UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dept. of Applied Statistics and Operational Research, and Quality

Non-invasive detection of bone fragments in poultry meat by the integration of ultrasound imaging, multivariate image analysis and machine learning

MASTER'S THESIS
MASTER'S DEGREE IN DATA ANALYSIS, PROCESS IMPROVEMENT AND DECISION SUPPORT ENGINEERING

AUTHOR: **GENTIL ANDRÉS COLLAZOS-ESCOBAR**

TUTOR: **JOSÉ MANUEL PRATS-MONTALBÁN**
EXTERNAL COTUTOR: **JOSÉ VICENTE GARCÍA-PÉREZ**

Academic year: 2024-2025

VALÈNCIA, SEPTEMBER 26, 2025

Abstract

Poultry meat industry requires intelligent systems for achieving non-invasive, non-destructive and real-time detection of bone fragments (BF). Therefore, the main aim of this study was to assess the feasibility of using ultrasound imaging, multivariate image analysis (MIA)-based multivariate statistical process control (MSPC) and Latent Variable based Machine Learning (LV-ML) to detect varying-size BF within boneless and skinless chicken breast fillets. BF of different sizes (2.0 × 1.5 cm, 2.0 × 1.0 cm, 1.5 × 0.3 cm, 1.0 × 0.3 cm, and 0.5 × 0.3 cm) were inserted into the chicken breast fillets in five different locations. Contact ultrasound images were acquired in the control (C) and out-control (OC, with bone) chicken breast fillets, by scanning the breast's surface, using transmission mode contact ultrasound sensors (1 MHz), following a pre-established pattern. Energy-magnitude and energy-distribution ultrasound parameters were computed at pixel level considering three approaches: time-domain (TDA), frequency domain (FDA), and the combination of time and frequency domains (TFDA). Additionally, the time-frequency domains block-scale hard (TFDABH) and time-frequency domains block-scale soft (TFDABS) were also assessed. For each approach, MIA-SPC procedure was followed considering the Principal Component Analysis (PCA) as basis latent-space. From PCA model, the Residual Sum Squares (RSS) and Hotelling's T-square (T^2) control statistics were used to classify the C and OC images projected on the PCA latent structure. In the case of LV-ML models, seven ML techniques were trained (75%) and validated (25%) to classify the C and OC images projected on the PCA latent eigenspace. Furthermore, the Mean-Decrease Accuracy (MDA)-Random Forest (RF)-Variable Selection (VS) framework was applied to identify and rank the most relevant LVs, thereby maximizing the detection performance of a new set of Latent Variable-based Machine Learning (RF-VS-LV-ML) classifiers. Partial Least Squares Regression (PLSR) was then used as a statistical tool for modeling and multi-objective optimization, enabling the exploration of the relationships between ML model hyperparameters, number of LVs tested, and time-frequency block-scaled domains data approaches. Results demonstrated that the presence of bone fragments within chicken breast fillets led to alterations in the energy-magnitude (avg. amplitude decrease from 81.6% to 52.6%, depending on the bone size) and energy-distribution ultrasound parameters (avg. variance decrease from 97.9% to 70.6% depending on the bone size). The RSS statistic achieved the best classification performance (accuracy of TDA, FDA and TFDA>95%) in distinguishing between C and OC images. Furthermore, the combined use of LV-ML models, RF-VS-LV-ML and PLSR multi-objective optimization led to determine the best performing RF-VS-LV-RF-FDA model with a classification accuracy above 99% in both training and validation datasets, while maintaining reasonably computational time for training process. These results highlight the potential of integrating ultrasound imaging with advanced MIA and ML modeling strategies for the rapid and accurate detection of BF in chicken breast fillets. This approach offers a non-invasive and computationally efficient tool, paving the way for real-time detection and the implementation of fast, intelligent, non-invasive and non-destructive quality inspection systems in poultry meat industry.

Keywords: chicken breast, foreign bodies, bone fragments, in-line and real-time quality monitoring, multivariate image analysis, multivariate statistical quality monitoring, unsupervised and supervised machine learning, variable selection, multi-objective optimization.

Resumen

La industria de la carne de ave requiere sistemas inteligentes que permitan la detección no invasiva, no destructiva y en tiempo real de fragmentos óseos (BF). Por lo tanto, el objetivo principal de este estudio fue evaluar la viabilidad del uso de imágenes por ultrasonido, análisis multivariante de imágenes (MIA) basado en control estadístico multivariante de procesos (MSPC) y modelos de aprendizaje automático basados en variables latentes (LV-ML) para detectar BF de diferentes tamaños en filetes de pechuga de pollo deshuesados y sin piel. BF de diferentes dimensiones (2.0×1.5 cm, 2.0×1.0 cm, 1.5×0.3 cm, 1.0×0.3 cm y 0.5×0.3 cm) fueron insertados en los filetes de pechuga en cinco ubicaciones distintas. Se adquirieron imágenes ultrasónicas de contacto en filetes control (C) y fuera de control (OC, con hueso) mediante el escaneo de la superficie de la pechuga en modo de transmisión, usando sensores ultrasónicos de contacto de 1 MHz y siguiendo un patrón preestablecido. Los parámetros ultrasónicos de magnitud y distribución de energía se calcularon a nivel de píxel considerando tres enfoques: dominio del tiempo (TDA), dominio de la frecuencia (FDA) y la combinación de ambos (TFDA). Adicionalmente, se evaluaron los enfoques de TFDA con escalado por bloques duros (TFDABH) y suave (TFDABS). Para cada enfoque, se aplicó el procedimiento MIA-MSPC considerando el Análisis de Componentes Principales (PCA) como espacio latente de base. A partir del modelo PCA, las estadísticas de Suma de Cuadrados Residuales (RSS) y T-cuadrado de Hotelling (T^2) se utilizaron para clasificar las imágenes C y OC proyectadas en la estructura latente del PCA. En el caso de los modelos LV-ML, se entrenaron siete técnicas de aprendizaje automático (75%) y se validaron (25%) para clasificar las imágenes C y OC proyectadas en el espacio latente del PCA. Además, se aplicó la Selección de Variables del Bosque Aleatorio basado en la Disminución de Precisión Media (MDA-RF-VS) para identificar y jerarquizar las variables latentes más relevantes, maximizando así el rendimiento de detección en un nuevo conjunto de clasificadores de aprendizaje automático basados en variables latentes (RF-VS-LV-ML). Posteriormente, la Regresión por Mínimos Cuadrados Parciales (PLSR) se utilizó como herramienta estadística para modelado y optimización multiobjetivo, lo que permitió explorar las relaciones entre los hiperparámetros de los modelos de ML, el número de LV evaluadas y los enfoques de datos en dominios tiempo-frecuencia con escalado por bloques. Los resultados demostraron que la presencia de fragmentos óseos en filetes de pechuga de pollo provocó alteraciones en los parámetros ultrasónicos de energía-magnitud (disminución promedio de la amplitud de 81.6% a 52.6%) y energía-distribución (disminución promedio de la varianza de 97.9% a 70.6%). La estadística RSS alcanzó el mejor desempeño de clasificación (precisión de TDA, FDA y TFDA >95%) al diferenciar entre imágenes C y OC. Asimismo, el uso combinado de modelos LV-ML, RF-VS-LV-ML y la optimización multiobjetivo mediante PLSR permitió determinar que el modelo RF-VS-LV-RF-FDA fue el de mejor desempeño, con una precisión de clasificación superior al 99%, manteniendo tiempos de cómputo razonables durante el proceso de entrenamiento. Estos resultados destacan el potencial de integrar imágenes por ultrasonido con estrategias avanzadas de MIA y ML para la detección rápida y precisa de BF en filetes de pechuga de pollo. Este enfoque ofrece una herramienta no invasiva y computacionalmente eficiente, allanando el camino para la detección en tiempo real y la implementación de sistemas de inspección de calidad rápidos, inteligentes, no invasivos y no destructivos en la industria avícola.

Palabras clave: pechuga de pollo, cuerpos extraños, fragmentos óseos, monitoreo de calidad en línea y en tiempo real, análisis de imágenes multivariado, monitoreo estadístico multivariado de calidad, aprendizaje automático supervisado y no supervisado, selección de variables, optimización multiobjetivo.

Resum

La indústria de la carn d'au requerix sistemes intel·ligents que permeten la detecció no invasiva, no destructiva i en temps real de fragments gose-us (BF). Per tant, l'objectiu principal d'este estudi va ser avaluar la viabilitat de l'ús d'imatges per ultrasò, anàlisi multivariant d'imatges (MIA) basat en control estadístic multivariant de processos (MSPC) i models d'aprenentatge automàtic basats en variables latents (LV-ML) per a detectar BF de diferents grandàries en filets de pit de pollastre desossats i sense pell. BF de diferents dimensions (2.0×1.5 cm, 2.0×1.0 cm, 1.5×0.3 cm, 1.0×0.3 cm i 0.5×0.3 cm) van ser inserits en els filets de pit de pollastre en cinc ubicacions distintes. Es van adquirir imatges ultrasòniques de contacte en filets control (C) i fora de control (OC, amb os) mitjançant l'escaneig de la superfície del pit de pollastre en mode de transmissió, usant sensors ultrasònics de contacte d'1 MHz i seguint un patró preestablit. Els paràmetres ultrasònics de magnitud i distribució d'energia es van calcular a nivell de píxel considerant tres enfocaments: domini del temps (TDA), domini de la freqüència (FDA) i la combinació dels dos (TFDA). Addicionalment, es van avaluar els enfocaments de TFDA amb escalat per blocs durs (TFDABH) i suau (TFDABS). Per a cada enfocament, es va aplicar el procediment MIA-MSPC considerant l'Anàlisi de Components Principals (PCA) com a espai latent de base. A partir del model PCA, les estadístiques de Suma de Quadrats Residuals (RSS) i T-quadrat de Hotelling (T^2) es van utilitzar per a classificar les imatges C i OC projectades en l'estructura latent del PCA. En el cas dels models LV-ML, es van entrenar set tècniques d'aprenentatge automàtic (75%) i es van validar (25%) per a classificar les imatges C i OC projectades en l'espai latent del PCA. A més, es va aplicar la Selecció de Variables del Bosc Aleatori basat en la Disminució de Precisió Mitjana (MDA-RF-VS) per a identificar i jerarquitzar les variables latents més rellevants, maximitzant així el rendiment de detecció en un nou conjunt de classificadors d'aprenentatge automàtic basats en variables latents (RF-VS-LV-ML). Posteriorment, la Regressió per Mínims Quadrats Parcials (PLSR) es va utilitzar com a ferramenta estadística per a modelatge i optimització multiobjectiu, la qual cosa va permetre explorar les relacions entre els hiperparámetros dels models de ML, el número de LV avaluades i els enfocaments de dades en dominis temps-freqüència amb escalat per blocs. Els resultats van demostrar que la presència de fragments ossis en filets de pit de pollastre va provocar alteracions en els paràmetres ultrasònics d'energia-magnitud (disminució mitjana de l'amplitud de 81.6% a 52.6%) i energia-distribució (disminució mitjana de la variància de 97.9% a 70.6%). L'estadística RSS va aconseguir el millor acompliment de classificació (precisió de TDA, FDA i TFDA >95%) en diferenciar entre imatges C i OC. Així mateix, l'ús combinat de models LV-ML, RF-VS-LV-ML i l'optimització multiobjectiu mitjançant PLSR va permetre determinar que el model RF-VS-LV-RF-FDA va ser el de millor acompliment, amb una precisió de classificació superior al 99%, mantenint temps de còmput raonables durant el procés d'entrenament. Estos resultats destaquen el potencial d'integrar imatges per ultrasò amb estratègies avançades de MIA i ML per a la detecció ràpida i precisa de BF en filets de pit de pollastre. Este enfocament oferix una ferramenta no invasiva i computacionalment eficient, aplanant el camí per a la detecció en temps real i la implementació de sistemes d'inspecció de qualitat ràpids, intel·ligents, no invasius i no destructius en la indústria avícola.

