IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

βDate of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2017.Doi Number*

# Estimation of the Microbiological Quality of Meat using Rapid and Non-Invasive Spectroscopic Sensors

**Lemonia-Christina Fengou[1], Iosif Mporas[2], Evgenia Spyrelli[1], Alexandra Lianou[1], and George-John Nychas[1]**

[1]Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, School of Food and Nutritional Sciences, Agricultural University of Athens, Athens 11855, Greece
[2]School of Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK

Corresponding author: George-John Nychas (e-mail: gjn@aua.gr), Iosif Mporas (e-mail: i.mporas@herts.ac.uk).

This work has been supported by the project "PhasmaFOOD", funded from the European Commission Horizon 2020 research and innovation programme under grant agreement No 732541.

**ABSTRACT** Spectroscopic methods in tandem with machine learning methodologies have attracted considerable research interest for the estimation of food quality. The objective of this study was the evaluation of Fourier transform infrared (FTIR) spectroscopy and multispectral imaging (MSI) coupled with appropriate machine learning regression algorithms for assessing meat microbiological quality. For this purpose, minced pork patties were stored aerobically and under modified atmosphere packaging (MAP) conditions, at isothermal and dynamic temperature conditions. At regular time intervals during storage, samples were subjected to (i) microbiological analysis, (ii) FTIR measurements and (iii) MSI acquisition. The collected FTIR data were processed by feature extraction methods to reduce dimensionality, and subsequently Support Vector Machines (SVM) regression models were trained using spectral features (FTIR and MSI) to estimate microbiological quality of meat (microbial population). The regression models were evaluated with different experimental replicates using distinct meat batches. The performance of the models was evaluated in terms of correlation coefficient (r), root mean square error (RMSE), mean absolute error (MAE) and residual prediction deviation (RPD). The RMSE values for the microbial population estimation models using FTIR were 1.268 and 1.024 for aerobic and MAP storage, respectively. The performance in terms of RMSE for the MSI-based models was 1.144 for aerobic and 0.923 for MAP storage, while the combination of FTIR and MSI spectra resulted in models with RMSE equal to 1.146 for aerobic and 0.886 for MAP storage. The experimental results demonstrated the potential of estimating the microbiological quality of minced pork meat from spectroscopic data.

**INDEX TERMS** Food technology, microbiological quality, Fourier transform infrared spectroscopy, multispectral imaging, machine learning, support vector regression.

## I. INTRODUCTION

In the current Food Safety approach, a wide range of audits and inspections are applied to evaluate the quality or safety of raw or processed materials and food products [1].This is largely based on good design of processes, products and procedures, where the *end or finished product testing* is considered to be the control measure of the production process. In the case of microbiological food safety, specific microbiological analyses should be performed. In specific, microbiological analyses can be implemented with conventional microbiology approaches (e.g., colony counting methods) or molecular techniques that are considered more reliable and accurate [2], [3]. Chemical analyses are also used to monitor safety and quality of foods. These analyses have certain disadvantages, as they are (i) time-consuming in providing retrospective results, (ii) costly, (iii) some require high-tech molecular tools and thus, experienced and specialized personnel and commonly (iv) destructive to test products, limiting thus their potential to be used on-, in- or at-line [4], [5]. Due to these disadvantages, such analysis approaches cannot sufficiently guarantee consumer protection, since 100% inspection and sampling

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



are technically, financially and logistically impossible.

In recent years, novel spectroscopic techniques have been used for assessing the quality and/or adulteration or other fraudulent practices of certain foods including raw meats and fish. Such techniques are based on visible/infrared (VIS/IR), Raman [6], nuclear magnetic resonance (NMR) spectroscopies and spectral imaging [7], [8]. Although the currently available data should allow for significant improvements in food chain traceability and food safety, such improvements have not been sufficiently fulfilled yet. Actually, novel spectroscopic techniques are not yet scientifically sound and food-specific data are not readily available for producers and food chain actors [9]-[10] [11] [12] [13] [14].

In this context, the potential use of spectroscopy-based sensors, in tandem with machine learning regression algorithms, in the estimation of microbial populations was investigated. The objective of this study was to evaluate the potential of spectral data for developing models that will allow accurate estimation of the microbial populations of minced pork samples. The spectral data evaluated for this purpose were Fourier transform infrared (FTIR) spectroscopy and multispectral imaging (MSI) data, while different parameters were taken into account in model development. Specifically, the collected pairs of data (spectroscopic measurements of each meat sample and the corresponding microbiological analyses results) utilized for model development and evaluation corresponded to (a) independent experimental replicates involving distinct meat batches (accounting for both experimental and biological variability) and (b) different packaging conditions (aerobic storage, modified atmosphere packaging). Moreover, with reference to the FTIR data, different feature extraction approaches were applied in order to extract useful information (due to the high dimensionality of the FTIR feature space compared to the number of samples).

## II. EXPERIMENTAL DESIGN

### A. PREPARATION OF MEAT SAMPLES AND STORAGE CONDITIONS

Minced pork was purchased from a local butchery in Athens, Greece and transported to the laboratory with minimal delay, i.e. within 30 min. Minced pork patties (*ca*. 100g each) were prepared and placed onto styrofoam trays (duplicate samples (two patties) were placed in each tray). The described procedure was repeated eight times and the experiments were chronically independent (eight experimental replicates). Four replicates were stored aerobically (AIR) and four replicates under modified atmosphere packaging (MAP) conditions. Data derived from experimental replicates 1 and 2 have been previously published [15]. Samples stored aerobically were wrapped with cling film, while samples packed under modified atmosphere packaging conditions were enclosed in plastic packages (length: 25cm, width: 25cm, thickness: 90μm, permeability of *ca.* 25, 90, 6 $cm^3\ m^{-2}day^{-1}bar^{-1}$(1 bar= $10^5$Pa), at 20°C and 50% RH for $CO_2$, $O_2$ and $N_2$, respectively) and flushed with 80%$O_2$ and 20%$CO_2$ using a HenkoVac 1900Machine (Howden Food Equipment BV, The Netherlands). The applied high-oxygen MAP conditions were selected so as the red colour of minced pork to be preserved and the microbiological shelf life to be extended [16]. Samples were stored at three different isothermal conditions, namely 4, 8 and 12°C and under dynamic temperature conditions (periodic temperature changes (4-8-12°C) every 8h) in high-precision (±0.5°C) programmable incubators (MIR-153, Sanyo Electric Co., Osaka, Japan) for a maximum time period of 15 days (i.e. during storage at 4°C under MAP storage). Duplicate samples were analyzed at 10-h or 14-h intervals until 110h, while after 110h samples were analyzed at 24-h intervals, until pronounced spoilage was observed. The incubation temperatures were recorded at 15-min intervals throughout storage using electronic temperature-monitoring devices (COX TRACER®, Cox Technologies Inc., Belmont, NC, USA). Meat samples corresponding to independent experimental replicates were evaluated in the context of this study, with duplicate samples (originating from different 100g portions) being analyzed at the defined time intervals for every storage condition (i.e. packaging and temperature). The sampling points (time intervals) were selected so as different microbiological quality levels of minced meat, corresponding to various microbial concentrations from low (fresh meat) to high populations (spoiled meat), to be covered; in this sense, the proposed methodology should allow for the development of an efficient meat quality assessment approach that is applicable in real-life setups such as food quality inspections in restaurants, butcheries and groceries where the storage conditions and time are unknown. The assessment of food quality and more specifically of microbiological quality is a rather difficult and complex task for food commodities, such as minced pork. A number of intrinsic (pH, $a_w$, chemical composition, Eh, food structure) and extrinsic (temperature, humidity, gaseous condition of packaging) factors influence the rate of food deterioration, loss of freshness and loss of quality across the food chain e.g. during the storage of the product from farm to fork. Among the intrinsic/extrinsic factors mentioned above, temperature is the most detrimental one. For this reason, in food science the simulation of conditions followed in the present study (i.e. different temperature and storage conditions) is essential. The more samples analyzed in different storage conditions the better the microbial modeling is. The applied analytical procedures included: (i) microbiological analysis, (ii) FTIR spectroscopy measurements and (iii) MSI acquisition. A total of 903 samples were analyzed in the course of this study, with 456 and 447 samples corresponding to storage under AIR and MAP conditions, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



### B. MICROBIOLOGICAL ANALYSIS

Minced pork patties were analyzed microbiologically on the day of arrival to the laboratory (time-zero) and at regular time intervals during storage as described in section II. A. For the purpose of microbiological analysis, 25-g minced pork portions were transferred aseptically to 400-ml sterile stomacher bags (Seward Medical, London, United Kingdom) containing 225 ml of quarter-strength Ringer's solution (Lab M Limited, Lancashire, United Kingdom) and homogenized in a Stomacher apparatus (Lab Blender 400, Seward Medical) for 60 sec at room temperature. Appropriate serial decimal dilutions in Ringer's solution were surface plated on tryptic glucose yeast agar (Plate Count Agar; Biolife, Milan, Italy) and incubated at 25°C for 72h, for the enumeration of total mesophiles (total viable counts, TVC). The obtained microbiological data were converted to log (colony forming units) per gram of meat (log CFU/g).

#### 1) AIR STORAGE

The evolution of TVC of minced pork stored aerobically at 4°C, 8°C, dynamic temperatures and 12°C for replicates 1, 2, 3 and 4 is illustrated in Fig. 1 (a), (b), (c) and (d), respectively. The initial TVC in minced pork ranged from 3.00 to 3.86 log CFU/g (Fig. 1), a concentration which is in accordance with the range of 2.0 to 4.2 log CFU/g reported in the scientific literature [17], [18]. Similar microbial growth trends were observed among independent replicates, with replicate 1 being associated with lower values of microbial counts compared to the other replicates in most sampling times, an observation that is most likely related to the better initial microbiological quality of meat (Fig. 1). It has been demonstrated that excessive quality deterioration of minced meat during aerobic storage, including off-odours' development and slime production, is evident when TVC is 7-8 log CFU/g [19], [20]. Mean (± standard deviation, n=8) TVC of samples reached 7-8 log CFU/g at 4°C in 158h (7.63 ± 0.79 log CFU/g), at 8°C in 96h (7.93 ± 0.48 log CFU/g), at dynamic temperatures in 96h (7.66 ± 0.43 log CFU/g) and at 12°C in 62h (7.57 ± 0.32 log CFU/g). The microbial growth recorded at 8°C was similar with that recorded at dynamic temperature conditions. The microbial groups that contributed mainly to the spoilage of minced pork were *Pseudomonas* spp., *Brochothrix thermosphacta* and lactic acid bacteria, with *Pseudomonas* spp. constituting the dominant spoilage microorganisms (data not shown),

