

## **Early diagnosis of NAFLD-NASH transition using mid infrared spectroscopy.**

Siouar Bensaid, Amar Kachenoura, Nathalie Costet, Victor De Ledinghen, Julien Vergniol, Fabrice Lainé, Bruno Turlin, Hugues Tariel, Lotfi Senhadji

► **To cite this version:**

Siouar Bensaid, Amar Kachenoura, Nathalie Costet, Victor De Ledinghen, Julien Vergniol, et al.. Early diagnosis of NAFLD-NASH transition using mid infrared spectroscopy.. Conference proceedings: .. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Institute of Electrical and Electronics Engineers (IEEE), 2016, pp.3602-3605. <10.1109/EMBC.2016.7591507>. <inserm-01476033>

**HAL Id: inserm-01476033**

**<http://www.hal.inserm.fr/inserm-01476033>**

Submitted on 24 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Early Diagnosis of NAFLD-NASH Transition using Mid Infrared Spectroscopy

Siouar Bensaid, Amar Kachenoura, Nathalie Costet, Victor De Ledinghen, Julien Vergniol, Fabrice Lainé, Bruno Turlin, Hugues Tariel and Lotfi Senhadji

**Abstract**—Non-alcoholic fatty liver disease (NAFLD) is defined as an excessive accumulation of fat in the liver in the absence of excessive drinking of alcohol. Initially considered as benign and self-limited, NAFLD may progress to the malignant stage of non-alcoholic steatohepatitis (NASH) characterized by degenerate hepatocellular ballooning and lobular inflammation. NASH can lead to hepatic fibrosis and ultimately to cirrhosis and hepatocellular carcinoma. Unfortunately, the transition from NAFLD to NASH is difficult to detect so far. In this paper, we propose to evaluate the characterization of NASH using mid infrared fiber evanescent wave spectroscopy on blood serum. We used an heuristic variable selection method and a generalized linear model to classify NAFLD and NASH spectra. The obtained results proved that this technique is a promising non-invasive and simple diagnosis tool for NASH.

## I. INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) can evolve to the necroinflammatory state of non-alcoholic steatohepatitis (NASH) that degenerates to mortal stages in 20% of the time [1]. Recent statistics warn of the high prevalence of NAFLD. Ten percent to 30% of the general population in the western world is affected [1]. Metabolic syndrome, obesity and insulin resistance are the most important factors responsible for the incidence of the disease. In fact, NAFLD incidence increases to 70% in type two diabetes and to 90% in obese people [2]. Other factors such as gender, genetics and ethnicity are proved influential [3]. Though only 1.5% to 6.5% of NAFLD patients evolves to NASH, this latter has been recognized as one of the leading causes of cirrhosis in adults and the second indication for liver transplants in the United States [1]. The high prevalence of NAFLD and its potential progression to NASH and mortal liver diseases make the screening task very challenging, especially that NAFLD patients are asymptomatic during the pre-cirrhotic stage.

Histologic diagnosis by liver biopsy is the gold standard to detect the transition from NAFLD to NASH. Nevertheless,

Siouar Bensaid, Amar Kachenoura, Nathalie Costet and Lotfi Senhadji are with INSERM, U1099, Rennes, F-35000, France and with Université de Rennes 1, LTSI, Rennes, F-35000, France. Nathalie Costet is with INSERM, UMR1085, IRSET, Rennes, France. Victor De Ledinghen and Julien Vergniol are with the Investigation Center of Hepatic Fibrosis, Haut-Lévêque Hospital, CHU Bordeaux, Pessac, France. Victor De Ledinghen is with INSERM U1053, Université de Bordeaux, Bordeaux, France. Fabrice Lainé is with CHU Rennes, Liver disease unit, F-35033 Rennes, France. Bruno Turlin is with INSERM U991 Rennes, France and with Université de Rennes 1, F-35000 Rennes, France. Hugues Tariel is with DIAFIR, Avenue Chardonnet, Parc Lorans 26J, 35000 Rennes, France.

