NMP), the lactate (Lac_NMP) and the pyruvate (Pyr_NMP) lower in the SCS group, but glycerol (Gly_NMP) for the NMP and SCS groups as means for the pigs measurements at each time-point. (b): Plot of L/P-ratio (Lac_Pyr_NMP) for the NMP and SCS groups as means for the pigs measurements at each time-point. (c): Plots of immunological markers for the NMP and SCS groups as means for the pigs measurements at each time-point.

All figures are made from the microdialysis (MD) data.

For the MD-data, the glucose (Glu_NMP), the lactate (Lac_NMP) and the pyruvate (Pyr_NMP) were lower in the SCS group, but glycerol (Gly_NMP) was lower in the NMP group, as seen in Fig. (3.6(a)). However, it can be claimed that they were rather similar in the groups. The L/P-ratio was also more beneficial for the NMP group as seen in Fig. (3.6(b)). Glucose, lactate, pyruvate, glycerol and the L/P-ratio are connected to the effects of the ischemia condition and it is advisable to keep it low (Nowak et al., 2002). We can also observe that the effects of WI are in general reversed as seen in Figures (3.6(a)) and (3.6(b)). The immunological markers, the IL's (interleukins) and TNFa (TNF-alpha) had a smaller mean for the SCS group than the NMP group. However, as seen in Fig. (3.6(c)) the NMP group had a negative slope at the end in ten of eleven markers. Thus, the main findings were that the SCS group was better regarding the hemodynamic variables, but for the MD variables the NMP group was somewhat better.

3.4.1 Future work

To increase the number of viable liver transplants it is important to increase the knowledge of both the NMP and the SCS methods. We will proceed with more experiments and research in these areas to try to accumulate more knowledge.

3.4.2 Contributions by the author

- Did all the statistical analysis presented in the draft.
- Wrote the statistical methodology and analysis in the draft.
- Wrote the result section in the draft.
- Wrote the discussion in the draft.

3.5 Concluding remarks

In this dissertation we have analyzed curves through the statistical lens of the two main methods, finding suitable critical decision regions and regression. These methods are very versatile techniques and have given successful solutions to the proposed problems from health-related fields presented here and will almost certainly be expanded and developed further in the future. In the field of food quality, we have proposed a novel method to detect change in quality, by the use of HSIs through scale-space maps. HSIs, also provided the data for the novel method developed to classify PSLs. This method gave us a rather efficient way to classify PSLs in malignant, benign and melanoma classes. In the diabetes field, we developed a new and more accurate estimation method of BG, which can be used by T1D-patients during physical exercise to reduce the risk of going into a diabetic coma as a result of hypoglycemia. The last application of statistical curve

analysis was through the analysis of data from two machine perfusion paths used with liver transplantation. We here tried to get a clearer picture of which path of machine perfusion was the most suitable, which could eventually help to increase the donor pool.

In conclusion, we have addressed the challenges presented in Chapter 1 of this dissertation through the contributions presented in the four included papers.

Bibliography

- Akaike, H. (1992). Information Theory and an Extension of the Maximum Likelihood Principle. In Kotz, S. and Johnson, N. L., editors, *Breakthroughs in Statistics*. Springer, New York, NY.
- American Cancer Society (2022). Cancer Facts & Figures 2022. Technical report, American Cancer Society.
- Andersen, R. (2008). *Modern Methods for Robust Regression*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America .
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1).
- Bevilacqua, A., Corbo, M. R., and Sinigaglia, M. (2016). *The Microbiological Quality of Food: Foodborne Spoilers*. Woodhead Publishing.
- Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10(1).
- Chaudhuri, P. and Marron, J. S. (1999). SiZer for Exploration of Structures in Curves. *Journal of the American Statistical Association*, 94(447):807–823.
- Christensen, R. (2019). *Advanced Linear Modeling*. Springer, Cham.
- Dicker, L., Sun, T., Zhang, C.-H., Keenan, D. B., and Shepp, L. (2013). Continuous blood glucose monitoring: A Bayes-hidden Markov approach. *Statistica Sinica*.
- Dumouchel, W. and O’Brien, F. (1991). Integrating a Robust Option into a Multiple Regression Computing Environment. *Computer Science*.
- Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association*, 56(293).
- Elder, D. E., Massi, D., Scolyer, R. A., and Willemze, R. (2018). *WHO Classification of Skin Tumours*, volume 11. IARC, 4 edition.

- Fabelo, H., Melian, V., Martinez, B., Beltran, P., Ortega, S., Marrero, M., Callico, G. M., Sarmiento, R., Castano, I., Carretero, G., Almeida, P., Garcia, A., Hernandez, J. A., and Godtlielsen, F. (2019). Dermatologic Hyperspectral Imaging System for Skin Cancer Diagnosis Assistance. In *2019 XXXIV Conference on Design of Circuits and Integrated Systems (DCIS)*. IEEE.
- Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. SAGE publishing, 1 edition.
- Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Goetz, A. F. H., Vane, G., Solomon, J. E., and Rock, B. N. (1985). Imaging Spectrometry for Earth Remote Sensing. *Science*, 228(4704):1147–1153.
- Hannig, J. and Marron, J. S. (2006). Advanced distribution theory for SiZer. *Journal of the American Statistical Association*.
- Hindberg, K., Hannig, J., and Godtlielsen, F. (2019). A novel scale-space approach for multi-normality testing and the k-sample problem in the high dimension low sample size scenario. *PLOS ONE*, 14(1).
- Hochberg, Y. and Tamhane, A. C. (1987). *Multiple comparison procedures*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley.
- Holland, P. W. and Welsch, R. E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics - Theory and Methods*, 6(9).
- Holmström, L. and Pasanen, L. (2017). Statistical Scale Space Methods. *International Statistical Review*, 85(1).
- Hurlbert, S. H. (1984). Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs*, 54(2).
- International Diabetes Federation (2021). IDF Diabetes Atlas. Technical report, International Diabetes Federation.
- James, G. M., Witten, D., Hastie, T., and Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R*. Springer, 2 edition.
- Jensen, D. J. and Elewski, B. E. (2015). The ABCDEF Rule: Combining the “ABCDE Rule” and the “Ugly Duckling Sign” in an Effort to Improve Patient Self-Screening Examinations. *The Journal of clinical and aesthetic dermatology*, 8(2).
- Kay, S. (1993). *Fundamentals of Statistical Signal Processing*. Prentice Hall.

- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598).
- Koller, M. (2016). *robustlmm* : An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, 75(6).
- Larsen, R. J. and Marx, M. L. (2018). *An Introduction to Mathematical Statistics and Its Applications*. Pearson.
- Li, Q., He, X., Wang, Y., Liu, H., Xu, D., and Guo, F. (2013). Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10):100901.
- Lindeberg, T. (1994). Scale-space theory: A basic tool for analyzing structures at different scales. *Journal of Applied Statistics*.
- Liu, J. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer.
- Mair, P. (2018). *Modern Psychometrics with R. Use R!* Springer.
- Marecki, H., Bozorgzadeh, A., Porte, R. J., Leuvenink, H. G., Uygun, K., and Martins, P. N. (2017). Liver ex situ machine perfusion preservation: A review of the methodology and results of large animal studies and clinical trials. *Liver Transplantation*, 23(5).
- Mian, Z., Hermayer, K. L., and Jenkins, A. (2019). Continuous Glucose Monitoring: Review of an Innovation in Diabetes Management. *The American Journal of the Medical Sciences*, 358(5):332–339.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.
- Norman, A. W. and Henry, H. L. (2015). *Hormones*. Elsevier.
- Nowak, G., Ungerstedt, J., Wernerman, J., Ungerstedt, U., and Ericzon, B.-G. (2002). Metabolic changes in the liver graft monitored continuously with microdialysis during liver transplantation in a pig model. *Liver Transplantation*, 8(5):424–432.
- Pinheiro, J. and Bates, D. M. (2000). *Mixed Effects Models in S and S PLUS*. Springer.
- R Development Core Team (2017). *R: A language and environment for statistical computing*.
- Robert, C. P. and Casella, G. (2010). *Monte Carlo Statistical Methods*. Springer.
- Sauer, T. (2018). *Numerical analysis*. Pearson, 3 edition.

- Schlegel, A., Muller, X., Kalisvaart, M., Muellhaupt, B., Perera, M. T. P., Isaac, J. R., Clavien, P.-A., Muiesan, P., and Dutkowski, P. (2019). Outcomes of DCD liver transplantation using organs treated by hypothermic oxygenated perfusion before implantation. *Journal of Hepatology*, 70(1):50–57.
- Sebastiani, G., Uteng, S., Godtlielsen, F., Polák, J., and Brož, J. (2020). Estimation of Blood Glucose Concentration During Endurance Sports. *International Journal of Biology and Biomedical Engineering*, 14.
- Seber, G. A. F. and Wild, C. J. (2003). *Nonlinear Regression*. Wiley, second edition.
- Sliney, D. H. (2016). What is light? The visible spectrum and beyond. *Eye*, 30(2):222–229.
- Stan Development Team (2016). RStan: the R interface to Stan.
- Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1).
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3):209–249.
- Uteng, S., Johansen, T. H., Zaballos, J. I., Ortega, S., Holmström, L., Callico, G. M., Fabelo, H., and Godtlielsen, F. (2020). Early Detection of Change by Applying Scale-Space Methodology to Hyperspectral Images. *Applied Sciences*, 10(7).
- Uteng, S., Numan, A., Godtlielsen, F., and Nedredal, G. (2022). Machine Perfusion of the Liver to Resuscitate and Reverse Ischemic Liver Injuries.
- Uteng, S., Quevedo, E., M. Callico, G., Castaño, I., Carretero, G., Almeida, P., Garcia, A., A. Hernandez, J., and Godtlielsen, F. (2021). Curve-Based Classification Approach for Hyperspectral Dermatologic Data Processing. *Sensors*, 21(3).
- Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5).
- Waltz, E. (2021). Pfizer’s Edge in the COVID-19 Vaccine Race: Data Science . *IEEE Spectrum*.
- Wand, M. P. and Jones, M. C. (1994). *Kernel Smoothing*. Monographs on Statistics & Applied Probability. Chapman & Hall CRC.
- Wei, C., Lunn, D. J., Acerini, C. L., Allen, J. M., Larsen, A. M., Wilinska, M. E., Dunger, D. B., and Hovorka, R. (2010). Measurement delay associated with the Guardian® RT continuous glucose monitoring system. *Diabetic Medicine*, 27(1).

- Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Wiley.
- Wilcox, R. (2017). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier, 4 edition.
- Wild, C. P., Weiderpass, E., and Stewart, B. W. (2020). World Cancer Report 2020. Technical report, World Health Organization.





Part II

The Papers

**Paper I: Early Detection of Change by
Applying Scale-space Methodology to
Hyperspectral Images**

Article

Early Detection of Change by Applying Scale-Space Methodology to Hyperspectral Images

Stig Uteng^{1,*†}, Thomas Haugland Johansen^{2,†} , Jose Ignacio Zaballos^{3,†}, Samuel Ortega^{3,†} , Lasse Holmström^{4,†}, Gustavo M. Callico^{3,†} , Himar Fabelo^{3,†}  and Fred Godtlielsen^{2,†}

¹ Department of Education and Pedagogy, UiT The Arctic University of Norway, 9019 Tromsø, Norway

² Department of Mathematics and Statistics, UiT The Arctic University of Norway, 9019 Tromsø, Norway; thomas.h.johansen@uit.no (T.H.J.); fred.godtlielsen@uit.no (F.G.)

³ Institute for Applied Microelectronics (IUMA), University of Las Palmas de Gran Canaria (ULPGC), 35001 Las Palmas de Gran Canaria, Spain; josegt@gmail.com (J.I.Z.); sortega@iuma.ulpgc.es (S.O.); gustavo@iuma.ulpgc.es (G.M.C.); hfabelo@iuma.ulpgc.es (H.F.)

⁴ Research Unit of Mathematical Sciences, University of Oulu, 90570 Oulu, Finland; lasse.holmstrom@oulu.fi

* Correspondence: stig.uteng@uit.no

† These authors contributed equally to this work.

Received: 30 December 2019; Accepted: 23 March 2020; Published: 27 March 2020



Abstract: Given an object of interest that evolves in time, one often wants to detect possible changes in its properties. The first changes may be small and occur in different scales and it may be crucial to detect them as early as possible. Examples include identification of potentially malignant changes in skin moles or the gradual onset of food quality deterioration. Statistical scale-space methodologies can be very useful in such situations since exploring the measurements in multiple resolutions can help identify even subtle changes. We extend a recently proposed scale-space methodology to a technique that successfully detects such small changes and at the same time keeps false alarms at a very low level. The potential of the novel methodology is first demonstrated with hyperspectral skin mole data artificially distorted to include a very small change. Our real data application considers hyperspectral images used for food quality detection. In these experiments the performance of the proposed method is either superior or on par with a standard approach such as principal component analysis.

Keywords: change detection; scale-space methodology; hyperspectral imaging

1. Introduction

For a time-varying system, detection of unexpected or unwanted change in its evolution can be of paramount importance. Examples include environmental monitoring, process control, or, referring to the examples considered in this article, identification of potentially malignant changes in skin moles or the onset of food quality deterioration (see, for example, [1–4]). The first changes may be small and manifest themselves in different scales and it may be crucial to detect them as early as possible. Statistical scale-space methodologies (see Section 2) can be very useful in such situations since exploring the measurements in multiple resolutions can help identify subtle changes. Examples of scale-space methods designed for change detection are the SiNos technique for capturing non-stationarities in a time series [5] and the iBSiZer method for detecting changes in images [6]. Our goal was to develop a method that can detect minor change while at the same time keeping the number of false alarms to a minimum. This is important in practical applications as a successful method must have both high sensitivity and high specificity.

Recently, Hindberg et al. proposed a scale-space method for testing whether k multivariate data sets of same dimension originate from the same distribution [7]. Thus, the proposed method solves the

classical k -sample problem using scale-space analysis and the method has proven successful in many applications. In the applications considered here the observed data consist of multivariate vectors obtained from spectral signatures and therefore changes in their characteristics can also be analyzed with this method. Unfortunately, it turns out that in this context the method suffers from two serious shortcomings: failing to detect very small changes and producing unacceptably high rates of false alarms in some situations (see Section 4). Our goal therefore is to design a scale-space method that would suffer less from these shortcomings.

As an illustration of the difficulty of detecting very small changes, consider the example in Figure 1 which is discussed in more detail in Sections 2 and 4. The original data set consists of a number of spectral signatures acquired by a push-broom hyperspectral camera, each signature corresponding to a particular spot in a skin mole. Several acquisitions of the mole are taken at the same time, and an example of one acquisition is given in Figure 1a where each curve corresponds to a specific spectral signature. To simulate a situation where the mole might begin to turn malignant, we manually distorted just one spectral signature (thus corresponding to a very small local change in the mole) in another acquisition of the same mole at spectral channel 80 on the horizontal axis in Figure 1a. In case of real moles, the first changes may be extremely hard to detect and a method with high sensitivity and specificity is therefore crucial. In our test, the distorted set of signatures in Figure 1b was compared with 14 other acquisitions and the goal was to detect the small change we manually introduced. It turned out that such a small change is indeed detected by our new methodology but not by the method suggested in [7] nor by a standard approach such as principal components analysis (PCA). We will return in more detail to this example in Section 4.

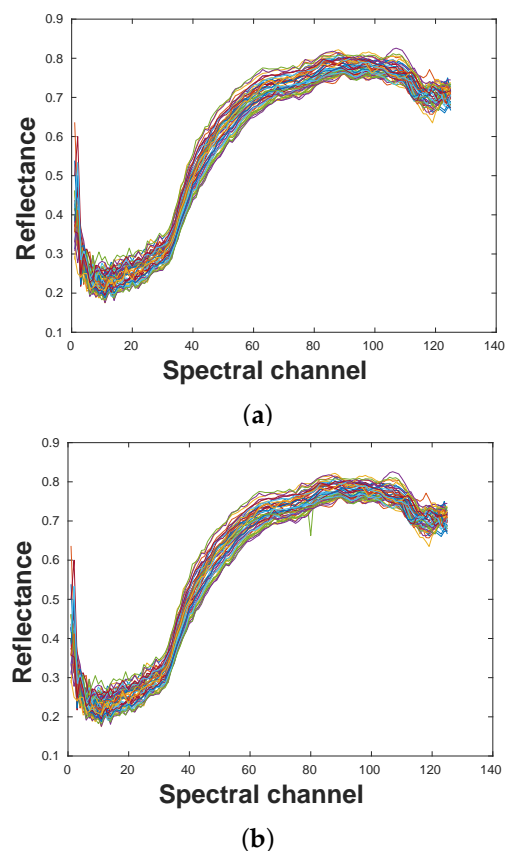


Figure 1. (a) The original undistorted curve families for the artificial example. (b) An example where a small artificial change has been introduced to the data set in Figure 1a at spectral channel 80.

2. Scale-Space Methodology

Scale-space theory is a framework for representing signals on multiple scales, developed by the computer vision, image processing and signal processing communities [8]. A recent review of statistical scale-space methodology can be found in [9]. The goal of statistical scale-space methodology is to extract statistically significant features from noisy data at several scales, often corresponding to different levels of resolution in the underlying object of interest. The data could be a set of observed curves where features at different levels of resolution might be of interest. These curves could, for example, correspond to spectral signatures from fish being frozen for different numbers of days, as is the case in our real data application. One acquisition of data consists of a number of p -dimensional vectors with unknown distribution, each vector representing the spectral signature at a particular pixel in the hyperspectral image. Thus, in our application, p represents the number of frequency bands (spectral channels) in the spectral signatures. In Section 4.2 we analyze three different acquisitions from the frozen fish. Under the null hypothesis, the number of days is assumed the same and the distributions are therefore assumed identical. In our approach, we perform several tests to flag when a new acquisition differs significantly from several previous acquisitions of day 0. The outcome of the tests is presented as a scale-space map, described in more detail below.

The core method of this paper is to test simultaneously for many different scales and positions (frequency bands). The scale s equals the number of different frequency bands being summed across. To be specific, this means that scale $s = 1$ corresponds to the situation where we test if the observed values at spectral frequency d are different between acquisitions of spectral signatures. At scale $s = 3$ and position d , a smoothing in terms of a weighted average of the observed values for spectral frequencies $d - 1$, d and $d + 1$ are used to test whether the acquisitions are different. The weights are calculated from an Epanechnikov kernel function (i.e., parabolic function) [10], the same as in [7]. For other scales, completely analogous smoothing over the frequency bands are made and used to perform the tests. Note that by applying this smoothing, we are able to test for differences in the acquisitions at all locations for a large number of scales. In fact, the tests are performed at all p spectral frequencies for a total number of n_s different scales. Instead of looking at a single location or a single scale, the described scale-space approach can help detect changes that appear at several levels of smoothing, i.e., resolution.

However, when testing for differences between spectral signature curves in different acquisitions it can be difficult to select the critical rejection thresholds due to multiple testing. One possibility is to use the Bonferroni correction method [11] designed to reduce false positives in testing multiple hypotheses. As an alternative, we also tried the statistical inference method described in [12] to find suitable critical rejection thresholds for the scale-space map. The critical values are used to test if a new acquisition differs from the existing acquisitions.

The training procedure at a location (d, s) is accomplished by comparing one acquisition to the others. To simplify the description we illustrate the methodology by testing for change in the sample mean, \bar{X} , over the pixels in the image. This training-procedure is the core difference between the method presented here and in [7], where there is a more direct comparison between curve families. Also, instead of the non-parametric Andersson-Darling test combined with either Bonferroni or False Discovery Rate correction for multiple hypothesis testing employed in [7], our novel method uses the t -test either with a Bonferroni correction or the inference approach suggested in [12]. Here, further, we assume that

$$\bar{X}_1 \sim N(\mu_1, \sigma^2) \tag{1}$$

for acquisition one and

$$\bar{X} = \frac{1}{n-1} \sum_{k=2}^n \bar{X}_k \sim N\left(\mu, \frac{\sigma^2}{n-1}\right), \tag{2}$$

for the remaining $n - 1$ of the acquisitions. Here n is the total number of acquisitions. The normal assumption makes sense due to the central limit theorem since all \bar{X}_k 's, $k = 1, \dots, n$ are averages over a large number of observations. Note that this means that we perform the training procedure by leaving

one out cross validation. In the description above, the mean is chosen as parameter, but we have also implemented and performed our testing procedure for the median, the standard deviation and the range. We do this since these parameters can better describe certain aspects of a distribution and may therefore capture different types of changes. In practice, we will therefore typically test all these parameters for potential changes. For parameters other than the mean, Equation (1) will be an approximation that may be violated in practice. Equation (2) will, however, still be a reasonable approximation for all parameters due to the central limit theorem, but will sometimes only hold approximately.