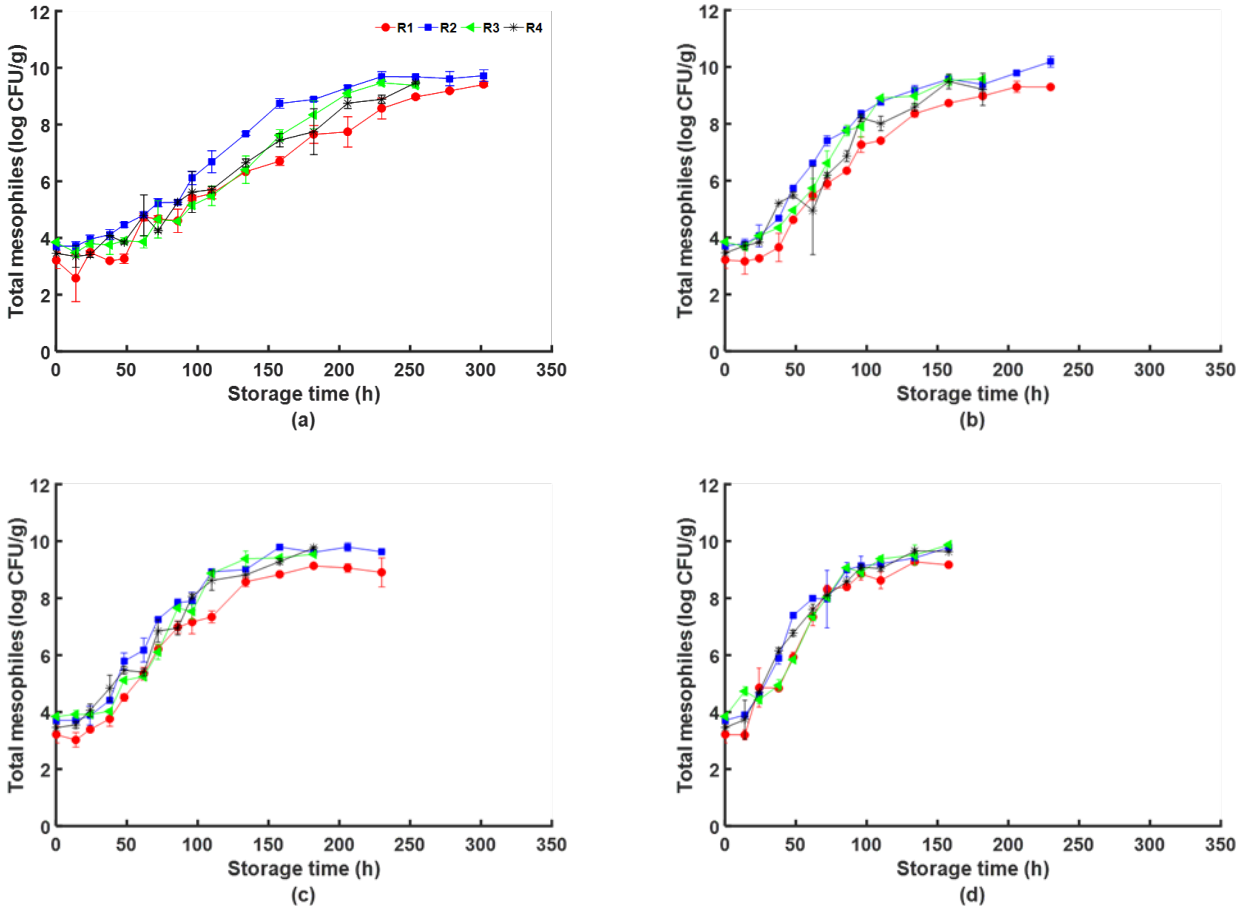


**FIGURE 1. Total mesophilic microbial populations (mean ± standard deviation, n=2) of minced pork stored aerobically at 4°C (a), 8°C (b), dynamic temperatures (c) and 12°C (d) for replicates 1 (R1), 2 (R2), 3 (R3) and 4 (R4).**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



findings that are in agreement with those previously reported in the scientific literature [21].

### 2) MAP STORAGE

The initial microbial populations in the minced pork samples corresponding to the four experimental replicates of MAP storage were generally higher compared to those recorded for the samples stored under AIR conditions. Indeed, the initial TVC in minced pork stored under MAP conditions ranged from 3.66 to 5.30 log CFU/g. The evolution of TVC during storage at 4°C, 8°C, dynamic temperatures and 12°C under MAP conditions for replicates 1, 2, 3 and 4 is illustrated in Fig. 2 (a), (b), (c) and (d), respectively. Since the initial microbiological quality of minced pork corresponding to replicates 1 and 2 was higher (i.e. lower initial microbial populations) compared to replicates 3 and 4, the duration of the applied storage in the latter case was shorter (Fig. 2). Similarly to the observations made under AIR storage conditions, the microbial evolution recorded during MAP storage at 8°C was comparable with that recorded at dynamic temperature conditions. Overall, the shelf life of minced pork did not appear to be considerably extended during MAP storage as compared to AIR storage under the conditions of this study, most likely due to the relatively high initial microbial population of the minced pork samples as mentioned above. The main spoilage bacteria expected to contribute to quality deterioration of minced pork during storage under MAP conditions are in decreasing order *Br. thermosphacta*, lactic acid bacteria and *Pseudomonas* spp. [21].

## C. FTIR SPECTROSCOPY

FTIR spectral data were collected using a ZnSe 45° HATR (Horizontal Attenuated Total Reflectance) crystal (PIKE Technologies, Madison, Wisconsin, United States) and an FTIR-6200 JASCO spectrometer (Jasco Corp., Tokyo, Japan) equipped with a standard sample chamber, a triglycine sulphate (TGS) detector and a Ge/KBr beam splitter. The crystal used has a refractive index of 2.4 and a depth of penetration of 2.0 μm at 1000 $cm^{-1}$. Using the Spectra Manager™ software version 2 (Jasco Corp.), spectra were collected over the wavenumber range of 4000 to 400 $cm^{-1}$, by accumulating 100 scans with a resolution of 4 $cm^{-1}$ within a period of 2 min. Prior to the measurements of the tested samples, reference spectra were acquired using the cleaned blank (no added sample) crystal. Small portions of minced

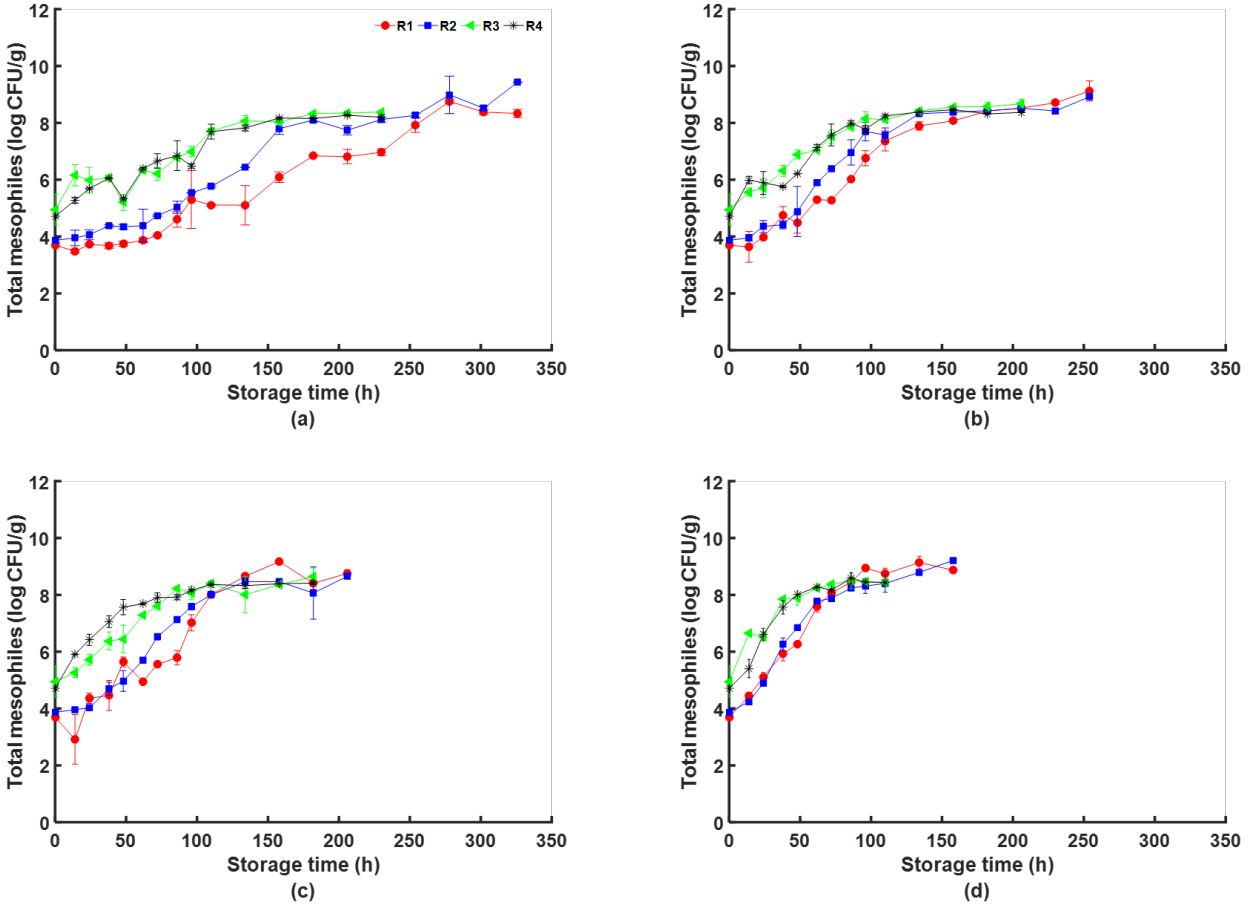


**FIGURE 2. Total mesophilic microbial populations (mean± standard deviation, n=2) of minced pork stored at 4°C (a), 8°C (b), dynamic (c) and 12°C (d) temperature conditions under modified atmosphere packaging conditions for replicates 1 (R1), 2 (R2), 3 (R3) and 4 (R4).**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



meat were placed on the crystal and measurements were performed. After each measurement, the crystal's surface was cleaned, first with detergent and distilled water and then with analytical grade acetone and dried using lint-free tissue. The FTIR spectra that were ultimately used in further analyses were in the approximate wavenumber range of 1800 to 900 $cm^{-1}$ [22], consisting of 934 wavenumbers.

### D. IMAGE ACQUISITION

Multispectral imaging data from portions (*ca.* 70g) of the minced pork samples were acquired using the VideometerLab system, originally developed by the Technical University of Denmark [23] and according to procedures that have been described in detail previously [24]. In brief, this instrument acquires multispectral images in 18 different, non-uniformly distributed wavelengths ranging from 405 to 970 nm (i.e. 405, 430, 450, 470, 505, 565, 590, 630, 645, 660, 850, 870, 890, 910, 920, 940, 950 and 970 nm). Prior to image acquisition, the system was subjected to a light set up procedure known as "autolight" and calibration. Each minced meat sample was placed in a Petri dish and the latter was placed inside an Ulbricht sphere, in which the camera is top-mounted and the corresponding multispectral image of the product's surface was taken.

