engineering proceedings

Proceeding Paper

Multivariate Analysis in Accelerated Shelf-life Assessment– An Overview †

Laíres Lima [1,2,3,4,*], Ana I. Pereira [2,4], Manuela Pintado [3], Márcio Carocho [1,4], Lillian Barros [1,4]

[1] Mountain Research Centre (CIMO), Polytechnic Institute of Bragança, Bragança, Portugal
[2] Research Center in Digitalization and Intelligent Robotics (CeDRI), Polytechnic Institute of Bragança, Bragança, Portugal
[3] Centre of Biotechnology and Fine Chemistry (CBQF), School of Biotechnology, Universidade Católica Portuguesa, Porto, Portugal
[4] Associate Laboratory for Sustainability and Technology in Mountains Regions (SusTEC), Polytechnic Institute of Bragança, Bragança, Portugal
* Correspondence: laires.lima@ipb.pt

† Presented at the 4th International Electronic Conference on Applied Sciences, 27 October–10 November 2023; Available online: https://asec2023.sciforum.net/.

Abstract: To meet the market demand for high-quality products, researchers and manufacturers have invested in the development of accurate methods for estimating shelf-life. Tests that consider the simultaneous effects of different parameters on food degradation are useful tools in shelf-life studies, as these parameters can directly influence quality and safety. With this in mind, the objective of this review is to gather pertinent information from recent studies (2006-2022) pertaining to multivariate analysis applied in accelerated shelf-life tests in order to facilitate a comprehensive understanding. The review focuses on multivariate techniques commonly employed in accelerated shelf-life modeling, namely, principal components analysis, partial least squares regression, orthogonal projections to latent structures discriminant analysis, and hierarchical cluster analysis. Through an extensive literature review, the collected data represent the evolution of these methods, taking into account current trends, advances in food shelf-life techniques, and future perspectives. It was observed that the recent literature provides limited information on the determination of shelf-life under multiple accelerated factors. However, the studies analyzed showed that multivariate analysis can be a useful tool in the interpretation of quality characteristics and can accurately predict the shelf-life of foods compared to univariate kinetic procedures. Multivariate statistical methods addressed in this work are presented as a promising method for foods tested, being applied together with different chemometric techniques. This comprehensive review contributes to the body of knowledge surrounding accelerated shelf-life testing, offering valuable insights for researchers, manufacturers, and stakeholders in the food industry.

Keywords: multivariate analysis; accelerate shelf-life; principal component analysis

Citation: Lima, L.; Pereira, A.I.; Pintado, M.; Carocho, M.; Barros, L. Multivariate Analysis in Accelerated Shelf-life Assessment– An Overview. *Eng. Proc.* **2023**, *52*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Received: date
Revised: date
Accepted: date
Published: date

1. Introduction

Shelf-life is an intrinsic property of any food product and is defined as the finite period after production during which the food product maintains a regulated level of quality and safety under well-defined storage conditions [1]. The main objective of accurate shelf-life measurement is to ensure that products are safe for consumption and of acceptable quality to meet consumer expectations [2].

Traditionally, real-time shelf-life tests (RSLTs) have been used to monitor quality changes over time, simulating the conditions experienced by the product on the shelf [1]. Although effective, RSLT can be time-consuming and arduous, especially for products with long durability. To address this challenge, researchers have turned to accelerated shelf-life tests (ALST), which subject the product to critical storage conditions for shorter durations. By leveraging data collected under these accelerated conditions, the results can be converted to reflect real-time storage scenarios [3].

The definition of the properties to be monitored, the storage conditions and the cut-off criteria are necessary for the correct performance of the two shelf-life tests mentioned. Several variables can affect sensory quality and food safety characteristics [1,3]. Therefore, multivariate analysis techniques are useful tools for shelf-life studies where different quality parameters must be monitored simultaneously.

The aim of this article is to present a comprehensive review of the application of multivariate analysis in determining the shelf-life of food products over the past 15 years. Through this investigation, we shed light on the advancements made in shelf-life evaluation and emphasize the relevance of these techniques. The findings presented here serve as a proof of concept concerning multivariate analysis applied to shelf-life tests, providing a basis for further exploration and utilization of this method in future research endeavors.