In our description below, we estimate σ^2 by the standard estimator for variance using all acquisitions apart from the one left out. In the case acquisition 1 is left out, this means that σ^2 is estimated by

$$S^2 = \frac{1}{n-2} \sum_{i=2}^n (\bar{X}_i - \bar{X})^2.$$

The critical quantile at location (d, s) , when only using the Bonferroni correction, is then given by

$$c(d, s) = t_{\frac{\alpha}{2p}, n-2}, \tag{3}$$

where $\alpha = 0.05$ is the significance-level and p the number of spectral channels of each spectral signature curve. In addition to the Bonferroni-corrected quantile in Equation (3), we tried here the so-called global quantile

$$c(d, s)_G = \Phi^{-1} \left(\left(1 - \frac{\alpha}{2} \right)^{\frac{1}{p \sum_{k=1}^{n_s} \theta_k}} \right), \tag{4}$$

proposed in [12]. Here Φ is the normal cumulative distribution function and n_s denotes the number of rows in the scale-space map. Moreover, θ_k is given by

$$\theta_k = 2\Phi \left(\frac{\sqrt{3 \log p}}{2s_k} \right) - 1,$$

where s_k is the scale in row k . In the testing procedure, we test

$$H_0 : \mu_1 = \mu \text{ against } H_1 : \mu_1 \neq \mu$$

using the test statistic

$$T = \frac{\bar{X}_{\text{test}} - \bar{X}_{\text{train}}}{S \sqrt{1 + \frac{1}{n-1}}},$$

where H_0 is rejected if

$$|T| > c(d, s) \text{ or } |T| > c(d, s)_G.$$

The algorithm is summarized in Algorithm 1 where *Par* is used to denote the parameter we are using in the tests.

The outcomes of the tests are graphically summarized in a scale-space map, where the horizontal and vertical axes correspond to spectral frequency bands and scales, respectively. Thus, at each location (d, s) we perform a test and the outcome is shown as a colored pixel, with red (blue) indicating a significant (not significant) difference at the position d for scale s .

To illustrate the method, consider the example introduced in Figure 1. Figure 2 shows the scale-space map produced by the procedure described above. The parameter used in this analysis was the range as it best detected the small change manually introduced to the data. Note how the map indicates a significant feature only for the smallest scales around the spectral channel given at point 80 on the horizontal axis. This is expected since the change is small and only present at one particular spectral channel for a single signature.

Algorithm 1 The SS_CC algorithm:

```

1: Initialization: Acquisitions that are correct under null hypothesis and the test-acquisitions are
   loaded.
2:
3: procedure SS_CC_TRAIN()
4:
5:   Input: The loaded acquisitions that are correct under the null hypothesis.
6:
7:   Initialization: The significance level  $\alpha$  is chosen.
8:
9:   for  $i = 1 : n$  do
10:
11:     procedure LEAVE ONE OUT( $k$ )
12:
13:       return index vector  $\mathbf{v}$  without  $k$ 
14:
15:        $\widehat{Par}(X_k) \sim N(Par, \sigma^2)$  from each  $(d, s)$  location.
16:
17:       for  $\mathbf{j}$  in  $\mathbf{v}$  do
18:
19:          $\widehat{Par}(X) \sim N\left(Par, \frac{\sigma^2}{n-1}\right)$  from each  $(d, s)$  location.
20:
21:       return  $mean(\widehat{Par}(X)), S$ 
22:
23:   return  $mean(\widehat{Par}(X)), S$ 
24:
25: procedure SS_CC_TEST()
26:
27:   Input: The new acquisitions.  $c(d, s), c(d, s)_G, mean(\widehat{Par}(X)), S$ 
28:
29:   Initialization: The significance level  $\alpha$  is chosen.
30:
31:    $T \leftarrow \frac{X_{test} - mean(\widehat{Par}(X))}{S\sqrt{1 + \frac{1}{n-1}}}$ 
32:
33:   return Significance matrix for scale-space map
34:

```

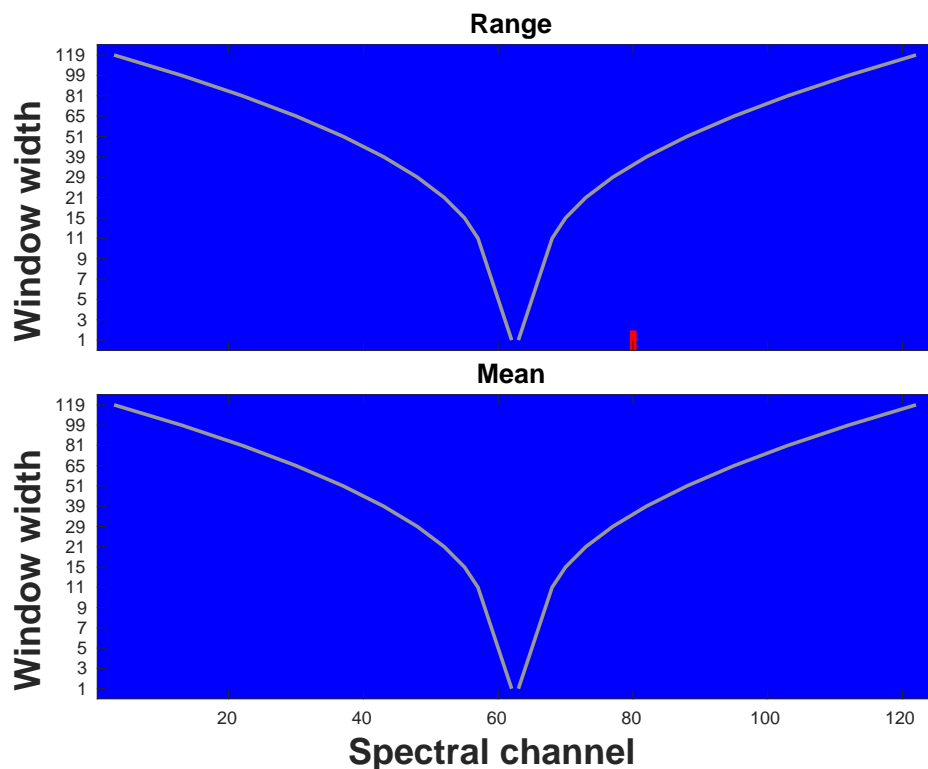


Figure 2. Scale-space significance map for the comparison between the hyperspectral image of a skin mole and an image obtained by manually distorting it. The original and distorted spectral signatures are shown in Figure 1. For the tests, the Hannig-Marron global rejection threshold was used both for the range and the mean.

3. Hyperspectral Acquisition System

In order to capture spectral signature curves from fish, a customized hyperspectral imaging (HSI) acquisition system was employed. Image acquisition was performed with a push-broom hyperspectral

camera with a spectral range of 410–1000 nm (see, for example, [3]) and spatial resolution of 0.3 mm across-track by 0.6 mm along-track (Norsk Elektro Optikk, model VNIR-1024). The camera was fitted with a lens focused at 1000 mm, mounted 1020 mm above a conveyor belt. Samples were illuminated using two custom made fiber optic line lights (Fiberoptics Technology inc., Pomfret, CT, USA), fitted with custom made collimating lenses yielding light lines approximately 5 mm wide (Optec S.P.A., Milano, Italy). Each line light was 400 mm wide, with six bundles of optical fibers. The light from 12 focused 150 W halogen lamps with aluminium reflectors (International Light Technologies, Peabody, MA, USA, model L1090) was fed into the fiberoptic bundles. The imaging and illumination setup is seen in Figure 3a. The optical power actually hitting the sample is approximately 0.16–0.79 Watt/(nm·sr·m²).

The illumination system is composed of a controller unit which allows controlling the brightness and the light source. This system permits us to regulate the light intensity according to the sample characteristics, such as color, size or other parameters dependent on light. The acquisition technique employed by this camera is the so-called push-broom method, which consists of an optical system capturing an image from a line in a plane as depicted in Figure 3b. The camera collects images as seen in Figure 3b.

To capture a hyperspectral image, either the camera or the sample must be moved synchronously with the shoot of the camera. In this case, the sample is moved using a linear actuator by a stepper motor along a line. The light used has been tested to emit in the whole spectral range. Before starting the capturing process, the camera must be focused and calibrated with a dark reference and a white reference. In this process, a tile with 99% of reflectance was chosen for the white reference.

The spectral signatures of the same frozen fish are taken on day 0, day 2, day 4, day 7 and day 10. On each day, we captured 912,082 signature curves in each of the four acquisitions made. The four acquisitions from day 0 were then compared to the other acquisitions in order to find significant differences as described in Section 4.

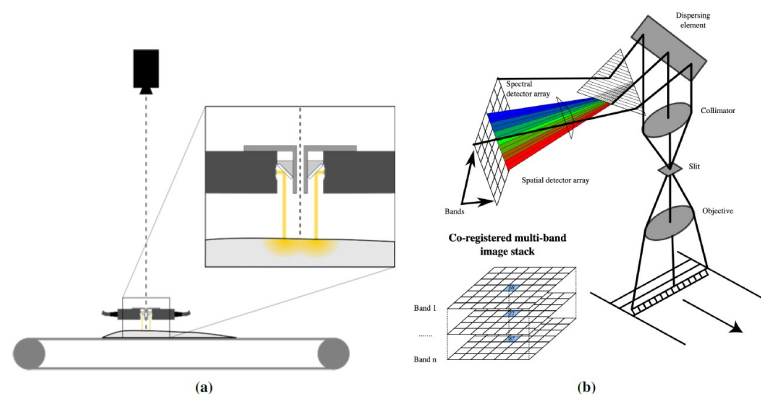


Figure 3. (a) The HSI setup with the Hyperspectral camera. © NOFIMA, Norway. (b) The HSI linear push-broom array. © <https://commons.wikimedia.org/wiki/User:Arbeck>.

After image acquisition, the data from the reference images were used to perform a radiometric calibration of the raw spectral signature of each pixel of the HSI cube as suggested in [13].

$$CI = \frac{RI - DI}{WI - DI} \quad (5)$$

where CI is the calibrated image, RI is the raw image and WI and DI are the white and dark reference images, respectively.

4. Results

4.1. Artificial Mole Example

Recall from Section 1 that the data in our artificial data example were first obtained by acquiring a hyperspectral image of a skin mole and then modifying it manually in order to introduce a small

distortion that could simulate a change in the mole itself. The HSI system used for the mole example differs from the one described in Section 3 and a detailed description can be found in [14]. In our analysis, we compared the proposed novel technique to the method of Hindberg et al. described in [7]. When applied in the present context, this method uses a two-sample test to decide if the test sample distribution differs from the distribution under the null hypothesis. We also experimented with principal component analysis (PCA, e.g., [15]). By examining the scatter plots of the most important principal component directions we concluded that PCA is unable to detect the distortion in the data as seen in Figure 4.

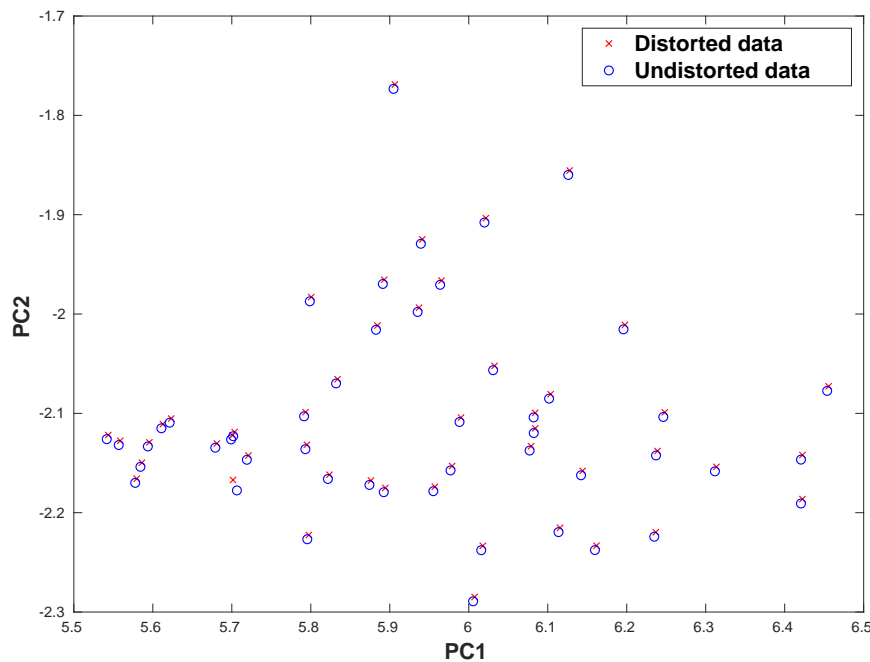


Figure 4. Plot of the first two principal components (PC) of the original mole spectral signatures (blue) and the spectral signatures of the distorted data (red) (see Figure 1b). Note the complete overlap of the two data sets, save for a small change at around $(PC1, PC2) = (5.7, -2.2)$.

Because of this, we focus on a comparison between the new methodology and the one described in [7]. In general, the method in [7] performed poorly in this challenging situation. This was also reflected through the false positive rate (FPR) reported in Figure 5.

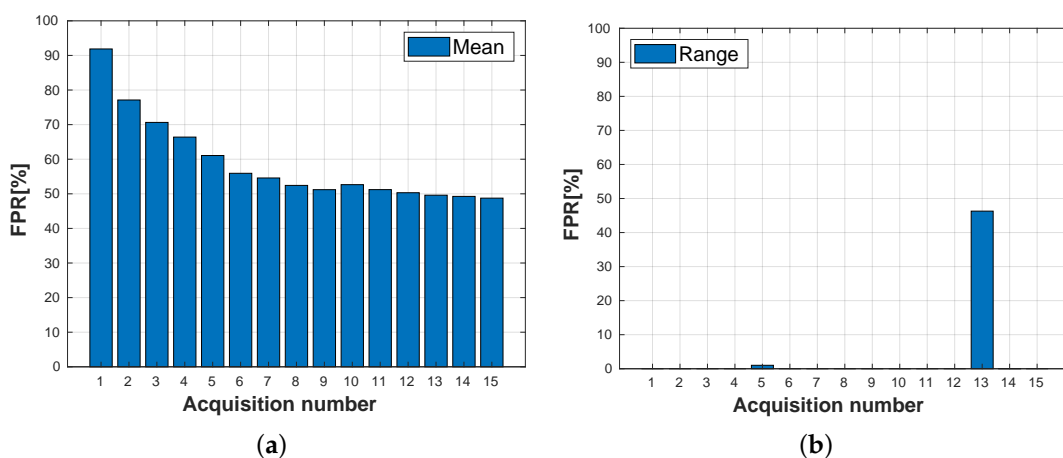


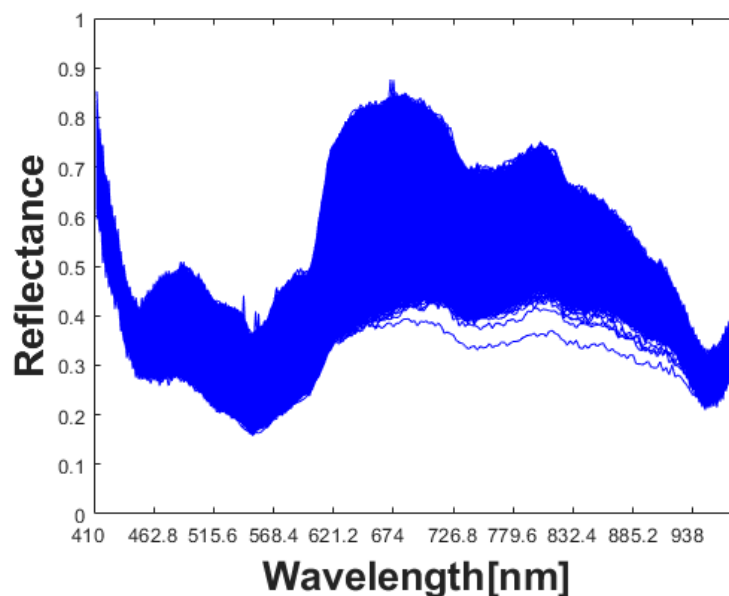
Figure 5. (a) The false positive rate (FPR), performing a leave-one-out test and using Bonferroni correction to account for multiple testing with the method described in [7]. (b) FPR obtained, performing a leave-one-out test and the Hannig-Marron global rejection threshold with the novel method proposed in this article.

The FPR results obtained by the scale-space methodology for the range are quite impressive. However, it should be noted that we were not able to get similar results for the mean, the median or the standard deviation. This is to be expected because the range is clearly best suited for detecting the type of distortion we introduced to the mole data reported in Figure 1. Still, while the range is the natural parameter to use in this situation, it is also clear that range based inference can be very sensitive to outliers, should the data include them.

4.2. Freshness of Fish

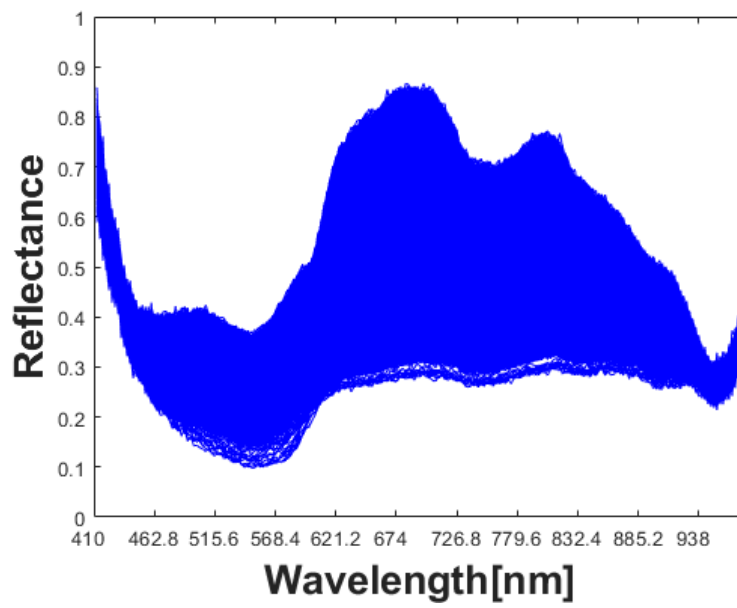
As a real data example, hyperspectral signatures of frozen fish were analyzed. The data acquisitions took place on several different days after a fish was captured. In our analysis, a subset of $301^2 = 90,601$ signatures were analyzed in each acquisition. This is a subset of the full HSI cube that consists of 912,082 spectral signature curves, chosen due to the upper limit of 130 GB of RAM available in the computer used to perform our experiments. An example taken from one fish at day 0 is given in Figure 6a and signatures for two different acquisitions for the same fish at day 4 are given in Figure 6b,c, respectively. Comparison of the signatures of Figure 6a with Figure 6b,c results in the significance maps in Figure 7. There the four panels show the results using two different parameters, the mean and the median. Changes are detected with the mean and the median, but with the standard deviation and the range, no statistically significant changes were detected. A careful examination of the curves in Figure 6 reveals that location parameters are expected to detect changes best in this case and this is indeed what happens. Typical FPR results are reported in Figure 8. From the results here we see that the Hannig-Marron critical value gives better performance. The difference is, however, not very clear in this example.

Due to computational challenges, results for the method of Hindberg et al. in [7] and PCA could not be obtained for the full data sets. In comparisons using only smaller subsets of the data, all three methods performed similarly in detecting changes while their FPRs were similar to Figure 8.

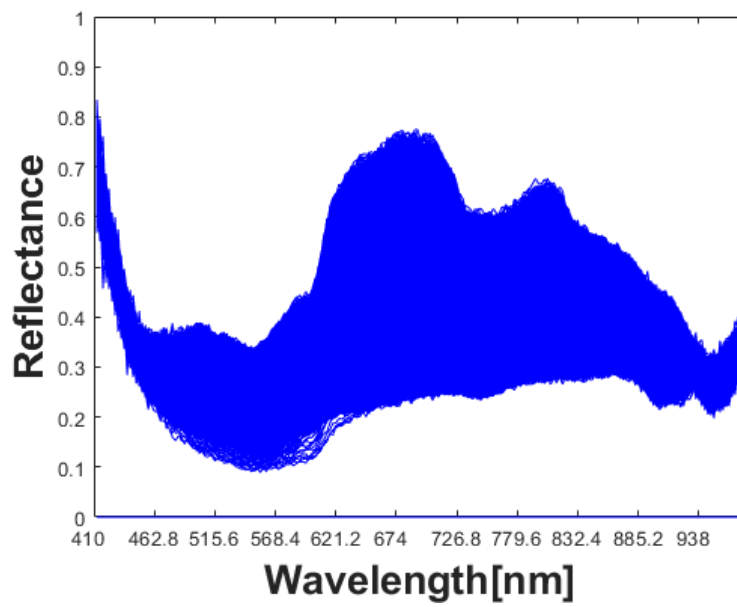


(a)

Figure 6. Cont.