The collected images were segmented using the accompanying software (version 2.12.39) of the VideometerLab system so as to remove information related to the background (i.e. Petri dish, fat). Specifically, first the contrast between the sample material and the other irrelevant objects was maximized, in order to enable a threshold operation. Then, Canonical Discriminant Analysis (CDA) was employed as a supervised transformation method so as to divide the image into regions of interest. In this way a segmented image was produced, where the region of interest was used to extract the spectral data. After the segmentation process of each image, the mean reflectance spectrum (i.e. mean intensity of pixels within the informative area) along with the corresponding standard deviation values were calculated. The MSI data that ultimately were used in further analyses were 18 mean reflectance values and 18 standard deviations, constructing a 36-dimentional feature vector representation for each MSI image, without using 2-dimensional image features.

The proposed framework is based on offline training and online operation (test phase) for the estimation of microbial population of meat using spectroscopic data (FTIR, MSI). The methodology is generalized and can be expanded by using or adding other spectroscopic techniques (e.g. Raman). The experimental design and methodology applied on minced pork samples in the framework of the present study is illustrated in Fig. 3.

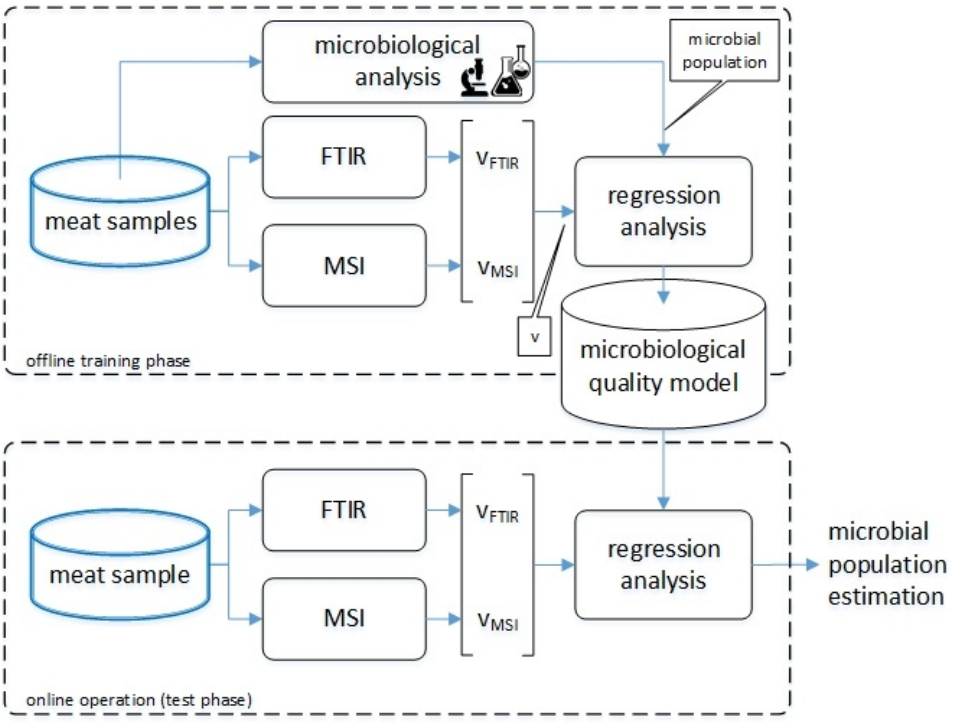


**FIGURE 3. Block diagram of the experimental design**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



## III. DATA ANALYSIS

Data (spectra, microbial populations) were collected and Support Vector Machines (SVM) regression models were developed for each sensor and packaging condition (FTIR_AIR, MSI_AIR, FTIR_MAP, MSI_MAP) and evaluated. Spectral data (FTIR, MSI) were used as input variables (X) and TVC as output (target) variables (Y). In the case of FTIR, where the number of variables is greater than the number of samples [25], different dimensionality reduction methods were applied (Principal Component Analysis, Partial Least Squares, ReliefF, Descriptive statistics, Autoencoders). The acquired FTIR variables were used as X variables. The training of the models was performed using three replicates, while the fourth replicate was used to evaluate the SVM regression models. A 4-fold cross validation experimental protocol was followed in order to avoid overlap between training and test replicates. The train and test sets used for the 4-fold cross validation experimental protocol are presented in Tables I and II.

Several well-known and widely used classifiers were evaluated and the results demonstrated that SVM offer competitive performance in most setups and was thus, selected for our analysis. Also, since SVM has good generalization ability and do not suffer from the curse of dimensionality [26], [27], it is regarded as an appropriate approach for comparison of datasets with different dimensionality. The performance of the regression models was evaluated in terms of Root Mean Square Error (RMSE). Except RMSE which was considered as the main numerical estimation metric as also in [10], [11], [12], [14], also were calculated the Mean Absolute Error (MAE), the correlation coefficient (r) and the Residual Prediction Deviation (RPD).

**Table I.** Train and test data sets used for model development, collected during AIR storage. R stands for replicate and TVC for total viable count (i.e. microbial population).

| Train Set (number of samples) | Train TVC Range log CFU/g | Test Set (number of samples) | Test TVC Range log CFU/g |
|---|---|---|---|
| R2, R3, R4 (336) | 3.08 – 10.32 | R1 (113) | 3.08 – 9.45 |
| R1, R3, R4 (336) | 2.00 – 9.93 | R2 (118) | 3.60 – 9.90 |
| R1, R2, R4 (348) | 2.00 – 10.32 | R3 (108) | 3.36 – 9.93 |
| R1, R2, R3 (348) | 2.00 – 10.32 | R4 (108) | 3.08 – 9.80 |

**Table II.** Train and test data sets used for model development, collected during MAP storage. R stands for replicate and TVC for total viable count (i.e. microbial population).

| Train Set (number of samples) | Train TVC Range log CFU/g | Test Set (number of samples) | Test TVC Range log CFU/g |
|---|---|---|---|
| R2, R3, R4 (326) | 3.76 – 9.45 | R1 (103) | 3.81 – 9.37 |
| R1, R3, R4 (325) | 2.30 – 9.37 | R2 (119) | 3.76 – 9.22 |
| R1, R2, R4 (341) | 2.30 – 9.45 | R3 (106) | 4.60– 8.90 |
| R1, R2, R3 (349) | 2.30 – 9.45 | R4 (98) | 4.70 – 8.70 |

### A. DIMENSIONALITY REDUCTION OF FTIR DATA

Several methods were used for the reduction of dimensions of the FTIR data. These methods are: (a) Principal Component Analysis (PCA), which is a method for reducing the dimensionality, but at the same time minimizing the information loss. A large number of potentially correlated factors are projected into a number of orthogonal (uncorrelated) factors (i.e. principal components) reducing thus the size of the initial dataset taking into account only X-variables [28]; (b) Partial Least Squares (PLS), an approach which, unlike PCA, constructs new factors (i.e. latent variables) that combine information about the variances of both the X-variables and Y-variables (find combinations of the predictors (X-variables) that have a large covariance with the response (target) values (Y-variables)) and is used for data with many, noisy and collinear variables [29], [30]; (c) ReliefF is a feature ranking method, which evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class and is able to detect conditional dependencies between attributes and to provide a unified view on the attribute estimation in regression [31]; (d) 22 descriptive statistics (Minimum, Maximum, Mean, Standard Deviation, Variance, Median, 10th Percentile, 25th Percentile, 75th Percentile, 90th Percentile, Kurtosis, Skewness, Mode, Sum, Range, Geometric Mean (geomean), Harmonic Mean (harmmean), Trimmean (10%), Trimmean (33%), Interquartile Range, Mean Absolute Deviation, Standard Error) were calculated for each sample's spectrum thus projecting the 934-dimensions of FTIR data to 22; (e) Autoencoder, which is a neural network trained to replicate its input at its output. When nonlinear activation functions are used, autoencoders provide nonlinear generalizations of PCA. Training an autoencoder is unsupervised in the sense that no labeled data are needed [32] - [33] [34] [35].

Spectral data were processed by each of the above described methods, (a)-(e), to reduce dimensionality either by projecting them to a new lower dimensional feature space (i.e. descriptive

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



statistics, autoencoders), or by projecting them to a new feature space with the same dimensionality and selecting a subset of the new feature space dimensions (PCA, PLS, ReliefF). All dimensionality reduction algorithms were trained using only the train sets presented in Tables I and II and the developed models for dimensionality reduction were then used to reduce the dimensionality of the corresponding test set, thus there was no overlap between train and test sets. The selection of the optimal number of features was based on the RMSE score minimization on the test datasets, except for the case of descriptive statistics in which the dimensionality was fixed to 22.

### B. SUPPORT VECTOR MACHINES REGRESSION

FTIR and MSI data were used for the development of SVM regression models. Data were mapped non-linearly into a higher dimensional feature space, so as a maximal separating hyper-plane to be constructed. Radial Basis Function (RBF) and Polykernel (POLY) were used as kernel functions and grid search was employed for finding the optimal cost (C) and gamma (γ) parameters, with the parameters being selected according to best performance of each test set [36]. Spectral data (FTIR, MSI) were concatenated in an early fusion setup. As regards FTIR data, they were used after the application of dimensionality reduction.

The evaluation results (RMSE, MAE, r, RPD) on the test sets are presented in Section IV. PCA, ReliefF, Descriptive Statistics and Autoencoders were implemented in MATLAB [37], while PLS was implemented in R 3.5.1 [38] and RStudio [39], using 'pls' package [30]. For the SVM regression modelling, the WEKA 3.8 software tool was utilized for machine learning using the SMOreg function [40].

## IV. RESULTS AND DISCUSSION

Spectral data obtained from FTIR spectroscopy in the range of 1800 to 900cm$^{-1}$ and from MSI in the range of 405 to 970 nm were used as X-variables for the development of the regression models (estimators of microbial population). FTIR data provide an overall fingerprint of the biochemical composition of food samples, also including the metabolic activity of microorganisms and are thus used as means of microbiological quality assessment. The abovementioned region has been widely used in model development for the ultimate purpose of food quality estimation [22], [41], [42]. On the other hand, MSI data cover the visible and a part of the near-infrared red (NIR) region. There are peaks in the abovementioned regions which are associated with myoglobin oxidation (visible region), as well as with protein and moisture content (near-infrared region) [43], [44].