Paraules clau: pit de pollastre, cossos estranys, fragments ossis, monitoratge de qualitat en línia i en temps real, anàlisi d'imatges multivariat, monitoratge estadístic multivariat de qualitat, aprenentatge automàtic supervisat i no supervisat, selecció de variables, optimització multiobjectiu.

PREFACE

Driven by both personal interest and professional commitment to providing objective information that supports decision-making in the meat industry and the broader food sector, this research was developed within the framework of the master's degree in data analysis, process improvement, and decision support engineering at Universitat Politècnica de València.

The main aim of this work was the development of quantitative tools (such as digital models of food manufacturing processes) for monitoring food quality and detecting foreign bodies in food products. During the master's dissertation period, significant efforts were devoted to calibrating robust multivariate statistical models tailored for industrial implementation. In addition, throughout the academic training provided by the master's program, particularly in those courses related to the use of advanced statistical techniques, the knowledge provided has laid the groundwork for the development of intelligent systems (digital models in combination with cutting-edge ultrasound technologies) that can contribute to statistical quality monitoring, process optimization, and informed decision-making within the food industry.

As a result, different tools (models, methodologies, patents, among others) have been developed and are now available to be applied in real-world scenarios, advancing the response to the challenges faced by the food industry in the context of Industry 4.0. These tools enable more effective quality monitoring and decision-making within the food sector. Furthermore, they provide valuable insights for optimizing non-invasive and non-destructive quality inspection systems, enhancing overall efficiency in production lines, and helping to overcome the challenges posed by the ongoing digitalization of the agrifood sector.

The main contributions of this study are outlined as follows:

Research patents:

1. **Collazos-Escobar, Gentil Andrés**., Peña Cerveró, Ramón., Bon Corbín, José., Benedito Fort, José Javier., **Prats-Montalbán, José Manuel**., **García-Pérez, José Vicente**., Cárcel, J. A., Fernández-Caballero-Fariñas, María Dolores. Dispositivo y Procedimiento para la Detección No Invasiva de Cuerpos Extraños en Alimentos Sólidos o Semisólidos con Ultrasonidos. UNIVERSIDAD POLITECNICA DE VALENCIA; AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIG. CIENTIFICAS, P202330154. 23 Feb 2023.

Research awards:

1. **Collazos-Escobar, Gentil Andrés**., **Prats-Montalbán, José Manuel**., Benedito Fort, José Javier., Giacomozzi, Soledad., Cárcel, Juan., **García-Pérez, José Vicente**. (2023). Detección no-invasiva de presencia de insectos en postres gelificados mediante ultrasonidos sin contacto. VIII Encuentro de Estudiantes de Doctorado de la Universitat Politècnica de València. València, España.

2. **Collazos-Escobar, Gentil Andrés**., Medina-Casas, Martha Patricia., **García-Pérez, José Vicente**. (2024). Sistemas avanzados e inteligentes para la monitorización no invasiva de la calidad de los alimentos (AI-FoodSafety). IV Edición de los Premios Aula Emprende en la Universitat Politècnica de València. València, España.

Research articles

Published:

1. **Collazos-Escobar, Gentil Andrés**., Gutiérrez-Guzmán, Nelson., Váquiro-Herrera, Henry., Bon, José., Cárcel, Juan., **García-Peréz, José Vicente**. (2023). Model-based investigation of the water adsorption in Achira (Canna edulis K.) biscuits. LWT, Food Science and Technology. DOI: https://doi.org/10.1016/j.lwt.2023.115472

2. **Collazos-Escobar, Gentil Andrés**., Barrios-Rodríguez Yeison., Bahamón-Monje Andres., Gutiérrez-Guzmán, Nelson. (2024). Mid-infrared spectroscopy and machine learning as a complementary tool for sensory quality assessment of roasted cocoa-based products. Infrared Physics & Technology, Volume 141, September 2024, 105482. DOI: https://doi.org/10.1016/j.infrared.2024.105482

3. **Collazos-Escobar, Gentil Andrés**., Lincetti, Elisa., Spilimbergo, Sara., **Prats-Montalbán, José Manuel**., **García-Pérez, José Vicente**., Benedito, José. (2025). Integrated use of ultrasound imaging and multivariate image analysis for detecting bone fragments in poultry meat. Food Research International, Volume 206, April 2025, Article: 16047. DOI: https://doi.org/10.1016/j.foodres.2025.116047

4. **Collazos-Escobar, Gentil Andrés**., Gutiérrez-Guzmán Nelson., Váquiro Henry A., **García-Pérez José Vicente**., Cárcel Juan A. (2025). Analysis of machine learning algorithms for the computer simulation of moisture sorption isotherms of coffee beans. Food and Bioprocess Technology. DOI: https://doi.org/10.1007/s11947-025-03785-x

5. **Collazos-Escobar, Gentil Andrés**., **Prats-Montalbán, José Manuel**., Giacomozzi, Soledad., Benedito, José., Gómez Álvarez-Arenas, Tomas E. and **García-Pérez**, **José Vicente**. (2025). Contactless detection of internal foreign bodies in foods using air-

coupled ultrasound: case studies on beef burger patties and jelly plates. Food Engineering. DOI: https://doi.org/10.1016/j.jfoodeng.2025.112777

Research articles from collaboration:

1. Sanchez-Jimenez, Virginia., **Collazos-Escobar, Gentil Andrés**., González-Mohino, Alberto., Gomez Alvarez-Arenas, Tomas., Benedito, José Javier., **García-Pérez, José Vicente**. (2023). Non-invasive monitoring of potato drying by means of air-coupled ultrasound. Food Control. DOI: 10.1016/j.foodcont.2023.109653

2. Llavata, Beatriz., **Collazos-Escobar, Gentil. Andrés**., **García-Pérez, José Vicente**., Cárcel, Juan. (2024). PEF pre-treatment and ultrasound-assisted drying at different temperatures as a stabilizing method for the up-cycling of kiwifruit: Effect on drying kinetics and final quality. Innovative Food Science & Emerging Technologies. DOI: https://doi.org/10.1016/j.ifset.2024.103591

3. Khanlar. Malikeh., **Collazos-Escobar, Gentil Andrés**., **García-Pérez, José Vicente**., Cárcel, Juan. (2025). Oleuropein extraction from olive leaves assisted by moderate electric fields and high-power ultrasound. A parametric study. Applied Food Research, Article 100654. DOI: https://doi.org/10.1016/j.afres.2024.100654

Book chapters:

1. **Collazos-Escobar, Gentil Andrés**, Gutiérrez-Guzmán, Nelson., Váquiro-Herrera, Henry., Bon, José., Cárcel, Juan., **García-Peréz, José Vicente**. (2023). Modelling of water sorption isotherms of dehydrated coffee beans using machine learning techniques. https://www.eurodrying2023.p.lodz.pl/travel-and-accomodation/

In addition, we have participated in various national and international conferences during the master's dissertation period:

International conferences:

1. **Collazos-Escobar, Gentil Andrés**., **García-Peréz, José Vicente**., **Prats-Montalbán, José Manuel**. (2021). Multivariate image analysis for detection of foreign bodies in burger meat. V Congreso Internacional de Investigación e Innovación en Ingeniería, Ciencia y Tecnología de Alimentos (IICTA 2021). Oral presentation.

2. **Collazos-Escobar, Gentil Andrés**., Gutiérrez-Guzmán, Nelson., Barrios-Rodriguez, Yeison. (2021). Mathematical sorption and machine learning modeling for predicting the equilibrium moisture content of specialty coffee beans (Coffee arabica L. Bourbon Rosado). V Congreso Internacional de Investigación e Innovación en Ingeniería, Ciencia y Tecnología de Alimentos (IICTA 2021). Oral presentation.

3. **Collazos-Escobar, Gentil Andrés**., **Prats-Montalbán, José Manuel**., Cárcel, Juan A., **García-Pérez, José Vicente**., Bon Corbín, José (2022). Convolutional Neural Networks for detecting foreign bodies in burger meat patties based on digital images. XIII Congreso Iberoamericano de Ingeniería de Alimentos (CIBIA 2022). Medellín, Colombia. Poster presentation.

4. **Collazos-Escobar, Gentil Andrés**., Bon, José., **García-Pérez, José Vicente**., Cárcel, Juan., Gutiérrez-Guzmán, Nelson. (2022). Artificial Neural Networks for predicting the water sorption isotherms of dehydrated cocoa and coffee products. XIII Congreso Iberoamericano de Ingeniería de Alimentos (CIBIA 2022). Medellín, Colombia. Oral presentation.

5. **Collazos-Escobar, Gentil Andrés**., Blanquer-Fernández, Maria., Benedito, José., Cárcel, Juan., **Prats-Montalbán, José Manuel**., Bon, José., **García-Pérez, José Vicente**. (2022). Rapid and non-destructive quality inspection of jelly products using acoustic ultrasonic imaging combined with machine learning techniques. 15th Conference of Food Engineering (CoFE'22). Raleigh, EEUU. Oral presentation.

6. **Collazos-Escobar, Gentil Andrés**., Sanchez-Jimenez, Virginia., Blanquer-Fernández, María., **García-Peréz, José Vicente**., Cárcel, Juan., **Prats-Montalbán, José Manuel**., Benedito, José Javier. (2023). Combination of acoustic imaging and machine learning algorithms for the rapid characterization of jelly-based products. 14th International Congress on Engineering and Food (ICEF14). Nantes, France. Oral presentation.

7. **Collazos-Escobar, Gentil Andrés**., Barrios-Rodríguez, Yeison., Bahamon-Monje, Andrés., Gutiérrez-Guzmán, Nelson., Cárcel, Juan., **García-Pérez, José Vicente**. (2023). Combination of machine learning with mid-infrared spectroscopy for sensory quality assessment of roasted cocoa-based products. 37th EFFoST International Conference. València, Spain. Poster presentation.

8. **Collazos-Escobar, Gentil Andrés**., Giacomozzi, Soledad., **Prats-Montalbán, José Manuel**., Benedito, José., Gosalbez, José., **García-Pérez, José Vicente**. (2023). Improvement in the detection of foreign bodies in jelly-based products through wavelet-based ultrasound-imaging. 37th EFFoST International Conference. València, Spain. Poster presentation.

9. **Collazos-Escobar, Gentil Andrés**., Giacomozzi, Soledad., **Prats-Montalbán, José Manuel**., Benedito, José., Gosalbez, José., **García-Pérez, José Vicente**. (2023). Use of air-coupled ultrasound in combination with machine learning to detect insects in gelled desserts. 37th EFFoST International Conference. València, Spain. Poster presentation.