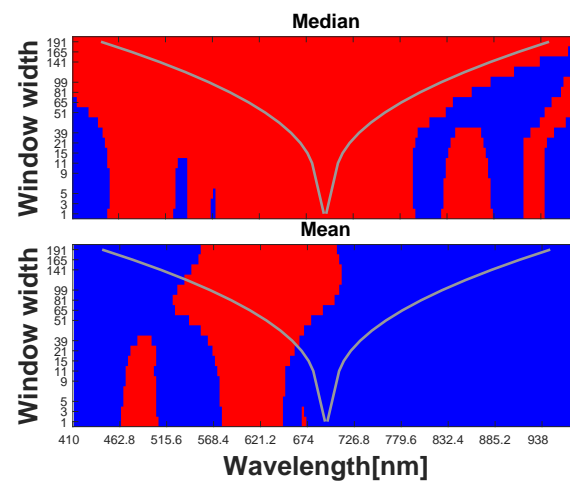


(b)

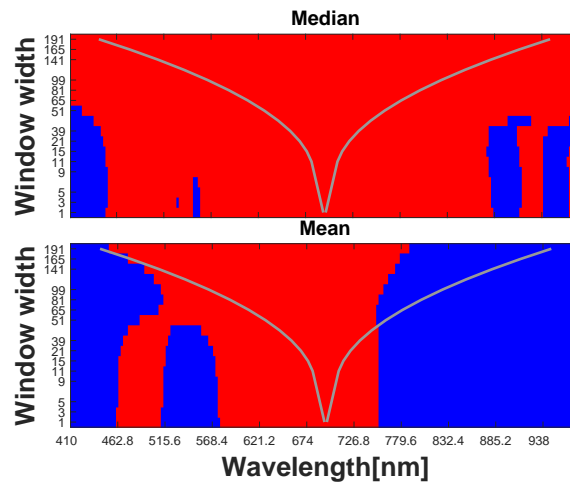


(c)

Figure 6. (a) Plot of HSI-curves from acquisition number one from frozen fish at day 0. (b) Plot of HSI-curves from acquisition number two from frozen fish at day 4. (c) Plot of HSI-curves from acquisition number four from frozen fish at day 4.



(a)



(b)

Figure 7. (a) Significance map for the comparison of the day 0 frozen fish acquisitions with day 4 acquisition number two using the median and the mean. (b) Significance map for the comparison of the day 0 frozen fish acquisitions with day 4 acquisition number four using the median and the mean. The Hannig-Marron rejection threshold was used in all maps.

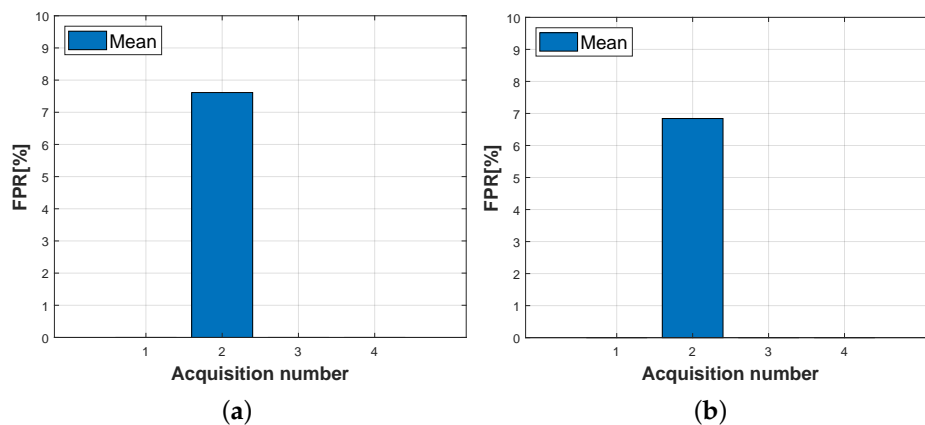


Figure 8. False positive rate (FPR) in the fish freshness example. (a) FPR from a leave-one-out test using Bonferroni correction to account for multiple testing. (b) FPR from a leave-one-out test and using the Hannig-Marron global rejection threshold.

5. Discussion and Future Research Directions

The experimental results of Section 4.1 suggest that the proposed scale-space methodology can be successful in detecting small changes in a hyperspectral image. To be useful in practice, such a method must have both high sensitivity and high specificity and our results for the artificial mole data clearly show promise in this respect. We are currently in the process of acquiring a large number of HSI data sets related to skin moles and lesions in collaboration with several hospitals in the Canary Islands, Spain. Our long term goal is to design a successful classifier for such data and the preliminary results obtained so far are promising [14]. However, we believe that a system capable of monitoring dynamical changes in a mole will be even more important as it is likely to be the best way to detect severe skin cancer at an early stage. In the future, we will therefore work on the development of such a system and our ultimate goal is to design a decision support tool based on just a few frequency bands so that an affordable version could be implemented on a smart phone and thereby be available for use on an individual basis.

One aspect of hyperspectral image data not utilized in the present study is its spatial structure. Taking spatial information into account is important because it can significantly improve the interpretation of the data when changes have been detected. Spatial information can be used both in the development of the change detection algorithm and in the interpretation of the results. Besides mole monitoring applications, a successful change detection method incorporating spatial information could perhaps also be used in the analysis of brain fMRI data for the detection of early signs of, for example, Alzheimer's disease [16].

Another area where the present methodology can be directly applied is in the design of robust controllers for Type 1 Diabetes patients. Successful results in this area are currently being obtained by using reinforcement learning (RL), see, for example, [17]. In the design of such machine learning algorithms, a good description of the patient's state space is needed for the algorithm to be able to learn better strategies. The state space contains information used to describe the patient's condition at a given time. Typically, the elements of state space in this context are time series of the most recent past blood glucose levels of the patient. At the beginning of the learning phase of an RL algorithm, the state space may be chosen reasonably coarse. During the learning process, the state space then may need to change because the algorithm encounters new states, that is, new glucose level time series, not included in the initial state space. The detection of such changes in the state space time series can be accomplished by the kind of methods discussed in this article. Research in this direction will therefore be pursued in the near future.

We also plan to further develop our approach to the analysis of fish freshness discussed in Section 4.2. For the design of a practical system that can be reliably used in fish industry one must first analyze data sets from several different fish at several time points after capturing. Then it is possible estimate both the within variance (of a day) and the between variances (between different time points) exhibited by the hyperspectral signatures. We will acquire such data in the future and the goal will be to perform an analysis that demonstrates how early changes in fish (or other types of food) quality can be detected in a reliable way.

Finally, we believe that the proposed methodology can be useful in combating problems in the so-called " $p > n$ " problems now commonly found in statistical data analyses. Here p and n refer to the number of model parameters and the number of available observations, respectively, and such problems are very common in applications that involve high dimensional data, see e.g., [18,19]. The methodology developed in this article was partly motivated by the need to improve the technique of Hindberg et al. [7] which was originally designed exactly for the $p > n$ situation where common covariance matrix based multivariate methods such as PCA are useless. Being clearly an improvement of the technique of Hindberg et al., the method developed in this article is potentially useful in the analysis of such high dimensional data.

6. Concluding Remarks

We have developed a scale-space methodology that can successfully detect small changes in curve data. In addition, the developed methodology has the potential to produce few false alarms, an important feature for any detection method. We analyzed the performance of the proposed method on data with artificial and real changes. In addition, we compared the new method to some natural competitors and demonstrated that it at least in some cases outperforms them. Finally, we outlined several future research directions for the new methodology that can lead to important new findings.

Author Contributions: S.U.: Method and writing, T.H.J.: Method and writing, J.I.Z.: Method and review, S.O.: Method and review, L.H.: Method and review, G.M.C.: Method and review, H.F.: Method and review, F.G.: Conceptualization and review. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by Tromsø Research Foundation through TFS project ID: 16_TF_FG.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Asokan, A.; Anitha, J. Change detection techniques for remote sensing applications: A survey. *Earth Sci. Inform.* **2019**, *12*, 143–160. [[CrossRef](#)]
- Truong, C.; Oudre, L.; Vayatis, N. Selective review of offline change point detection methods. *Signal Process.* **2019**, *167*. [[CrossRef](#)]
- Lu, G.; Fei, B. Medical hyperspectral imaging: A review. *J. Biomed. Opt.* **2014**, *19*. [[CrossRef](#)] [[PubMed](#)]
- Alander, J.T.; Bochko, V.; Martinkauppi, B.; Saranwong, S.; Mantere, T. A Review of Optical Nondestructive Visual and Near-Infrared Methods for Food Quality and Safety. *Int. J. Spectrosc.* **2013**, *2013*. [[CrossRef](#)]
- Olsen, L.R.; Sørbye, S.H.; Godtliebsen, F. A scale-space approach for detecting non-stationarities in time series. *Scand. J. Stat.* **2007**, *35*, 119–138. [[CrossRef](#)]
- Holmström, L.; Pasanen, L. Bayesian scale space analysis of differences in images. *Technometrics* **2012**, *54*, 16–29. [[CrossRef](#)]
- Hindberg, K.; Hannig, J.; Godtliebsen, F. A novel scale-space approach for multinormality testing and the k-sample problem in the high dimension low sample size scenario. *PLoS ONE* **2019**, *14*. [[CrossRef](#)] [[PubMed](#)]
- Lindeberg, T. Scale-space theory: A basic tool for analyzing structures at different scales. *J. Appl. Stat.* **1994**, *21*, 225–270. [[CrossRef](#)]
- Holmström, L.; Pasanen, L. Statistical Scale Space Methods. *Int. Stat. Rev.* **2017**, *85*, 1–30. [[CrossRef](#)]
- Wand, M.P.; Jones, M.C. Kernel Smoothing. In *Monographs on Statistics & Applied Probability*; Chapman & Hall: London, UK, 1994; ISBN 9780412552700.
- Hochberg, Y.; Tamhane, A.C. *Multiple Comparison Procedures*; Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics; Wiley: Hoboken, NJ, USA, 1987; ISBN 9780471822226.
- Hannig, J.; Marron, J.S. Advanced distribution theory for SiZer. *J. Am. Stat. Assoc.* **2006**, *101*, 484–499. [[CrossRef](#)]
- Lawrence, K.C.; Park, B.; Windham, W.R.; Mao, C. Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. *Trans. Am. Soc. Agric.* **2003**, *46*, 513–521. [[CrossRef](#)]
- Fabelo, H.; Melián, V.; Martínez, B.; Beltrán, P.; Ortega, S.; Marrero, M.; Callicó, G.M.; Sarmiento, R.; Castaño, I.; Carretero, G.; et al. Dermatologic Hyperspectral Imaging System for Skin Cancer Diagnosis Assistance. In Proceedings of the 2019 XXXIV Conference on Design of Circuits and Integrated Systems (DCIS), Bilbao, Spain, 20–22 November 2019.
- Johnson, R.A.; Wichern, D.W. *Applied Multivariate Statistical Analysis (Classic Version)*; Pearson Modern Classics for Advanced Statistics Series; Pearson: London, UK, 2018; ISBN 9780134995397.
- Malik, F.; Farhan, S.; Fahiem, M.A. An ensemble of classifiers based approach for prediction of Alzheimer's disease using fMRI images based on fusion of volumetric, textural and hemodynamic features. *Adv. Electr. Comput. Eng.* **2018**, *18*, 61–71. [[CrossRef](#)]
- Tejedor, M.; Woldaregay, A.Z.; Godtliebsen, F. Reinforcement learning application in diabetes blood glucose control: A systematic review. *Artif. Intell. Med.* **2020**, *104*. [[CrossRef](#)]
- Giraud, C. *Introduction to High-Dimensional Statistics*. In *Monographs on Statistics & Applied Probability*; Chapman & Hall: London, UK, 2014; ISBN 9781482237955.

19. Aoshima, M.; Shen, D.; Shen, H.; Yata, K.; Zhou, Y.H.; Marron, J.S. A survey of high dimension low sample size asymptotics. *Aust. New Zealand J. Stat.* **2018**, *60*, 4–19. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Paper II: Curve-based Classification
Approach for Hyperspectral Dermatologic
Data Processing**

Article

Curve-Based Classification Approach for Hyperspectral Dermatologic Data Processing

Stig Uteng ^{1,*} , Eduardo Quevedo ² , Gustavo M. Callico ² , Irene Castaño ³, Gregorio Carretero ³, Pablo Almeida ⁴, Aday Garcia ⁵ , Javier A. Hernandez ⁴ and Fred Godtlielsen ⁶

- ¹ Department of Education and Pedagogy, UiT the Arctic University of Norway, 9019 Tromsø, Norway
- ² Institute for Applied Microelectronics, Universidad de Las Palmas de Gran Canaria, 35016 Las Palmas de Gran Canaria, Spain; eduardo.quevedo@ulpgc.es (E.Q.); gustavo@iuma.ulpgc.es (G.M.C.)
- ³ Department of Dermatology, Hospital Universitario de Gran Canaria Doctor Negrín, 35016 Las Palmas de Gran Canaria, Spain; irenecastano@hotmail.com (I.C.); gcarher@gobiernodecanarias.org (G.C.)
- ⁴ Department of Dermatology, Complejo Hospitalario Universitario Insular-Materno Infantil, 35016 Las Palmas de Gran Canaria, Spain; pjalmmar75@gmail.com (P.A.); jhersant@gobiernodecanarias.org (J.A.H.)
- ⁵ Department of Electromedicine, Complejo Hospitalario Universitario Insular-Materno Infantil, 35016 Las Palmas de Gran Canaria, Spain; agarcia@iuma.ulpgc.es
- ⁶ Department of Mathematics and Statistics, UiT the Arctic University of Norway, 9019 Tromsø, Norway; fred.godtlielsen@uit.no
- * Correspondence: stig.uteng@uit.no



Citation: Uteng, S.; Quevedo, E.; M. Callico, G.; Castaño, I.; Carretero, G.; Almeida, P.; Garcia, A.; A. Hernandez, J.; Godtlielsen, F. Curve-Based Classification Approach for Hyperspectral Dermatologic Data Processing. *Sensors* **2021**, *21*, 680. <https://doi.org/10.3390/s21030680>

Received: 15 December 2020

Accepted: 15 January 2021

Published: 20 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This paper shows new contributions in the detection of skin cancer, where we present the use of a customized hyperspectral system that captures images in the spectral range from 450 to 950 nm. By choosing a 7×7 sub-image of each channel in the hyperspectral image (HSI) and then taking the mean and standard deviation of these sub-images, we were able to make fits of the resulting curves. These fitted curves had certain characteristics, which then served as a basis of classification. The most distinct fit was for the melanoma pigmented skin lesions (PSLs), which is also the most aggressive malignant cancer. Furthermore, we were able to classify the other PSLs in malignant and benign classes. This gives us a rather complete classification method for PSLs with a novel perspective of the classification procedure by exploiting the variability of each channel in the HSI.

Keywords: hyperspectral; curve fit; statistical discrimination; melanoma; benign; malignant

1. Introduction

Hyperspectral (HS) imaging (HSI) combines conventional imaging and spectroscopy methods in a single imaging technique providing both spatial and spectral information of the captured display [1,2]. It is also a fitting method for medical applications due to its non-invasive, non-ionizing, and label-free nature [3]. Dermatology is one of the medical fields where HSI have shown its potential as an appropriate imaging technique [4].

Currently, the most common form of cancer, with more than 1.3 million new cases worldwide in 2018, is skin cancer [5]. There are several types of skin lesions. Pigmented skin lesions (PSLs) contain a wide variety of types, including cancerous and non-cancerous PSLs [6]. There are two types of PSLs depending on the type of growth of the tissue: benign and malignant. Nevus, which is benign, has a slow growth rate and is noncancerous, while e.g., melanomas, which is malignant, are invasive and potentially metastatic tumors [6]. Other types of skin cancer produced by different types of cells include squamous cell carcinoma and basal cell carcinoma. However, melanomas are much more dangerous than other types of skin cancer, and an early detection of this skin lesion can be extremely

important in improving the patient survival [7]. The diagnosis of PSLs is performed by dermatologists employing their naked eye or dermoscopic cameras, which enhance the morphological visualization of the PSL. After the inspection, an analysis of the lesion following the ABCDE (Asymmetry, Border irregularity, Color, Diameter and Evolving size, shape, or color) rule is applied in order to establish a preliminary diagnosis [7]. A suspicious lesion will accordingly get a histopathological analysis carried out to determine the definitive diagnosis. These diagnostic tools based on conventional imaging have been used to help dermatologists in preliminary diagnosis. However, conventional imaging has limitations that could be surpassed by the enriched spectral information provided by the HSIs.

HSI technology has already been explored as a target technology to aid in PSLs diagnosis. A recent study by Leon et al. showed promising results in discriminating PSLs by random forests and artificial neural networks [8]. Another recent study by Hosking et al. used digital dermoscopy HSI with machine learning methods in order to detect melanomas [9]. Kazianka et al. used HSIs to differentiate between normal skin, melanomas, and moles [10]. They used unsupervised approaches to segment the images and moreover evaluated the performance of several supervised classifiers aimed to retrieve the diagnosis of each pixel. The results indicated that it was possible to differentiate melanomas from moles with high specificity and sensitivity. In another study, Nagaoka et al. [11] collected a dataset for discriminating between melanomas and other PSLs, evaluating the statistical significance of an index defined to this end. Tomatis et al. used a classifier based on a neural network model [12]. There have also been developed some commercial systems as SIAscope/SIAscopy [13] and MelaFind [14–16]. MelaFind is being used in several studies. In [15], Elbaum et al. conduct a leave-one-out cross-validation procedure using a database composed of 183 melanocytic nevus and 63 melanomas. Monheit et al. [16] applied it in a multicenter study. At last, Song et al. performed a paired comparison between MelaFind and a reflectance confocal microscopy system to differentiate between melanoma and non-melanoma PSLs by comparing MelaFind with a confocal microscopy system. In the study by Stamnes et al., they discriminate between malignant and benign PSLs, i.e., no specific melanoma detection [17].

The current default method of data analysis (DA) is deep learning (DL). DL emerged as a competing DA method around 2009 and solves a wide range of problems exemplified by the versatile TensorFlow-package [18]. DL is also applied in the HSI field. However, DL often needs a huge amount of data in order to train the parameters enough in the neural network [19]. Thus, in the proposed method, this presents a difficult problem due to the lack of data. We are pursuing this problem by an ongoing data collection project, aiming for future applications of DL systems.

Most of the research in the use of HSI for skin analysis is focused on the automatic diagnosis of PSLs. Clearly, a successful methodology of this type would be an extremely useful decision support tool for general practitioners (GPs) and dermatologists. In our contribution in this direction, we were able to find clear patterns for the melanoma PSLs, the other malignant PSLs and the benign PSLs. In the novel exploitation of the variability of each channel in the HSI cubes, we made curves of the mean and standard deviation of sub-images in these channels. Furthermore, using these curves and their inherent patterns, we were able to discriminate between malignant and benign PSLs, but also between the aggressive melanoma PSL and the others.

The outline of the paper is as follows. First, we describe the developed dermatological HSI acquisition system, and thereafter, we describe the datasets we are using in this curve-based classification method. Next, we point out how the data are pre-processed and labeled before they are used as input in our new classification method, which consists of a training, validation, and testing phase.

2. Materials and Methods

2.1. Hyperspectral In Vivo Dermatologic Data

The system used to acquire the HS dermatologic images is described in [20], using the same database as in Leon et al. [8]. The database consists of 76 images of PSLs from 61 patients: 36 cancerous and 40 noncancerous. The dataset is divided into a training set, a test set, and a validation set. The training set is subsequently divided into melanoma PSLs, malignant PSLs, and benign PSLs. Each mentioned set is mutually exclusive. The acquisition system is composed of an HS snapshot camera able to capture HS images in the very near infrared (VNIR) range, between 450 and 950 nm, with a spatial resolution of 50×50 pixels and 125 spectral bands. This acquisition system employs a customized dermoscopic contact structure and a halogen source light (150 W) coupled to a fiber optic ring light guide for cold light emission. The effective capturing area of the system is 12×12 mm, and the acquisition time is lower than 1 s. The system was applied to create an HS database composed of 76 images of PSLs from 61 patients. The data acquisition was done at the Hospital Universitario de Gran Canaria Doctor Negrín and Complejo Hospitalario Universitario Insular-Materno Infantil (Spain). The research protocol and consent procedures were approved by the Comité Ético de Investigación Clínica-Comité de Ética en la Investigación (CEIC/CEI) at the University Hospital Doctor Negrin.

2.2. Curve-Based Classification Approach

In the pre-processing step, a calibration is done of the HSI using the white and dark reference images following Equation (1), where CI is the calibrated image, RI is the raw image obtained from the HS camera, and DI and WI are the dark and white reference images, respectively. The white reference image is obtained capturing an image of a white reference tile that reflects 99% of the incident light. The dark reference image is captured by keeping the tap in the lens of the camera.

$$CI = \frac{RI - DI}{WI - DI} \quad (1)$$

Finally, the data is normalized in the range [0,1] to avoid differences in the spectral signature intensities caused by possible different illumination conditions.