\*This work was supported by the AMNIFIR project through a grant from the French National Research Agency (ANR), TecSan Programme, grant number ANR-12-TECS-0012

this procedure is invasive with inherent risks of complications at the biopsy site (bleeding, infection and accidental injury to other organs) and a high possibility of hospitalization afterward. Therefore, patients are generally reluctant to undergo this test especially when they are still asymptomatic. Several non-invasive imaging-based diagnosis tools are proposed. Ultrasonography is the most popular technique thanks to its availability, low cost and absence of radiation exposure. However, it has many flaws. It cannot detect mild stages of NASH, it is operator dependent, interfered by intra-abdominal gas and returns poor image quality in obese patients [2]. Computed tomography (CT) scan, magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS) are also used and perform better than the ultrasound modality [4]. Yet, they can only detect steatosis higher than 20-30% [4]. They also suffer from cost and availability limitations, radiation exposure risks for CT, and sequence dependency for MRI and MRS. Other non-invasive strategies are oriented to look for potential serum biomarkers, mainly proteins and cytokines, to build prediction models and scores [5]. Physical parameters, e.g. age, sex, body mass index and weight are also incorporated into some models to improve their performances. Nevertheless, most of the proposed clinical models suffer from low sensitivity (< 66%), lack of generalisability and discrepancies between the published results and the external validations [6].

In addition to high performance, a perfect screening method should be non-invasive, simple, objective and cheap. Unfortunately, this is not fully satisfied by the existing techniques. In our work, we investigated the direction of optical methods, mainly the technique of fiber evanescent wave spectroscopy (FEWS) [7]. FEWS is an infrared remote spectroscopy that combines the principle of attenuated total reflection (ATR) and the technology of fiber sensors to allow rapid and *in-situ* analysis of samples. In FEWS, an infrared signal is transmitted in an optical fiber by total internal reflection. ATR occurs when a chemical sample is brought into contact with the fiber, thereby generating a partial absorption of the infrared signal in the fiber interface. The spectrum of the attenuated infrared wave picked up at the fiber output provides the metabolic characteristics of the sample. The development of chalcogenide glass optical fibers played a key role in implementing FEWS for biomedical and clinical applications [8]. These bio-sensors are characterized by a large optical window covering the mid-infrared (MIR) range where are located the fundamental vibrational modes of most of the interesting biomolecules [8]. In [9], authors

proved that analyzing blood serum using this technique provides useful fingerprints to diagnose patients with diseases having metabolic consequences. In this study, we explored the potential use of FEWS to diagnose NASH via blood serum. In our experiments, we are the first to test the new patented MIR bio-sensors developed by the company Diafir [10].

A description of the materials and the discrimination methods is presented in section II. In section III, classification scores are computed and discussed. Finally, we conclude in section IV.

## II. MATERIALS AND METHODS

### A. Dataset

A multi-centric cohort of 58 patients from Rennes and Bordeaux hospitals was composed. All of them were cognitizant volunteers for clinical study, aged over 18 (men or women), with a suspicion of steatosis, but free of viral hepatitis, liver cancer or other inflammatory disease. Liver biopsies were collected on each patient to establish gold standard decision. Liver biopsies were embedded in paraffin, slices were sectioned and stained with hematoxylin-eosin-safran (HES) for global observation, and with Sirius Red to assess fibrosis. NASH severity was assessed by pathologists via histological analysis using NAS Score. This method assesses three parameters: steatosis level, inflammatory status and ballooning. Patients with positive score in all categories are considered affected by NASH [11]. The final verdict resulted in 35 patients diagnosed with NASH and 23 diagnosed with NAFLD. Blood samples were collected from these patients for routine analysis in the laboratory. They were frozen at  $-80^{\circ}\text{C}$  before the analysis. After clotting, they were centrifuged to extract the serum. A partial amount of this serum was used for MIR spectroscopy examination.