10. **Collazos-Escobar, Gentil Andrés**., Lincetti, Elisa., Giacomozzi, Soledad., Spilimbergo, Sara., Benedito, José., **García-Pérez, José Vicente**. (2023). Detection of bone fragments in chicken breast using non-invasive air-coupled ultrasound imaging. 37th EFFoST International Conference. València, Spain. Poster presentation.

11. **Collazos-Escobar, Gentil Andrés**., Giacomozzi, Soledad., **Prats-Montalbán, José Manuel**., Benedito, José., **García-Pérez, José Vicente**. (2023). Non-invasive detection of foreign bodies in burger meat patties by contactless ultrasound imaging. 37th EFFoST International Conference. València, Spain. Oral presentation.

12. **Collazos-Escobar, Gentil Andrés**., Llavata, Beatriz., **García-Pérez, José Vicente**., Simal, Susana., Zhang, Hongwei., Cárcel, Juan. (2025). Computer modeling of ultrasonically assisted drying of PEF-pretreated kiwifruit using machine learning techniques. 39th EFFoST International Conference. Porto, Portugal. Poster presentation.

13. **Collazos-Escobar, Gentil Andrés.,** Bahamón-Monje, Andrés., Salas-Calderón, Karen., Barrios-Rodriguez, Yeison., Gutiérrez-Guzmán, Nelson. (2025). Detection of coffee defects in green and roasted beans using Deep Learning and infrared spectroscopy. 39th EFFoST International Conference. Porto, Portugal. Poster presentation.

14. **Collazos-Escobar, Gentil Andrés**., Hussaim, Tassadaq., Benedito, José, Zhang, Hongwei., Cárcel, Juan., **García-Pérez José Vicente**. (2025). Latent-based machine learning and contact ultrasound for detection of foreign bodies in jelly-based products. 39th EFFoST International Conference. Porto, Portugal. Poster presentation.

15. **Collazos-Escobar, Gentil Andrés**., Morales-Angulo, Ever., Bahamón-Monje, Andres., Gutiérrez-Guzmán, Nelson. (2025). Non-destructive quality assessment of green and roasted specialty coffee using machine learning and FT-NIR spectroscopy. 39th EFFoST International Conference. Porto, Portugal. Poster presentation.

National conferences:

1. **Collazos-Escobar, Gentil Andrés**., Cárcel, Juan A., **García-Pérez, José Vicente**., Lincetti E., Benedito Fort, José Javier., **Prats-Montalbán, José Manuel**., Bon Corbín, José. (2022). Aplicación de la transformada wavelet discreta para la rápida detección de cuerpos extraños en pechuga de pollo a partir de imágenes acústicas. XI Congreso Nacional de Ciencia y Tecnología de los Alimentos (CyTA/CESIA 2022). Zaragoza, España: Servicio de Publicaciones Universidad de Zaragoza. Poster presentation.

2. **Collazos-Escobar, Gentil Andrés**., Bon, José., Benedito, José., Cárcel, Juan., **García-Pérez, José Vicente**., **Prats-Montalbán, José Manuel**., Lincetti, Elisa. (2022). Uso de técnicas de análisis multivariante para mejorar la detección de fragmentos óseos en pechuga de pollo mediante ultrasonidos por contacto. XI Congreso Nacional de Ciencia y Tecnología de los Alimentos (CyTA/CESIA 2022). Zaragoza, España: Servicio de Publicaciones Universidad de Zaragoza. Oral presentation.

3. **Collazos-Escobar, Gentil Andrés**., Giacomozzi, Soledad., **García-Pérez, José Vicente**., **Prats-Montalbán, José Manuel**., Benedito Fort, José Javier. (2024). Ultrasonidos sin contacto para la detección de fragmentos óseos en pechugas de pollo. EN AGROALNEXT 2024. Innovación y transferencia en el sector agroalimentario español. Gandía, España.

4. **Collazos-Escobar, Gentil Andrés**., Giacomozzi, Soledad., **Prats-Montalbán, José Manuel**., Benedito Fort, José Javier., **García-Pérez, José Vicente**., Cárcel, Juan. (2024). Uso de imágenes digitales y redes neuronales convolucionales para la detección automática de cuerpos extraños en postres gelificados comerciales. XII Congreso Nacional de Ciencia, Tecnología e Ingeniería de los Alimentos (CyTA/CESIA 2024). Barcelona, España.

Master lectures:

1. **Collazos-Escobar, Gentil Andrés**. (2022). Uso de herramientas de análisis multivariante para la monitorización y control estadístico de procesos agroindustriales. I Workshop Investigación en Ciencia e Ingeniería de los Alimentos. Neiva-Huila, Colombia. Oral presentation.

Given course:

1. **Collazos-Escobar, Gentil Andrés**. (2021). Uso de Técnicas de Previsión para el Análisis de Series Temporales. Given course in the subject of Statistics and Experimental Design in the master's degree in coffee science and technology of the Universidad Surcolombiana. Neiva Huila, Colombia. Oral presentation.

***Experimental direction of final master's projects*:**

1. Blanquer Fernández, María, **Collazos-Escobar, Gentil Andrés**., Dalmau-Estelrich, María Esperanza, Benedito, José, **García-Pérez, José Vicente**. (2022). Uso de tecnologías ultrasónicas para la caracterización de gelatinas. Máster Universitario en Gestión de la Seguridad y Calidad Alimentaria. Defense date in 22/09/2022, Course 2022/2023. Grade 10.

***Experimental direction of final degree's projects*:**

1. García Gómez, Ricardo, **Collazos-Escobar, Gentil Andrés**., **García-Pérez, José Vicente**. (2024). Detección de cuerpos extraños en postres gelificados mediante ultrasonidos sin contacto y técnicas de aprendizaje automático supervisado. Grado en ingeniería agroalimentaria y del medio rural. Defense date in X/09/2024, Course 2023/2024. Grade 9.5.

Besides, the candidate has participated in a project that provided financial support for this master's final project: The ULTRADIGITAL project (AGROALNEXT/2022/045) as part of the AGROALNEXT program. This program is supported by the MCIN, with funding from the European Union NextGenerationEU (PRTR-C17.I1), as well as the Generalitat Valenciana.

INDUSTRIAL RELEVANCE

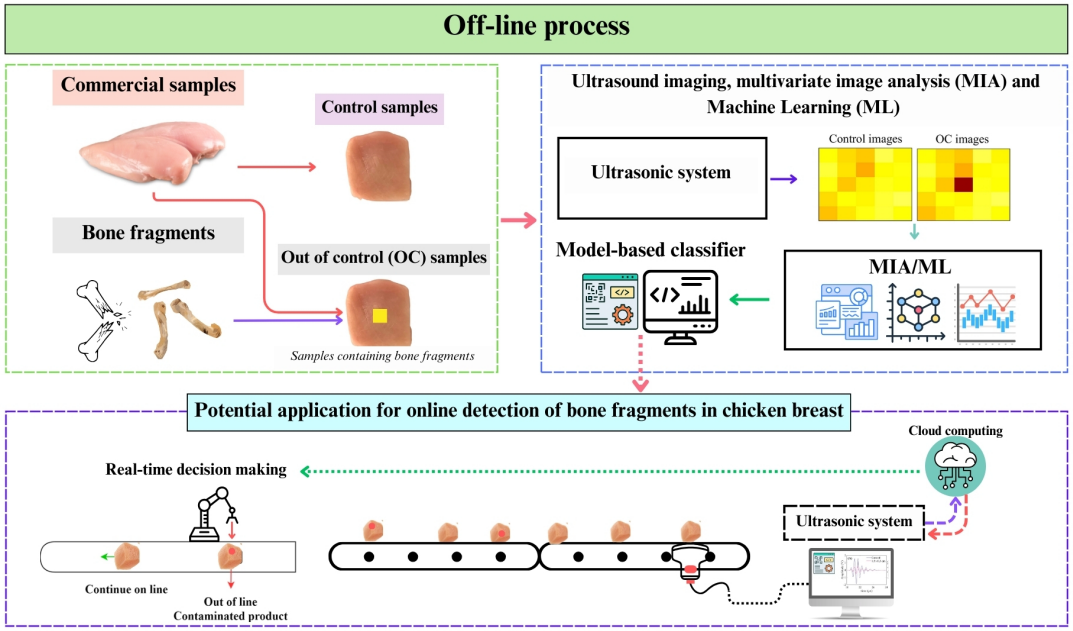
In recent years, the poultry meat industry has experienced significant growth and has become the most widely produced type of meat globally. Consumption continues to increase, primarily due to its affordability and high nutritional value.

One persistent challenge in the poultry industry is the detection of bone fragments. During the rapid mechanical deboning process, industrial machinery separates chicken breasts from the skeleton, which can lead to small bone fragments being embedded in the fillets. To ensure consumer safety, it is crucial to detect these fragments in-line and in real time. As a result, the poultry industry requires intelligent systems capable of non-invasive, real-time bone fragment detection.

In the context of Industry 4.0, the integration of non-invasive and non-destructive technologies with advanced digital models such as multivariate statistical techniques, machine learning, and deep learning offers innovative solutions for real-time quality monitoring of food products. In this regard, the combination of ultrasound imaging, multivariate image analysis, and machine learning is essential. The development of intelligent monitoring systems based on these technologies can enable fast and accurate inline detection of bone fragments, reducing contamination risks and ensuring consumer safety.

Digital twins, a digital representations of physical poultry quality monitoring systems, facilitate interactive communication between ultrasound systems and data-driven models (e.g., statistical, machine learning, and deep learning). Implementing digital twins supports continuous monitoring and optimization of detection systems, ensuring high standards of quality control while enhancing both production efficiency and product quality. Therefore, this study highlights the role of multivariate statistical models and machine learning as digital models in improving bone fragment detection in chicken breasts. By data-driven modeling and model optimization, the inline-implementation of these models can significantly reduce the occurrence of bone fragments, resulting in safer and higher-quality poultry products. Thus, the adoption of these technologies demonstrates their potential to transform food safety protocols not only within the poultry sector, but across the entire food processing industry.

Graphical abstract



Highlights

- Non-invasive and non-destructive methodology is proposed to detect bone fragments (BF) within chicken breast products.
- Ultrasound imaging, multivariate image analysis (MIA) and machine learning (ML) methods were combined to detect BF within poultry meat.
- The presence of BF reduced ultrasonic energy-related parameters.
- Ultrasound imaging allowed detecting bone fragments in chicken breast fillets.
- Integrated used of MIA and Multivariate Statistical Process Control (MIA-MSPC) models achieved >95% accuracy in BF detection.
- The improvement of MIA-MSPC was explored through Latent-Variable-based ML (LV-ML) models and the Variable-Selection-based Random Forest LV-ML (RF-VS-LV-ML) strategy.
- Partial Least Squares Regression (PLSR) combined with the Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization method enabled multi-objective hyperparameter tuning and facilitated the selection of the best ML model for maximizing BF detection performance.
- The use of LV-ML and RF-VS-LV-ML significantly enhanced the detection of BF in poultry breast samples.
- The RF-VS-LV-RF-FDA model achieved >99% accuracy in BF detection.
- Ultrasound technology combined with MIA/ML provides a rapid, non-invasive, and non-destructive tool for statistical quality monitoring in foods.
- Future research should focus on implementing this intelligent system for real-time quality inspection of chicken-based products in industrial inline settings.