2.3. Region of Interest Curves

The region of interest (ROI) was obtained through the extraction of 7×7 sub-images for all the channels. Subsequently, the mean (x-coordinate of the curves) and standard deviation (y-coordinate of the curves) of these ROIs from each HS channel were plotted against each other, thus creating ROI curves for each HSI cube. These curves were rather different, some almost linear, other more convoluted. Thus, we applied a morphological dilation filter to the more complex curves to ease the interpretation and subsequently use these filtered curves in function fits. Morphological dilation is a common way to enhance or change digital images in some way. Thus, often used in bridging gaps, our morphological filter can be seen as an inverse dilation, a pruning.

There were some issues of this method that needed to be addressed. In particular, how the ROIs are chosen. So, the first issue is where to collect this ROI in the HSI, i.e., which coordinates to choose? To shed some light on this, we established some criteria for each of the classes of PSLs and further picked the ROIs best suited for creating the curves with adequate classification properties. In each channel, we started from $(x,y) = (16, 16)$ and went to $(x,y) = (35, 35)$. This gave us $20 \times 20 = 400$ HS ROI curves for each PSL HSI cube collected from different locations. However, with some overlap and since each ROI area is 7×7 pixels, we cover the HSI from the start at $(x,y) = (13, 13)$ to the end at $(x,y) = (38, 38)$: thus, an area of $26 \times 26 = 676$ pixels. We tried a wider area, but the results did not improve, so we kept this size due to computational cost. Thus, we tested how the ROI extraction affected the mean square error (MSE) of the curves chosen by suitable criteria for the melanoma as well as the other malignant and benign PSLs.

The CPU used to run the code was an Intel Xeon 2.8 GHz running on a 16 GB RAM computer.

Curve-Fitting

By trying to classify the PSLs through fits of the ROI curves, all the HSI ROI curves were fitted by a quartic (fourth degree) polynomial function, where the coefficients are found through the `nlinfit` function of MATLAB[®], which minimizes the weighted least squares equation,

$$\sum_{i=1}^n w_i (y_i - f(x_i, b))^2 \quad (2)$$

where the weights were chosen through a weight function, $w(y) = 1/(0.011 + 0.011y)^2$ and $n = 125$, the number of data points. Further is $f(x, b) = b_0 x^4 + b_1 x^3 + b_2 x^2 + b_3 x + b_4$, the quartic regression function, with parameters $b_j, j = 0, \dots, 4$. Then, the `nlinfit`-function uses the iteratively reweighted least-square algorithm, which is designed to deal with outliers [21]. Furthermore, the use of a quartic polynomial with a weight function proved to be the best option with regard to flexibility and robustness to outliers. We also tested higher degree polynomial functions; however, they were prone to overfitting.

Figure 1 gives an overview of the algorithm. The constants a, b, c, d, e, f , and g are found through the training phase. The features are as follows: df (first-order derivative), ddf (second-order derivative), $totm = a \cdot df + b \cdot ddf$, $mabdf = \text{mean} + \text{abs}(df)$ and mean .

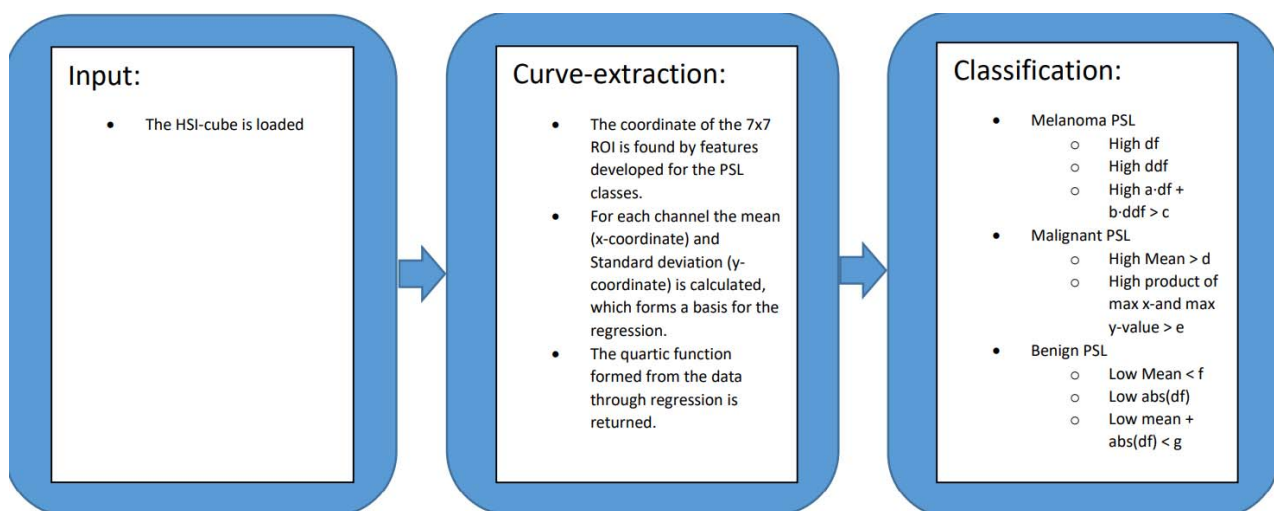


Figure 1. Schematic overview of the algorithm.

3. Results

Here, we will represent the obtained experimental results. We start the representation of the training phase.

3.1. The Training Phase of the Method

As mentioned above, we acquired an ROI from each channel of the HSIs, where the curves were formed through the calculation of the mean and standard deviation of the ROIs from each channel. First, we investigated the melanoma PSLs. The features coming most natural from earlier experiments are a high mean first-order derivative (Max df) and a high mean curvature given by the mean of the second derivative (Max ddf). However, during the training phase, we found that a high sum of mean df (multiplied by 1, a in Figure 1), mean ddf (multiplied by 0.1, b in Figure 1), and the derivative at the maximum value gave the most successful feature (Max $totm > 2.86$, c in Figure 1). In Table 1 and Figure 2a, we can observe that this feature produced rather good fits according to the MSEs: the MSE ddf following this measure and the MSE df lower for two melanomas. Thus, all

the plots of the HSI ROI curves from the melanoma PSLs are from the fits of the Max totm curves. Moreover, the recurring theme in Figure 2c–f is that all the HSI ROIs came from the interior and/or near the edges of the melanomas. In Table 1, we can also observe the most distinct outlier, the P82C1000 PSL, which did not have the same characteristics as the other melanomas, behaving more similar to the malignant PSLs. In Figure 2a, we can observe that this PSL is the one with an MSE ddf diverging most clearly from the MSE totm.

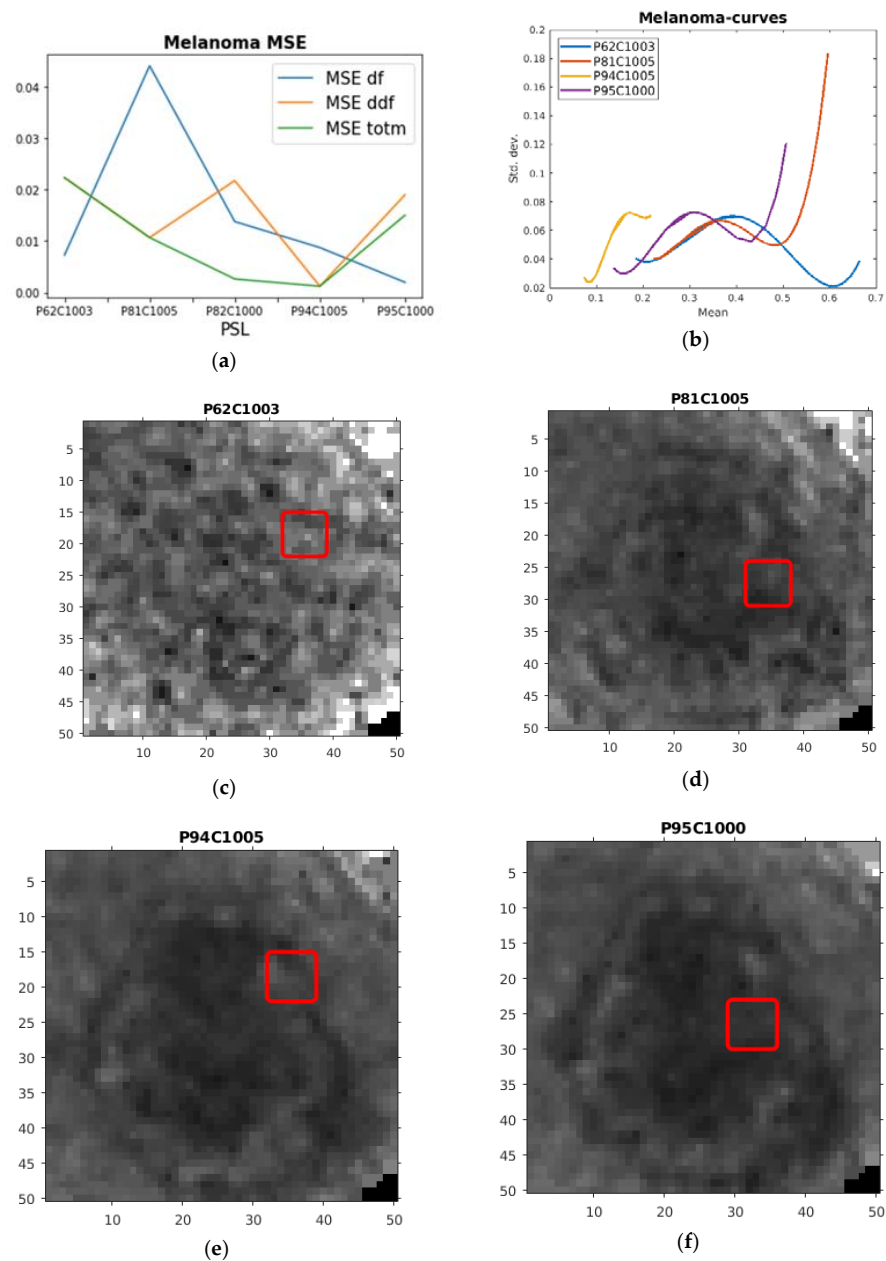


Figure 2. (a) Plot of mean square error (MSE) first-derivative feature (MSE df), second-derivative feature (MSE ddf) and sum feature (MSE totm) for the hyperspectral image (HSI) ROI curves from the training melanoma PSLs. (b) Plot of quartic polynomial fits of melanoma HSI ROI curves chosen according to the low MSE totm. (c) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P62C1003 PSL. (d) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P81C1005 PSL. (e) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P94C1005 PSL. (f) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P95C1000 PSL.

Table 1. The results for the region of interest (ROI) investigation for the melanoma pigmented skin lesions (PSLs) with the corresponding curves with maximum mean first derivative, Max df, maximum curvature, Max ddf, maximum weighted sum of mean df, mean ddf and df at the maximum point, Max totm and the corresponding MSEs of the chosen curves.

PSL	Max df	Max ddf	Max totm	MSE df	MSE ddf	MSE totm
P62C1003	0.19959	5.2261	1.2716	0.00729	0.022353	0.022353
P81C1005	0.90111	14.465	5.0123	0.044138	0.010691	0.010691
P82C1000	0.17383	−0.1158	−0.3373	0.01381	0.021781	0.0026241
P94C1005	1.1388	45.812	5.0485	0.008742	0.001214	0.0012136
P95C1000	0.60491	19.092	3.3212	0.001983	0.019006	0.015033

Regarding the malignant PSLs, we inferred based on earlier experiments that the HSI ROI curves with a high mean (Max Mean > 0.05, d in Figure 1) or a high product of max x and max y (Max Prod > 0.395, e in Figure 1) from the fitted curve are plausibly originating from this type of PSL. This proved successful for some of the HSI ROI curve fits from the malignant PSLs. As seen in Table 2 and Figure 3a, this was the most difficult type PSL to fit, giving rise to rather high MSEs. The Max Mean was the best feature accordingly. Some of the curves with least MSE Mean are plotted in Figure 3b, where all have means lying above the standard deviation (y-value) of 0.05 (d in Figure 1). From these HSI ROI mean standard deviation curves, the most complex data occurred—thus, as mentioned above, giving rise to rather high MSEs. The corresponding ROIs of the plotted curves are also here lying on the inside or at the edge of PSLs, as seen in Figure 3c–f.

Table 2. The results for the ROI investigation for the malignant PSLs with the corresponding curves with maximum product of max x times max y, Max Prod, maximum mean, Max Mean, and the corresponding MSEs of the chosen curves.

PSL	Max Prod	Max Mean	MSE Prod	MSE Mean
P67C1003	0.20058	0.15051	3.4133	2.0846
P32C1000	0.05982	0.054024	0.02244	0.006436
P87C1000	0.33499	0.2143	11.323	0.41494
P106C1000	0.9194	0.35846	15.474	11.997
P21C1000	0.13709	0.11865	0.039819	0.16209
P56C1002	0.73114	0.17111	5.6309	3.7864
P66C1001	0.17669	0.11542	5.3595	0.64848
P75C1000	0.36819	0.18003	4.1902	4.234
P77C1000	0.31525	0.15402	1.3532	0.65383
P80C1003	0.094159	0.13404	0.2015	0.033497
P88C1000	0.16356	0.14231	1.5906	0.038208
P89C1001	0.23628	0.15301	3.1098	3.8229
P90C1002	0.33155	0.19675	3.872	0.087464
P91C1003	0.1763	0.12843	0.039625	0.039625
P101C1000	0.63537	0.24154	25.639	1.0538
P104C1000	0.77046	0.20304	8.3724	8.3684
P110C1000	0.18187	0.097568	0.96777	0.96777
P112C1000	0.34958	0.33396	0.19203	0.045852
P116C1004	0.26348	0.14593	2.9284	2.9284
P92C1004	1.4734	0.28289	19.805	21.99
P109C1000	0.38193	0.23467	0.23374	0.48969

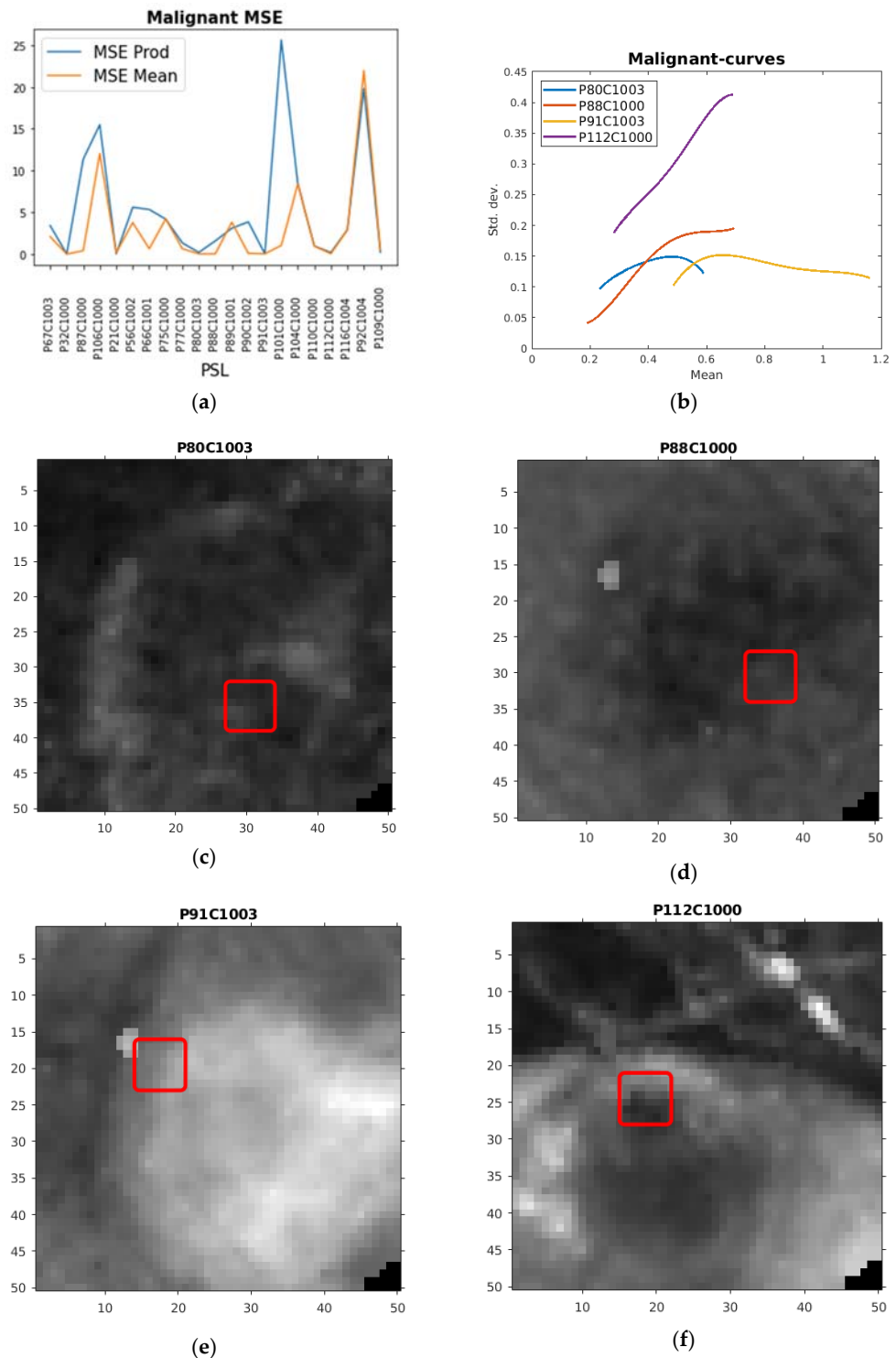


Figure 3. (a) Plot of MSE product of max x and max y (MSE Prod) and maximum mean (MSE Mean) for the fitted HSI ROI curves from the malignant PSLs. (b) Plot of quartic polynomial fits of malignant HSI ROI curves chosen according to low MSE mean. (c) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P80C1003 PSL. (d) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P88C1000 PSL. (e) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P91C1003 PSL. (f) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P112C1000 PSL.

The benign PSLs HSI ROI curves proved easier to fit than the malignant ones. From earlier experiments, two useful features came to be low mean (Min Mean < 0.05, g in Figure 1), low absolute value first order derivative (Min abs df), and the sum of these (Min mabdf < 0.109, f in Figure 1). This proved rather successful with a few exceptions, as can be observed in Table 3 and Figure 4a, with Min abs df and Min mabdf as similar rather successful features. In Figure 4b, four of the ROIs with lowest MSE abs df are plotted with their corresponding placements of the ROIs given in Figure 4c–f. We can here see that for these PSLs, the ROIs are somewhat different placed than for the melanomas and malignant ones—not all inside the PSLs, but at some sort of feature, nevertheless.

Table 3. The results for the ROI investigation for the benign PSLs with the corresponding curves with minimum mean, Min Mean, minimum absolute value of the first derivative Min abs df, minimum sum of mean and absolute value of the first derivative Min mabdf, and the corresponding MSEs of the chosen curves.