### A. AIR STORAGE

SVM models were developed and evaluated after dimensionality reduction methods (PCA, PLS, ReliefF, Descriptive statistics and Autoencoder) were applied to FTIR data of AIR storage. A visualization of the relation between the measured (via microbiological analysis) and estimated (by the model) microbial population is illustrated in the scatterplots shown in Fig. 4. As can be seen in Fig. 4, the two variables (measured and estimated population) have positive association with a slight exception for the case of Autoencoders. A moderate linear relationship can be seen in all methods with PCA, PLS and ReliefF methods presenting stronger relationship than descriptive statistics and Autoencoders, which is in agreement with the results of Table III. The best performance of the SVM models on FTIR data was obtained using PCA and PLS methods, as shown in Table III. In a review of the literature [14] about spectroscopic data related to food science and more specifically to microbial evaluation, it was concluded that PLS is the most frequently used method, followed by ANN, PCA and SVM methods. In more detail according to RMSE, MAE and RPD values, the best performance on the testing data sets was observed using PCA with the scores being 1.268, 1.028 and 1.759, respectively. Slightly worse performance was obtained when applying the PLS feature selection method with RMSE, MAE and RPD being 1.277, 1.050 and 1.746, respectively. r was slightly better in the case of PLS by 0.003 compared to PCA, which is not a considerable difference. The use of ReliefF method for dimensionality reduction resulted in a lower performance compared to the two aforementioned methods (r=0.774, RMSE=1.417, MAE=1.174 and RPD=1.574). Moreover, the Descriptive Statistics and the Autoencoder exhibited a considerably lower performance than the PCA and PLS feature selection methods, most probably due to the limited size of the training data set, which was not enough to robustly train an autoencoder and estimate descriptive statistical values. When microbial population was estimated using SVM and Descriptive Statistics, the performance metrics of r, RMSE, MAE and RPD were equal to 0.388, 2.231, 1.868 and 0.999, respectively, while in the case of Autoencoders the corresponding values were 0.177, 2.314, 2.035 and 0.964.

**Table III.** Performance of microbial population estimation under AIR storage conditions using SVM regression modelling and different dimensionality reduction methods applied on the FTIR data.

| Dimensionality reduction methods | r | RMSE | MAE | RPD |
|---|---|---|---|---|
| PCA | 0.828 | 1.268 | 1.028 | 1.759 |
| PLS | 0.831 | 1.277 | 1.050 | 1.746 |
| ReliefF | 0.774 | 1.417 | 1.174 | 1.574 |
| Descriptive statistics | 0.388 | 2.231 | 1.868 | 0.999 |
| Autoencoder | 0.177 | 2.314 | 2.035 | 0.964 |

r: correlation coefficient; RMSE: root mean square error; MAE: mean absolute error; RPD: residual prediction deviation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



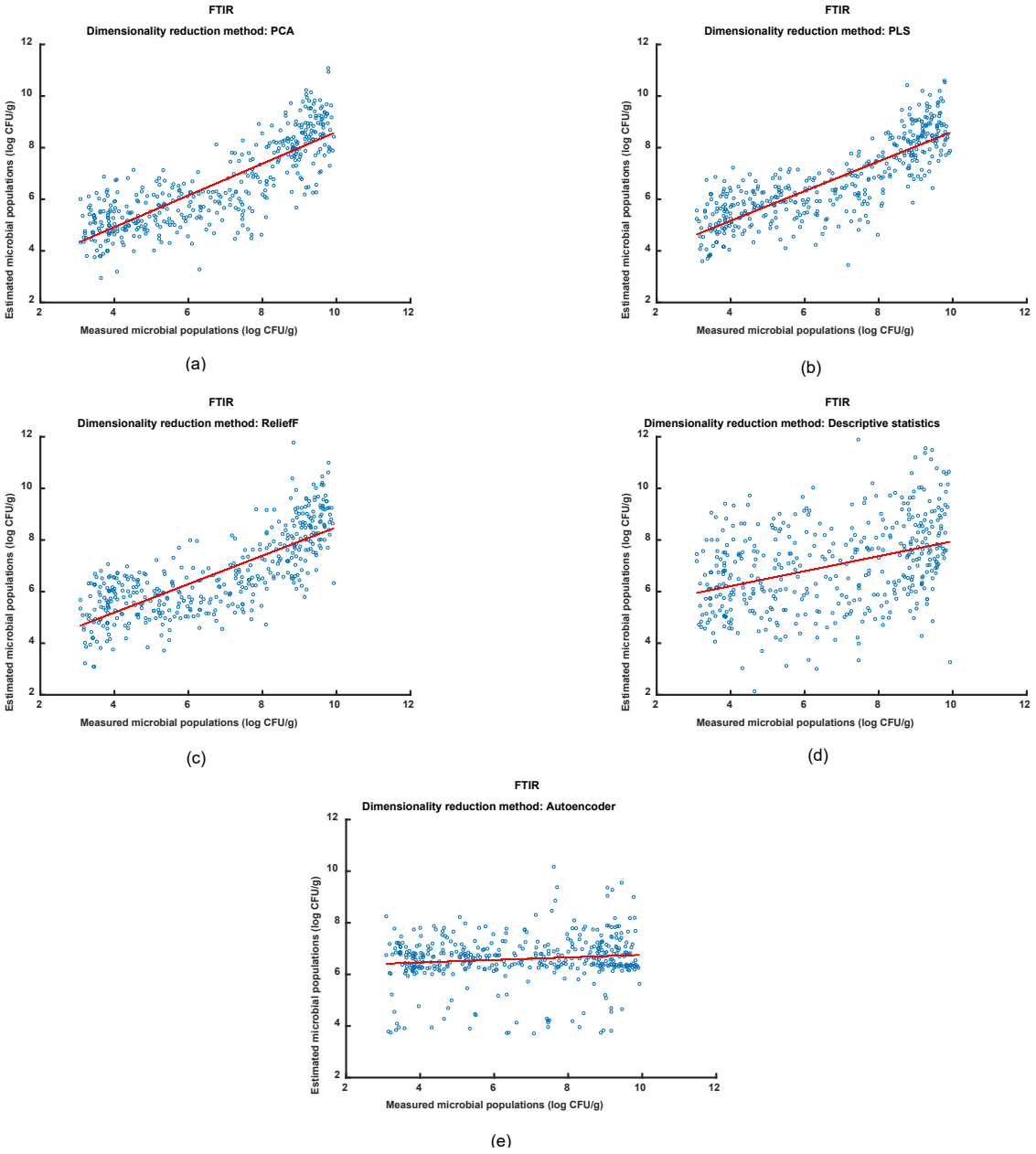


**FIGURE 4. Scatterplots of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on FTIR features for different dimensionality reduction methods, namely (a) PCA, (b) PLS, (c) ReliefF, (d) Descriptive statistics and (e) Autoencoder for data acquired under AIR storage conditions. Red line corresponds to linear fit of the data.**

The relation between the measured (via microbiological analysis) and estimated microbial population by the SVM models using MSI and MSI+FTIR features is visualized in Fig. 5 and 6, respectively. As can be seen in both scatterplots, the two variables (measured and estimated population) have positive association and stronger linear relationship compared to Fig. 4, which is in agreement with the results shown in Table IV. The performance metrics of r, RMSE, MAE and RPD (Table IV) for models developed using MSI data were estimated to be 0.859, 1.144, 0.921 and 1.949, respectively. The performance of the SVM regression models using MSI data was better compared to FTIR. Models derived from early fusion had comparable performance with models derived from MSI (r=0.863, RMSE=1.146, MAE=0.937 and RPD=1.946).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



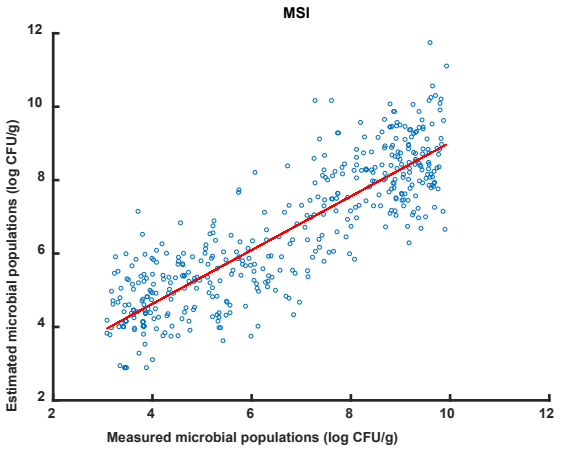


**FIGURE 5. Scatterplot of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on MSI features for data acquired under AIR storage conditions. Red line corresponds to linear fit of the data.**

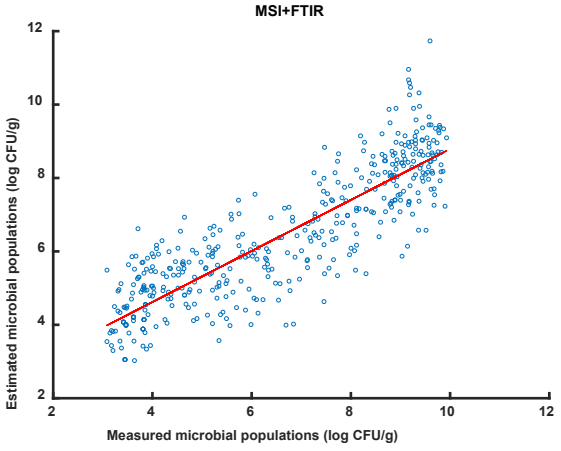


**FIGURE 6. Scatterplot of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on MSI+FTIR features for data acquired under AIR storage conditions. Red line corresponds to linear fit of the data.**

**Table IV.** Performance of microbial population estimation under AIR storage conditions using MSI and MSI+FTIR, for the best performing FTIR setup, as features.

| Features | r | RMSE | MAE | RPD |
|---|---|---|---|---|
| MSI | 0.859 | 1.144 | 0.921 | 1.949 |
| MSI+FTIR | 0.863 | 1.146 | 0.937 | 1.946 |

r: correlation coefficient; RMSE: root mean square error; MAE: mean absolute error; RPD: residual prediction deviation.

The performance (in terms of RMSE scores) of the four test sets using different dimensionality reduction methods under AIR storage is shown in Fig. 7. The highest performance, as indicated by the lowest RMSE value, was observed when 100 principal components from PCA method were used. In contrast to PCA, the best performance when using the PLS method was observed with fewer components (10 to 30 latent variables), which is in agreement with the scientific literature [30], where it has been reported that when using PLS fewer components are needed to achieve similar prediction accuracies compared to PCA. On the other hand, PCA achieved slightly better performance in terms of RMSE due to the good generalization ability of SVM which do not suffer from the curse of dimensionality [26], [27], thus making the SVM regression model not quite sensitive to dimensionality variations. In the case of ReliefF method, the RMSE was reduced and minimized when over 300 variables were used. On the other hand, RMSE values of SVM models applied to Autoencoders were stable at about 2.150 to 2.549 (Fig. 7).