2. Multivariate analysis in ASLT

Shelf-time tests that involve complex foods can comprise a large amount of discrete data. This is because the quality attributes monitored may vary with the passage of time and changes in temperature. To deal with the large amount of data, researchers usually resort to multivariate analysis techniques for data compression, since the variables studied in the experiment may be associated with the same physical, biochemical, and microbiological phenomena, among others [4].

Multivariate analyzes have been studied since the beginning of the 20th century [5,6], but were incorporated into computational methods only in the 1970s [7,8]. Since then, multivariate techniques have been applied to solve large problems with several variables in different areas [4]. Multivariate approaches basically comprise two stages. The first step consists of data compression by reducing the complexity of the original data. Then, a modeling step is performed to select the most significant features. Multivariate analyzes applied to ASLTs are based on the collection of sample data under different storage conditions to understand the nature and types of deteriorative events that may occur and identify their corresponding indicators [9].

This section emphasizes the multivariate analysis techniques that are usually applied in accelerate shelf-life modeling, namely Principal Components Analysis (PCA) [10,11], Partial Least Squares Regression (PLSR) [12–14], Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) [15], and Hierarchical Cluster Analysis (HCA) [15,16].

2.1. Principal Components Analysis (PCA)

As most food products have quality parameters that change simultaneously over time, it is recommended to use multivariate techniques such as MASLT for a more accurate assessment, which is commonly associated with the PCA technique for a better selection of the most important parameters in degrading behavior [5].

In the context of shelf-life assessment, the first step consists of constructing matrices that represent the variability of quality parameters collected during the experiment under different conditions of temperatures T and storage times t. The arrangement of these matrices allows the construction of a multivariate space that explains the dependence of quality parameters in relation to time and temperature [5].

In cases where the quality attributes of the multivariate space have different scales, the matrices must be normalized by centering the mean of the data in order to standardize the scale.

After structuring the dataset, the PCA analysis is performed to identify the PCs, loadings and scores. Correlating this information with time, kinetic graphs (variation of PC scores during storage time) are obtained, and the constants and orders of the degradation reactions involved are determined. The loading matrix can be applied to calculate the cut-off criteria, corresponding to maximum acceptable scores for each time-related PC. Subsequently, the temperature dependence of the rate constants is defined using kinetic models (*e.g.* Arrhenius equation) and the lifetime is then calculated from the extrapolation of the accelerated data. This methodology is briefly described in Fig. 2.

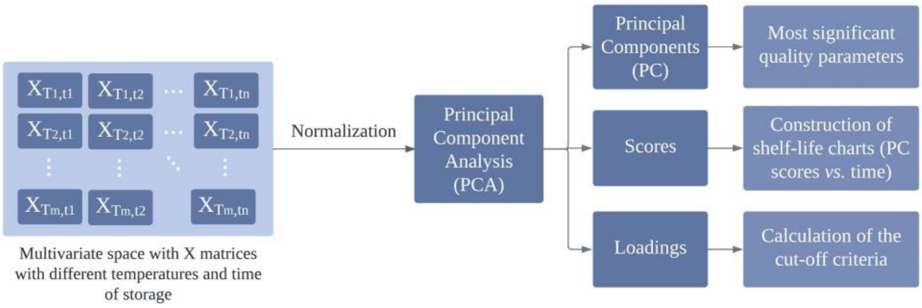

Figure 1. MASLT algorithm for shelf-life estimation. Adapted from [10].

Thenceforth, shelf-life methodologies that use PCA proved to be efficient in identifying and selecting attributes that directly influence food degradation [11,17–27].

2.2. Partial Least Squares Regression (PLSR)

Partial Least Squares Regression (PLSR) is a multivariate statistical data analysis approach that combines the benefits of PCA and regression analysis, which can establish predictive models with high accuracy when there are linear correlations between variables [29].

In its simplest form, PLSR creates vectors of orthogonal scores by maximizing the covariance between different sets of variables (predictors and responses) [30]. This technique is very useful when there is a large number of independent variables that can influence a characteristic of the food matrix. Because of this, this method has been used in several studies to evaluate the impact of different features (temperature, pH, light, microbiological growth, among others) on the shelf-life of food products and in the development of regression models that simulate food degradation [12,28–31].