PSL	Min Mean	Min absdf	Min mabdf	MSE Mean	MSE absdf	MSE mabdf
P60C3004	0.035875	0.04211	0.078126	0.43249	0.49182	0.49182
P24C1000	0.047758	0.13751	0.20262	5.7292	2.9709	2.9709
P61C1004	0.037392	0.019541	0.070525	0.11543	0.12084	0.049766
P78C3000	0.054088	0.14207	0.22912	2.0738	0.98825	0.98825
P83C1003	0.053259	0.051972	0.15544	2.2384	0.006806	0.20734
P29C1000	0.017986	0.037797	0.067169	0.37593	0.056122	0.056122
P29C2000	0.016499	0.028643	0.050963	0.12963	0.029577	0.029577
P13C2000	0.018738	0.025028	0.068659	0.004459	0.018544	0.073912
P13C3000	0.020795	0.058914	0.086964	0.097557	0.033792	0.020257
P16C1000	0.017759	0.049908	0.073101	0.18981	0.013001	0.013001
P17C1001	0.019861	0.021036	0.055137	0.15354	0.013007	0.028137
P17C2002	0.023828	0.012184	0.047559	0.079709	0.020721	0.018856
P18C1000	0.023887	0.010399	0.054889	0.11762	0.007339	0.0073387
P25C2000	0.053336	0.21251	0.27597	4.1742	0.91204	0.91204
P25C3000	0.057866	0.13919	0.22013	2.3664	13.174	9.6771
P26C1000	0.050246	0.18733	0.2605	4.5807	1.4319	1.4319
P27C1000	0.023408	0.052952	0.08173	0.12196	0.039168	0.047658
P27C2000	0.020774	0.045139	0.069434	0.24855	0.16042	0.16042
P27C3000	0.02075	0.058085	0.086034	0.029003	0.029591	0.016113
P27C4000	0.019193	0.026195	0.051928	0.070661	0.037739	0.037739
P29C3000	0.017297	0.042738	0.068037	0.14874	0.01073	0.022491
P30C1000	0.02094	0.00574	0.045026	0.09903	0.001885	0.0032779

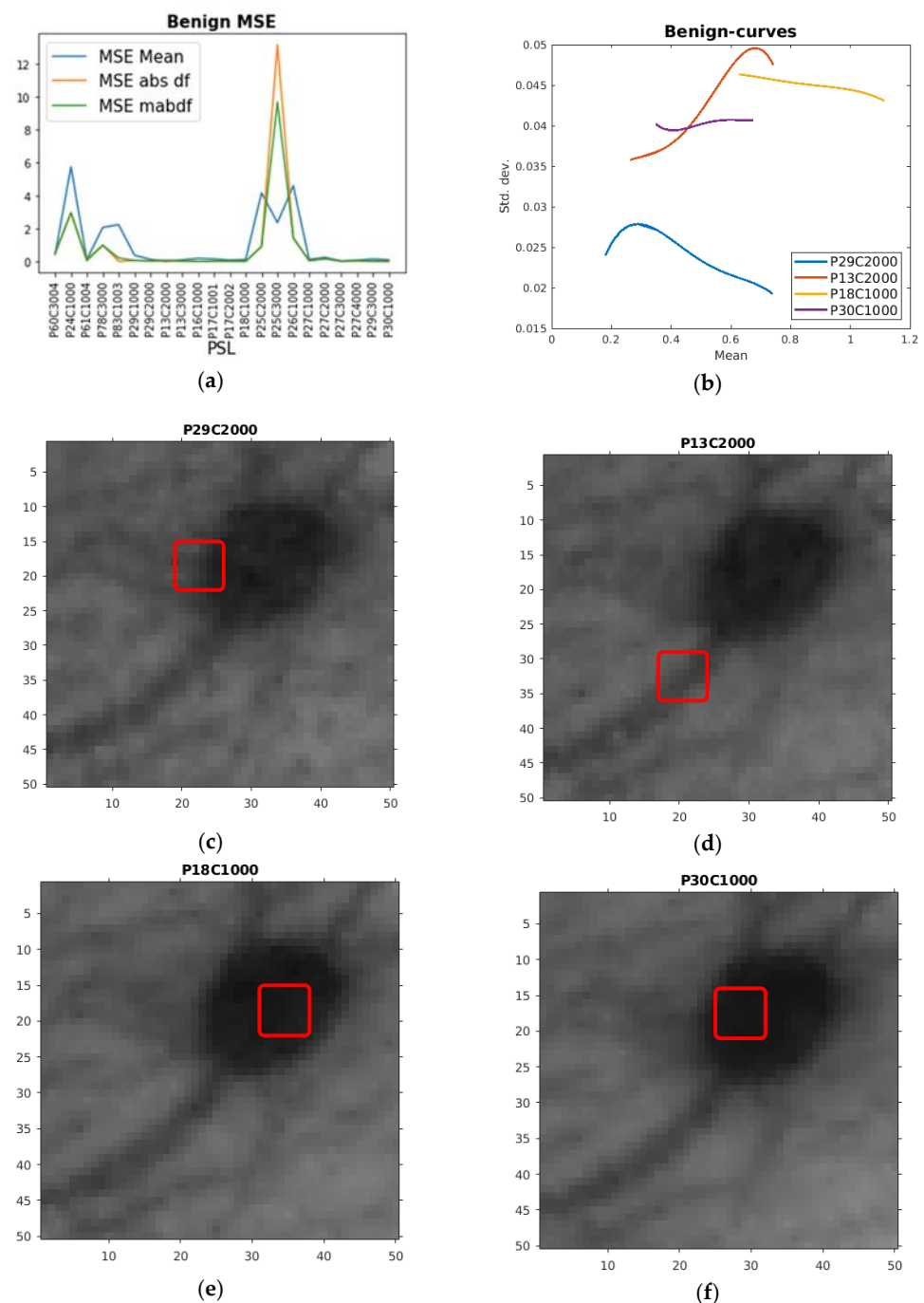


Figure 4. (a) Plot of MSE mean feature (MSE Mean), absolute value of the first derivative feature (MSE abs df), and the sum of these (MSE mabdf) for the HSI ROI curves from the benign PSLs. (b) Plot of quartic polynomial fits of benign HSI ROI curves chosen according to low MSE abs df. (c) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P29C2000 PSL. (d) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P13C2000 PSL. (e) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P18C1000 PSL. (f) The corresponding ROI in red drawn on the HSI for one of the channels used in the creation of the HSI ROI curve from the P30C1000 PSL.

Classification

From the training phase, we inferred that the classification procedure can consist of two consecutive steps.

First, we identify the curves where the first derivative increases, possibly converging at maximum value of the HSI ROI curves mean. Additionally, the curves with high mean positive curvature and less double inflexion points were also identified as melanomas. In addition, the sum $df + 0.1ddf$ proved successful (totm).

Then, we find the two clusters of the rest of the curves, such that they can be classified as benign or malignant. This is mainly done by choosing what curves have mean above $y = 0.05$ (malignant) and below $y = 0.05$ (benign). For the benign curves, an additional useful feature was the sum of the mean of the fitted curve and the mean absolute value of the first-order derivative in each point (mabdf).

3.2. Validation and Test Experiment

We also performed an experiment in order to validate the training phase with the same type of ROI curves as in the training part but different PSLs. This resulted in the usual 125 points mean and standard deviation curve; then, a quartic polynomial function was fitted, as mentioned earlier. Based on the features developed in the training phase, then, we were able to first discriminate between melanomas and the rest of the PSLs, where a typical melanoma fit is given in Figure 5a, and then discriminate between malignant and benign PSLs. Examples of their respective fits given in Figure 5b,c.

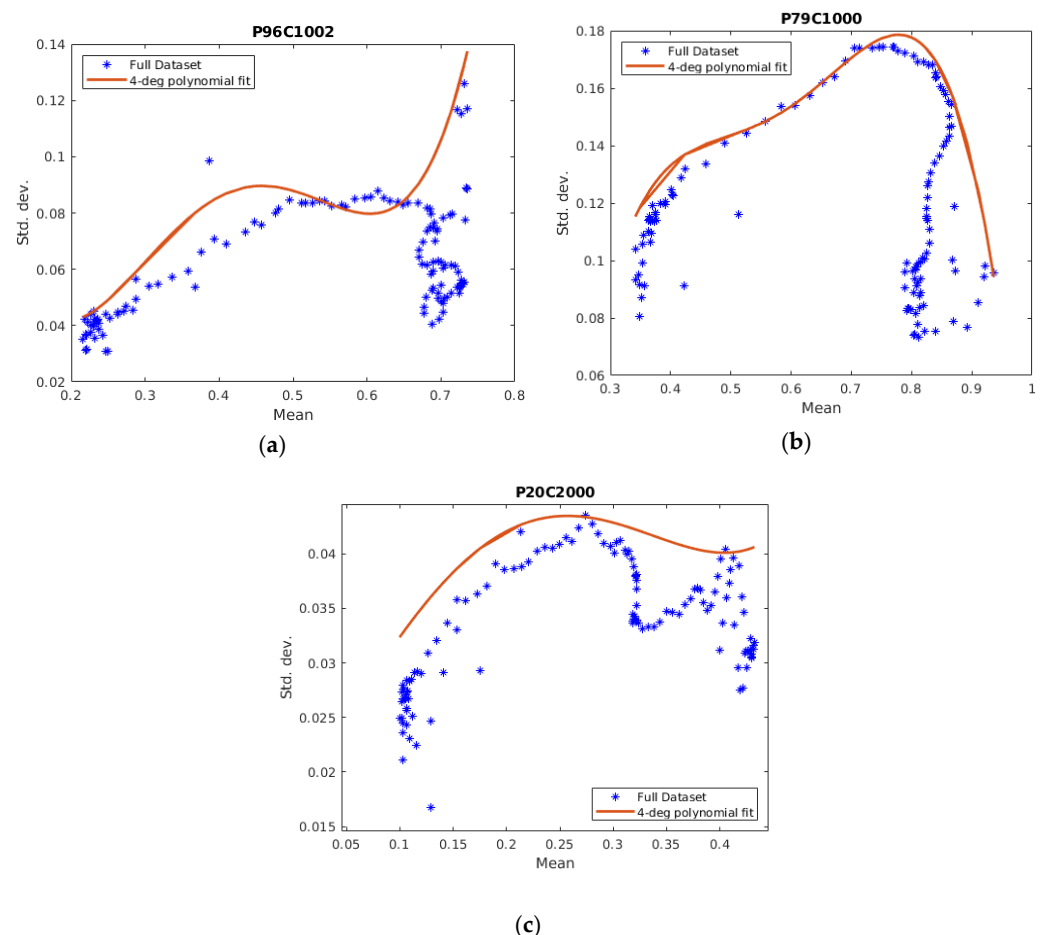


Figure 5. Examples of the fitted curve plots with data points from the validation phase: (a) Plot of a quartic polynomial fit of a melanoma HSI ROI curve and the data points. (b) Plot of a quartic polynomial fit of a malignant HSI ROI curve and the data points. (c) Plot of a quartic polynomial fit of a benign HSI ROI curve and the data points.

When looking at the results, we denote correctly classifying a malignant PSL as a true positive (tp) and correctly classifying a benign PSL as a true negative (tn). Furthermore,

falsely classifying a malignant PSL is denoted as a false positive (fp) and at last, falsely classifying a benign PSL is denoted as a false negative (fn). By using the validated classification criteria, we were able to test the method with a test set. The results are given in Table 4, where the P23C1001 gave a false positive due to some regression difficulties regarding outliers.

Table 4. The PSLs in the test set and their classification correctness.

PSL	Correctness
P100C1000	tp
P23C1001	fp
P102C1000	tp
P28C1000	tn
P107C1003	tn
P69C1003	tp
P13C1000	tn
P74C1002	tp
P14C1000	tn
P97C1004	tp

The accuracy of this proposed method is 90%. The sensitivity is 100% and the specificity is 80%. Considering the melanoma discrimination, both sensitivity and specificity are 100%. In Table 5 results from other researchers are provided.

Table 5. Comparison with results from other studies. (* Sensitivity for melanoma detection. ¥ Only reported sensitivity for three melanoma lesions. α Only reported sensitivity for four melanoma lesions).

References	Patients	Images	Bands	Range (nm)	Sensitivity (%)	Specificity (%)
Tomatis et al. [12]	1278	1391	15	483–950	80.4 *	75.6
Kazianka et al. [10]	-	310	300	-	95 *	-
Moncrieff et al. [13]	311	348	8	400–1000	100 *, ¥	5.5
Fink et al. [14]	111	360	10	430–950	100 *, ¥	5.5
Song et al. [22]	55	36	10	430–950	71.4 *, α	25
Monheit et al. [16]	1257	1612	10	430–950	98.2 *	9.5
Nagaoka et al. [11]	97	134	124	380–780	96.0 *	87
Stamnes et al. [17]	-	157	10	365–1000	97	97
Stamnes et al. [17]	-	712	10	365–1000	99	93
Hosking et al. [9]	100	52	21	350–950	36 *	100 *
Leon et al. [8]	61	76	116	450–950	87.5/100 *	100
Proposed	61	76	125	450–950	100	80/100 *

From Table 5, we can discern that most research studies have been discrimination of the aggressive melanoma PSLs, where Nagaoka et al. [11] and Leon et al. [8] had the highest sensitivity and specificity. Stamnes et al. [17] also had high such measures; however, these were only discriminating between benign and malignant PSLs. Thus, the sensitivity and specificity in the method developed here is comparable to the others.

The works listed in Table 5 mostly use an array of machine learning methods, while the proposed method only use a regression method, making it a simpler approach than the others.

The computation time was on average 0.16 s for each PSL. However, including the best ROI search, this jumps to 70 s. Thus, it is a rather slow algorithm due to rather complex calculations. Compared to the previous work in [8], in which six of the 10 patients were correctly identified with an average execution time of 0.5 s, the proposed methodology identifies nine out of 10 patients with an average of 160 s. However, this is only the core

algorithm; an efficient segmentation algorithm has to be included also, raising the execution time.

4. Limitations and Future Directions

Further studies must be conducted in order to improve the limitations in this research. The first limitation is the rather low number of samples in particular for the melanoma class. In addition, in the benign and malignant class, it would have been desirable of an increase of samples. However, the number of samples was enough to find and validate the patterns for each class of PSLs considered here. Another limitation is the computing time. However, this will be decreased by the automatic choosing of ROIs by segmentation of the PSLs, which have several known potential solutions. We will also try to improve the regression of the curves by further experimentation of the weight function or other types of regression. The developed system in this paper is at an experimental level; however, it has the potential to assist in the diagnosis of skin cancer and reduce the number of biopsies relating to this.

5. Conclusions

The work presented here shows that HSIs can be used to discriminate between melanoma, malignant, and benign PSLs as a potential future diagnostic tool for dermatologists or GPs. The proposed curve-based method is able to do this discrimination by using the inherent patterns of these curves. The presented method was able to correctly classify nine out of 10 in the test set, identifying all cancerous cases accurately (100% of sensitivity). Thus, it is a rather good result. In addition, in the test phase, it had a sensitivity of 80% overall and 100% for the melanomas. Since this method is relatively experimental, we will improve it further in order to increase its accuracy and lower its computation time. The limitation of the proposed method is mainly the lack of data, which is remedied in an ongoing project of data collection. The merits of the proposed method is the utilization of the richer structure provided by the HSI technology in a novel way.

Author Contributions: Conceptualization, F.G. and G.M.C.; methodology, S.U.; software, S.U.; validation, S.U., formal analysis, S.U.; investigation, S.U.; resources, G.M.C.; data curation, I.C., G.C., A.G., P.A., J.A.H.; writing—original draft preparation, S.U.; writing—review and editing, E.Q.; visualization, S.U.; project administration, E.Q., F.G., G.M.C.; funding acquisition, F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the Canary Islands Government through the ACIISI (Canarian Agency for Research, Innovation and the Information Society), ITHACA project “Hyperspectral Identification of Brain Tumors” under Grant Agreement ProID2017010164 and it has been partially supported also by the Spanish Government and European Union (FEDER funds) as part of support program in the context of Distributed HW/SW Platform for Intelligent Processing of Heterogeneous Sensor Data in Large Open Areas Surveillance Applications (PLATINO) project, under contract TEC2017-86722-C4-1-R.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to patient confidentiality.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hong, D.; Wu, X.; Ghamisi, P.; Chanussot, J.; Yokoya, N.; Zhu, X.X. Invariant Attribute Profiles: A Spatial-Frequency Joint Feature Extractor for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3791–3808. [[CrossRef](#)]
2. Huang, H.; Luo, F.; Liu, L.; Yang, Y. Dimensionality reduction of hyperspectral images based on sparse discriminant manifold embedding. *ISPRS J. Photogramm. Remote Sens.* **2015**, *106*, 42–54. [[CrossRef](#)]
3. Lu, G.; Fei, B. Medical hyperspectral imaging: A review. *J. Biomed. Opt.* **2014**, *19*, 010901. [[CrossRef](#)] [[PubMed](#)]

4. Johansen, T.H.; Møllersen, K.; Ortega, S.; Fabelo, H.; Garcia, A.; Callico, G.M.; Godtliebsen, F. Recent advances in hyperspectral imaging for melanoma detection. *Wiley Interdiscip. Rev. Comput. Stat.* **2019**, *12*, e1465. [[CrossRef](#)]
5. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)]
6. LeBoit, P.E.; Burg, G.; Weedon, D.; Sarasin, A. *Pathology & Genetics Skin Tumors*; World Health Organization: Geneva, Switzerland, 2006.
7. Tsao, H.; Olazagasti, J.M.; Cordoro, K.M.; Brewer, J.D.; Taylor, S.C.; Bordeaux, J.S.; Chren, M.-M.; Sober, A.J.; Tegeler, C.; Bhushan, R.; et al. Early detection of melanoma: Reviewing the ABCDEs American Academy of Dermatology Ad Hoc Task Force for the ABCDEs of Melanoma. *J. Am. Acad. Dermatol.* **2015**, *72*, 717–723. [[CrossRef](#)]
8. Leon, R.; Martinez, B.; Fabelo, H.; Ortega, S.; Melian, V.; Castaño, I.; Carretero, G.; Almeida, P.; Garcia, A.; Quevedo, E.; et al. Non-Invasive Skin Cancer Diagnosis Using Hyperspectral Imaging for In-Situ Clinical Support. *J. Clin. Med.* **2020**, *9*, 1662. [[CrossRef](#)] [[PubMed](#)]
9. Hosking, A.-M.; Coakley, B.J.; Chang, D.; Talebi-Liasi, F.; Lish, S.; Lee, S.W.; Zong, A.M.; Moore, I.; Browning, J.; Jacques, S.L.; et al. Hyperspectral imaging in automated digital dermoscopy screening for melanoma. *Lasers Surg. Med.* **2019**, *51*, 214–222. [[CrossRef](#)] [[PubMed](#)]
10. Kazianka, H.; Leitner, R.; Pilz, J. Segmentation and classification of hyper-spectral skin data. In *Studies in Classification, Data Analysis, and Knowledge Organization*; Springer: Berlin/Heidelberg, Germany, 2008.
11. Nagaoka, T.; Nakamura, A.; Kiyohara, Y.; Sota, T. Melanoma Screening System Using Hyperspectral Imager Attached to Imaging Fiberscope. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012.
12. Tomatis, S.; Carrara, M.; Bono, A.; Bartoli, C.; Lualdi, M.; Tragni, G.; Colombo, A.; Marchesini, R. Automated melanoma detection with a novel multispectral imaging system: Results of a prospective study. *Phys. Med. Biol.* **2005**, *50*, 1675–1687. [[CrossRef](#)] [[PubMed](#)]
13. Moncrieff, M.D.; Cotton, S.; Claridge, E.; Hall, P. Spectrophotometric intracutaneous analysis: A new technique for imaging pigmented skin lesions. *Br. J. Dermatol.* **2002**, *146*, 448–457. [[CrossRef](#)] [[PubMed](#)]
14. Fink, C.; Jaeger, C.; Jaeger, K.; Haenssle, H.A. Diagnostic performance of the MelaFind device in a real-life clinical setting. *J. Dtsch. Dermatol. Ges.* **2017**, *15*, 414–419. [[CrossRef](#)] [[PubMed](#)]
15. Elbaum, M.; Kopf, A.W.; Rabinovitz, H.S.; Langley, R.G.B.; Kamino, H.; Mihm, M.C.; Sober, A.J.; Peck, G.L.; Bogdan, A.; Gutkowitz-Krusin, D.; et al. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study. *J. Am. Acad. Dermatol.* **2001**, *44*, 207–218. [[CrossRef](#)] [[PubMed](#)]
16. Monheit, G.; Cognetta, A.B.; Ferris, L.; Rabinovitz, H.; Gross, K.; Martini, M.; Grichnik, J.M.; Mihm, M.; Prieto, V.G.; Googe, P.; et al. The performance of MelaFind: A prospective multicenter study. *Arch. Dermatol.* **2011**, *147*, 188–194. [[CrossRef](#)] [[PubMed](#)]
17. Stamnes, J.J.; Ryzhikov, G.; Biryulina, M.; Hamre, B.; Zhao, L.; Stamnes, K. Optical detection and monitoring of pigmented skin lesions. *Biomed. Opt. Express* **2017**, *8*, 2946–2964. [[CrossRef](#)] [[PubMed](#)]
18. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
19. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: An overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709. [[CrossRef](#)]
20. Fabelo, H.; Carretero, G.; Almeida, P.; Garcia, A.; Hernandez, J.A.; Godtliebsen, F.; Melian, V.; Martinez, B.; Beltran, P.; Ortega, S.; et al. Dermatologic Hyperspectral Imaging System for Skin Cancer Diagnosis Assistance. In Proceedings of the 2019 34th Conference on Design of Circuits and Integrated Systems, Bilbao, Spain, 20–22 November 2019.
21. Holland, P.W.; Welsch, R.E. Robust regression using iteratively reweighted least-squares. *Commun. Stat. Theory Methods* **1977**, *6*, 813–827. [[CrossRef](#)]
22. Song, E.; Grant-Kels, J.M.; Swede, H.; D’Antonio, J.L.; Lachance, A.; Dadras, S.S.; Kristjansson, A.K.; Ferenczi, K.; Makkar, H.S.; Rothe, M.J. Paired comparison of the sensitivity and specificity of multispectral digital skin lesion analysis and reflectance confocal microscopy in the detection of melanoma in vivo: A cross-sectional study. *J. Am. Acad. Dermatol.* **2016**, *75*, 1187–1192. [[CrossRef](#)] [[PubMed](#)]

**Paper III: Estimation of Blood Glucose
Concentration During Endurance Sports**

Estimation of blood glucose concentration during endurance sports

Giovanni Sebastiani^{1,2,4}, Stig Uteng³, Fred Godtliebsen⁴, Jan Polák⁵ and Jan Brož⁶,

1) Istituto per le Applicazioni del Calcolo "Mauro Picone", CNR, Rome, Italy,

2) Department of Mathematics "Guido Castelnuovo", "Sapienza University of Rome", Italy,

3) Department of Education UIT The Arctic University of Norway, Tromsø, Norway,

4) Department of Mathematics and Statistics UIT The Arctic University of Norway, Tromsø, Norway,

5) Department of Pathophysiology, Third Faculty of Medicine, Charles University, Prague, Czech Republic,

6) Department of Internal Medicine, Second Faculty of Medicine, Charles University, Prague, Czech Republic

Received: May 22, 2020. Revised: June 19, 2020. Accepted: July 9, 2020. Published: July 21, 2020.