Spectroscopic methods, and particularly FTIR data, have also been used in other studies for the estimation of microbial quality of meat stored under aerobic conditions. For example, in [45] minced pork was stored at different isothermal conditions, where a total of 134 samples (67 samples for training and 67 for testing) were used. In the case of this latter study, r (correlation coefficient) for train and test set was 0.895 and 0.880, respectively. In [46] beef fillets stored at different temperatures (0, 5, 10, 15 and 20°C), models were developed using ANN, and the RMSE value for train and test set was 1.821 and 1.978, respectively, and the corresponding RMSE values for PLS modeling were 1.073 and 1.993. In the study [47], microbiological quality of goat and fallow meat was studied using PLSR, with samples being stored at 3°C in the context of four experimental replicates (60 samples per meat type); the RMSE values estimated for external validation of models developed for fallow and goat were 0.75 and 0.74, respectively. Promising results regarding the estimation of microbiological spoilage based on FTIR data has been also reported for chicken breast [48]. Moreover, MSI data have been successfully used for the quantitative assessment of meat spoilage [49], [50], [24]. Actually, in the latter study and regarding the correlation of MSI and microbiological data, r values as high as 0.928, 0.918 and 0.783 were reported for model calibration, cross-validation and prediction, respectively. Going beyond the parameters studied in the abovementioned studies, different experimental replicates and temperature conditions were taken into account in the present study. The developed models were full-cross validated using independent experiments (corresponding to distinct meat batches) as test sets for the purpose of model optimization and method evaluation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



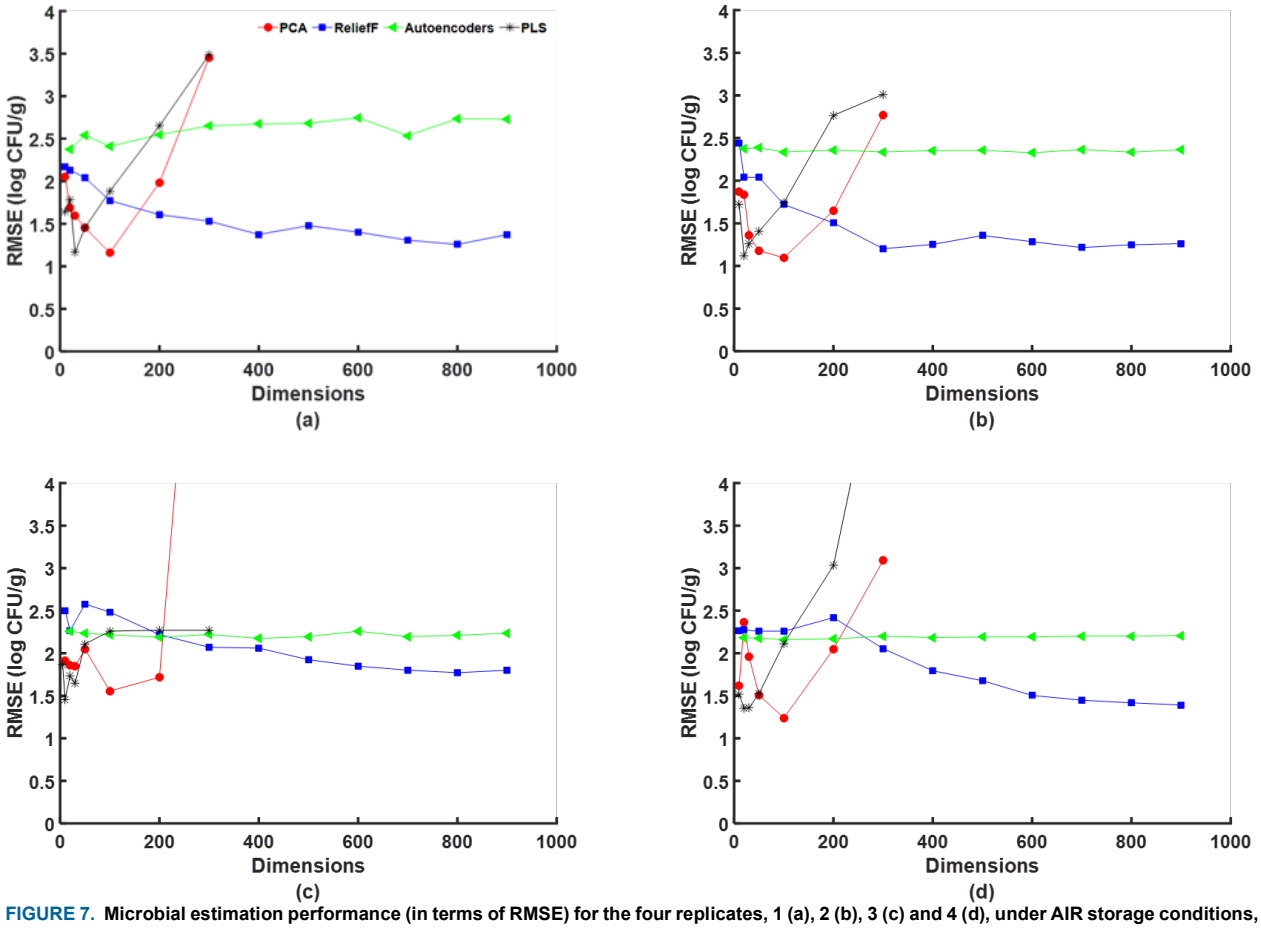


**FIGURE 7. Microbial estimation performance (in terms of RMSE) for the four replicates, 1 (a), 2 (b), 3 (c) and 4 (d), under AIR storage conditions, for different dimensionality reduction methods with respect to the number of components (dimensions) used.**

### *B. MAP STORAGE*

The SVM regression models developed from data acquired during storage under MAP conditions, had comparable performance with those obtained from samples stored under AIR conditions. In Fig 8., the relation between the measured (via microbiological analysis) and estimated (by the model) microbial population is illustrated in the form of scatterplots. A moderate linear relationship can be seen in all methods with PCA, PLS and ReliefF methods presenting stronger relationship than descriptive statistics and Autoencoders, which is in agreement with the results shown in Table V. In Table V, r, RMSE, MAE and RPD scores when using FTIR data are shown; the best performance of the MAP models was obtained using the PLS feature selection method with r, RMSE, MAE and RPD being equal to 0.765, 1.024, 0.809 and 1.538, respectively. Models developed using the PCA and ReliefF methods were slightly worse compared to the PLS method. PCA-based models had better performance (r=0.759, RMSE=1.047, MAE=0.787 and RPD=1.504) compared to ReliefF-based models (r=0.706, RMSE=1.141, MAE=0.951 and RPD=1.381). The methods of Descriptive Statistics and Autoencoders resulted in poor model performance, similarly to the AIR storage case. RMSE scores of the SVM regression models developed using FTIR data with Descriptive Statistics and Autoencoders were equal to 1.429 and 1.563, respectively.

**Table V.** Performance of microbial population estimation under MAP storage conditions using SVM regression modelling and different dimensionality reduction methods applied on the FTIR data.

| Dimensionality reduction methods | r | RMSE | MAE | RPD |
|---|---|---|---|---|
| PCA | 0.759 | 1.047 | 0.787 | 1.504 |
| PLS | 0.765 | 1.024 | 0.809 | 1.538 |
| ReliefF | 0.706 | 1.141 | 0.951 | 1.381 |
| Descriptive statistics | 0.457 | 1.429 | 1.188 | 1.102 |
| Autoencoder | 0.245 | 1.563 | 1.309 | 1.007 |

r: correlation coefficient; RMSE: root mean square error; MAE: mean absolute error; RPD: residual prediction deviation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



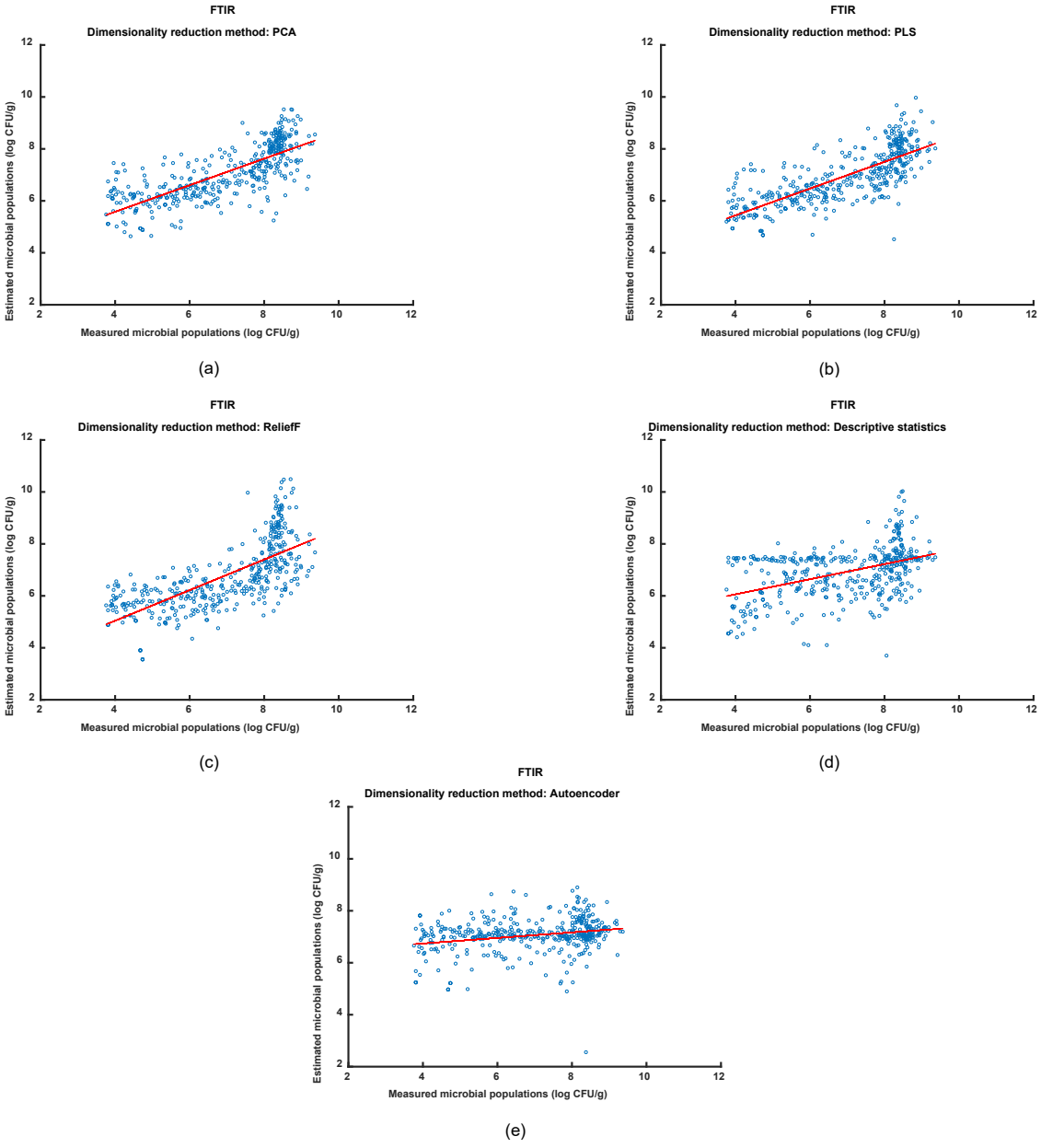


**FIGURE 8. Scatterplot of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on FTIR features for different dimensionality reduction methods, namely (a) PCA, (b) PLS, (c) ReliefF, (d) Descriptive statistics and (e) Autoencoder for data acquired under MAP storage conditions. Red line corresponds to linear fit of the data.**

The performance of the SVM regression models was better when MSI data concatenated with FTIR data were used as shown in Table VI. A visualization of the relation between the measured (via microbiological analysis) and estimated microbial population by the SVM models using MSI and MSI+FTIR features is provided in Fig. 9 and 10, respectively. As can be seen in both scatterplots, the two variables (measured and estimated population) have positive association and stronger linear relationship compared to Fig. 8 which is in agreement with the results shown in Table VI. In Table VI, the r, RMSE, MAE and RPD values of the microbial population estimation models developed using MSI data were 0.810, 0.923, 0.727 and 1.705, respectively. With reference to models derived from early fusion of MSI and FTIR data, slightly better performance was achieved (r=0.834, RMSE=0.886, MAE=0.697 and RPD=1.778).