The reviewed studies emphasize the importance of sensory evaluation and instrumental analysis in predicting food quality and shelf-life. The use of statistical methods such as partial least squares regression (PLSR) allows for the identification of key factors influencing spoilage and establishing relationships between sensory attributes and instrumental measurements.

2.3. Orthogonal Projections to Latent Structures - Discriminant Analysis (OPLS-DA)

Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) is a modification of the Nonlinear Iterative Partial Least Squares (NIPALS) algorithm applied to PLS. The development of OPLS was driven by the large amount of uncorrelated variation in data sets, particularly in multivariate sets. This method can be seen as a pure preprocessing method to remove systematic orthogonal variation from a data set X, or it can be an integrated part of the PLS method to provide simpler and easier to interpret models [32].

In shelf-life studies, OPLS-DA is used as a supervised analysis technique to identify significant patterns and visualize metabolic changes under different storage and processing conditions [33–35]. OPLS-DA has emerged as a valuable technique in shelf-life studies, particularly for identifying significant patterns under various storage and processing conditions.

2.4. Hierarchical Cluster Analysis (HCA)

Hierarchical cluster analysis is a mathematical method that treats each sample as a point in a multidimensional space based on selected variables. To calculate the proximity between the samples, the distance between each point and all other points is determined, resulting in a matrix that describes the proximity between all the samples studied. The Euclidean distance is commonly used to calculate this distance between two points [15]. Based on this proximity matrix, a similarity dendrogram can be constructed, showing the similarity between samples [36].

There are several specific algorithms to group the points in multidimensional space and form hierarchical clusters based on proximity matrix information. The interpretation of the similarity dendrogram is intuitive: close samples must have similar values for the measured variables and must be mathematically close to each other in the multidimensional space. The greater the proximity of the measurements between the samples, the greater the similarity between them [36]. The dendrogram establishes a hierarchy in this similarity, allowing for a two-dimensional visualization of the similarity or dissimilarity of the entire set of samples. However, when the variables have different scales, as in the case of food analysis, data pre-processing is necessary. The Z transformation is commonly used to equalize the statistical importance of all variables by transforming the measures of each variable into a mean of zero and a variance of one [37]. This allows for a more accurate analysis of the data.

Among the reviewed studies, it is noted that the use of HCA is often associated with other multivariate methods for pattern recognition and classification purposes, allowing the identification of groups or clusters within a dataset based on similarities or differences between samples [16,38].

3. Potentials and limitations

Since it is a recent technique, the application of multivariate analysis in shelf-life studies is still very simple and is often performed in comparison to the already established techniques. The shelf-life estimates obtained with MASLT (using PCA, HCA, PLS and derivatives) have been shown to be more reliable than ASLT or RSLT, but the choice of multivariate method is limited by the difficulty of performing specific analyzes and by the collection of data involving more experimental points. Nevertheless, the technique should not be underestimated: In all studies analyzed in this review, the multivariate approach provided conclusive, accurate, and more complete results. In addition, MASLT allows modifications to the resources used, such as the inclusion of additional steps depending on the type of food and constraints in PCA, to optimize the selection of PCs. Increased use of MASLT in food shelf-life prediction seems inevitable, as more quality parameters are considered in these studies as the complexity of the composition and the quality and safety guidelines increase.

4. Concluding remarks and future perspectives

Multivariate shelf-life analysis takes a more holistic approach and considers the cumulative effect of deterioration caused by multiple factors. Recent literature provides limited information on the determination of shelf-life under multiple accelerated factors, but the studies evaluated in this work showed that multivariate analysis can be a useful tool in the interpretation of quality characteristics and can accurately predict shelf-life of foods compared to univariate kinetic procedures in comparative studies. The multivariate statistical methods addressed in this work are presented as promising for the foods tested, being applied together with different chemometric techniques. However, in the literature, no complete standard methodology for all food products or at least for broader food groups was found.

Although some authors have explored different quality limiting parameters, the temperature is still the most used for this purpose, as temperature variations directly interfere with degradation processes in real situations. As for the most used model, the Weibullian distribution and the Arrhenius equation have been the most applied, considering kinetics of different orders and in relation to the specifications of the food product. Developing multivariate predictive models to determine shelf-life is certainly an arduous task. On the other hand, it leads to a reduction in the time and cost needed to estimate shelf-life.