Abstract- In this paper, we describe a new statistical approach to estimate blood glucose concentration along time during endurance sports based on measurements of glucose concentration in subcutaneous interstitial tissue. The final goal is the monitoring of glucose concentration in blood to maximize performance in endurance sports. Blood glucose concentration control during and after aerobic physical activity could also be useful to reduce the risk of hypoglycemia in type 1 diabetes mellitus subjects. By means of a low invasive technology known as "continuous glucose monitoring", glucose concentration in subcutaneous interstitial tissue can now be measured every five minutes. However, it can be expressed as function of blood glucose concentration along time by means of a convolution integral equation. In the training phase of the proposed approach, based on measurements of glucose concentration in both artery and subcutaneous interstitial tissue during physical activity, the parameters of the convolution kernel are estimated. Then, given a new subject performing aerobic physical activity, a deconvolution problem is solved to estimate glucose concentration in blood from continuous glucose monitoring measurements.

Keywords- Endurance sports, blood glucose concentration, continuous glucose monitoring

I. INTRODUCTION

PHYSICAL activity (PA) in endurance sports is characterized by prolonged muscular work at high intensity (high heart frequency). The energy needed to maintain muscular fiber contraction during endurance sports is mainly produced by means of chemical transformations of fatty acids and glucose. Although energy production obtained from lipids is necessary in endurance PA, glucose as source of energy plays a very important role. Glucose is stored in the human body in muscles and liver in the form of glycogen, and several grams of glucose are diluted in blood. During PA, glucose is produced in muscles by glycogenolysis. In addition, it is also transported there by means of blood circulation: the sources are liver glycogenolysis, gluconeogenesis in liver and kidneys

and ingested food [11]. Demand of glucose in high intensity exercise is substantial and could potentially lead to hypoglycemia, i.e. drop of blood glucose (BG) concentration below 3.6 mmol/l (glycemia) [10]. Thus, to keep glycemia within the normal range (3.6-5.6 mmol/l) (see [10]) the influx of glucose from blood to the muscles is regulated by several hormones among which insulin plays the major role. Insulin production is decreased during exercise to prevent excessive leak of glucose into the muscles [17]. From the above considerations, it is clear that monitoring glucose concentration in tissues and blood can be very relevant to maximize performance in endurance sports. Moreover, it is absolutely essential for subjects with type 1 diabetes mellitus (T1D) in which insuline is not produced and must be injected subcutaneously. Thus, the risk of hypoglycemia in T1D subjects during PA is relevant because the physiological system to minimize it is disrupted. To minimise such a risk it is recommended that the insulin dose administered before and after an exercise period should be decreased, carbohydrates should be ingested regularly and glycemia should be measured as frequently as possible (at least every 30 min) [20].

Recently, a new and low invasive technology has been introduced, known as continuous glucose monitoring (CGM), by which glucose concentration in subcutaneous interstitial tissue can be measured every five minutes [15]. This technique is mainly used by T1D subjects, who greatly benefit of it to monitor the level of glucose in tissues and blood in order to control it. Studies proved that T1D patients can significantly benefit of regular aerobic PA, which reduces glucose concentration in both blood and tissues [4]. Unfortunately, the fear of the potentially severe effects of getting hypoglycemia during or after performing PA is a major reason why most T1D subjects do not perform PA [3].

We point out that by CGM we can only measure glucose concentration in subcutaneous interstitial tissue $G_s(t)$ at sampling times while we are interested more in temporal glucose level in tissues and blood. Fast and large changes in glucose concentration in blood $G_b(t)$, can be observed during PA. However, it is well known that related changes in $G_s(t)$ are seen after a time delay (lag time) [9], [23]. The quantity $G_s(t)$ can be modeled as a function of $G_b(t)$ by means of a

convolution integral equation with asymmetric kernel involving only the past due to causality [6].

Here, we focus on the estimation of glucose concentration dynamics in blood during aerobic PA from CGM measurements. To this aim, we propose to first perform some PA experiments on one or more subjects. During training, CGM measurements are performed together with those of glucose concentration in both artery and capillaries, that reveal to be very close to each other. These data are then used to estimate the convolution kernel parameters. Given a new subject, a deconvolution problem is then solved to estimate $G_b(t)$ from CGM measurements during PA. This task could be introduced within a general tool, which also includes glucose concentration forecasting either by machine learning or by mathematical models [1], in order to monitor glucose concentration in blood aiming to maximize performance in endurance sports. In connection to the control of BG concentration during PA and after it could also be useful to reduce the risk of hypoglycemia in T1D subjects, and recent work, [2, 7, 8, 12, 13, 16, 18, 19, 21, 24], indicate the importance of handling PA for T1D patients. The current paper makes progress in this direction by obtaining a more precise description of BG levels in artery. Utilizing this new information may be the key to better BG control during PA.

Here we present initial results that indicate that the well-known delay in CGM measurements can be significantly reduced. This delay may be life threatening for T1D patients, in particular during PA where the BG level can be lowered rapidly. Future utilization of these results are therefore expected to improve the BG behavior in T1D patients in general and in particular during PA.

II. PA EXPERIMENTS AND MODELS

PA experiments are performed using an electrically braked ergometer at the target heart rate (THR) according to an intensity I corresponding to 50% of individual heart rate reservoir, calculated following the Karvonen equation: $THR = (HR_{max} - HR_{rest}) * I/100 + HR_{rest}$ [14]. Measurements of CGM and glucose concentration in artery and capillaries are performed every 5 and 10-15 minutes, respectively both during PA and for 60 minutes after its end. The training begins approximately 120 minutes since breakfast and after 10 minutes warm up phase. The exercise is interrupted when either symptoms of hypoglycemia occurred or BG concentration of 3.5 mmol/l is measured. Immediately at end of training, 20 g of glucose diluted in 150 ml of water is ingested.

We focus here on experiments with a PA phase followed at its end by glucose solution ingestion. To describe both $G_b(t)$ and $G_s(t)$ we therefore adopted a model with two sine functions. The first sine function decreases until its local minimum is reached. Then, the second sine function increases starting from this point which corresponds to its local minimum. This ensures that also right and left derivatives at common local minimum point are equal. The model used has the following parameters: minimum location, 3 parameters for each sine function, but only six of them are free because of the continuity constraint of the two sine functions at their local minimum point.

The quantity $G_s(t)$ as a function of $G_b(t)$ is commonly modeled [6] by means of asymmetric convolution, due to causality, with exponential kernel $h(t)$:

$$G_s(t) = \int_0^t G_b(\tau)h(t - \tau)d\tau + \epsilon, \quad \epsilon \sim N(0, \sigma^2), \quad (1)$$

which is discretized to:

$$\mathbf{G}_s = \mathbf{H}\mathbf{G}_b + \epsilon, \quad \epsilon \sim N(0, \sigma^2\mathbf{I}). \quad (2)$$

However, due to the two phases of glucose decrease and increase with different dynamics, here we use two different exponential functions for the convolution kernel, one before the minimum location of $G_b(t)$ and the other one after it. Given the functions $G_b(t)$ and $G_s(t)$ estimated for one subject or more, the convolution kernel parameters are estimated by minimizing the mean squared error between $G_s(t)$ and its estimate by numerical convolution of $G_b(t)$. In a new subject, model parameters of $G_s(t)$ are estimated from CGM measurements by minimizing mean squared error between data and model values at sampling times. Deconvolution of $G_s(t)$ is then performed to estimate model parameters for $G_b(t)$ by minimization of mean squared errors between estimated $G_s(t)$ and the numerical convolution of $G_b(t)$. Here the convolution kernel is assumed to be known since we use the one already estimated before. This minimization is performed here by Simulated Annealing, which will be the dominating algorithm in terms of complexity, $O(k^5)$, as given in [22], where k is the length of the data vector for each patient. The complexity of least square estimation is $O(k^3)$, due to matrix inversion and for the convolution (i.e. matrix-vector multiplication), $O(k^2)$, both given in [5]. The method is summarized in Algorithm (1).

Algorithm 1 The CGMPA algorithm

Input: Measurements of CGM and arterial glucose concentration
Training:
 For $i = 1 : n$
 Model fitting to CGM data of subject i
 Model fitting to arterial data of subject i
 End
Estimation: Kernel parameters via convolution
Input: CGM measurements for a new subject
Estimation:
 $G_s(t)$ parameters via CGM data fitting
 $G_b(t)$ parameters via deconvolution of $G_s(t)$

III. RESULTS

In Fig. 1 we can observe the data of both CGM and arterial glucose concentration measured for one subject during the PA experiment. A double sine model is fitted by minimizing the mean square error for both the $G_s(t)$ -data and the $G_b(t)$ -data. As seen, there is a very good agreement between model and data. In the initial phase, $G_b(t)$ decreases due to PA. The ingestion of a glucose solution at the end of PA, makes $G_b(t)$ to increase again. The changes of $G_s(t)$ due first to PA and then to glucose ingestion appear clearly delayed w.r.t. those

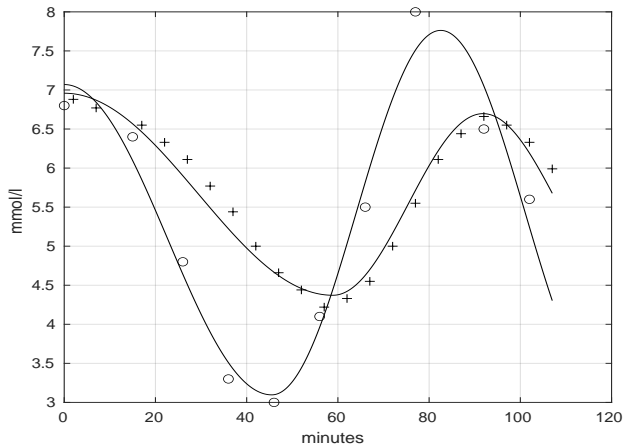


Fig. 1: Example of data for CGM (+) and arterial glucose concentration (o) measured for one subject during PA experiment. Continuous lines show best data fits by double sine model for $G_s(t)$ and $G_b(t)$, respectively.

of $G_b(t)$. This can be claimed since it is clear, for e.g. the upper panel of Fig. 2, that the variation in CGM is delayed compared to BG in artery during both PA and after intake of glucose at the end of PA. To see this, draw a horizontal line for e.g. a value of 8 mmol/l. During PA, $G_b(t)$ arrives to this value after around 35 minutes while $G_s(t)$ arrives to this value after around 50 minutes. This means a delay of around 15 minutes, a length that may be life threatening for a T1D patient. The same behavior for this horizontal line is observed during the glucose ingestion where $G_b(t)$ arrives to 8 mmol/l after around 95 minutes while $G_s(t)$ arrives there after 110 minutes. Note that also here the delay is around 15 minutes. Given the functions $G_b(t)$ and $G_s(t)$ estimated for this subject, the convolution kernel parameters are estimated by minimizing the mean squared error between $G_s(t)$ and its estimate by numerical convolution of $G_b(t)$. This is done by exhaustive search on a finite 2d-grid of values for the time constants (in minutes) of the two exponential functions from 1 minute to the maximum value of sampling times (in minutes) with steps of 1 minute.

Fig. 2 shows similar data of CGM and arterial glucose concentration for two other subjects. The best model fit for $G_b(t)$ is superimposed on the arterial glucose concentration data. Based on the convolution kernel estimated before, an estimate for $G_s(t)$ is obtained by numerical convolution of $G_b(t)$. A good agreement is observed between CGM data and estimated $G_s(t)$ computed at sampling times.

In Fig. 3 the same data of CGM and arterial glucose concentration as in Fig. 2 appear. The best model fit for $G_s(t)$ is superimposed on the CGM data. Based on the convolution kernel estimated before, an estimate for $G_b(t)$ is obtained by numerical deconvolution of $G_s(t)$. A good agreement is observed between arterial glucose concentration data and estimated $G_b(t)$ computed at sampling times.

The percentage total mean square error for the mean square error minimization fitting for the convolution and de-

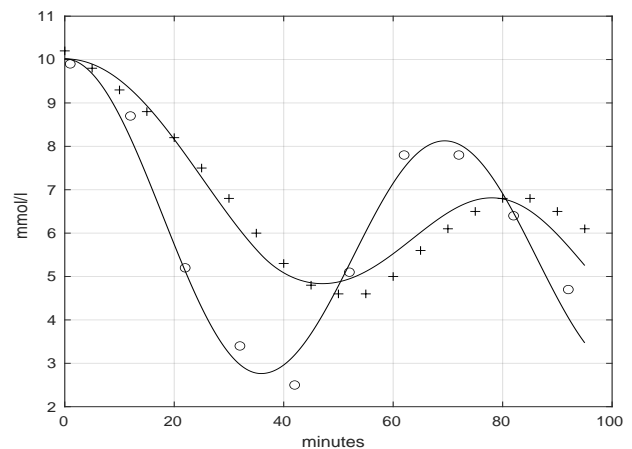
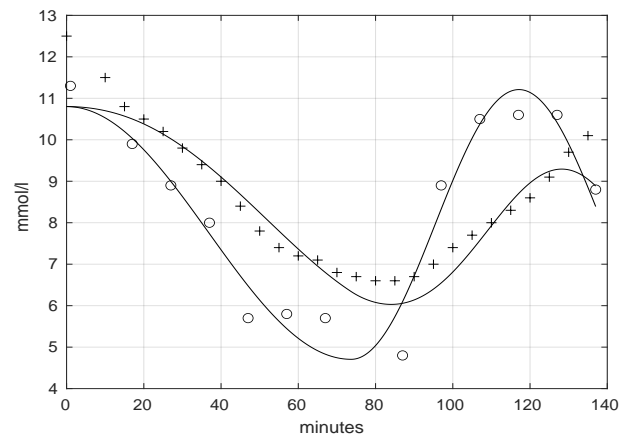


Fig. 2: Data for CGM (+) and arterial glucose concentration (o) measured for two additional persons performing a similar PA experiment. The best model for $G_b(t)$ fitted to arterial glucose concentration data is also superimposed on them. For each of the two subjects, estimate of $G_s(t)$, as obtained by numerical convolution of $G_b(t)$, is superimposed on the CGM data.

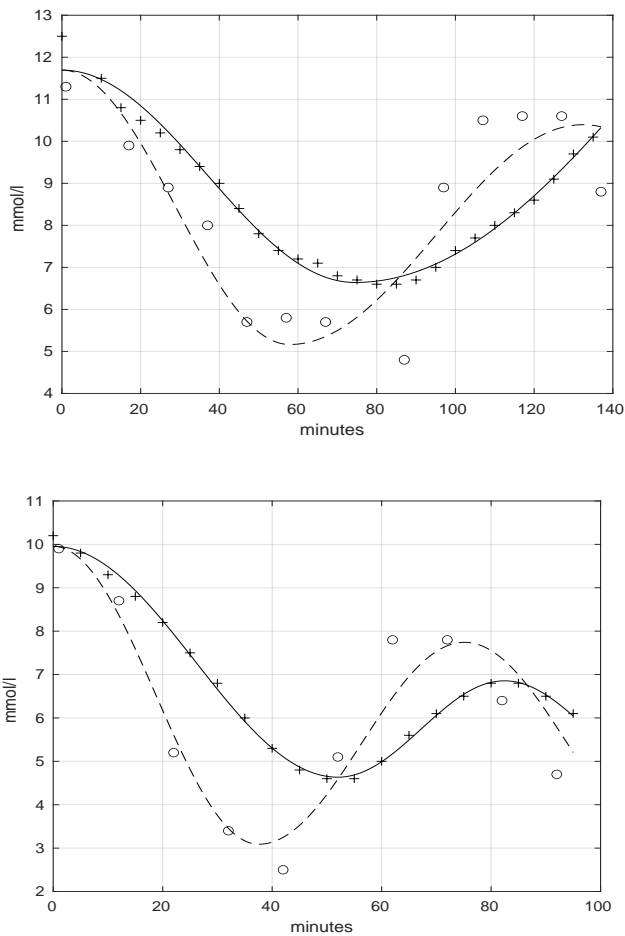


Fig. 3: Data for CGM (+) and arterial glucose concentration (o) as in Fig. 2. Best model for $G_s(t)$ fitted to CGM data is also superimposed on them. For each of the two subjects, estimate of $G_b(t)$ obtained by numerical deconvolution of $G_s(t)$ is superimposed on the arterial glucose concentration data.

convolution estimates for the two subjects in Figs. 2 and 3 are shown in Table 1.

Subject	Convolution	Deconvolution
1	$G_b : 7\%$ $G_s : 5\%$	$G_s : 2\%$ $G_b : 10\%$
2	$G_b : 6\%$ $G_s : 5\%$	$G_s : 0.3\%$ $G_b : 11\%$

Table 1: Percentage total mean square error for the mean square error minimization fitting for the convolution and deconvolution estimates for the two subjects in Figs. 2 and 3.

IV. CONCLUSION

The present work shows that a statistical approach can be used to successfully estimate BG concentration along time during PA from CGM measurements. This task could be introduced within a general tool to monitor glucose concentration in blood aiming to maximize performance in endurance

sports. This could also be used to reduce the risk of hypoglycemia in T1D subjects performing PA.

REFERENCES

- [1] Ajmera, I., Swat, M., Laibe, C., Novère, N. L., and Cheliah, V. (2013). The impact of mathematical modeling on the understanding of diabetes and related complications. *CPT: Pharmacometrics and Systems Pharmacology*, 2(e54).
- [2] Beneyto, A., Bertachi, A., Bondia, J., and Vehi, J. (2020). A New Blood Glucose Control Scheme for Unannounced Exercise in Type 1 Diabetic Subjects. *IEEE Transactions on Control Systems Technology*, 28(2):593–600.
- [3] Brazeau, A.-S., Rabasa-Lhoret, R., Strychar, I., and Mircescu, H. (2008). Barriers to Physical Activity Among Patients With Type 1 Diabetes. *Diabetes Care*, 31(11):2108–2109.
- [4] Chimen, M., Kennedy, A., Nirantharakumar, K., Pang, T. T., Andrews, R., and Narendran, P. (2012). What are the health benefits of physical activity in type 1 diabetes mellitus? A literature review. *Diabetologia*, 55(3):542–551.
- [5] Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (2009). *Introduction to Algorithms, Second Edition*. MIT Press.
- [6] Dicker, L. H., Sun, T., Zhang, C. H., Keenan, D. B., and Shepp, L. (2013). Continuous blood glucose monitoring: A Bayes-hidden Markov approach. *Statistica Sinica*, 23(4):1595–1627.
- [7] Frank, S., Jbaily, A., Hinshaw, L., Basu, R., Basu, A., and Szeri, A. J. (2018). Modeling the acute effects of exercise on insulin kinetics in type 1 diabetes. *Journal of Pharmacokinetics and Pharmacodynamics*, 45(6):829–845.
- [8] Garcia-Tirado, J., Corbett, J. P., Boiroux, D., Jørgensen, J. B., and Breton, M. D. (2019). Closed-loop control with unannounced exercise for adults with type 1 diabetes using the Ensemble Model Predictive Control. *Journal of Process Control*, 80:202–210.
- [9] Garg, S. K., Voelmlle, M., and Gottlieb, P. A. (2010). Time lag characterization of two continuous glucose monitoring systems. *Diabetes Research and Clinical Practice*, 87(3):348–353.
- [10] Güemes, M., Rahman, S. A., and Hussain, K. (2016). What is a normal blood glucose? *Archives of Disease in Childhood*, 101(6):569–574.
- [11] Hargreaves, M. (2000). Skeletal muscle metabolism during exercise in humans. *Clinical and Experimental Pharmacology and Physiology*, 27(3):225–228.
- [12] Hobbs, N., Hajizadeh, I., Rashid, M., Turksoy, K., Breton, M., and Cinar, A. (2019). Improving Glucose Prediction Accuracy in Physically Active Adolescents With Type 1 Diabetes. *Journal of Diabetes Science and Technology*, 13(4):718–727.