### B. MIR spectroscopy

The experimental setup consisted of an SPID™ non cooled spectrometer, coupled with a chalcogenide glass fiber enclosed in LS23™ sensor developed by Diafir (Fig. 1). SPID™ is a Fourier transformed infrared (FTIR) spectrometer that was especially conceived to accommodate biosensors developed by Diafir. It is made compact compared to classical spectrometers so that it can be used in point of care application. Chalcogenide glass fibers enjoy rheological properties that made them an optimal choice for FEWS implementation. They exhibit a large optical transparency in the MIR spectral range from 2 to 12  $\mu\text{m}$  with optical losses below 1  $\text{dB}\cdot\text{m}^{-1}$  in the 6-9  $\mu\text{m}$  region [12]. Diafir managed to develop tapered fibers that have compact looped sensing heads with diameter reaching 2  $\text{mm}$ , making them the most sensitive and robust sensors so far. Their fabrication procedure is detailed in [13]. These sensors were used in our experiments, they are single use and require roughly 7  $\mu\text{l}$  volume of serum per analysis.

The FTIR spectra were acquired in the absorption mode in the [4000,400]  $\text{cm}^{-1}$  frequency range. The nominal spectral resolution was set to 4  $\text{cm}^{-1}$  and a zero-filling factor of 2

was employed, yielding a discrete spectral point spacing of 2  $\text{cm}^{-1}$ . A Blackman Harris three-term apodisation function was used for Fourier transformation. Sixty-four scans were recorded and averaged to yield the sample absorption intensity. The optical signal was recorded at the output extremity of the fiber, providing the infrared single beam spectra. A single beam reference spectrum was obtained for the background (without sample) before each sample analysis, yielding for every wavenumber a reference intensity. This step is necessary to take into consideration possible side effects in the experiment such as the entrance/exit conditions of the infrared beam, the interaction and attenuation along the optical signal transportation section, the transition of the modes during the taper to the sensing zone, and effects related to fiber bending or surface roughness [14]. The sample and reference spectra of each patient were processed to derive only one “intensity ratio” spectrum that was used as input data in the statistical analysis presented afterward.

### C. Discrimination methods

The resulting NAFLD and NASH raw spectra are presented in Fig. 2. Initial portion before 900  $\text{cm}^{-1}$  was removed to avoid artifacts. The portion between 1800  $\text{cm}^{-1}$  and 2800  $\text{cm}^{-1}$  was also omitted to eliminate the contribution of the environmental  $\text{CO}_2$ . We notice that spectra from both categories are similar in average. Therefore we expect that intergroup correlations will be important. The hydrophobic property of chalcogenide glass fibers helps to minimize the effect of water, but it does not remove it completely, especially when the analyzed samples are liquids as in our case. We can see the influence of water in the presence of a baseline (slow wave) and a significant lobe after 3000  $\text{cm}^{-1}$  in all spectra. This internal artifact is very disturbing since it may screen small details that are important for discrimination. Moreover, it increases the intergroup correlations and brings a spurious similarity into the intragroup samples. In Fig. 3, we notice clearly the homogeneity between the diagonal blocks (intragroup correlations) and the off diagonal blocks (intergroup correlations) of the correlation matrix, which makes the discrimination task very complicated.

For signal analysis, we proposed three main steps: *i*) we computed the standardized form of signals (mean centered and divided by standard deviation) in order to ensure their comparability and detect the potential presence of outliers. In our case, all spectra were eligible for tests, *ii*) we then considered baseline correction. A classical method consists in using the first derivative of signals that is computed by a simple difference between two consecutive samples. In some cases, the first derivative is not good enough to reduce the baseline effect correctly, second order derivative is thus used. The classification procedure was separately applied on raw data and also on both first and second derivatives, *iii*) we implemented three different classifiers. The first classifier, k-means, is an unsupervised method where a dataset is partitioned into  $k$  clusters by minimizing a certain distance [15]. In our case, we had two clusters ( $k = 2$ ) and the criterion was set to the Euclidean distance. The second

classifier is a support vector machine (SVM) that is based on learning models (kernel functions) from a control dataset to build a hyperplane separating data in two classes [15]. In our tests, we set a linear kernel and applied a leave-one-out cross-validation (LOOCV) scheme to evaluate classification scores. In this method, one patient is removed from the given dataset, in turn, and used as a test set. The LOOCV scheme is justified by the small dataset available (58 patients). The third classifier consisted in using a generalized linear model (GLM) which is a generalization of linear regression. In GLM, the response variable is related to the linear model via a link function. In the case of binary response variable with a Bernoulli distribution, the typical choice of link function is the logit link [16]:

$$\mathbf{X}^T \boldsymbol{\beta} = \ln \frac{\mathbf{p}}{\mathbf{1} - \mathbf{p}}, \quad (1)$$

where  $\mathbf{X}$  is the  $M \times N$  matrix of dataset, here  $M$  and  $N$  denote the size of one spectrum and the number of patients, respectively.  $\boldsymbol{\beta}$  is the  $M$ -length vector of coefficients estimated with the control set, and  $\mathbf{p}$  is the  $N$ -length probability vector such that  $p_n$ , the  $n^{\text{th}}$  component of  $\mathbf{p}$ , is the probability of classifying patient  $n$  as positive (NASH in our context).  $p_n$  is compared to a threshold learned with the control set. A labeling is then assigned to the corresponding patient. We used the same LOOCV scheme as with SVM. However, the spectrum size in our dataset is  $M = 1027$  wavenumbers (variables). Hence, the use of GLM classifier implies the estimation of 1027 coefficients for  $N = 58$  leave-one-out sweeps, which makes computations very cumbersome. Moreover, the more variables we use the higher the risk of correlated data is, and GLM method fails for correlated data. To alleviate this problem, we proposed a simple variable selection pretreatment based on statistical significance measurement. We run the GLM test individually for each variable (each column of  $\mathbf{X}$ ) and compute the corresponding p-value. The variable is kept when its p-value is lower than a fixed threshold  $\sigma$ . In general,  $\sigma$  is tuned in the interval  $(0, 0.05]$ . In our simulations, we used an heuristic fashion to tune  $\sigma$  by varying it in  $(0, 0.05]$  and computing each time the corresponding classification scores. The optimal threshold was set to the value with the best classification scores, provided that GLM converges.

In summary, We used three classifiers (k-means, SVM and GLM) to test separately three inputs (raw spectra, first derivative and second derivative), which resulted in running nine tests.

### III. RESULTS AND DISCUSSION

We used the classical classification scores: sensitivity (Se), specificity (Sp) and accuracy (Acc) rates to evaluate and compare the different tests. All results are summarized in table I. As expected, the use of derivatives improved classification scores as they removed the fallacious correlation between the variables. The second derivative gave better results than the first derivative probably because the latter was not efficient enough to remove the baseline properly.

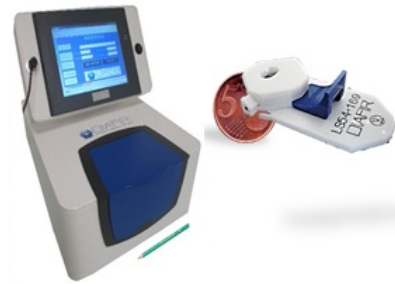


Fig. 1. SPID™ spectrometer and LS23™ sensor patented by Diafir.

k-means gave poor scores with the three types of inputs. SVM was more sensitive especially with the derivatives but unfortunately the specificity is still poor. We notice the superiority of GLM classifier that gave the best performance in all cases. The optimal result was obtained with the second derivative where an accuracy of 81% is achieved. The variable selection pretreatment selected only 4 variables when directly run on spectra ( $\sigma = 5.5 \cdot 10^{-3}$ ), 5 variables on the first derivative ( $\sigma = 4.32 \cdot 10^{-3}$ ) and finally 12 variables on second derivative ( $\sigma = 1.53 \cdot 10^{-2}$ ). The increasing number of selected variables shows once again the role of the second derivative to bring out more discriminative patterns screened by the presence of water beforehand. The selected wavenumbers lie close to the small fluctuations around  $3000 \text{ cm}^{-1}$  and in the interval  $(1800, 900] \text{ cm}^{-1}$ . In addition to reducing computations, variable selection pretreatment proved that the sought discriminative patterns can be confined in only few wavenumbers, which paves the ground for prospective exploratory researches about the biochemical interpretation of these wavenumbers.

Results obtained by the heuristic variable selection method and GLM are promising and incite us to test our dataset using more sophisticated variable selection techniques. In this work, we used simple baseline correction methods, namely, first and second derivatives. However, these methods introduce some artifacts. Therefore, more optimized techniques could be investigated in the future.