TABLE OF CONTENS

LIST OF FIGURES

Fig. 10. Example of the frequency spectrum of chicken breast samples with and without bone fragments. Bone fragments of size 2.0 × 1.5 cm (A), size of 2.0 × 1.0 cm (B), size 1.5 × 0.3 cm (C), size of 1.0 × 0.3 cm (D) and size of 0.5 × 0.3 cm (E).

Fig. 11. Classification performance of the multivariate control statistics used for detection of bone fragments in chicken breast. Average A_{cc} for both RSS and T^2 considering TDA (A, C), FDA (E, G) and TFDA (I, K) approaches. Average Se and Sp for both RSS and T2 considering TDA (B, D), FDA (F, H) and TFDA (J, L) approaches. TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency domain approach), A_{cc} (overall accuracy), S_e (sensibility), S_p (specificity), RSS (Residual Sum Squares) and T^2 (Hotelling's T-squared).

Fig. 12. Average A_{cc} performance of RSS and T^2 control statistics used for detection of bone fragments in chicken breast using different number of ultrasound images. Results for TDA (A), FDA (B) and TFDA (C). TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency domain approach), A_{cc} (overall accuracy), RSS (Residual Sum Squares) and T^2 (Hotelling's T-squared).

Fig. 13. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 14. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 15. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 16. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 17. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 18. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Fig. 19. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Fig. 20. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Fig. 21. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Fig. 22. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Fig. 23. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-domain approach (TDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 24. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the frequency-domain approach (FDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 25. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach (TFDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 26. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 27. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 28. Computational time (CT) of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of CT are reported for the training (75%) dataset.

Fig. 29. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

Fig. 30. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

Fig. 31. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

Fig. 32. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

Fig. 33. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

Fig. 34. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the RF hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process. DTe (decision tree), NTs (number of trees).

Fig. 35. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Naïve Bayes (RF-VS-LV-NB) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the NB hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process. LS (Laplace Smoothing).

Fig. 36. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Linear Discriminant Analysis (RF-VS-LV-LDA), Random Forest-Variable Selection-Latent Variable-Quadratic Discriminant Analysis (RF-VS-LV-QDA) and Random Forest-Variable Selection-Latent Variable-Generalized Linear Model Analysis (RF-VS-LV-GLM) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS). Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process.

Fig. 37. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Support Vector Machines (LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the SVM's hyperparameters and number of latent variables (NLVs). NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset),

A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 38. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Support Vector Machines (LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the SVM's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 39. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Random Forest (LV-RF) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the RF's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), DTe (decision tree), NTs (number of trees; 50, 500 1000 5000 10000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 40. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Naïve Bayes (LV-NB) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the NB's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 41. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Linear Discriminant Analysis (LV-LDA) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 42. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Quadratic Discriminant Analysis (LV-QDA) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 43. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Generalized Linear Model (LV-GLM) considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 44. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Random Forest-Variable Selection-Latent Variable-Support Vector Machines (RF-VS-LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS) and the SVM's hyperparameters. Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 45. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS) and the RF's hyperparameters. Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), DTe (decision tree), NTs (number of trees; 50, 500, 1000, 5000 and 10000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 46. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of the optimized Latent Variable-Support Vector Machines (LV-SVM), Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA), Latent Variable-Generalized Linear Model (LV-GLM), Random Forest-Variable Selection-Latent Variable-Support Vector Machines (RF-VS-LV-SVM), Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF), Random Forest-Variable Selection-Latent Variable-Naïve Bayes (RF-VS-LV-NB), Random Forest-Variable Selection-Latent Variable- Linear Discriminant Analysis (RF-VS-LV-LDA), Random Forest-Variable Selection-Latent Variable- Quadratic Discriminant Analysis (RF-VS-LV-QDA) and Random Forest-Variable Selection-Latent Variable- Generalized Linear Model (RF-VS-LV-GLM) models. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset),

AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

Fig. 47. Optimal Latent Variables (LVs) selected using the Mean Decrease Accuracy (MDA, %) criterion from the Random Forest (RF) model. This eigenspace was employed for calibration and validation of the Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) framework. Using frequency-domain (FDA) features to feed the RF-VS-LV-RF yielded the best-performing optimized model (RF-VS-LV-RF-FDA). Variable importance of each LV in maximizing sample classification (with vs. without bone fragments) according to RF accuracy (A). Loading plots of the 30 most important FDA energy-magnitude-distribution ultrasound parameters ranked by MDA (B to H). Parameters include M_0 (zero-order moment), F_r (center frequency of the phase spectrum), MP (maximum peak of the frequency spectrum), VAR_{sp} (spectral variance of the phase spectrum), SKE_{sp} (spectral skewness of the phase spectrum), KUR_{sp} (spectral kurtosis of the phase spectrum), and ENT_{sp} (spectral entropy of the phase spectrum).

Fig. 48. Three-dimensional score plots of the nine most important Latent Variables (LVs; Fig. 47) selected by the Mean Decrease Accuracy (MDA) criterion from the Random Forest-Variable selection (RF-VS) strategy. Panels (A to C) show representative combinations of principal components (PCs) derived from the frequency-domain (FDA) features, while panel (D) presents the two-dimensional score projection of PC70 vs PC1 for clustering comparison. Sample groups correspond to Control and different bone fragment defect sizes (2.0×1.5 cm, 2.0×1.0 cm, 1.5×0.3 cm, 1.0×0.3 cm, and 0.5 × 0.3 cm). This selected eigenspace was used to feed the RF-VS-LV-RF yielded the best-performing optimized model (RF-VS-LV-RF-FDA).

Supplementary material

Fig. 1S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 2S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 3S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 4S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

Fig. 5S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

LIST OF TABLES

Table 10. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using time-domain approach (TDA).

Table 11. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using frequency-domain approach (FDA).

Table 12. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using time-frequency domain approach (TFDA).

Table 13. Classification performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics in the detection of varying-size bone fragments using the time-domain (TDA), frequency-domain (FDA) and time-frequency domain (TFDA) approaches.

Table 14. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Support Vector Machine (LV-SVM) and Latent Variable-Random Forest (LV-RF) and the number of latent variables (NLVs) used in model tunning.

Table 15. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Naïve Bayes (LV-NB) and Latent Variable-Linear Discriminant Analysis (LV-LDA) and the number of latent variables (NLVs) used in model tunning.

Table 16. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) and the number of latent variables (NLVs) used in model tunning.

Table 17. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach and hyperparameters belonging to each Machine Learning model.

Table 18. Multi-objective optimized Latent Variable-Machine Learning (LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Table 19. Multi-objective optimized Latent Variable-Machine Learning (LV-ML) and Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Table 20. Multi-objective optimized Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Table 21. Classification performance of the optimized Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) model using frequency-domain (FDA) features, yielding the best-performing configuration (RF-VS-LV-RF-FDA) for detecting bone fragments (BF) of varying sizes in chicken breast samples.

Supplementary material

Table 1S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction time-domain approach (feTDA).

Table 2S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction frequency-domain approach (feFDA).

Table 3S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction time-frequency domain approach (feTFDA).

1. Introduction

The poultry meat industry has undergone a rapid expansion in recent years, and it is currently the most produced meat worldwide (Aggrey et al., 2023; Fang et al., 2023). The consumption of poultry meat is increasing, due to its affordability, high nutritional value, and the large variety of derived processed products (Jiang et al., 2018). However, poultry meat production encounters several challenges, primarily related to disease management and the assurance of product quality and safety. Moreover, another significant concern is the presence of foreign bodies (FBs) in the final manufactured products.

Physical contamination resulting from the presence of FBs in food products poses significant health risks to consumers and can harm a company's reputation and legal compliance, since it has been recognized as the primary source of consumers' complaints received by food manufacturing companies (Edward and Stringer, 2007). In fact, consumer tolerance for any form of food contamination is decreasing, particularly for FBs such as wood, metal fragments, plastic particles, and bone fragments, as FBs may carry pathogens and microorganisms or cause physical harm when ingested (Djekic et al., 2017).

The detection of FBs represents a crucial bottleneck in the management of food safety and quality within poultry meat industry (Nielsen et al., 2013), and, in particular, bone fragment (BF) detection is a persistent problem. During the rapid mechanical deboning process, industrial machinery separates chicken breasts from the skeleton, which can result in bone fragments becoming embedded in the fillets. Detecting these fragments in-line and real-time is critical to ensure consumer' safety. Therefore, there is an urgent need for non-invasive, cost-effective, and intelligent systems for real-time meat product quality and safety monitoring; but still, the automation of the food manufacturing process remains a formidable challenge (Ali and Hashim, 2021). In this context, achieving automatic and dependable detection of FBs stands as a primary objective for the meat industry among the ongoing digital revolution, the incorporation of resilient industrial sensors and computer-assisted algorithms to facilitate real-time decision-making is essential (Belaud et al., 2019).

Traditional analytical techniques relying on electromagnetic radiation, including magnetic detectors, X-rays, and hyperspectral sensors, have been extensively used in the detection of FBs within food products (Yaqoob et al., 2021). These methods come with certain limitations for food inspection, such as the high cost of the equipment and its maintenance, challenges associated with their integration into food processing lines, and, in some cases, limited penetration capability to thoroughly analyze the internal structure of food (Pérez-Santaescolástica et al., 2019).

Regarding the detection of BFs in the meat industry, Yoon et al. (2007) employed a system based on Near-Infrared (NIR) spectroscopy for detecting BFs in chicken breast samples, while McFarlane et al. (2003) successfully utilized the X-ray backscatter technique to identify chicken clavicles and near-surface bone pieces in chicken breast pieces. Further, Lim et al. (2022) tested the X-ray imaging technique to detect soft plastic bullets within chicken breast samples, representing an improvement over conventional X-ray methods, which cannot

discern soft FBs. Nevertheless, X-ray-based methods, in general, are characterized by a significant drawback: they are expensive to operate, need costly equipment, pose risks to operators, and require complex post-image processing (McFarlane et al., 2003).

Ultrasound technology (US) has been employed as a valuable tool for the non-destructive testing of food materials. Ultrasound offers advantages over the aforementioned technologies: it enables faster inspection, it is cost-efficient, versatile, easy to manipulate, safe for personnel, and suitable for real-time in-line application (Fariñas et al., 2021), which is in accordance with the goals set by the Fourth Industrial Revolution (Industry 4.0) (Fariñas et al., 2023).

Consequently, US has emerged as a promising technology for detecting FB in foods. In the food industry, the conventional method for analyzing food products and processes relies on the contact ultrasonics (CUS) technology. In CUS, sensors require close contact with the food material to eliminate air gaps at the sensor-sample interface and enhance energy transfer into the sample. This contact is achieved through the use of coupling materials such as water, oil, or glycerine (Sánchez-Jiménez et al., 2023).

More recently, non-contact ultrasonics (NCUS) technology has gained recognition in the food industry and is considered highly suitable for non-destructive analysis of food products (Fariñas et al., 2021). Although this technology is still in the development phase for its industrial application, it shows great promise for future in-line applications, particularly in the detection of FBs.