Data fusion of spectroscopic data in tandem with data analysis has been studied for several applications related to food quality assessment and food fraud [51], [52], but only few of the conducted research studies are related to the evaluation of meat quality [53], [54].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



**Table VI.** Performance of microbial population estimation under MAP storage conditions using MSI and MSI+FTIR, for the best performing FTIR setup, as features.

| Features | r | RMSE | MAE | RPD |
|---|---|---|---|---|
| MSI | 0.810 | 0.923 | 0.727 | 1.705 |
| MSI+FTIR | 0.834 | 0.886 | 0.697 | 1.778 |

r: correlation coefficient; RMSE: root mean square error; MAE: mean absolute error; RPD: residual prediction deviation.

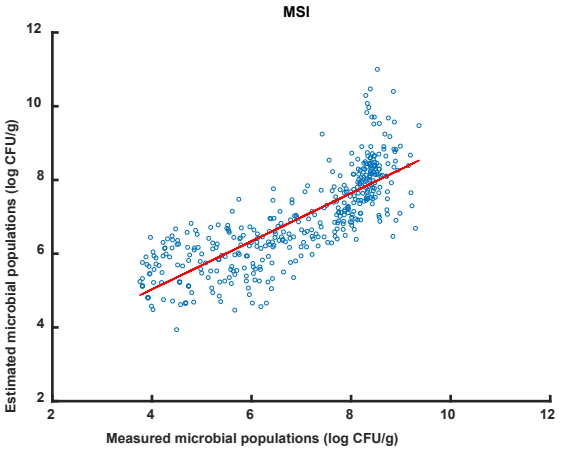


**FIGURE 9. Scatterplot of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on MSI features for data acquired under MAP storage conditions. Red line corresponds to linear fit of the data.**

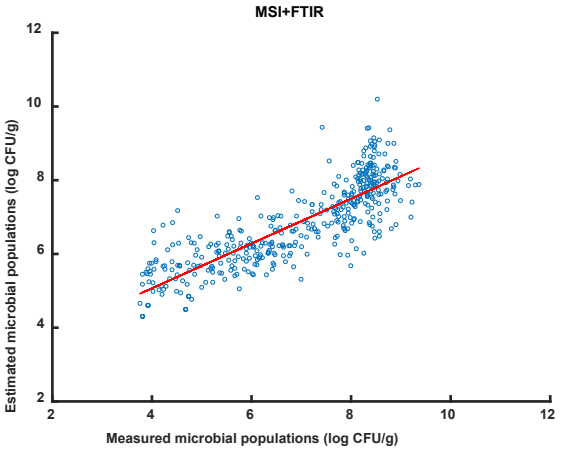


**FIGURE 10. Scatterplot of the measured (via microbiological analysis) and estimated by the SVM regression model microbial populations based on MSI+FTIR features for data acquired under MAP storage conditions. Red line corresponds to linear fit of the data.**

The performance (in terms of RMSE scores) of the four test data sets using different dimensionality reduction methods under MAP storage conditions is shown in Fig. 11. Similarly, to the results observed under AIR conditions, the PCA and PLS methods outperformed all other evaluated dimensionality reduction methods. It can be seen that PCA and PLS dimensionality reduction methods exhibited a similar trend regarding the RMSE scores varying across different components (dimensions) used. It is interesting that ReliefF and Autoencoders presented similar RMSE score curves as well, however both performing significantly worse than PLS and PCA. The highest performance (i.e. the lowest RMSE score value) was observed when 10 latent variables from the PLS dimensionality reduction method were used (see Fig. 11 (c)). RMSE scores obtained using Autoencoders were almost stable, independently of the number of components used.

Although MAP is a common packaging technology for extending the shelf life of meat, the utilization of spectroscopic data for the evaluation of microbiological quality of meat under such storage conditions has not been extensively studied. Argyri et al. [22] reported findings pertinent to the development and evaluation of regression models based on FTIR spectral data and corresponding to beef samples (n=98) stored at 5°C and under different packaging conditions (aerobic and MAP). Different methods (i.e. PLS regression, genetic programming, genetic algorithm, ANN and SVM regression including different kernel functions) were applied, and the minimum RMSE score reported was 0.504 [22]. In another study [55], beef samples (n=186) were stored under aerobic, MAP and active packaging conditions at temperatures from 0 to 15 °C, and an RMSE value of 0.71 was reported for cross-validation of a model correlating microbiological and FTIR spectral data. Similarly promising were the results of a more recent study [56], in which meat samples (n=105) were stored under different packaging conditions (aerobic, MAP), and different instruments (electronic nose, HPLC, GC-MS, FTIR and MSI) and algorithms were evaluated for the prediction of meat spoilage, as this is manifested via the growth of distinct microbial groups.

The rapid assessment of meat quality is of vital importance for industries and authorities in order to ensure quality of perishable food for consumers. Conventional techniques for microbial enumeration are reliable, certified methods and commonly acceptable, but are time-consuming and laborious. For example, the result of probably the most commonly used technique (Standard Plate Count) for the enumeration of microbial populations in food are available after 48-72h. In this context, results are not directly available in order for timely corrective actions to be applied and economic losses and consumer complaints to be minimized. Hence, there is an increasing research interest in rapid and non-invasive spectroscopic techniques.

In the present study, the RMSE scores (expressing in part the performance of the developed models) estimated for FTIR were 1.268 and 1.024 for AIR and MAP storage, respectively, whereas the corresponding values for MSI were 1.144 and 0.923, while for FTIR+MSI these values were 1.146 and 0.886. Since the estimated RMSE values were, in most of the cases, within the “acceptable” microbiological variability (i.e.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



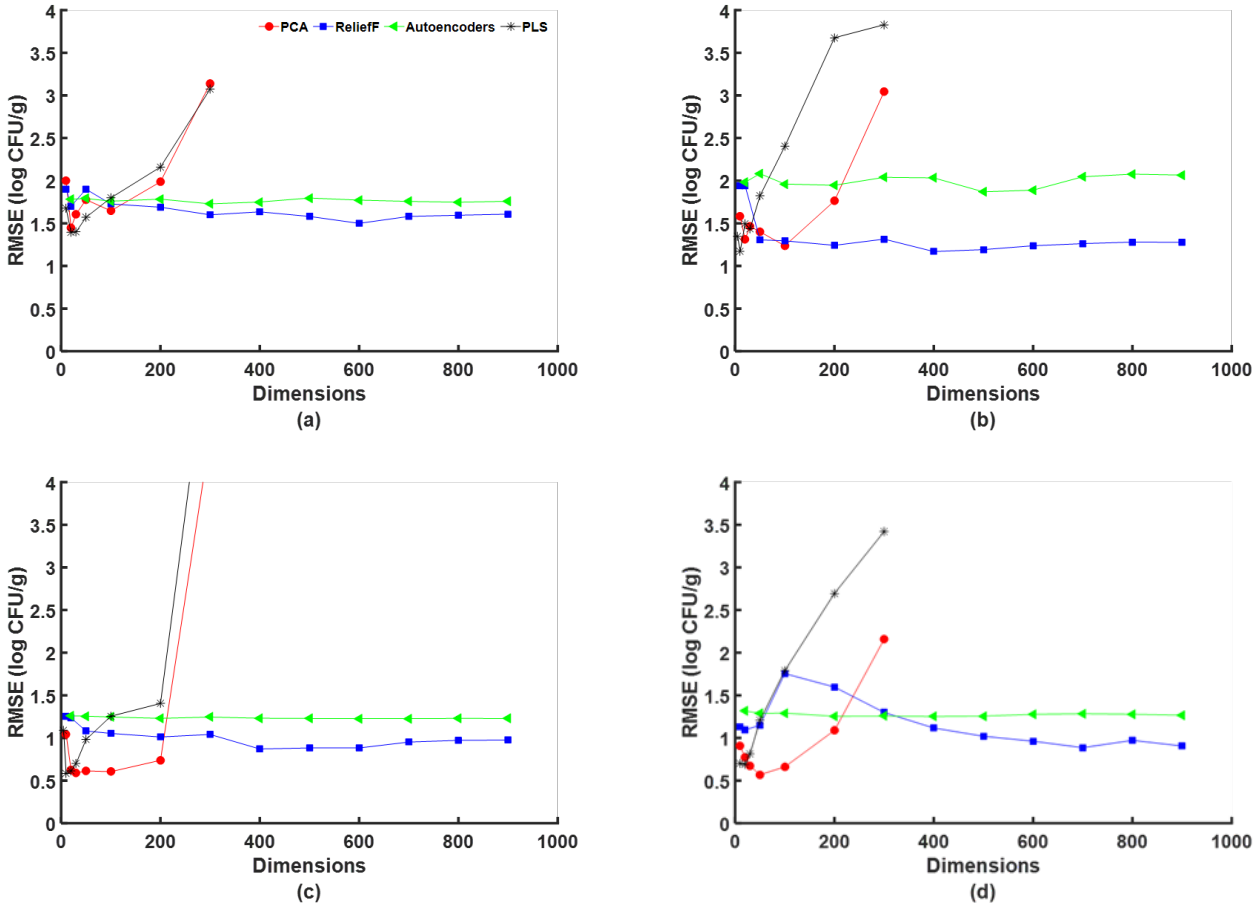


**FIGURE 11. Microbial estimation performance (in terms of RMSE) for the four replicates, 1 (a), 2 (b), 3 (c) and 4 (d), under MAP storage conditions, for different dimensionality reduction methods with respect to the number of components (dimensions) used.**

within the 1-log cfu range commonly encountered in microbiological studies due to the extensive variability characterizing biological systems), the predictive information provided by the developed models can be regarded as rather promising for future application. Indeed, as also commented previously, a sample is considered as "correctly predicted" from a microbiological standpoint when the difference between predicted and observed values is < 1 log cfu [57]. As demonstrated by the data collected in this study, the RMSE scores estimated for the models corresponding to MAP storage were systematically lower than those referring to AIR storage; such difference may be associated with the different organisms dominating in these two packaging conditions. It is also worth mentioning, that the application of a 4-fold cross validation experimental protocol demonstrated the differences among experimental replicates, and how extensive may the encountered biological variability be; for instance, the RMSE values estimated for the models developed in the case of FTIR+MSI and for samples stored under MAP ranged from 0.461 to 1.142 (data not shown).