According to the panorama perceived in the literature review, future trends point to the practical implementation of promising chemometric approaches combined with computational techniques and statistical analysis of multivariate data.

Funding: This research was funded by FCT- Foundation for Science and Technology, P.I., (FCT, Portugal) through national funds FCT/MCTES to CIMO (UIDB/00690/2020 and UIDP/00690/2020), CeDRI (UIDB/00319/2020 and UIDB/05757/2020), SusTEC (LA/P/0007/2021) and UIDB/50006/2020; national funding by FCT, P.I., through the institutional and individual scientific employment program-contract for L.B. and M.C. contracts. Fundação para a Ciência e Tecnologia (FCT) supported this work through Ph.D. grant (2022.13393.BD) for L.L.

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

References

1. Corradini, M.G.; Peleg, M. Shelf-Life Estimation from Accelerated Storage Data. *Trends Food Sci Technol* 2007, *18*, 37–47.
2. Labuza, T.P. *Shelf-Life Dating of Foods*; Food & Nutrition Press: Westport, UK, 1982;
3. Martín-Torres, S.; Ruiz-Castro, L.; Jiménez-Carvelo, A.M.; Cuadros-Rodríguez, L. Applications of Multivariate Data Analysis in Shelf Life Studies of Edible Vegetal Oils – A Review of the Few Past Years. *Food Packag Shelf Life* 2022, *31*.
4. Kemsley, E.K.; Defernez, M.; Marini, F. Multivariate Statistics: Considerations and Confidences in Food Authenticity Problems. *Food Control* 2019, *105*, 102–112.
5. Wishart, J. The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population. *Biometrika* **1928**, *20A*, 32–52, doi:10.1093/biomet/20A.1-2.32.
6. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J Educ Psychol* **1933**, *24*, 417–441, doi:10.1037/h0071325.
7. Chambers, J.M. Computers in Statistical Research: Simulation and Computer-Aided Mathematics. *Technometrics* **1970**, *12*, 1, doi:10.2307/1267347.
8. Lewis, P.A.W. *Large-Scale Computer-Aided Statistical Mathematics*; Monterey, California, 1972;
9. Nicoli, M.C. An Introduction to Food Shelf Life: Denitions, Basic Concepts, and Regulatory Aspects. In *Shelf Life Assessment of Food*; CRC Press, 2012; Vol. 1.
10. Pedro, A.M.K.; Ferreira, M.M.C. Multivariate Accelerated Shelf-Life Testing: A Novel Approach for Determining the Shelf-Life of Foods. *J Chemom* **2006**, *20*, 76–83, doi:10.1002/cem.995.
11. Córdova, A.; Quezada, C.; Saavedra, J. A MASLT Method Comparison over Univariate Kinetic Modeling for Determination of Shelf-Life in Cereal Snack of Dried Apples. *Procedia Food Sci* **2011**, *1*, 1045–1050, doi:10.1016/j.profoo.2011.09.156.
12. Hejri-Zarifi, S.; Ahmadian-Kouchaksaraei, Z.; Pourfarzad, A.; Khodaparast, M.H.H. Dough Performance, Quality and Shelf Life of Flat Bread Supplemented with Fractions of Germinated Date Seed. *J Food Sci Technol* **2014**, *51*, 3776–3784, doi:10.1007/s13197-013-0929-7.
13. Wibowo, S.; Buvé, C.; Hendrickx, M.; Van Loey, A.; Grauwet, T. Integrated Science-Based Approach to Study Quality Changes of Shelf-Stable Food Products during Storage: A Proof of Concept on Orange and Mango Juices. *Trends Food Sci Technol* **2018**, *73*, 76–86, doi:10.1016/j.tifs.2018.01.006.
14. Kebede, B.T.; Grauwet, T.; Magpusao, J.; Palmers, S.; Michiels, C.; Hendrickx, M.; Van Loey, A. Chemical Changes of Thermally Sterilized Broccoli Puree during Shelf-Life: Investigation of the Volatile Fraction by Fingerprinting-Kinetics. *Food Research International* **2015**, *67*, 264–271, doi:10.1016/j.foodres.2014.10.017.