In the meat sector, the CUS measurements have been satisfactorily employed for monitoring the physicochemical modifications in beef steaks during the dry salting process (Fariñas et al., 2023), for on-line monitoring of the ham salting process (García-Pérez et al., 2019) or for the characterization of dry-cured ham (Corona et al., 2013), among other applications. Regarding the detection of BF, Correia et al. (2008) designed and assembled an ultrasonic system based on CUS in pulse-echo mode to detect bone fragments in mechanically deboned chicken breasts. The effectiveness of their system for detecting BFs of different sizes was evaluated, showing an acceptable detection of fragments ranging from 6 mm^2 to 16 mm^2, based on attenuation values. However, the authors claimed important limitations in the application of their apparatus, being the primary drawback the significant variation in the obtained amplitude ratio values, leading to inconsistent and unreliable measurements. To solve this problem, another ultrasonic sensing modes such as through-transmission and pitch and catch can also be assessed (Mohd Khairi et al., 2015).

An additional benefit of US lies in its capacity for spatial analysis of food products by creating ultrasound imaging (USI). USI serves as a valuable non-destructive tool for inspecting food by scanning the surface of the product (Gan, 2020). It offers a spatial representation of internal characteristics, facilitating the evaluation of physicochemical attributes related to composition, texture or internal irregularities. The applicability of ultrasound imaging (USI) for identifying internal gas pockets and defects (such as cracks) was assessed in Swiss-Type Cheese by Eskelinen et al. (2007). Consequently, the potential of USI can also be explored for the detection of BF within chicken breasts, regardless of their specific location.

The massive volume of data generated by using US sensors in real-time requires robust mathematical models to extract relevant information from ultrasonic signals. In this sense, pattern recognition techniques constitute an advanced tool that is embraced by Industry 4.0 in the context of digitalization purposes (Ozturk et al., 2023). Many complex decision-making processes are involved in the food manufacturing processes. Thus, food industry can definitely benefit from these mathematical tools (Ni et al., 2020).

Pattern recognition based-models are mainly classified into two categories: unsupervised techniques such as the Principal Component Analysis (PCA) and hierarchical cluster analysis (HCA); and supervised techniques such as Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Generalized Linear Model (GLM), Partial Least Squares (PLS), and Soft Independent Modelling by Class Analogy (SIMCA). Additionally, in the machine learning (ML) field, supervised algorithms for regression and classification (decision trees-DTe, Random Forest-RF, Support Vector Machine-SVM, k-nearest-neighbors-kNN, Naïve Bayes-NB, among others) are also used (Jiménez-Carvelo et al., 2019).

The Principal Component Analysis (PCA) is one of the most important mathematical technique used in the manufacturing and process industries (Macgregor and Kourti, 1995). It is a statistical method used to simplify complex datasets by reducing their dimensionality. It achieves this by applying an orthogonal transformation that converts a group of possibly correlated variables into a new set of variables that are linearly independent from one another, referred to as principal components (Li et al., 2025). PCA is one of the most established techniques for exploratory data analysis in chemometrics (Godoy et al., 2014). Its ability to capture the main sources of variance within multivariate datasets, especially those with collinear variables or ill-conditioned matrices, makes it particularly suitable for applications involving high-dimensional ultrasonic or spectroscopic data (Sánchez-Jiménez et al., 2023). By projecting the original observations into a reduced latent space defined by uncorrelated principal components, PCA simplifies complex datasets and enhances interpretability, while also enabling the development of a statistical model for process monitoring and defect detection (Villalba et al., 2019).

In this sense, Statistical Process Control (SPC) has become a key methodology in manufacturing and process industries for tracking process behavior over time. Its main goal is to ensure that critical variables remain within acceptable limits, indicating that the process operates under statistical control (where only inherent), common-cause variation is present. Tools such as Shewhart, Cumulative Sum (CUSUM), and Exponentially Weighted Moving Average (EWMA) control charts are employed to detect deviations caused by special or assignable causes. Identifying and addressing these causes enables sustained process and quality improvements through corrective actions or operational adjustments. Conventional Multivariate Statistical Process Control (MSPC) techniques aim to track the stability of the process mean by constructing control charts based on Hotelling's T-square statistic derived from the original set of measured variables. This method assumes a fixed covariance structure and requires the inversion of the estimated covariance matrix, which becomes problematic when the number of variables approaches or exceeds the number of observations (common

in modern industrial data-rich environments). Additionally, these approaches rely on complete, noise-free datasets, a condition often unmet in automated industrial settings. Therefore, the MSPC faces limitations in scalability and robustness when applied to high-dimensional or incomplete process data, restricting its practical deployment in complex manufacturing systems (Villalba et al., 2019).

The use of PCA in the MSPC framework represents an important advantage since it avoids the computational and practical limitations associated with traditional SPC that require inversion of high-dimensional covariance matrices. By projecting the original correlated variables onto a reduced latent space composed of a few principal components, PCA simplifies the monitoring task while preserving the essential variance structure of the process. Control charts developed in this low-dimensional subspace function as multivariate process performance indices, offering a clear and interpretable representation similar to univariate SPC charts. At the same time, they leverage the full multivariate information, resulting in greater sensitivity to process anomalies. Furthermore, this approach is inherently more robust to missing or noisy data, making it highly suitable for complex, automated environments (Babamoradi et al., 2013).

In the PCA-MSPC, the monitoring of process deviations is performed using a limited number of principal components to construct the Hotelling's T-square statistic in the reduced latent space. This statistic quantifies deviations in the most informative principal variables , which capture the dominant sources of variation within the process. However, since Hotelling's T-square at each component reflects only the variability within the subspace defined by the retained principal components, it is complemented by the Residual Sum Square statistic, which measures the residual variation orthogonal to this subspace. By using both Hotelling's T-square and Residual Sum Square in a dual-chart monitoring strategy, enhances the fault detection in multivariate systems (Lemaigre et al., 2016). PCA-MSPC has been successfully applied to various food-related monitoring tasks, including the supervision of wine quality during fermentation (Cavaglia et al., 2020), the detection of melamine adulteration in milk through vibrational spectroscopy (Fernández Pierna et al., 2016), and the authentication of food products based on multivariate profiles (Preys et al., 2007). The authors reported that the PCA-MSPC approach demonstrated high computational efficiency suitable for online implementation, providing a robust and practical tool for real-time process monitoring and detection of anomalies in agro-food systems.

The data used in food industry analysis ranges from unstructured (text, audio, video, images and among others) to highly structured data (relational databases, spreadsheets, CSV files, among others) (Jin et al., 2020). Due to the demonstrated robustness of PCA in food quality inspection, this method facilitates the analysis of unstructured datasets (such as USIs) within the framework of Multivariate Image Analysis (MIA). MIA involves the application of multivariate techniques to extract both spectral and spatial information from images (Prats-Montalbán et al., 2011). This approach involves applying PCA to an unfolded multivariate image, resulting in an unsupervised classification of image pixels based on their spectral characteristics, which are represented in the PCA score space. The extracted information is subsequently analyzed by iteratively segmenting regions of interest within the PCA score

space and mapping the corresponding pixels back onto the original image. This procedure, commonly known as masking, enables the identification and extraction of relevant spectral features from the image (Duchesne et al., 2012).

MIA is useful as an explorative technique, clustering, defect detection and MSPC (Prats-Montalbán et al., 2009). MIA-MSPC can be applied twofold depending on the goal in hand: the first approach operates at the pixel level, requiring the monitoring and/or detection of defects or phenomena occurring within the individual captured images and the second approach entails process control based on the analysis of the image as a whole. Regardless the approach and following the MIA-MSPC framework, the first step involves calibrating a PCA model using a set of images corresponding to Normal Operating Conditions (NOC or Control), those acquired under the assumption that the process is functioning correctly. From this reference model, two key statistics such as Hotelling's T-square and Residual Sum Squares can be computed from the scores and residuals, respectively. Then, once a PCA model has been built on some NOC images and the Hotelling's T-square and Residual Sum Squares control limits have been established, the pixels of new images can be projected onto the calibrated PCA model. In this way, it becomes possible to identify which pixels exceed the statistical thresholds of model. Within this context, rather than visualizing the full Hotelling's T-square and Residual Sum Squares images, it is more informative to highlight only those pixels that exhibit extreme behavior within the model (exceeding the Hotelling's T-square limit) or that fall outside the model (surpassing the Residual Sum Squares limit) (Prats-Montalbán et al., 2011). The feasibility of MIA-MSPC has been demonstrated in several works including the monitoring of colour random texture (Reis, 2015), the on-line monitoring of a freeze-drying process for pharmaceutical products in vials (Colucci et al., 2019) and the defect Detection in Random Colour Textures (López et al., 2006).

The PLS model is another well-kwon advanced multivariate statistical tool widely used in chemometrics for performing multivariate regression/classification. PLS is an iterative method for finding latent variables that maximize covariance between the input and response variables (Duma et al., 2024). Although PLS is also commonly latent-structure based model used in multivariate process monitoring (especially when there is a known response variable), PCA is often preferred when the objective is unsupervised (no response variable is available) defect detection, data visualization, or latent space modeling without requiring prior class information images (Prats-Montalbán et al., 2011). In any case, if both process variables and product quality data are available, multivariate statistical predictive models based on projection to latent structures, such as PLS, can also be employed (Prats-Montalbán et al., 2012; Villalba et al., 2019).

Since the detection of BFs in the poultry meat industry is an unsupervised, data-driven task (due to the lack of prior labeling of samples) MIA-MSPC based on PCA emerges as the most suitable approach. In fact, MIA-MSPC based on PCA results in the detection of BF of different sizes within chicken breast samples using USI have been previously reported by Collazos-Escobar et al. (2025). The authors reported that USI obtained by CUS enable the detection of BF through statistically significant changes in both energy-magnitude and energy-distribution ultrasound parameters at the USI pixel level, highlighting their influence

on signal attenuation, spectral features, and variability depending on fragment size. The MIA-MSPC demonstrated the feasibility of using both Hotelling's T-square and Residual Sum Squares statistics for the detection of BF, achieving an overall accuracy greater than 95%.

Additionally, the authors acknowledged that supervised pattern recognition approaches could potentially be explored to enhance detection performance. Although these techniques have shown promise in related applications (Bowler et al., 2023, 2020; Caladcad et al., 2020; Conde et al., 2008; Sánchez-Jiménez et al., 2023; Velásquez et al., 2021b, 2019), their implementation requires additional steps, such as the definition of a latent variable space (from PCA as example) and careful tuning of hyperparameters to optimize model performance. As the BF detection problem is an inherently data-driven task, exploring alternative modeling strategies (including unsupervised/supervised machine learning techniques) may provide valuable insights and further improvements (Collazos-Escobar et al., 2023b). Thus, the analysis of whether supervised ML techniques can improve the results of an unsupervised model is a relevant matter of scientific interest. Nonetheless, every effort in exploring the balance between computational complexity and model's likelihood should be considered in these modeling strategies.