### IV. CONCLUSION

In this paper, we investigated the use of FEWS and chalcogenide glass bio-sensors to discriminate NASH from NAFLD patients. We were the first to test the new bio-sensors innovated by Diafir in this application. We tested three classifiers and obtained encouraging results with GLM associated with an heuristic variable selection pretreatment. Preliminary results show that this non-invasive technology is a promising simple diagnosis tool for NASH. However, there are some limitations that should be investigated in future work. The most important point is the validation of these results in a larger patients cohort. There are also potential variable selection algorithms that should be investigated to find out more optimized discriminative patterns.

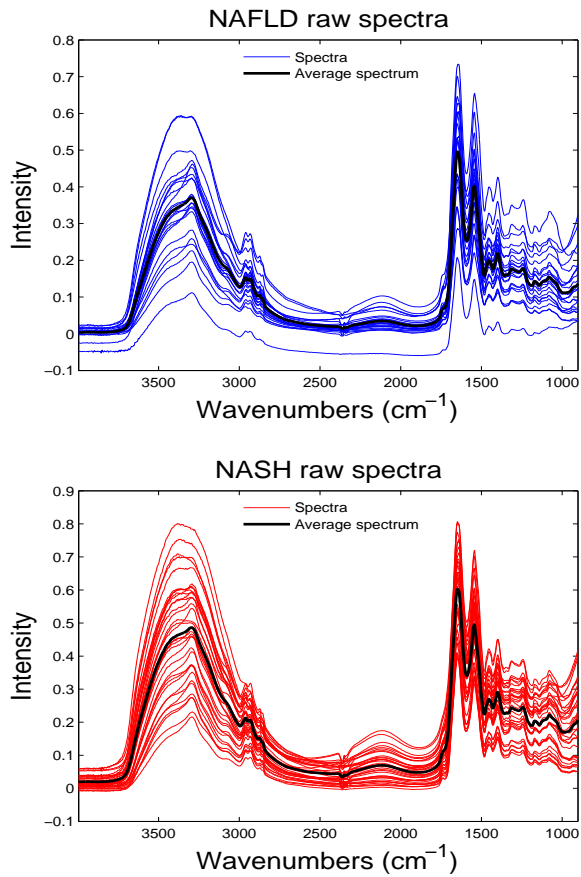


Fig. 2. MIR spectra: NAFLD (blue), NASH (red) and their corresponding averages (black).

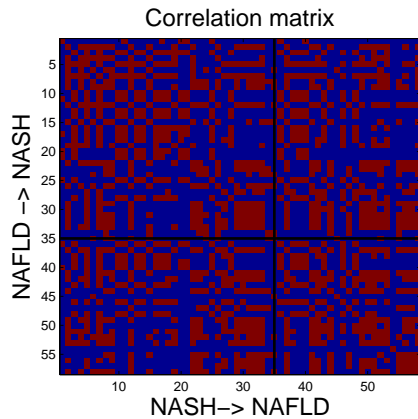


Fig. 3. Correlation matrix of spectra, a threshold is set to 0.98 where correlation coefficients  $> 0.98$  are set to 1 (red), 0 otherwise (blue).

## REFERENCES

- [1] Z. M. Younossi, A. B. Koenig, D. Abdelatif, Y. Fazel, L. Henry, and M. Wymer, "Global epidemiology of non-alcoholic fatty liver disease meta-analytic assessment of prevalence, incidence and outcomes," *Hepatology*, 2015.
- [2] M. Ahmed, "Non-alcoholic fatty liver disease in 2015," *World J Hepatol*, vol. 7, no. 11, pp. 1450–1459, 2015.
- [3] D. N and F. M, "Nonalcoholic Fatty Liver Disease A Major Public Health Challenge for the 21st Century," *JSM Gastroenterology and Hepatology*, 2014.

TABLE I  
CLASSIFICATION SCORES FOR ALL TESTS.