- [13] Jagers, J. R., King, K. M., Watson, S. E., and Wintergerst, K. A. (2019). Predicting Nocturnal Hypoglycemia with Measures of Physical Activity Intensity in Adolescent Athletes with Type 1 Diabetes. *Diabetes Technology and Therapeutics*, 21(7):406–408.
- [14] Karvonen, M. J., Kentala, E., and Mustala, O. (1957). The effects of training on heart rate; a longitudinal study. *Annales medicinae experimentalis et biologiae Fenniae*, 35(3):307–315.
- [15] Mian, Z., Hermayer, K. L., and Jenkins, A. (2019). Continuous Glucose Monitoring: Review of an Innovation in Diabetes Management. *American Journal of the Medical Sciences*, 5(5):377–390.
- [16] Paoletti, N., Liu, K. S., Smolka, S. A., and Lin, S. (2017). Data-driven robust control for type 1 diabetes under meal and exercise uncertainties. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10545, pages 23–39.
- [17] Raguso, C. A., Coggan, A. R., Gastaldelli, A., Sidossis, L. S., Bastyr, E. J., and Wolfe, R. R. (1995). Lipid and carbohydrate metabolism in IDDM during moderate and intense exercise. *Diabetes*, 44(9):1066–1074.
- [18] Resalat, N., Youssef, J. E., Reddy, R., and Jacobs, P. G. (2016). Design of a dual-hormone model predictive control for artificial pancreas with exercise model. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 2270–2273.
- [19] Resalat, N., Youssef, J. E., Reddy, R., and Jacobs, P. G. (2017). Evaluation of model complexity in model predictive control within an exercise-enabled artificial pancreas. *IFAC-PapersOnLine*, 50(1):7756–7761.
- [20] Riddell, M. C., Gallen, I. W., Smart, C. E., Taplin, C. E., Adolfsson, P., Lumb, A. N., Kowalski, A., Rabasa-Lhoret, R., McCrimmon, R. J., Hume, C., Annan, F., Fournier, P. A., Graham, C., Bode, B., Galassetti, P., Jones, T. W., Millán, I. S., Heise, T., Peters, A. L., Petz, A., and Laffel, L. M. (2017). Exercise management in type 1 diabetes: a consensus statement. *The Lancet Diabetes and Endocrinology*, 5(3):542–551.
- [21] Riddell, M. C., Zaharieva, D. P., Tansey, M., Tsalikian, E., Admon, G., Li, Z., Kollman, C., and Beck, R. W. (2019). Individual glucose responses to prolonged moderate intensity aerobic exercise in adolescents with type 1 diabetes: The higher they start, the harder they fall. *Pediatric Diabetes*, 20(1):99–106.
- [22] Sasaki, G. H. and Hajek, B. (1988). The time complexity of maximum matching by simulated annealing. *Journal of the ACM (JACM)*, 35:387–403.
- [23] Wei, C., Lunn, D. J., Acerini, C. L., Allen, J. M., Larsen, A. M., Wilinska, M. E., Dunger, D. B., and Hovorka, R. (2010). Measurement delay associated with the Guardian® RT continuous glucose monitoring system. *Diabetic Medicine*, 27(1):117–122.
- [24] Yardley, J. E. (2020). Exercise and the Artificial Pancreas: Trying to Predict the Unpredictable in Patients With Type 1 Diabetes? *Canadian Journal of Diabetes*, 44:119–120.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US

**Paper IV: Machine Perfusion of the Liver to
Resuscitate and Reverse Ischemic Liver
Injuries**

Machine Perfusion of the Liver to Resuscitate and Reverse Ischemic Liver Injuries

Stig Uteng¹, Anmar Numan², Geir Nedredal² and Fred Godtliobsen³,

1) Department of Education UIT The Arctic University of Norway, Tromsø, Norway,

2) Surgical Research Lab, Institute of Clinical Medicine, UIT The Arctic University of Norway and UNN, Tromsø, Norway,

3) Department of Mathematics and Statistics UIT The Arctic University of Norway, Tromsø, Norway.

Abstract- Liver transplantation is an important medical procedure. However, to increase the donor pool, a new source of organs is patients with irreversible brain damage in the intensive care unit. These patients do not fulfill the brain death criteria. Withdrawal of life support will however cause the fulfillment of the death criteria. This type of donation is referred to as donation after circulatory death (DCD) liver transplantation as opposed to donation after brain death (DBD). Organs from DCD donors imply the deleterious effects of warm ischemia (WI) in the withdrawal phase until the declaration of death. In this context, the length of the WI time is critical, which occurs from withdrawal of support to the cold flush of organs and subsequent topical cooling. We will in this study compare two groups with two different treatments of livers removed from pigs. Further, we conduct a statistical analysis to find differences between these groups.

Keywords- DCD, normothermic perfusion, static cold storage, liver, ischemia,

I. INTRODUCTION

To increase the donor pool, a new source of organs is patients with irreversible brain damage in the intensive care unit. These patients do not fulfill the brain death criteria. Withdrawal of life support will however cause the fulfillment of the death criteria. This type of donation is referred to as donation after circulatory death (DCD) as opposed to donation after brain death (DBD). Organs from DCD donors imply the deleterious effects of warm ischemia (WI) in the withdrawal phase until the declaration of death. In this context, the length of the WI time is critical, which occurs from withdrawal of support to the cold flush of organs and subsequent topical cooling. Accordingly, DCD liver grafts are associated with a high incidence of dysfunction, primary non-function, and ischemic cholangiopathy [1, 2]. The consensus in the clinic as of today is a maximum of 30 minutes of WI [2, 3]. Due to inferior results with DCD liver grafts after conventional static cold storage (SCS), interest in liver machine preservation is renewed [4]. Machine perfusion is believed to overcome the three major shortcomings associated with conventional SCS by:

- allowing viability assessment,

- reducing the ischemia/reperfusion injury, and
- avoiding cold injury [5].

Although this paper is most relevant to the discipline of transplantation surgery, its potential findings can have an impact on hepatobiliary surgery [6]. Since the indications for major hepatectomies have expanded during the last decades, and also high-risk patients with steatosis (i.e. fatty livers), fibrosis/cirrhosis, and chemotherapy-induced liver injury are included for major liver surgery [7].

There is currently an ongoing debate on the perfusion route and the mode of perfusion. In the clinic and research today, either dual perfusion (portal vein and hepatic artery) or single perfusion (portal vein) are utilized. We have chosen dual perfusion with a roller-pump giving pulsatile flow in the hepatic artery and centrifugal pump giving laminar flow in the portal vein [8]. This we believe is a more physiological approach. The bile ducts are mainly supplied by the hepatic artery. This is an important argument for dual perfusion in machine preservation since dual perfusion supplies oxygen-rich blood supply to the peribiliary vascular plexus. Bile duct strictures can be classified as anastomotic strictures or non-anastomotic strictures [9]. Non-anastomotic strictures were first described due to hepatic artery thrombosis after liver transplantation [10].

There are two major schools of thought in machine perfusion: normothermic machine perfusion (NMP) and hypothermic machine perfusion (HMP) [8, 11]. Neither has yet proven its superiority. Machine perfusion may reverse the negative responses caused by the warm ischemia, and in addition, allow a performance assessment of the liver before transplantation. Thus, we will investigate experimentally:

- What are the effects of ischemia/reperfusion injuries beyond the consensus of today (i.e. WI>30 min) in the liver parenchyma including the bile ducts and their potential reversal?
- Do the ischemic changes reflected in tissue metabolism improve once the organ is connected to the machine perfusion?
- Is controllable hemodynamics a prognostic marker to deem whether the livers can be transplanted or not?

II. METHODS

A. Materials

The chemicals: Tacrolimus, Lanoxin, mycophenolate mofetil, midazolam. Prostacycline Cefotaxim, epoprostenol, and THAM were obtained from Sigma-Aldrich (St. Louis, MO). Paracetamol was obtained from Norsk Medisinaldepot (Oslo, Norway).

B. Animal preparation

Anesthesia induction was with IM ketamine (30 mg/kg), and IM atropine (0.02 mg/kg) injection. Further induction was done with IV pentobarbital sodium (10 mg/kg) and IV fentanyl (10 microg/kg) via an ear vein cannula. Thereafter, maintenance was with IV midazolam 0.15mg/kg/h and IV morphine (2 mg/kg/h). Heart rate and corneal reflex assessment were used as markers of the depth of sedation. The temperature was kept at 38°C (core temperature for pigs). The skin was prepared and draped from the upper neck to the pubic bone and bilaterally to each mid-axillary line. A tracheostomy was then performed with an incision of the skin over the trachea. A breathing tube was placed into the trachea, secured with ties, and connected to a mechanical ventilator. Thereafter, a central venous line was placed in the right external jugular vein and an artery line in the left common carotid artery. An incision was made from the sternal notch to the pubic bone with cautery and extended down to the sternum and through the linea alba and peritoneum. A median sternotomy was performed, and a sternal retractor was inserted. The abdominal incisions were held open with towel clips. An microdialysis (MD) catheter was then placed in the liver segments 8/4a with a modified Seldinger technique. The MD catheter was secured with a 5-0 polypropylene suture to the hepatic connective tissue. IV heparin is thereafter administered to reach a level of activated clotting time (ACT) of 450-500. The distal aorta was identified and dissected free from the surrounding tissues just above its bifurcation. The gastroduodenal artery was identified and dissected superiorly to expose the common hepatic artery. The hepatic artery was then dissected proximally to the celiac trunk and aorta. The distal common bile duct and surrounding tissues were mobilized. The portal vein will be dissected free from the surrounding tissues. The right triangular and coronary ligaments of the liver were divided and the bare area of the liver was carefully dissected free from the diaphragm. The posterior surface of the vena cava was identified and cut free from the retroperitoneal tissues. The right adrenal vein on the posterior surface of the vena cava was ligated with 2-0 silk suture and then divided. Thereafter, blood-letting for the priming of the extracorporeal circuit was started. The blood was let from the arterial line, while the lost volume will be replaced with the crystalloid Ringer's acetate solution. The mean arterial pressure is kept above 50 mmHg to avoid hypotension during the blood-letting. In a clinical DCD setting, when support is withdrawn, death is declared after the presence of cardiac arrest with pulseless electrical activity and loss of blood pressure. The experimental design was as similar as possible to the clinical situation. However, to standardize the warm ischemia time, the supra celiac aorta was cross-clamped. In addition, the portal vein was cross-clamped to increase the re-

producibility of the animal model. Thereafter, both the total IV anesthesia and mechanical ventilation were discontinued. An overdose of IV pentobarbital and IV potassium chloride was given to induce cardiac arrest and thereby euthanize the animal. The warm ischemia time is the period from circulatory arrest (i.e. cross-clamping) to cannulation and flushing via the abdominal aorta and the portal vein. Accordingly, a 45 minutes wait period was initiated after the circulatory arrest. After this period has elapsed, the flushing will consist of in situ gravity-driven flushes through the portal vein and the celiac trunk (hepatic artery) with cold 2 liters 0.9% saline (4°C), followed by 2 liters of cold (4°C) IGL-1 (Institute Georges Lopez)-1 preservation solution. Before cannulation, the tubing was flushed with 0.9% saline solution and all air was removed. The celiac trunk was cut and cannulated with an artery cannula and secured with silk ties. Saline ice slush was applied to the surface of the liver to aid in surface cooling. A venotomy in the vena cava in the chest was made and this will function as a vent. After completion of the hepatectomy, the liver was placed in a basin with cold IGL-1 solution on the back table. Once the liver was connected to the machine perfusion, a cannula was placed in the common bile duct, secured with a silk tie, and the cystic duct was ligated.

C. ECMO, Tubing, Oxygenators, Pumps, and Measurements

The extracorporeal membrane oxygenator (ECMO) circuit consisted of PVC tubing, two pediatric oxygenators (Maquet, Rastatt, Germany), and a centrifugal pump house. The tubings and centrifugal pumps were obtained from Medtronic (Minneapolis, MN). Maquet pediatric oxygenators were obtained from Maquet (Rastatt, Germany). The centrifugal pump (model, manufacturer). Roller pump (model, manufacturer). The extracorporeal circuits were primed with Ringer's acetate solution 38°C. Blood gases were maintained using a membrane oxygenator.

D. Biochemical Analyses

All samples were centrifuged at 1.900 g at +4°C immediately after sampling and frozen at -70°C for later analysis.

E. Oxygen Consumption Measurements

The hepatic arterial partial pressure of oxygen (11–12 kPa), and portal vein arterial pressure, and outflow from the reservoir. The partial pressures of oxygen in samples were analyzed with a blood gas analyzer (Rapidlab 865; Chiron Diagnostics). The sampling was performed with a plastic syringe and immediately analyzed to avoid equilibration with atmospheric air. O_2 consumption in BAL (nmol/min) arteriovenous PO_2 flow = $1.18 \cdot 10^{-3}$, where arteriovenous PO_2 (mmHg) arterial PO_2 venous PO_2 , flow was 5 ml/min, and $1.18 \cdot 10^{-3}(\text{mol ml}^{-1} \text{mmHg}^{-1})$ was a constant derived from the Bunsen solubility coefficient for oxygen ($0.03 \text{ mol } O_2 \text{ ml}^{-1} \text{mmHg}^{-1}$) and the volume of 1 mole of oxygen (25.4 l) at 37°C (32). O_2 consumption in NPCs cultivated alone (mol/l) PO_2 $1.18 \cdot 10^{-3}$, where PO_2 equals venous PO_2 , and there is no flow through the bioreactor. The values for consumption of oxygen by NPCs were corrected for the fall in oxygen concentration observed in the cell-free controls.

F. Approach

We applied flow sensor and pressure sensors in the hepatic artery, portal vein, and vena cava. The pumps were pressure-controlled: 75% of the total flow will be perfused via the portal vein, and 25% will be perfused via the hepatic artery [12]. The following algorithm of organ perfusion was followed (1 mL/min/g liver; 25%/75% in hepatic artery/ portal vein, respectively) [13]. The hepatic artery was perfused with a mean arterial pressure of 70 mmHg and the portal vein of 10 mmHg. The partial pressures were measured with a blood gas analyzer (ABL800, Radiometer). Two separate pediatric oxygenators (Medtronic) were applied to obtain desired oxygen tensions of 7.0 kPa in the portal vein and 12.0 kPa in the hepatic artery.

The MD catheters were placed in the liver (in segments 4/8) with the Seldinger technique. MD is a technique that enables the monitoring of the organs and tissues in real-time. The double-lumen MD catheter is perfused with a fluid containing dextran and electrolytes by microinjection pumps at a velocity of 1 microL/minute (CMA 107, M Dialysis AB, Sweden). The fluid (i.e. the microdialysate) was collected in micro vials. The micro vials were analyzed every 30 minutes, with an MD analyzer (Iscus, M Dialysis AB, Sweden). The analyzer analyzed tissue glucose, lactate, pyruvate, and glycerol. Ischemia causes increased lactate levels (but not pyruvate levels), thereby giving an increased lactate/pyruvate ratio (L/P-ratio). Thus, monitoring changes in the ischemia-exposed liver in real-time. The micro vials will be stored cold for later analysis of inflammatory markers. Several inflammatory mediators will be analyzed en bloc in a multiplex assay (Bio-Rad Laboratories, Inc., Hercules, CA), these include among others: interleukin-1 receptor antagonist (IL1ra), and interleukin-6 (IL6). Also, TNF-alpha (TNFa) is an important marker for systemic inflammation [14].

G. Data Acquisition and Experimental Design

The two groups are called SCS and NMP, where the total treatments for both these groups last in 720 min, including WI in the first 30 min. Both groups consist of seven pigs that have their livers removed. In the first group, the livers are in WI and SCS in 240 min before NMP, and in the second group WI and NMP in 720 min.

During the experiment were several hemodynamic parameters(variables) were monitored. These include Systemic Vascular Resistance (SVR), Portal vein pressure (P_porta), Portal vein Flow (Flow_porta), Arterial pressure (P_art_hep), Arterial Flow (Flow_art_hep), Total flow (Tot_flow), and Temperature (Temp). These parameters were sampled each minute. The monitored MD parameters (variables) include bile production (bile_prod), several immunological parameters as the interleukins (IL1a, IL1beta, IL1ra, IL2, IL4, IL6, IL8, IL10, IL12, and IL18), and TNF-alpha (TNFa) (all these were sampled only four times with even frequency for each pig). Further, were these Lactate, Pyruvate, Glucose, and Glycerol variables monitored: Glu_NMP, Lac_NMP, Pyr_NMP, Gly_NMP, Lac_Pyr_NMP (L/P-ratio), and Lac_Glu_NMP. These last MD parameters were sampled, with the same frequency, thirteen times during the experiment for each pig.

H. Statistics

One of the goals of the statistical analysis is to investigate the degree of difference between the SCS and NMP groups. The default method is the analysis of variance (ANOVA) when there are more than two groups and the t-test if there are only two groups. However, due to the problem of pseudoreplication, i.e. the observations on the response variable are not independent, we have to define dependence within and between the groups [15]. This is called a random effect, which is elaborated rather thoroughly in [16], where they develop models and software packages to do inference in the resulting linear mixed model (LMM). The current default software package doing this inference is lme4 presented in [17], which is written for R [18]. However, this package assumes Gaussian distributed error components, which was not fulfilled in our inference. Hence, we applied the R-package brms [19], which utilizes a Bayesian approach. This also means that two of the three assumptions of repeated measurements ANOVA are violated: Normality of the error term and independence of measurements (sphericity is not relevant with only two groups), which means that LMM is a more appropriate option [20]. Further, do the study presented in this paper has a multilevel two-way nested design and we wrote the model as:

$$y_{ijk} = Intercept + a_i + b_{ij} + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), \quad (1)$$

$$a_i \sim N(0, \sigma_{Protocol}^2), b_{ij} \sim N(0, \sigma_{Pig}^2),$$

where ϵ_{ijk} is the residual, and $N()$ a Gaussian distribution with mean zero and standard deviation σ_ϵ , for the residuals. a_i and b_{ij} are the random effects with standard deviations $\sigma_{Protocol}$ and σ_{Pig} , for Protocol and Pig, respectively, the *Intercept* is a constant, i , the groups, j , the individual pig and k , the measurement time points. The brms-package is a backend to RStan, the R interface to Stan [21], which applies the Hamiltonian Monte Carlo algorithm [22]. This means that the intercept and the standard deviations were given priors. The prior for the intercept was chosen to be a Gaussian distribution with mean zero and standard deviation ten. For all the standard deviations a half-Cauchy distribution with location zero and scale one truncated below zero, was chosen. We compared these computations with computations done by the lmer-function from the lme4-package, presented in [17], which gave rather similar results as the brms-package. However, the assumption of Gaussian-distributed residuals was in general not fulfilled. To determine if there were any random effects we compared the fit of Eq. (1) with

$$y_{ijk} = Intercept + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), \quad (2)$$

i.e. only an intercept. The lmer-function gave a significantly improved model by including the random effects according to the likelihood ratio test (LRT) for all the parameters. We also observed some random slope effects for some of the parameters and this was subsequently included in another model as *Time* for each Pig. Due to computational issues, we could not use the brms-package for this analysis. Thus, we had to use the lmer function from the lme4-package for this more involved model. Also, we compared this with the robust lmer

package [23]. The results for these computations also showed clear evidence for random effects and were rather similar to the first calculations without the random slope effect.

III. RESULTS

A. Statistical Results

The intraclass correlations (ICC) for Protocol and Pig show the correlation of the measurements within each group and within each pig, and can be seen as an effect measure for this within-effects [24]. The values of the variables Flow_porta, Flow_art_hep, Tot_flow, Gly_NMP, and Lac_NMP show a somewhat stronger within-group correlation than for the other variables as seen in Tab. (1). Further, the model fits are indicated by the widely accepted information criterion (WAIC), introduced in [25], which gives a measure of the predictive error of the random-effects models and the baseline or nullmodels. This relation between the WAIC value and predictive errors is discussed in [26], and is currently the best information criterion. As seen in Tab. (1) there are random effects for all the hemodynamics variables and most of the MD variables, with the clearest evidence of this for SVR, P_porta, Glu_NMP, and bile_prod, which have the lowest relative difference between the WAIC of the nullmodel and the random-effects model.

Hemodynamics							
Variables	WAIC _{0mod}	WAIC _{mod}	σ_{Proto}	σ_{Pig}	σ_{ϵ}	ICC _{Proto}	ICC _{Pig}
SVR	25008.6	14832.0	1.38	1.27	0.71	0.47	0.87
P_porta	41407.9	27417.1	1.89	4.77	1.79	0.12	0.89
P_art_hep	62639.2	50864.8	21.10	21.46	9.89	0.44	0.90
Flow_porta	87705.4	75453.8	724.95	142.26	59.42	0.96	0.99
Tot_flow	90912.4	83637.5	1001.89	356.86	108.06	0.88	0.99
Flow_art_hep	84525.5	78219.7	551.17	98.66	72.82	0.95	0.98
Temp	42359.9	41426.3	2.64	2.09	4.96	0.19	0.32

MD data with measurements from t0 to t12							
Variables	WAIC _{0mod}	WAIC _{mod}	σ_{Proto}	σ_{Pig}	σ_{ϵ}	ICC _{Proto}	ICC _{Pig}
Glu_NMP	1137.7	1000.8	1.92	9.59	7.13	0.03	0.65
bile_prod	997.6	919.6	1.15	5.15	5.42	0.02	0.49
Lac_NMP	677.4	648.8	0.73	1.27	2.16	0.08	0.32
Pyr_NMP	1746.1	1678.7	24.85	83.46	74.16	0.05	0.58
IL2	1118.4	1114.5	1.18	1.25	9.51	0.01	0.03
IL1beta	1717.9	1719.6	11.61	20.69	80.81	0.02	0.08
Gly_NMP	1925.9	1928.5	83.71	3.45	177.81	0.18	0.18

Table 1: The results of the statistical model fitting of the variables for the hemodynamic data and microdialysis (MD) data with the model given by Eq. (1) using the brms-package.