To elucidate whether the use of supervised ML techniques can improve upon the results previously obtained with MIA-MSPC based on PCA in the detection of BFs, various algorithms such as SVM, DTe, RF, NB, LDA, QDA, GLM and among others, can be assessed. The assessment of these models in the improvement of MIA-MSPC can be conducted using the MIA as the basis for ML's model calibration. This procedure is commonly carried out by the use of PCA's scores as features in the ML model tunning. This strategy has been successfully done in both prediction and classification tasks (Caladcad et al., 2020; Collazos-Escobar et al., 2024, 2023a; Conde et al., 2008; Barrios-Rodríguez et al., 2021; Sánchez-Jiménez et al., 2023). Thereby, this procedure paves the way for the analysis and comparison between the two categories of pattern recognition based-models to address the same data-driven problem.

In this sense, SVM have emerged as pivotal tools in ML. Its ability to handle high-dimensional data efficiently, coupled with their applicability to both classification and regression tasks, makes SVM reliable for complex analytical challenges and data-driven tasks. SVM functions by identifying an optimal hyperplane that maximally separates data points belonging to different classes (Scatigno and Festa, 2022). In real-world datasets, which often exhibit overlapping classes, the soft-margin SVM formulation is employed. This approach introduces slack variables and a regularization parameter, allowing a trade-off between maximizing the margin and minimizing classification errors (Chauchard et al., 2004). To handle non-linear patterns, SVMs utilize kernel functions that implicitly map input features into higher-dimensional spaces where linear separation becomes more feasible. Commonly used kernels include linear, polynomial, and radial basis function kernels. The choice of kernel and its associated hyperparameters significantly affects model performance and must be carefully optimized for the specific application (Scatigno and Festa, 2022).

SVMs have been widely applied in the field of food science and engineering. For instance, SVM models have combined with near infrared spectroscopy techniques for acidity

prediction in grapes (Chauchard et al., 2004), in determining coconut maturity level integrating SVM and acoustic signals (Caladcad et al., 2020), in the classification of the maturity stage of coffee cherries (Velásquez et al., 2021b) and to quantify the influence of maturity stage on drying kinetics of coffee cherries (Velásquez et al., 2021a), in the rapid prediction and description of the moisture content changes in achira biscuits (Collazos-Escobar et al., 2023b) and dried coffee beans during storage (Collazos-Escobar et al., 2025a, 2025b). In all of these applications, the authors claimed that SVM was able in handling high dimensionality, complex and non-linearities datasets, shown promise in process improvement and real time decision-making in the food industry.

Despite SVM's advantages, these models have notable limitations. One of the most important limitation is that the training process on large-scale datasets can be computationally intensive, although recent advances in hardware and optimization algorithms have mitigated some of these concerns (Xu et al., 2025). Moreover, model performance is highly sensitive to the tuning of hyperparameters (such as the regularization parameter and kernel-specific parameters) which requires careful optimization, commonly achieved through grid search, cross-validation and strategies based on design of experiments (Collazos-Escobar et al., 2025a). Another significant limitation of SVM is their limited interpretability, as they are often regarded as "black-box" models. This poses challenges in domains where understanding the contribution of individual features is essential for informed decision-making. Nevertheless, their robustness, flexibility, and strong predictive performance continue to make SVMs a popular choice in analytical and chemometric applications (Zhu et al., 2024).

Another relevant technique widely used in the field of supervised ML is the RF (Gholizadeh et al., 2020). RF have become a cornerstone in ML due to their robustness, versatility, and ability to handle high-dimensional data. RF has proven particularly effective for modeling complex datasets, supporting tasks such as classification, regression, and feature selection (Otchere, 2023). RF is a classification algorithm composed of an ensemble of DTe for both classification and regression tasks. A DTe is a hierarchical, tree-structured model that recursively partitions the feature space into disjoint regions based on feature-based decision rules. At each internal node, the algorithm selects a feature and a corresponding threshold that maximizes a splitting criterion, such as Gini impurity or information gain, thereby creating two or more nodes. This process continues recursively until a stopping condition is achieved (such as reaching a maximum depth, a minimum number of samples per node, or achieving complete purity in the leaf nodes). Each terminal node (leaf) assigns a class label in classification tasks or a numerical value in regression tasks, based on the majority class or average of the training samples contained in that node (Sun and Hu, 2017).

As an ensemble method, RF constructs multiple DTe and aggregates their outputs (using majority voting for classification and averaging for regression) which enhances predictive accuracy and reduces the risk of overfitting (Yeap et al., 2020). This makes it especially suitable for the intricate and noisy datasets commonly encountered in real-scenarios. During training process, each tree is built from a bootstrap sample of the original dataset, and at each split, a random subset of features is considered. This randomized approach introduces

diversity among the trees, which helps reduce model variance and improve generalization performance (Fu et al., 2014).

One of the most relevant RF advantages is its capacity to estimate feature importance of independent variables in the tunning of a predictive/classification model (Collazos-Escobar et al., 2024). During the training process of RF, variable importance scores are computed to quantify the predictive contribution of each feature. These scores guide the stratified sampling of the feature subspace when building the forest, enabling the algorithm to prioritize highly informative independent variables while still incorporating those with lower relevance, thereby preserving potential complementary information (Wang et al., 2023).

The assessment of variable importance can be approached through strategies such as measuring the reduction in node impurity (e.g., Gini impurity) or evaluating the decrease in predictive accuracy (using metrics like mean decrease accuracy or the increase in mean squared error) when a feature is permuted (Collazos-Escobar et al., 2023a). These measures can inform the calibration of ensemble strategies, enabling the construction of more parsimonious and interpretable models based on feature selection results derived from RF. This approach (tuning a model using the variable importance criterion for feature selection from a trained RF) has been successfully applied across several applications (Chen et al., 2018; He et al., 2021; Liberda et al., 2021; Wang et al., 2019). Additionally, RF has been effectively employed in the food industry (Caladcad et al., 2020; Malaslı et al., 2025), demonstrating its potential as a robust ML technique for quality inspection in the food industry, including the detection of BF in chicken breast products.

The NB is another pattern recognition-supervised ML technique used in the data analytics. It is a probabilistic (stochastic) model commonly applied to classification tasks. NB operates under the simplifying assumption that the features within each class are independent of one another and follow a Gaussian (normal) distribution (An and Zhang, 2025). The algorithm calculates the conditional probability that a given data instance belongs to each class based on its features, then assigns it to the class with the highest probability. Despite its "Naïve" assumption of feature independence, NB often performs remarkably well when this assumption is reasonably valid. Its simplicity, computational efficiency, and effectiveness with high-dimensional data make it a popular choice in many practical applications (Fink et al., 2025).

In the literature there is often reported the use of the LDA, QDA and GLM as supervised pattern recognition techniques to address classification problems in real industrial scenarios (Nibouche et al., 2024). The idea behind of these techniques is the calibration of a classifier based on a training dataset of labeled instances that can accurately predict new feature observations into one of the known groups (Cabana and Lillo, 2022).

LDA is a parametric classification method based on the assumption that the classes are drawn from multivariate normal distributions sharing an identical variance (covariance matrix). The decision boundary is linear; a straight line in two dimensions or a hyperplane in higher-dimensional space, and is derived using the Mahalanobis distance such that all points on the boundary have equal posterior probability of belonging to either class. LDA is

computationally efficient and performs well when its underlying assumptions hold, but its accuracy deteriorates when the class distributions differ substantially in dispersion or shape (Dixon and Brereton, 2009). Similarly, QDA assumes multivariate normality of class distributions but relaxes the requirement of equal variance/covariance matrices across classes. This allowance produces a quadratic decision boundary capable of accommodating classes with distinct shapes and dispersion patterns. While QDA's flexibility enables it to model more complex class structures, it entails estimating a greater number of parameters, which increases the risk of overfitting, particularly in scenarios with limited training data (Vranckx et al., 2021).

Finally, GLM extends traditional linear regression (LR) to handle non-normal response variables such as binary, count, or skewed data. In the same way as LR, GLMs assume low correlation among predictors, but in practice multicollinearity can inflate the variance of the maximum likelihood estimator (Algamal, 2018). The GLM model achieves its performance via three key components: the random component, which specifies the probability distribution of the response variable from the exponential family (such as normal, binomial, or Poisson); the systematic component, which represents a linear combination of the predictor variables; and the link function, which connects the expected value of the response variable to the linear predictor. This structure enables the model to capture non-linear relationships between predictors and the mean of the response while maintaining interpretability (Guisset et al., 2019).

One of the main limitations in the calibration of these supervised pattern recognition/ML techniques (DTe, RF, SVM, NB, LDA, QDA and GLM) is the multicollinearity between input variables in high dimensional and low-sample size real world datasets (Chen et al., 2025). For instance, NB classifier assumes that the variables used in the calibration process are independent (van Herwerden et al., 2022). Further, LDA, QDA and GLM also require independent input variables in the calibration process.

In datasets where the number of samples exceeds the number of variables and the variables exhibit low collinearity; it is feasible to compute the inverse of the variance-covariance matrix. Nevertheless, such conditions are not common in real industrial settings, where the inherently multivariate nature of processes tends to generate high-dimensional variable spaces coupled with a limited number of observations, often due to the cost and complexity of data acquisition. This poses a challenge for supervised models (DTe, RF, SVM, NB, LDA, QDA and GLM) as these methods require the computation of covariance matrix for whole and/or each class, and these matrices must be invertible. When the number of variables approaches or surpasses the number of observations within a class, the covariance matrix becomes singular or nearly singular, rendering the model computationally infeasible (Siqueira et al., 2017). In such scenarios, dimensionality reduction techniques such as PCA (unsupervised) and PLS become essential. These methods reduce the number of variables, enabling the inversion of the variance-covariance matrix while retaining most of the relevant information comprising in the new obtained latent variables (Sánchez-Jiménez et al., 2023). A notable advantage of PCA, in particular, is its ability to compress large multivariate datasets into a small number of orthogonal principal components that preserve the majority

of the variance present in the original data (Dixon and Brereton, 2009). Thus, the strategy of MIA-ML (use of latent space from PCA on unfolded images for ML model tunning) becomes promising in enhancing the predictive capability of a model based on latent structures, in exploring improvements, reducing computational time, and achieving better detection capability through the efficient calculation and interpretation of the most informative latent variables (Collazos-Escobar et al., 2024; Ramtanon et al., 2025).

The integration of US technology and pattern recognition ML techniques has been previously employed in various applications, including the monitoring of yogurt fermentation process (Bowler et al., 2023), tracking the drying of potato slices (Sánchez-Jiménez et al., 2023), the detection of internal cracks in Manchego cheese (Conde et al., 2008), and has also been used for assessing the coconut maturity (Caladcad et al., 2020).

Regarding the detection of FBs, Zhao et al. (2006) implemented a monitoring system based on CUS and Artificial Neural Networks for the detection of glass fragments in filled glass containers. The authors report that integrating CUS and artificial intelligence techniques, they successfully detected FBs of different sizes (accuracy≥95%) within the canned beverages. However, their application is limited to a single point ultrasound measurement per foodstuff (Mohd Khairi et al., 2018). This issue is overcome by using USI since an ultrasonic image enables the assessment of spatial and internal properties of a food product.