15. Patras, A.; Brunton, N.P.; Downey, G.; Rawson, A.; Warriner, K.; Gernigon, G. Application of Principal Component and Hierarchical Cluster Analysis to Classify Fruits and Vegetables Commonly Consumed in Ireland Based on in Vitro Antioxidant Activity. *Journal of Food Composition and Analysis* **2011**, *24*, 250–256, doi:10.1016/j.jfca.2010.09.012.
16. Elhadef, K.; Smaoui, S.; ben Hlima, H.; Ennouri, K.; Fourati, M.; Chakchouk Mtibaa, A.; Ennouri, M.; Mellouli, L. Effects of Ephedra Alata Extract on the Quality of Minced Beef Meat during Refrigerated Storage: A Chemometric Approach. *Meat Sci* **2020**, *170*, doi:10.1016/j.meatsci.2020.108246.
17. Saavedra, J.; Córdova, A.; Gálvez, L.; Quezada, C.; Navarro, R. Principal Component Analysis as an Exploration Tool for Kinetic Modeling of Food Quality: A Case Study of a Dried Apple Cluster Snack. *J Food Eng* **2013**, *119*, 229–235, doi:https://doi.org/10.1016/j.jfoodeng.2013.05.036.
18. Techer, C.; Baron, F.; Delbrassinne, L.; Belaïd, R.; Brunet, N.; Gillard, A.; Gonnet, F.; Cochet, M.F.; Grosset, N.; Gautier, M.; et al. Global Overview of the Risk Linked to the Bacillus Cereus Group in the Egg Product Industry: Identification of Food Safety and Food Spoilage Markers. *J Appl Microbiol* **2014**, *116*, 1344–1358, doi:10.1111/jam.12462.
19. Richards, M.; De Kock, H.L.; Buys, E.M. Multivariate Accelerated Shelf-Life Test of Low Fat UHT Milk. *Int Dairy J* **2014**, *36*, 38–45, doi:https://doi.org/10.1016/j.idairyj.2013.12.012.
20. Richards, M.; Buys, E.M.; De Kock, H.L. Survival Analysis, Consumer Perception and Physico-Chemical Analysis of Low Fat UHT Milk Stored for Different Time Periods. *Int Dairy J* **2016**, *57*, 56–61, doi:https://doi.org/10.1016/j.idairyj.2016.02.037.
21. Derossi, A.; Mastrandrea, L.; Amodio, M.L.; de Chiara, M.L. V; Colelli, G. Application of Multivariate Accelerated Test for the Shelf Life Estimation of Fresh-Cut Lettuce. *J Food Eng* **2016**, *169*, 122–130, doi:https://doi.org/10.1016/j.jfoodeng.2015.08.010.
22. Kumar, C.T.M.; Raju, P.N.; Singh, A.K. Effect of Packaging Materials and Storage Temperatures on Shelf Life of Micronutrient Fortified Milk-Cereal Based Complementary Food. *J Packag Technol Res* **2017**, *1*, 135–148, doi:10.1007/s41783-017-0020-3.
23. Giménez, A.; Gagliardi, A.; Ares, G. Estimation of Failure Criteria in Multivariate Sensory Shelf Life Testing Using Survival Analysis. *Food Research International* **2017**, *99*, 542–549, doi:10.1016/j.foodres.2017.04.031.
24. Sehwag, S.; Upadhyay, R.; Das, M. Optimization and Multivariate Accelerated Shelf Life Testing (MASLT) of a Low Glycemic Whole Jamun (Syzygium Cumini L.) Confection with Tailored Quality and Functional Attributes. *J Food Sci Technol* **2018**, *55*, 4887–4900, doi:10.1007/s13197-018-3423-4.
25. Cruz-Tirado, J.P.; Oliveira, M.; de Jesus Filho, M.; Godoy, H.T.; Amigo, J.M.; Barbin, D.F. Shelf Life Estimation and Kinetic Degradation Modeling of Chia Seeds (Salvia Hispanica) Using Principal Component Analysis Based on NIR-Hyperspectral Imaging. *Food Control* **2021**, *123*, 107777, doi:https://doi.org/10.1016/j.foodcont.2020.107777.
26. Jiang, Y.; Yang, X.; Jin, H.; Feng, X.; Tian, F.; Song, Y.; Ren, Y.; Man, C.; Zhang, W. Shelf-Life Prediction and Chemical Characteristics Analysis of Milk Formula during Storage. *LWT* **2021**, *144*, 111268, doi:https://doi.org/10.1016/j.lwt.2021.111268.
27. Li, Y.; Ding, S.; Wang, Y. Shelf Life Predictive Model for Postharvest Shiitake Mushrooms. *J Food Eng* **2022**, *330*, 111099, doi:https://doi.org/10.1016/j.jfoodeng.2022.111099.
28. Blixt, Y.; Borch, E. Comparison of Shelf Life of Vacuum-Packed Pork and Beef. *Meat Sci* **2002**, *60*, 371–378, doi:10.1016/S0309-1740(01)00145-0.
29. Zhang, Y.; Zhu, D.; Ren, X.; Shen, Y.; Cao, X.; Liu, H.; Li, J. Quality Changes and Shelf-Life Prediction Model of Postharvest Apples Using Partial Least Squares and Artificial Neural Network Analysis. *Food Chem* **2022**, *394*, 133526, doi:10.1016/j.foodchem.2022.133526.
30. Chaudhry, M.M.A.; Amodio, M.L.; Babellahi, F.; de Chiara, M.L. V; Amigo Rubio, J.M.; Colelli, G. Hyperspectral Imaging and Multivariate Accelerated Shelf Life Testing (MASLT) Approach for Determining Shelf Life of Rocket Leaves. *J Food Eng* **2018**, *238*, 122–133, doi:https://doi.org/10.1016/j.jfoodeng.2018.06.017.