	Se %	Sp %	Acc %
Raw spectra			
k-means	37.1	69.5	50
SVM	62.8	39.1	53.4
GLM	71.4	56.5	65.5
1 <sup>st</sup> Derivative spectra			
k-means	51.4	30.4	43.1
SVM	71.4	39.1	58.6
GLM	71.4	78.2	74.1
2 <sup>nd</sup> Derivative spectra			
k-means	60	69.5	63.8
SVM	74.2	47.8	63.8
GLM	<b>82.8</b>	<b>78.2</b>	<b>81</b>

- [4] V. Ratziu, S. Bellentani, H. Cortez-Pinto, C. Day, and G. Marchesini, "A position statement on NAFLD/NASH based on the EASL 2009 special conference," *J. Hepatol.*, vol. 53, no. 2, pp. 372–384, 2010.
- [5] M. Obika and H. Noguchi, "Diagnosis and Evaluation of Nonalcoholic Fatty Liver Disease," *Exp Diabetes Res*, 2012.
- [6] R. Kwok, Y.-K. Tse, G. L.-H. Wong, Y. Ha, A. U. Lee, M. C. Ngu, H. L.-Y. Chan, and V. W.-S. Wong, "Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease the role of transient elastography and plasma cytokeatin-18 fragments," *Alimentary Pharmacology & Therapeutics*, vol. 39, no. 3, pp. 254–269, 2014.
- [7] R. Chahal, F. Starecki, C. Boussard-Plédel, J.-L. Doualan, K. Michel, L. Brilland, A. Braud, P. Camy, B. Bureau, and V. Nazabal, "Fiber evanescent wave spectroscopy based on IR fluorescent chalcogenide fibers," *Sensors and Actuators B: Chemical*, vol. 229, pp. 209–216, 2016.
- [8] S. Hocdé, O. Loréal, O. Sire, C. Boussard-Plédel, B. Bureau, B. Turlin, J. Keirsse, P. Leroyer, and J. Lucas, "Metabolic imaging of tissues by infrared fiber-optic spectroscopy: an efficient tool for medical diagnosis," *J. Biomed. Opt.*, vol. 9, no. 2, pp. 404–407, 2004.
- [9] M.-L. Anne, C. Le Lan, V. Monbet, C. Boussard-Plédel, M. Ropert, O. Sire, M. Pouchard, C. Jard, J. Lucas, J. L. Adam, P. Brissot, B. Bureau, and O. Loréal, "Fiber evanescent wave spectroscopy using the mid-infrared provides useful fingerprints for metabolic profiling in humans," *J Biomed Opt*, vol. 14, no. 5, 2009.
- [10] H. Tariel and F. Charpentier, "Device generating evanescent waves, and method for the implementation thereof," FR Patent WO/2015/110 767, July 31, 2015.
- [11] P. Bedossa, C. Poitou, N. Veyrie, J.-L. Bouillot, A. Basdevant, V. Paradis, J. Tordjman, and K. Clement, "Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients," *Hepatology*, vol. 56, no. 5, pp. 1751–1759, 2012.
- [12] J. Keirsse, B. Bureau, C. Boussard-Plédel, P. Leroyer, M. Ropert, V. Dupont, M. L. Anne, C. Ribault, O. Sire, O. Loréal, and J. L. Adam, "Chalcogenide glass fibers used for in situ infrared spectroscopy in biology and medicine," in *Optical Sensing*, B. Culshaw, A. G. Mignani, and R. Riesenberger, Eds., 2004, pp. 61–68.
- [13] P. Houizot, M.-L. Anne, C. Boussard-Plédel, O. Loréal, H. Tariel, J. Lucas, and B. Bureau, "Shaping of Looped Miniaturized Chalcogenide Fiber Sensing Heads for Mid-Infrared Sensing," *Sensors*, vol. 14, no. 10, pp. 17905–17914, 2014.
- [14] B. Bureau, C. Boussard-Plédel, J. L. Adam, and J. Lucas, "Infrared optical fiber as evanescent wave bio-sensors," in *Biomedical Optics 2005*, I. Gannot, Ed. International Society for Optics and Photonics, March 2005, pp. 1–8.
- [15] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [16] Annette J. Dobson, *An Introduction to Generalized Linear Models*, 2nd ed., F. Boca Raton, Ed. London: Chapman and Hall/CRC, 2002.