B. Hemodynamic Changes as Prognostic Markers for Irreversible Ischemic Injuries

We found a significant difference between the two groups for the hemodynamic data as can be observed in Fig. (1) and Tab. (2), which also was established through the LMM analysis above. The SVR variable is of some importance, indicating a lower probability of injuring the organ during machine perfusion in the SCS group compared to the NMP group. Vasoconstriction in the rewarming phase at the start of the machine perfusion period (from 4°C to 38°C) is typical. This vasoconstriction and the high pressures especially in the hepatic artery (P_art_hep) will commonly decline when the organ becomes normothermic (38°C). Beyond the rewarming phase, the pressures in the hepatic artery and the portal vein (P_porta) will deem whether an organ is transplantable or not. As seen in Fig. (2), in the SCS group these parameters show a better prospect.

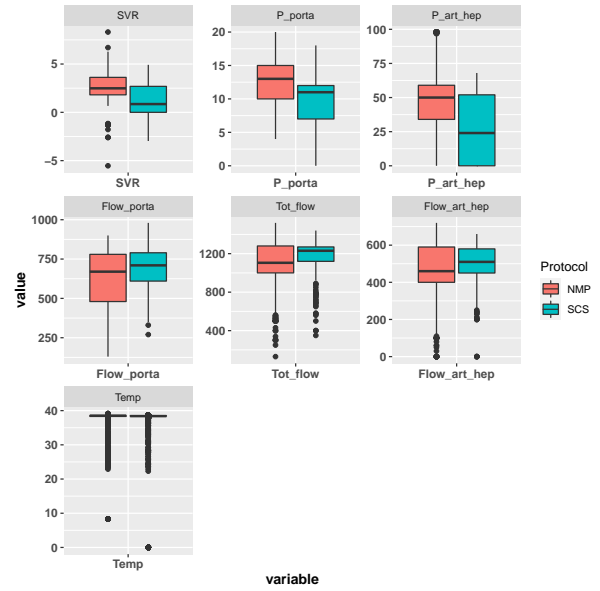


Fig. 1: Boxplots of the hemodynamic data for the NMP and SCS groups.

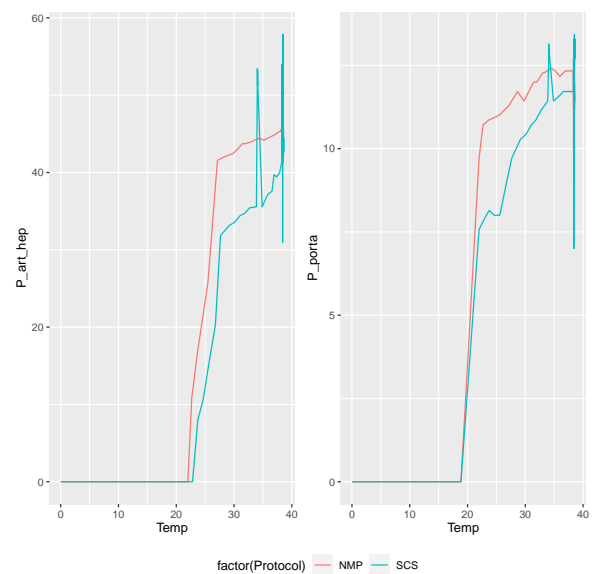


Fig. 2: Plot of P_art_hep and P_porta vs. Temperature for the NMP and SCS groups.

Variables	Mean NMP	Var NMP	Mean SCS	Var SCS
Flow_porta	625.07	25568.95	701.61	13878.63
P_porta	12.70	13.30	8.66	28.00
Flow_art_hep	471.50	14265.94	492.76	12192.81
P_art_hep	49.07	275.96	26.02	555.38
Temp	37.78	14.26	37.09	42.31
Tot_flow	1096.57	46481.11	1194.37	16284.77
SVR	2.72	1.48	1.20	1.87
Glu_NMP	12.29	1.83e+02	5.42	60.20
Lac_NMP	2.63	6.26	2.46	5.77
Pyr_NMP	93.75	1.06e+04	43.30	4.75e+03
Gly_NMP	92.56	2.76e+04	144.69	8.17e+04
Lac_Pyr_NMP	94.70	1.12e+05	217.09	1.10e+06
Lac_Glu_NMP	0.40	0.58	0.48	0.21
bile_prod	7.96	81.70	4.95	32.40
IL1a	4.21	1.37e+02	4.34	59.80
IL1beta	46.59	2.27e+04	63.06	4.58e+04
IL1ra	17101.60	2.46e+08	8367.27	3.07e+08
IL2	2.90	1.00e+02	8.43	8.74e+02
IL4	2.80	21.60	30.79	2.72e+04
IL6	2020.49	6.64e+06	1368.09	7.54e+06
IL8	9792.97	9.93e+08	727.52	1.98e+06
IL10	117.63	1.23e+04	55.80	1.16e+04
IL12	7.18	2.51e+02	11.98	2.23e+03
IL18	720.34	2.53e+06	472.99	1.06e+06
TNFa	153.52	3.30e+04	59.13	1.36e+04
Pre_Weight	1312.71	8.41e+03	1233.86	5.53e+03
Post_Weight	1391.14	5.26e+04	1335.14	3.87e+03
Weight_change_prct	3.49	2.25e+02	7.37	52.40

Table 2: The mean and variance of the hemodynamic, the MD data and weight data.

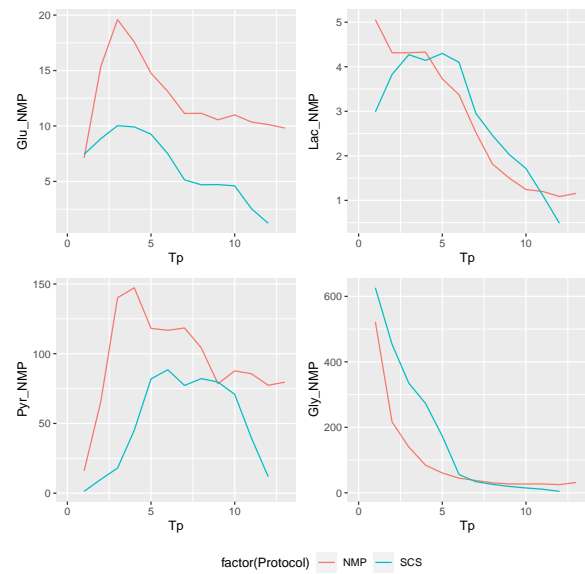


Fig. 3: Plot of glucose (Glu_NMP), the lactate (Lac_NMP) and the pyruvate (Pyr_NMP) lower in the SCS group, but glycerol (Gly_NMP) for the NMP and SCS groups as means for the pigs measurements at each time-point.

C. MD as a Reproducible Marker of Assessment of Resuscitated Ischemically Injured Liver Grafts

For the MD-data were the glucose (Glu_NMP), the lactate (Lac_NMP), and the pyruvate (Pyr_NMP) lower in the SCS group, but glycerol (Gly_NMP) was lower in the NMP group, as seen in Tab. (2) and Fig. (3). However, it can be claimed that they were rather similar in these groups. The L/P-ratio was also more beneficial for the NMP group as seen in Fig. (4). Glucose, lactate, pyruvate, glycerol, and the L/P-ratio are connected to the effects of the ischemia condition and it is advisable to keep it low [27]. We can also observe that the effects of WI are in general reversed as seen in figures (3) and (4). For the interleukins (the IL's) the means were lower for the SCS group, as seen in Tab. (2). However, in Fig. (5) we can observe that the immunological markers in the SCS group, have seven increasing variables of the total eleven. This means that in this group there is an increasing amount of inflammatory states, which is negative for the graft (liver) condition.

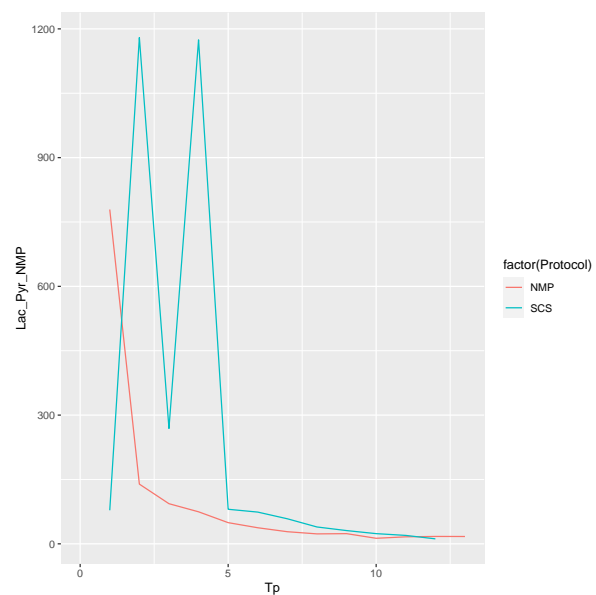


Fig. 4: Plot of L/P-ratio (Lac_Pyr_NMP) for the NMP and SCS groups as means for the pigs measurements at each time-point.

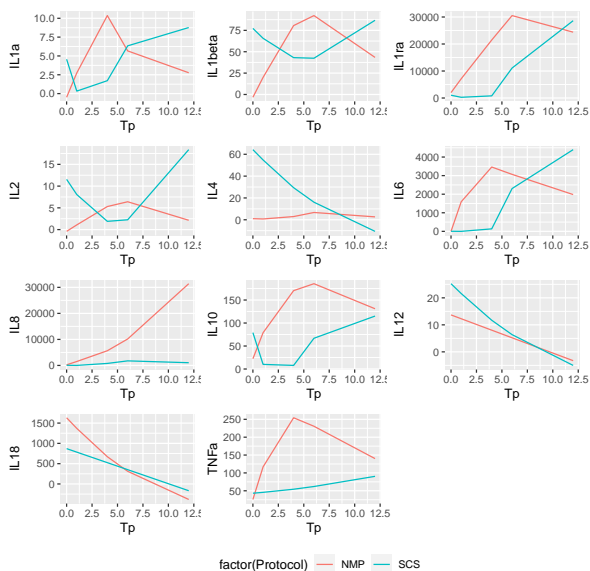


Fig. 5: Plots of immunological markers for the NMP and SCS groups as means for the pigs measurements at each time-point.

IV. DISCUSSION

We have conducted a rather comprehensive experiment to test the performance of two different machine perfusion paths, given by the groups NMP and SCS, in connection to liver transplantation. We did in general find that the SCS method was somewhat better than the NMP method. This was the case in particular for the hemodynamic data. For these variables, it could be observed that the SVR, P_art_hep, and P_porta variables had lower means for the SCS group than the NMP group. This signifies less probability for damage during the machine perfusion phase. For the MD data, and in particular, for the immunological markers, the NMP group showed a more promising result, where most of the interleukins and TNF-alpha were decreasing, showing a lower inflammatory response. Also, the L/P-ratio was lower and better for the NMP group. Thus, none of the groups were deemed better than the other overall.

There is an increase in indications for liver transplantation and accordingly, the need for organs is increasing (www.scandiarttransplant.org) [28]. One emerging indication is liver transplantation for metastatic nonresectable colorectal cancer. A study from Norway has shown a survival benefit in patients with metastatic nonresectable colorectal cancer who undergo liver transplantation compared to standard chemotherapy [29]. Moreover, the incidence of non-alcoholic fatty liver disease is increasing worldwide, which can progress to cirrhosis requiring liver transplantation [30]. This project is a collaboration between the University Hospital of Northern Norway (UNN), and Rikshospitalet University Hospital. Solid-organ transplantation in Norway is performed at Rikshospitalet. The procurement of organs is performed countrywide, where UNN is one of the main suppliers of organs for transplantation. The travel time of the procurement team from UNN to Rikshospitalet is approximately three hours. Traditionally, organs are procured from donors fulfilling the brain death criteria, also known as donation after DBD. There is now however an interest in increasing the

donor pool, therefore organs from donation after DCD are becoming utilized. Also, in Norway, we started transplanting livers (and kidneys) from DCD donors in 2015 [31]. This has proved very useful and will be continued.

REFERENCES

- [1] Srikanth Reddy, Miguel Zilvetti, Jens Brockmann, Andrew McLaren, and Peter Friend. Liver transplantation from non-heart-beating donors: Current status and future prospects. *Liver Transplantation*, 10(10), oct 2004. ISSN 1527-6465. doi: 10.1002/lt.20268.
- [2] Emre A. Eren, Nicholas Latchana, Eliza Beal, Don Jr. Hayes, Bryan Whitson, and M. Sylvester Black. Donations After Circulatory Death in Liver Transplant. *Experimental and clinical transplantation*, 2016.
- [3] David P. Foley, Luis A. Fernandez, Glen Levenson, L Thomas Chin, Nancy Krieger, Jeffery T. Cooper, Brian D. Shames, Yolanda T. Becker, Jon S. Odorico, Stuart J. Knechtle, Hans W. Sollinger, Munci Kalayoglu, and Anthony M. D'Alessandro. Donation After Cardiac Death. *Annals of Surgery*, 242(5), nov 2005. ISSN 0003-4932. doi: 10.1097/01.sla.0000186178.07110.92.
- [4] J. V. Guarrera, S. D. Henry, B. Samstein, R. Odeh-Ramadan, M. Kinkhabwala, M. J. Goldstein, L. E. Ratner, J. F. Renz, H. T. Lee, R. S. Brown, Jr., and J. C. Emond. Hypothermic Machine Preservation in Human Liver Transplantation: The First Clinical Series. *American Journal of Transplantation*, 10(2), feb 2010. ISSN 16006135. doi: 10.1111/j.1600-6143.2009.02932.x.
- [5] Thomas Vogel, Jens G. Brockmann, Constantin Cousios, and Peter J. Friend. The role of normothermic extracorporeal perfusion in minimizing ischemia reperfusion injury. *Transplantation Reviews*, 26(2), apr 2012. ISSN 0955470X. doi: 10.1016/j.trre.2011.02.004.
- [6] Kwan Man, Sheung-Tat Fan, Irene O L Ng, Chung-Mau Lo, Chi-Leung Liu, and John Wong. Prospective Evaluation of Pringle Maneuver in Hepatectomy for Liver Tumors by a Randomized Study. *Annals of Surgery*, 226(6), oct 1997. ISSN 0003-4932. doi: 10.1097/00000658-199712000-00007.
- [7] Jarnagin William R., Gonen Mithat, Fong Yuman, DeMatteo Ronald P, Ben-Porat Leah, Little Sarah, Corvera Carlos, Weber Sharon, and Blumgart Leslie H. Improvement in perioperative outcome after hepatic resection: analysis of 1,803 consecutive cases over the past decade. *Annals of Surgery*, 236(4), 2002. doi: 10.1097/01.SLA.0000029003.66466.B3.
- [8] Hazel Marecki, Adel Bozorgzadeh, Robert J Porte, Henri G Leuvenink, Korkut Uygun, and Paulo N Martins. Liver ex situ machine perfusion preservation: A review of the methodology and results of large animal studies and clinical trials. *Liver Transplantation*, 23(5), oct 2017. ISSN 15276465. doi: 10.1002/lt.24751.

- [9] Carlijn I Buis, Harm Hoekstra, Robert C Verdonk, and Robert J Porte. Causes and consequences of ischemic-type biliary lesions after liver transplantation. *Journal of Hepato-Biliary-Pancreatic Surgery*, 13(6), oct 2006. ISSN 0944-1166. doi: 10.1007/s00534-005-1080-2.
- [10] A B Zajko, W L Campbell, G A Logsdon, K M Bron, A Tzakis, C O Esquivel, and T E Starzl. Cholangiographic findings in hepatic artery occlusion after liver transplantation. *American Journal of Roentgenology*, 149(3), oct 1987. ISSN 0361-803X. doi: 10.2214/ajr.149.3.485.
- [11] Andrea Schlegel, Xavier Muller, Marit Kalisvaart, Beat Muellhaupt, M. Tamara P.R. Perera, John R. Isaac, Pierre-Alain Clavien, Paolo Muiesan, and Philipp Dutkowski. Outcomes of DCD liver transplantation using organs treated by hypothermic oxygenated perfusion before implantation. *Journal of Hepatology*, 70(1):50–57, jan 2019. ISSN 01688278. doi: 10.1016/j.jhep.2018.10.005.
- [12] W. W. Lautt. Mechanism and role of intrinsic regulation of hepatic arterial blood flow: hepatic arterial buffer response. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 249(5), nov 1985. ISSN 0193-1857. doi: 10.1152/ajpgi.1985.249.5.G549.
- [13] Ahmed Nassar, Qiang Liu, Kevin Farias, Laura Buccini, William Baldwin, Ana Bennett, Martin Mangino, Samuel Irefin, Jacek Cywinski, Toshihiro Okamoto, Teresa Diago Uso, Giuseppe Iuppa, Basem Soliman, Charles Miller, and Cristiano Quintini. Impact of Temperature on Porcine Liver Machine Perfusion From Donors After Cardiac Death. *Artificial Organs*, 40(10), oct 2016. ISSN 0160564X. doi: 10.1111/aor.12699.
- [14] Geir I Nedredal, Kjetil Elvevold, Marcio F Chedid, Lars M Ytrebø, Christopher F Rose, Sambit Sen, Bård Smedsrød, Rajiv Jalan, and Arthur Revhaug. Pulmonary vascular clearance of harmful endogenous macromolecules in a porcine model of acute liver failure. *Annals of Hepatology*, 15(3), oct 2016. ISSN 16652681. doi: 10.5604/16652681.1198821.
- [15] Stuart H. Hurlbert. Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs*, 54(2), jun 1984. ISSN 0012-9615. doi: 10.2307/1942661.
- [16] J.C. Pinheiro and D. M. Bates. *Mixed Effects Models in S and S PLUS*. Springer, 2000.
- [17] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 2015. ISSN 1548-7660. doi: 10.18637/jss.v067.i01.
- [18] R Development Core Team. R: A language and environment for statistical computing., 2017. ISSN 16000706.
- [19] Paul-Christian Bürkner. Advanced Bayesian Multi-level Modeling with the R Package brms. *The R Journal*, 10(1), 2018. ISSN 2073-4859. doi: 10.32614/RJ-2018-017.
- [20] Andy Field, Jeremy Miles, and Zoe Field. *Discovering Statistics Using R*. SAGE publishing, 1 edition, 2012.
- [21] Stan Development Team. RStan: the R interface to Stan, 2016. URL <http://mc-stan.org/>.
- [22] J.S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, 2001.
- [23] Manuel Koller. robustlmm : An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, 75(6), 2016. ISSN 1548-7660. doi: 10.18637/jss.v075.i06.
- [24] David C. Howell. *Statistical Methods for Psychology*. Cengage Learning, 8 edition, 2012.
- [25] Sumio Watanabe. *Algebraic Geometry and Statistical Learning Theory*. Cambridge University Press, 2009.
- [26] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), sep 2017. ISSN 0960-3174. doi: 10.1007/s11222-016-9696-4.
- [27] Grzegorz Nowak, Johan Ungerstedt, Jan Wernerman, Urban Ungerstedt, and Bo-Göran Ericzon. Metabolic changes in the liver graft monitored continuously with microdialysis during liver transplantation in a pig model. *Liver Transplantation*, 8(5):424–432, may 2002. ISSN 15276465. doi: 10.1053/jlts.2002.32943.
- [28] Kenji Kihara, Shinichi Ueno, Masahiko Sakoda, and Takashi Aikou. Effects of hyperbaric oxygen exposure on experimental hepatic ischemia reperfusion injury: Relationship between its timing and neutrophil sequestration. *Liver Transplantation*, 11(12), oct 2005. ISSN 1527-6465. doi: 10.1002/lt.20533.
- [29] Svein Dueland, Morten Hagness, Pål-Dag Line, Tor-mod Kyrre Guren, Kjell Magne Tveit, and Aksel Foss. Is Liver Transplantation an Option in Colorectal Cancer Patients with Nonresectable Liver Metastases and Progression on All Lines of Standard Chemotherapy? *Annals of Surgical Oncology*, 22(7), oct 2015. ISSN 1068-9265. doi: 10.1245/s10434-014-4137-0.
- [30] Iliana Doycheva, Kymberly D. Watt, and Naim Alkhoury. Nonalcoholic fatty liver disease in adolescents and young adults: The next frontier in the epidemic. *Hepatology*, 65(6), jun 2017. ISSN 02709139. doi: 10.1002/hep.29068.
- [31] Nedredal Geir, Foss Stein, Aandahl Erik M, Dagness Morten, Fosby Bent, Sørensen Daniel W, and Line Pål D. Pilot study: Transplantation of controlled donation after circulatory death (cDCD) livers. *Scandinavian Transplantation Society*, 2016.