Beyond biological variability, the variability related to the environmental conditions, and more specifically to storage temperature and atmosphere, was also taken into consideration in this study. Specifically, different temperature profiles, both isothermal and dynamic, were applied during storage of minced pork patties in an attempt to account for the variable temperature conditions likely to be encountered in the farm-to-fork continuum. A wide range of temperatures, varying from well-controlled refrigeration (4°C in this study) to slightly or excessively abusive (8 or 12°C, respectively in this study) temperatures may be observed in the food supply chain, and there is a high likelihood of food commodities being exposed to changing temperature conditions throughout distribution. In this sense, the value of the results of this study is, among others, delineated by the fact that different sources of variability are taken into account in model development, with the resulting models expected to allow for robust and realistic predictions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



## V. CONCLUSION

In conclusion, a methodology for microbial population estimation in meat portions based on non-invasive spectroscopic sensors was presented. The proposed methodology is using FTIR and MSI sensors and the extracted data are processed by machine learning regression models to estimate the microbial population. The methodology can perform estimation of microbial populations in less than few minutes, in contrast to traditional microbial population analysis techniques which require 48-72h as well as food technology expert staff, thus allowing real-time food quality inspection onsite at food industries and giving the potential of inspections even in food retail and service premises (e.g., supermarkets and restaurants).

As demonstrated by the findings of the present study, FTIR and MSI data, as well as early fusion of such spectral data, appear to hold considerable potential for the estimation of the microbiological quality of meat stored under different temperature and packaging conditions. Further investigation with respect to the (inevitably encountered) biological variability and collection of more data to train machine learning models for regression can result in the development of more precise estimation models thus allowing the onsite and real-time assessment of the microbiological quality of meat and meat products.

## REFERENCES


[1] G. J. Nychas, E. Z. Panagou, and F. Mohareb, "Novel approaches for food safety management and communication," *Current Opinion in Food Science*, vol. 12, pp.13-20, Dec. 2016, DOI. 10.1016.j.cofs.2016.06.005.

[2] A. K. Bhunia, "One day to one hour: how quickly can foodborne pathogens be detected?," *Future microbiology*, vol. 9, no. 8, pp. 935-946, Oct. 2014, DOI. 10.2217/fmb.14.61.

[3] A. I. Doulgeraki and G. J. E. Nychas, "Monitoring the succession of the biota grown on a selective medium for pseudomonads during storage of minced beef with molecular-based methods," *Food microbiology*, vol. 34, no. 1, pp. 62-69, May. 2013, DOI. 10.1016/j.fm.2012.11.017.

[4] M. T. Munir, W. Yu, B. R. Young, D. I. Wilson, and, Industrial Information and Control Centre, "The current status of process analytical technologies in the dairy industry," *Trends in Food Science & Technology*, vol. 43, no. 2, pp. 205-218, June 2015, DOI. 10.1016/j.tifs.2015.02.010.

[5] F. van den Berg, C. B. Lyndgaard, K. M. Sørensen, and S. B. Engelsen, "Process analytical technology in the food industry," *Trends in food science & technology*, vol. 31, no. 1, pp. 27-35, May 2013, DOI. 10.1016/j.tifs.2012.04.007.

[6] F. Corvucci, L. Nobili, D. Melucci, and F. V. Grillenzoni, "The discrimination of honey origin using melissopalynology and Raman spectroscopy techniques coupled with multivariate analysis," *Food chemistry*, vol. 169, pp. 297-304, Feb. 2015, DOI. 10.1016.j.foodchem.2014.07.122.

[7] A. I. Ropodi, E. Z. Panagou, and G. J. E. Nychas, "Multispectral imaging (MSI): A promising method for the detection of minced beef adulteration with horsemeat," *Food Control*, vol. 73, pp. 57-63, March 2017, DOI. 10.1016/j.foodcont.2016.05.048.

[8] A. I. Ropodi, E. Z. Panagou and G. J. E. Nychas.. "Rapid detection of frozen-then-thawed minced beef using multispectral imaging and Fourier transform infrared spectroscopy," *Meat science*, vol. 135, pp. 142-147, Jan. 2018, DOI. 10.1016/j.meatsci.2017.09.016.

[9] L. Zhou, C. Zhang, F. Liu, Z. Qiu, and Y. He, "Application of Deep Learning in Food: A Review," *Comprehensive Reviews in Food Science and Food Safety*, Sep. 2019, DOI. 10.1111/1541-4337.12492.

[10] L. C. Fengou, A.Lianou, P. Tsakanikas, E. N. Gkana, E. Z. Panagou, and G. J. Nychas,"Evaluation of Fourier transform infrared spectroscopy and multispectral imaging as means of estimating the microbiological spoilage of farmed sea bream," *Food microbiology*, vol. 79, pp. 27-34, Jun. 2019, DOI. 10.1016.j.fm.2018.10.020.

[11] P. Tsakanikas, L. C. Fengou, E. Manthou, A. Lianou, E. Z. Panagou, and G. J. Nychas, "A unified spectra analysis workflow for the assessment of microbial contamination of ready-to-eat green salads: Comparative study and application of non-invasive sensors," *Computers and electronics in agriculture*, vol. 155, pp. 212-219, Dec. 2018, DOI. 10.1016.j.compag.2018.10.025.

[12] A. I. Ropodi, E. Z. Panagou, and G. J. Nychas, "Data mining derived from food analyses using non-invasive/non-destructive analytical techniques; determination of food authenticity, quality & safety in tandem with computer science disciplines," *Trends in Food Science & Technology*, vol. 50, pp. 11-25, Apr. 2016, DOI. 10.1016.j.tifs.2016.01.011.

[13] J. U. Porep, D. R. Kammerer, and R. Carle, "On-line application of near infrared (NIR) spectroscopy in food production," *Trends in Food Science & Technology,* vol. 46, no. 2, pp. 211-230, Dec. 2015, DOI. 10.1016.j.tifs.2015.10.002.

[14] H. J. He and D. W. Sun, "Microbial evaluation of raw and processed food products by Visible/Infrared, Raman and Fluorescence spectroscopy," *Trends in food science & technology*, vol. 46, no. 2, pp. 199-210, Dec. 2015, DOI. 10.1016.j.tifs.2015.10.004.

[15] L. C. Fengou, E. Spyrelli, A. Lianou, P. Tsakanikas, E. Z. Panagou, and G. J. E. Nychas, "Estimation of Minced Pork Microbiological Spoilage through Fourier Transform Infrared and Visible Spectroscopy and Multispectral Vision Technology," *Foods*, vol. 8, no. 7, pp. 238, July 2019, DOI. 10.3390/foods8070238.

[16] M. A. Tørngren, M. T. Darré, and, L. Kristensen, 2013, August. Optimising the eating quality and shelf life of retail pork cuts using muscle specific three-gas MA-packaging. Presented in 59th International Congress of Meat Science and Technology, Izmir, Turkey.

[17] C. Zweifel, R. Fischer, and R. Stephan, "Microbiological contamination of pig and cattle carcasses in different small-scale Swiss abattoirs," *Meat science*, vol. 78, no. 3, pp. 225-231, Mar. 2008, DOI. 10.1016/j.meatsci.2007.06.025.

[18] R. A. Holley, M. D. Peirson, J. Lam, and K. B. Tan, "Microbial profiles of commercial, vacuum-packaged, fresh pork of normal or short storage life," *International journal of food microbiology*, vol. 97, no. 1, pp. 53-62, Dec. 2004, DOI. 10.1016/j.ijfoodmicro.2004.03.029.

[19] R. H. Dainty and B. M. Mackey, "The relationship between the phenotypic properties of bacteria from chill-stored meat and spoilage processes," *Journal of Applied Bacteriology*, vol. 73, no. s21, pp. viis-172s, Dec. 1992, DOI. 10.1111/j.1365-2672.1992.tb03630.x.

[20] J. C. Ayres, "Temperature relationships and some other characteristics of the microbial flora developing on refrigerated beef," *Journal of* Food Science, vol. 25, no. 1, Jan. 1960, DOI.10.1111/j.1365-2621.1960.tb17930.x.

[21] A. I. Doulgeraki, D. Ercolini, F. Villani, and G. J. Nychas, "Spoilage microbiota associated to the storage of raw meat in different conditions," *International journal of food microbiology*, vol. 157, no. 2, pp. 130-141, Jul. 2012, DOI. 10.1016/j.ijfoodmicro.2012.05.020.

[22] A. A. Argyri, R. M. Jarvis, D. Wedge, Y. Xu, E. Z. Panagou, R. Goodacre, and G. J. Nychas, "A comparison of Raman and FT-IR spectroscopy for the prediction of meat spoilage," *Food Control*, vol. 29, no. 2, pp. 461-470, Feb. 2013, DOI. 10.1016.j.foodcont.2012.05.040.

[23] J. M. Carstensen and J. F. Hansen, "An Apparatus and a Method of Recording an Image of an Object," U.S. Patent 7 119 930, Oct. 10, 2006.

[24] E. Z. Panagou, O. Papadopoulou, J. M. Carstensen, and G. J. Nychas, "Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage," *International journal of food microbiology*, vol. 174, pp. 1-11, Mar. 2014, DOI. 10.1016.j.ijfoodmicro.2013.12.026.

[25] R. G. Brereton, "Consequences of sample size, variable selection, and model validation and optimisation, for predicting classification ability from analytical data," *TrAC Trends in Analytical Chemistry*, vol. 25, no. 11, pp. 1103-1111, Dec. 2006, DOI. 10.1016.j.trac.2006.10.005.

[26] C. Burges, "A tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, June 1998, DOI. 10.1023/A:1009715923555.

[27] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to Platt's SMO Algorithm for SVM Classifier Design," *Neural Computation*, vol. 13, no. 3, pp. 637–649, March 2001, DOI. 10.1162/089976601300014493.

[28] I. T. Jolliffe and and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A:*

*Mathematical, Physical and Engineering Sciences,* vol. 374, no. 2065, Apr. 2016, DOI. 10.1098.rsta.2015.0202.

[29] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics and intelligent laboratory systems*, vol. 58, no. 2, pp. 109-130, Oct. 2001, DOI. 10.1016.S0169-7439(01)00155-1.

[30] R. Wehrens and B. H. Mevik, "The pls package: principal component and partial least squares regression in R," Journal of Statistical Software, vol. 18, no. 2, Jan. 2007, DOI. 10.18637.jss.v018.i02.

[31] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," Machine learning. vol. 53, no. 1-2, pp. 23-26, Oct. 2003, DOI. 10.1023/A:1025667309714.

[32] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232-242, April, 2016, DOI. 10.1016/j.neucom.2015.08.104.

[33] J. Wang, H. He and D. V. Prokhorov, "A Folded Neural Network Autoencoder for Dimensionality Reduction", *Procedia Computer Science*, vol. 13, pp. 120-127, 2012, DOI. 10.1016/j.procs.2012.09.120.