31. Mohammad, H.N.B.; Pourfarzad, A.; Zahedi, H.; Ahmadian-Kouchaksaraie, Z.; Haddad Khodaparast, M.H. Development of Sourdough Fermented Date Seed for Improving the Quality and Shelf Life of Flat Bread: Study with Univariate and Multivariate Analyses. *J Food Sci Technol* **2015**, *53*, 209–220, doi:10.1007/s13197-015-1956-3.
32. Bylesjö, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J.K.; Holmes, E.; Trygg, J. OPLS Discriminant Analysis: Combining the Strengths of PLS-DA and SIMCA Classification. *J Chemom* **2006**, *20*, 341–351, doi:10.1002/cem.1006.
33. Lee, D.H.; Kang, B.S.; Park, H.J. Effect of Oxygen on Volatile and Sensory Characteristics of Cabernet Sauvignon during Secondary Shelf Life. *J Agric Food Chem* **2011**, *59*, 11657–11666, doi:10.1021/jf200759d.
34. Rocchetti, G.; Bernardo, L.; Pateiro, M.; Barba, F.J.; Munekata, P.E.S.; Trevisan, M.; Lorenzo, J.M.; Lucini, L. Impact of a Pitanga Leaf Extract to Prevent Lipid Oxidation Processes during Shelf Life of Packaged Pork Burgers: An Untargeted Metabolomic Approach. *Foods* **2020**, *9*, 1668, doi:10.3390/foods9111668.
35. Rangel-Huerta, O.D.; Uhlig, S.; Ivanova, L.; Dang, T.T.; Rode, T.M.; Noriega Fernández, E.; Fæste, C.K. Metabolomics Workflow for Quality Control of Differently-Processed Pre-Cooked Chicken Fillets. *Food Chem* **2022**, *370*, 131006, doi:10.1016/j.foodchem.2021.131006.
36. Schonlau, M. Visualizing Non-Hierarchical and Hierarchical Cluster Analyses with Clustergrams. *Comput Stat* **2004**, *19*, 95–111, doi:10.1007/BF02915278.
37. Torres, E.A.F. da S.; Garbelotti, M.L.; Moita Neto, J.M. The Application of Hierarchical Clusters Analysis to the Study of the Composition of Foods. *Food Chem* **2006**, *99*, 622–629, doi:10.1016/j.foodchem.2005.08.032.
38. Custódio, F.B.; Vasconcelos-Neto, M.C.; Theodoro, K.H.; Chisté, R.C.; Gloria, M.B.A. Assessment of the Quality of Refrigerated and Frozen Pork by Multivariate Exploratory Techniques. *Meat Sci* **2018**, *139*, 7–14, doi:10.1016/j.meatsci.2018.01.004.