There is a notable gap in the existing literature regarding of an intelligent system for real-time quality inspection of chicken based products. Even worst, there are no systems based on non-invasive, non-destructive and cost-efficient technologies for the detection of FBs and BFs in poultry meat industry. Thus, the proposed research introduces a groundbreaking approach for the non-invasive detection of internal BFs in poultry meat through the integration of USI, MIA, and ML. While Industry 4.0 has driven the development of automated, real-time monitoring systems for the food industry, current industrial solutions remain limited in detecting food contaminants embedded within the product matrix. As we previously stated, conventional methods such as X-rays and magnetic detectors are effective for high-density or ferrous foreign bodies (metal pieces) but fail to identify low-density FBs. This work overcomes these limitations by leveraging ultrasound technology capability in the analysis of internal structure of foods, the advanced MIA framework for feature extraction, and ML for accurate detection performance of uncontaminated and contaminated foodstuffs, enabling robust, rapid, and automated detection of BFs within poultry meat. To our knowledge, this is the first approach of a fully integrated USI, MIA-MSPC and MIA-ML for real-time, non-invasive detection of internal food contaminants in the poultry industry, representing a significant advancement in food safety monitoring under the framework of Industry 4.0. Additionally, the analysis of the same data-driven task using different strategies based on unsupervised and supervised pattern recognition techniques also represents an improvement towards the development of computational efficient, parsimonious and high likelihood models to be further used in real-time industrial applications.

Comprehensive efforts should be directed toward elucidating the statistical results of implementing a non-invasive food-inspection system for the detection of BFs founded on a hybrid methodology that sequentially integrates unsupervised latent-structure modeling with

supervised learning techniques, coupled with the multi-objective optimization of statistical models and their comparative assessment in terms of feasibility for deployment in real-time industrial environments. Consequently, contributions that advance the state of the art in chemometrics and intelligent laboratory systems warrant thorough consideration. Therefore, if the abovementioned ideas are considered, the purpose of this work was threefold: (i) to experimentally determine the feasibility of USI obtained by CUS technology for detecting BFs of varying sizes, (ii) to assess the capability integrate MIA-MSPC and USI for the calibration of a digital model for detecting BFs within chicken breast and (iii) to analyze the improvement in the detection capacity of statistical models by using a hybrid strategy based MIA-ML.

2. Materials and methods

2.1 Chicken breast samples

Skinless and boneless chicken breast samples were purchased from a local grocery store in Valencia (Spain) and kept in the fridge at 4 °C until use (Fig. 1A). Since the ultrasound images could not be taken in refrigerated conditions, due to the ultrasound equipment was not adapted to work inside a refrigeration chamber, the entire chicken breast fillets were left out of the refrigerator until they reached room temperature. The breasts were then cut into 5 × 5 cm samples with a thickness of about 1.5 cm (Fig. 1B). Samples with no BF were considered as the control samples.

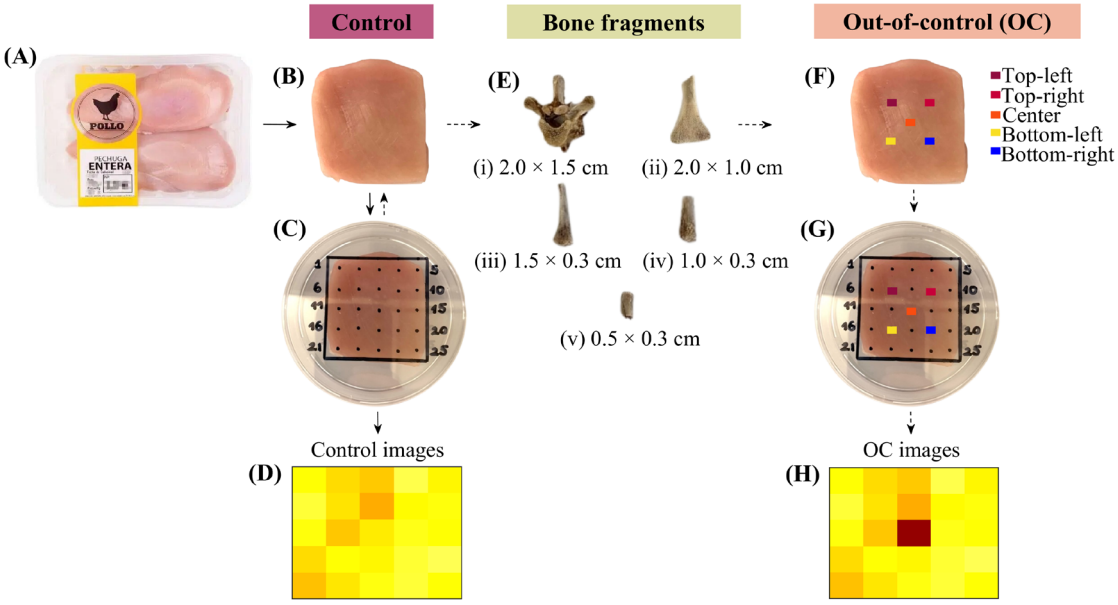


Fig. 1. Flowchart of the main steps to prepare control and OC (out-of-control) samples. A, B, C and D for control images, furthermore C, B, E, F, G and H for OC images.

2.2 Bone fragments

A set of bone fragments extracted from different parts of chicken skeleton was considered. For this purpose, a whole chicken was purchased, boiled for 20 minutes at 80 ºC, and then manually deboned, in order to extract these bone fragments. The set of bone pieces used in the experiments (Fig. 1E) consisted of a bone obtained from dorsal vertebrae with dimensions of 2.0 × 1.5 cm (Fig. 1Ei), a fragment taken from the chest bone of 2.0 × 1.0 cm (Fig. 1Eii), and three different fragments extracted from the chicken rib with sizes of 1.5 × 0.3 cm (Fig. 1Eiii), 1.0 × 0.3 cm (Fig. 1Eiv) and 0.5 × 0.3 cm (Fig. 1Ev).

2.3 Ultrasound experimental set-up

The ultrasound images were acquired using the experimental set-up described in Fig. 2. The equipment consisted in a computer (Fig. 2A), an oscilloscope (Fig. 2B, MDO3024, Tektronix, WA, USA), an ultrasonic generator-receiver (Fig. 2C, 5077 PR, Olympus, Houston, TX, USA), a pair of commercial ultrasound transducers (Fig. 2D, A314S-SU model, Panametrics, Waltham, MA, USA) of 1 MHz central frequency and 1 cm of diameter, operating in through-transmission mode, the food sample (Fig. 2E) and a digital caliper (Fig. 2F, 192-633 Serie, Mitutoyo, Japan). A program was developed in LabVIEW® 2018 (National Instruments, Austin, TX, USA) to record the ultrasonic signals.

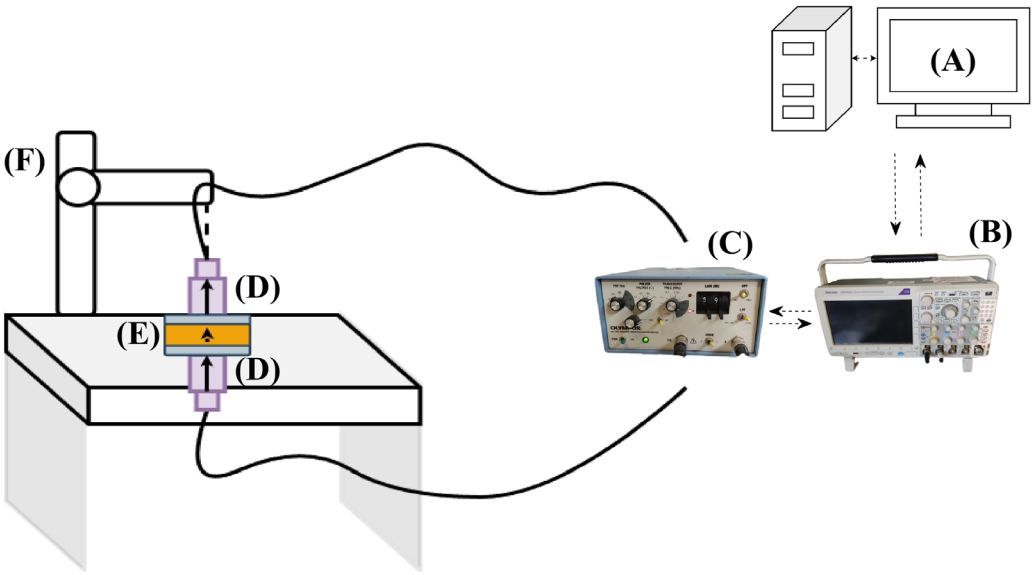


Fig. 2. Equipment used for ultrasound image acquisition. It consisted of one computer (A), one oscilloscope (B), one generator-receiver (C), two ultrasonic transducers (D), food sample (E) and digital caliper (F).

2.4 Experimental procedure

Control samples (without BF) were placed in polystyrene plates (86.4 ± 0.1 mm diameter, 14 ± 0.1 mm thickness; Fig. 1C) in order to measure the ultrasound signals in the same locations for each sample, thus obtaining the USI. To achieve this, a pre-established matrix of 25 points (5 × 5 cm) separated every 1 cm (Fig. 1C) was previously drawn on the surface of the polystyrene plates. Each point of this matrix corresponded to a pixel of the image (Fig. 1D). Tap water was used to wet both the transducers and the polystyrene plate's surface, to improve the transmission of the transducer's energy through the polystyrene lids.

After the imaging process, each type of bone fragment was inserted into the previously measured control samples, using a laboratory forceps, trying to place it equidistant from each face of the chicken breast sample. Each bone type was tested in five different locations (Fig. 1F), namely, the top-left, top-right, center, bottom-left and bottom-right, corresponding to the

position 7, 9, 13, 17 and 19 of the pre-established matrix (Fig. 1G), respectively. Thus, OC ultrasound images (Fig. 1H) were obtained. Each type of bone fragment (n = 5, Fig. 1E) was analyzed in triplicate (n = 3) at each location within food samples (n = 5, Fig. 1F). Thus, a total of 75 chicken samples (5 × 5 × 3) were obtained. However, 81 chicken breast samples (Fig. 1B) were analyzed, since six additional samples were considered to ensure consistent data; due to some tests had to be repeated because of measurement uncertainty. Thereby, a dataset of 81 images of control samples and 81 images of OC were obtained.

Ultrasound images were obtained using the experimental set-up (Fig. 2) by hand scanning the sample surface following the preestablished pattern drawn on the polystyrene plates (Fig. 1C). For each measured point, two types of ultrasound signals (each of 10k points, average of 128 acquisitions) in the time-domain were obtained. The first one was acquired with a receiver gain of –20 dB and used to compute the energy-related ultrasound parameters in the time (section 2.5.1) and frequency (section 2.5.2) domains. Then, a second type of ultrasound signal was acquired with a gain of 0 dB (Fig. 2C) to calculate the ultrasound velocity (section 2.5.1). Thus, two types of 3D images of 5 × 5 cm (spatial dimensions-2D of scanned product's surface) × 10k points (measured ultrasound signal at each point-1D) were acquired in every run. Moreover, the thickness of the samples was gathered for each pixel using the digital caliper.