[34] B. A. Olshausen and D. J. Field. "Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1," *Vision Research*, vol. 37, no. 23, Dec. 1997, pp. 3311–3325, DOI. 10.1016/S0042-6989(97)00169-7.

[35] M. F. Moller, "A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning," *Neural Networks*, vol. 6, 1993, pp. 525–533, DOI. 10.1016/s0893-6080(05)80056-5.

[36] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273-297, Sep. 1995, DOI. 10.1007/BF00994018.

[37] MATLAB, Release 2019a, The MathWorks, Inc., Natick, Massachusetts, United States

[38] R, R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[39] RStudio, Rstudio Team (2016). Rstudio: Integrated Development for R. Rstudio, Inc., Boston, MA URL http://www.rstudio.com/.

[40] E. Frank, M. A. Hall, and I. H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

[41] D. I. Ellis, D. Broadhurst, and R. Goodacre, "Rapid and quantitative detection of the microbial spoilage of beef by Fourier transform infrared spectroscopy and machine learning," *Analytica Chimica Acta*, vol. 514, no. 2, pp. 193-201, Jul. 2004, DOI. 10.1016/j.aca.2004.03.060.

[42] O. Al-Jowder, M. Defernez, E. K. Kemsley, and R. H. Wilson, "Mid-infrared spectroscopy and chemometrics for the authentication of meat products," *Journal of agricultural and food chemistry*, vol. 47, no. 8, pp. 3210-3218, Jul. 1999, DOI. 10.1021/jf981196d.

[43] F. Tao, Y. Peng, C. L. Gomes, K. Chao, and J. Qin, "A comparative study for improving prediction of total viable count in beef based on hyperspectral scattering characteristics," *Journal of Food Engineering*, vol. *162*, pp. 38-47, Oct. 2015, DOI. 10.1016/j.jfoodeng.2015.04.008.

[44] D. F. Barbin, G. ElMasry, D. W. Sun, and P. Allen, "Non-destructive determination of chemical composition in intact and minced pork using near-infrared hyperspectral imaging," *Food chemistry*, vol. *138*, no. 2-3, pp. 1162-1171, Jun., 2013, DOI. 10.1016/j.foodchem.2012.11.120.

[45] O. Papadopoulou, E. Z. Panagou, C. C.Tassou, and G. J. Nychas, "Contribution of Fourier transform infrared (FTIR) spectroscopy data on the quantitative determination of minced pork meat spoilage," *Food Research International,* vol. 44, no. 10, pp. 3264-3271, Dec., 2011, DOI. 10.1016/j.foodres.2011.09.012.

[46] E. Z. Panagou, F. R. Mohareb, A. A.Argyri, C. M. Bessant, and G. J. Nychas, "A comparison of artificial neural networks and partial least squares modelling for the rapid detection of the microbial spoilage of beef fillets based on Fourier transform infrared spectral fingerprints," *Food Microbiology,* vol. 28, no. 4, pp. 782-790, Jun. 2011, DOI. 10.1016/j.fm.2010.05.014.

[47] M. J. P. Moreira, A. C. Silva, J. M. de Almeida, and C. Saraiva, "Characterization of deterioration of fallow deer and goat meat using microbial and mid infrared spectroscopy in tandem with chemometrics," *Food Packaging and Shelf Life*, vol. 15, pp. 169-180, Mar. 2018, DOI. 10.1016/j.fpsl.2018.01.010.

[48] D. Alexandrakis, G. Downey, and A. G. Scannell, "Rapid non-destructive detection of spoilage of intact chicken breast muscle using near-infrared and Fourier transform mid-infrared spectroscopy and multivariate statistics," *Food and Bioprocess Technology*, vol. 5, no. 1, pp. 338-347, Jan. 2012, DOI. 10.1007/s11947-009-0298-4.

[49] . Tsakanikas, D. Pavlidis, E. Panagou, and G. J. Nychas, "Exploiting multispectral imaging for non-invasive contamination assessment and mapping of meat samples," *Talanta*, vol. 161, pp. 606-614, Dec. 2016, DOI. 10.1016/j.talanta.2016.09.019.

[50] B. S. Dissing, O. S. Papadopoulou, C. Tassou, B. K. Ersbøll, J. M. Carstensen, E. Z.Panagou, and G. J. Nychas, "Using multispectral imaging for spoilage detection of pork meat," *Food and Bioprocess Technology,* vol. 6, no.9, pp. 2268-2279, Sep. 2013, DOI. 10.1007/s11947-012-0886-6.

[51] C. Márquez, M. I. López, I. Ruisánchez, and M. P. Callao, "FT-Raman and NIR spectroscopy data fusion strategy for multivariate qualitative analysis of food fraud," *Talanta*, vol. 161, pp. 80-86, Dec. 2016, DOI. 10.1016/j.talanta.2016.08.003.

[52] E. Borràs, J. Ferré, R. Boqué, M. Mestres, L. Aceña, and, O. Busto, "Data fusion methodologies for food and beverage authentication and quality assessment–A review," *Analytica Chimica Acta*, vol. 891, pp. 1-14, Sep. 2015, DOI. doi.org/10.1016/j.aca.2015.04.042.

[53] F. Y. Kutsanedzie, Z. Guo, and Q. Chen, "Advances in Nondestructive Methods for Meat Quality and Safety Monitoring," *Food Reviews International*, vol. 35, no.6, pp. 536-562, Mar. 2019, DOI. 10.1080/87559129.2019.1584814.

[54] F. Qu, D. Ren, Y. He, P. Nie, L. Lin, C. Cai, and T. Dong, "Predicting pork freshness using multi-index statistical information fusion method based on near infrared spectroscopy," *Meat science*, vol. 146, pp. 59-67, Dec. 2018, DOI. 10.1016/j.meatsci.2018.07.023.

[55] M. S. Ammor, A. Argyri, and G. J. E. Nychas. "Rapid monitoring of the spoilage of minced beef stored under conventionally and active packaging conditions using Fourier transform infrared spectroscopy in tandem with chemometrics," *Meat Science*, vol. 81, no. 3, pp. 507-514, March 2009, DOI. 10.1016/j.meatsci.2008.10.015.

[56] L. Estelles-Lopez, A. Ropodi, D. Pavlidis, J. Fotopoulou, C. Gkousari, A. Peyrodie, E. Panagou, G. J. Nychas, and F. Mohareb. "An automated ranking platform for machine learning regression models for meat spoilage prediction using multi-spectral imaging and metabolic profiling," *Food research international,* vol. 99, pp. 206-215, Sep. 2017, DOI. 10.1016/j.foodres.2017.05.013.

[57] F. Mohareb, O. Papadopoulou, E. Panagou, G. J. Nychas and, C. Bessant, "Ensemble-based support vector machine classifiers as an efficient tool for quality assessment of beef fillets from electronic nose data," *Analytical Methods*, vol. 8, no. 18, pp. 3711-3721, April 2016, DOI. 10.1039/C6AY00147E.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3000690, IEEE Access



**FENGOU LEMONIA-CHRISTINA** received the integrated master in Food Science and Human Nutrition, School of Food, Biotechnology, and Development, from the Agricultural University of Athens, Athens, Greece, in 2016. She is currently pursuing the Ph.D. degree in food microbiology, in the Laboratory of Microbiology and Biotechnology of Foods at the aforementioned University Department, with her research activity being oriented towards the assessment of spectroscopic methods coupled with data analysis and machine learning methodologies as means of evaluating food spoilage and adulteration.

**IOSIF MPORAS** received his Diploma and PhD degree in Electrical and Computer Engineering from the University of Patras, Greece in 2004 and 2009, respectively. Since 2016 he holds the post of Senior Lecturer in Information Engineering at the University of Hertfordshire, UK. Before that he was Adjunct Assistant Professor at TEI of Western Greece and postdoctoral researcher at the University of Patras. His research interests include signal processing and machine learning. Dr Mporas has participated in more than 10 European R&D projects as Researcher, Senior Researcher and Principal Investigator. He has served as Chair and PC member in international conferences and workshops and as Guest Editor in special issues of journals. Dr Iosif Mporas has published more than 100 papers in international conferences and journals and his papers have been cited more than 900 times.

**EVGENIA SPYRELLI** received the B.S. degree in chemical engineering from Aristotle University, Thessaloniki, Greece, in 2015 and the M.S. degree in food science and technology & human nutrition from Agricultural University, Athens, Greece, in 2018. She is currently pursuing the Ph.D. degree in microbiology and biotechnology in foods at Agricultural University, Athens, Greece.
From 2018 until now, she is working as a PhD student at the laboratory of microbiology and biotechnology in foods in Agricultural University of Athens. Her research interest includes the implementation of rapid, no destructive spectroscopic methods on foods in tandem with microbiological experiments and performance of machine learning algorithms for quality and/or quantitative assessment in foods.

**ALEXANDRA LIANOU** holds a B.Sc. degree in Agricultural Biotechnology (Agricultural University of Athens, Greece), a M.Sc. degree in Animal Sciences (Meat Science/Food Safety program, Colorado State University, USA) and a Ph.D. degree in Food Microbiology (Aristotle University of Thessaloniki, Greece).
She is working as a senior research associate at the Agricultural University of Athens (Department of Food Science and Human Nutrition, Laboratory of Microbiology and Biotechnology of Foods). Her research interests are related to the microbiological safety and quality of foods, and more specifically to the development and application of quantitative microbiology approaches in these areas as well as to the implementation of rapid, non-invasive spectroscopy-based methods in the evaluation of food spoilage. Up to now, Dr. Lianou has participated in 10 competitive research projects (eight international and two Greek national projects), covering a wide range of safety/quality issues of various foods including meat/poultry and their products, produce commodities, dairy products and fish. During her research activity, she has been author/co-author in 35 articles published in peer-reviewed international journals, in six international publishers' book chapters and in several abstracts/papers published in the proceedings of international and Greek national conferences.

**GEORGE-JOHN NYCHAS** was born in Athens, Greece in November 1954. He holds a BSc in Food Science and PhD in Food Microbiology (University of Bath, UK, 1984). Since 1994 is the head of the Laboratory of Microbiology and Biotechnology of Foods (LMBF) of AUA. From March 2010 till July 2011 he was the President of the Greek National Food Authority (www.efet.gr). He was member of the EFSA – Biohazard group, expert in Predictive modelling / Quantitative Risk Assessment (QRA) while recently he was member of the ERWG of EFSA (2012-2015). He has coordinated 4 projects and participated in more than 20 EU projects and 10-concerted Action.
Prof Nychas through these projects, his team has acquired extensive experience on; a) the assessment of food safety and spoilage through microbiological analysis (pathogens spoilage organisms), and physicochemical analysis (metabolomics) in combination with machine learning, b) on responses of stress adapted pathogens grown planktonically or attached on stainless steel surfaces (biofilms), and c) on modelling the behaviour of microbial populations throughout the food chain to assist reliable estimation of microbial food safety risk. d) Implementation of *Process analytical technology (PAT)* in Food Industry. Prof Nychas has published 269 papers (Scopus, March 2020) with *ca. 13000 citations and h=65.*