2.5 Feature extraction

Different parameters related to energy (Fig. 3) and also the ultrasonic velocity were computed at pixel level. Previous to the calculations, the ultrasound signals (signal contain in every pixel) were baseline-corrected to eliminate any bias associated with electrical noise from the network. This correction entailed determining the mean value of each signal within the range of 1300–1800 points. If the mean value of the signal was below 0 V, the absolute mean value was added to the entire signal (Fig. 3A). Conversely, if the mean value was above 0 V, it was subtracted from the signal. Thus, each parameter summarized a channel of the image. e.g., if six parameters were estimated from a 3D image, a new image of 5 × 5 (spatial dimensions) × 6 (computed parameters) was obtained. The flowchart illustrating the procedure conducted to obtain the USI, considering different feature extraction strategies, is summarized in Fig. 4.

2.5.1 Time domain analysis

Energy-magnitude ultrasound parameters such as peak-to-peak distance (PP, V; Fig. 3B), energy (ENG, V^2; Fig. 3B) and integral of signals (INT, V μs; Fig. 3B) (Bowler et al., 2023) and ultrasound velocity (V_e, m/s) were computed in the time-domain from the ultrasound 3D images. The PP (Eq. 1), ENG (Eq. 2) and INT (Eq. 3) were obtained from the 3D images with signals acquired at –20 dB, while the V_e was calculated from signals acquired at 0 dB. The INT was computed by using the trapezoidal numerical method "trapz" of MATLAB® R2023a (The MathWorks Inc., Natick, MA, USA).

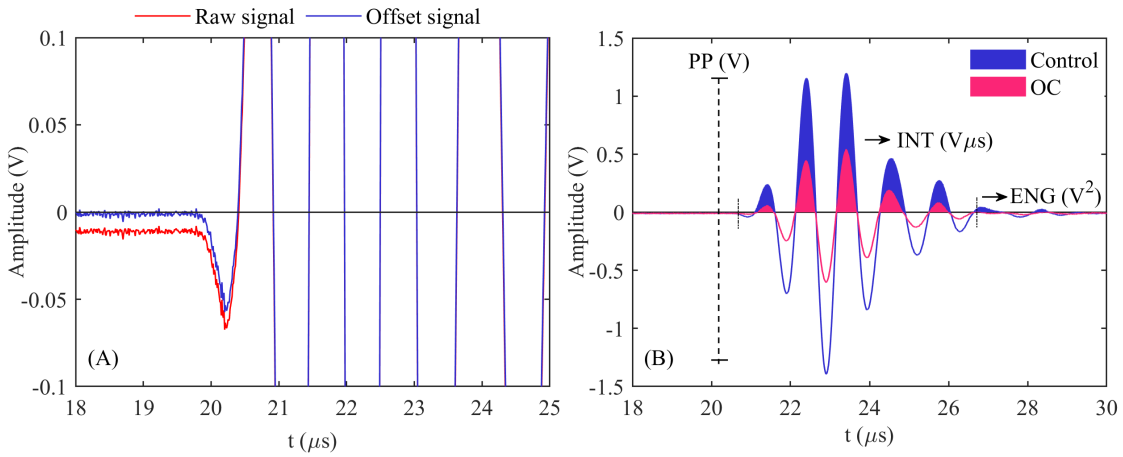


Fig. 3. Baseline correction of time-domain ultrasound signals (A) and energy-magnitude ultrasound parameters (B).

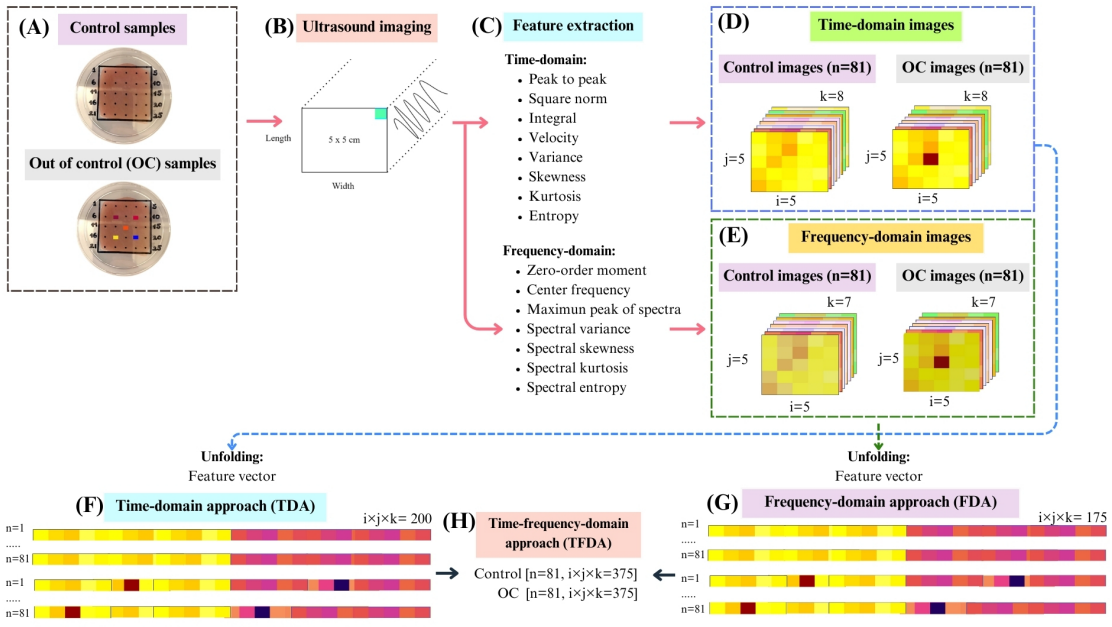


Fig. 4. Flowchart illustrating the ultrasound-based feature extraction procedure and the methodological strategies implemented to identify the presence of bone fragments in out-of-control (OC) samples using time-frequency domain approaches.

According to Caesarendra and Tjahjowidodo (2017), the variance (VAR_t, V^2), skewness (SKE_t), kurtosis (KUR_t) and entropy (ENT_t) time-domain energy-distribution parameters were also considered. The VAR_t (Eq. 4), SKE_t (Eq. 5), KUR_t (Eq. 6) and ENT_t (Eq. 7) were computed by using "var", "skewness", "kurtosis" and "entropy" MATLAB functions for each pixel of USI.

$$PP = \max(X_t) - \max|\min(X_t)| \quad (1)$$

$$ENG = \|X_t\|^2 \quad (2)$$

$$\mathrm{INT} = \sum_{i=1}^{N} \mathrm{Xz}_{t_i} t_i \quad (3)$$

$$\mathrm{VAR}_t = \frac{\sum_{i=1}^{N} (X_{t_i} - \overline{X_t})^2}{(N-1)} \quad (4)$$

$$\mathrm{SKE}_t = \frac{\left[\sum_{i=1}^{N} (X_{t_i} - \overline{X_t})^3\right]/(N-1)}{{\sigma_t}^3} \quad (5)$$

$$\mathrm{KUR}_t = \frac{\left[\sum_{i=1}^{N} (X_{t_i} - \overline{X_t})^4\right]/(N-1)}{{\sigma_t}^4} \quad (6)$$

$$\mathrm{ENT}_t = -\sum_{i=1}^{N} p(X_{t_i}) \log_2 p(X_{t_i}) \quad (7)$$

Where X_t is the ultrasound signal in the time-domain, Xz_t corresponded with the positive values of X_t, $\overline{X}_t$ is the mean of each ultrasound signal in the time-domain, N is the number of elements of each ultrasound signal, t is the vector which registered the ultrasound signal's traveling time (μs), σ_t the standard deviation of each ultrasound signal in time-domain and $p(X_{ti})$ is the probability of the occurrence of the *i*-th amplitude value in the discretized time-domain ultrasound signal.

To assess the V_e (Eq. 8), the time of flight (TOF, μs) was firstly calculated. TOF (Eq. 9) was computed by using Eq. 9 following the energy threshold method (ETM) described by García-Pérez et al. (2019).

$$\mathrm{Vel} = \frac{L}{\mathrm{TOF}} \quad (8)$$

$$\mathrm{TOF} = \frac{(\mathrm{TOA} - T_r)}{ae} \quad (9)$$

Where L (m) is the thickness of every pixel of ultrasound images, which was measured using a digital caliper (Fig. 2F). TOA is the time-of-arrival (points) which measured the number of points corresponding to the time required for an ultrasound wave to propagate between the emitter transducer to the receiver transducer. The T_r (number of points) is the trigger signal and ae (100 Mpoints/s) is the acquisition speed.

2.5.2 Frequency domain analysis

The Fast Fourier Transform (FFT) was applied on the time-domain ultrasound signals of each pixel of 3D ultrasound images to obtain the ultrasound frequency spectrum (computed via "fft", MATLAB function). The phase-spectrum (phs, Eq. 10) served as the basis for the calculation of the energy-related ultrasound parameters in the frequency domain. Zero-order moment (M_0, MHz, Eq. 11) corresponded with the integral of the area under curve of the phs, which quantified its energy (García-Pérez et al., 2019).

The first-order moment (M_1, MHz) was calculated using Eq. 12. Plus, by dividing M_0 by the M_1, the center frequency of the phs (F_r) was then computed with Eq. 13. Additionally, the maximum peak of the frequency spectrum (MP, Eq 14) was also considered. As in section 2.5.1, the spectral-variance (VAR_{sp}, Eq. 15), spectral-skewness (SKE_{sp}, Eq. 16), spectral-kurtosis (KUR_{sp}, Eq. 17) and spectral-entropy (ENT_{sp}, Eq. 18) were also determined (Caesarendra and Tjahjowidodo, 2017).

$$\mathrm{phs} = |\mathrm{FFT}| \tag{10}$$

$$M_0 = \sum_{f=1}^{N=f_{FFT}} \mathrm{phs(f)}\ \Delta f \tag{11}$$

$$M_1 = \sum_{f=1}^{N=f_{FFT}} \mathrm{phs(f)}\ f^r \Delta f \tag{12}$$

$$F_r = \frac{M_1}{M_0} \tag{13}$$

$$\mathrm{MP} = \max(\mathrm{phs}) \tag{14}$$

$$\mathrm{VAR_{sp}} = \frac{\sum_{i=1}^{N}(\mathrm{phs_i} - \overline{\mathrm{phs}})^2}{(N-1)} \tag{15}$$

$$\mathrm{SKE_{sp}} = \frac{\left[\sum_{i=1}^{N}(\mathrm{phs_i} - \overline{\mathrm{phs}})^3\right]/(N-1)}{{\sigma_{sp}}^3} \tag{16}$$

$$\mathrm{KUR_{sp}} = \frac{\left[\sum_{i=1}^{N}(\mathrm{phs_i} - \overline{\mathrm{phs}})^4\right]/(N-1)}{{\sigma_{sp}}^4} \tag{17}$$

$$\mathrm{ENT_{sp}} = -\sum_{i=1}^{N} p(\mathrm{phs_i}) \log_2 p(\mathrm{phs_i}) \tag{18}$$

Where f is the vector of spectral frequencies (MHz), f_{FFT} is the maximal frequency obtained by using the FFT, r represents the order of the moment, $\overline{\mathrm{phs}}$ is the mean of each phs belonging to each pixel of ultrasound images, σ_{sp} the standard deviation of each phs belonging to each pixel of ultrasound images and $p(\mathrm{phs_i})$ is the probability of the occurrence of the *i*-th value in the discretized phs.

2.6 Statistical analysis

In order to assess the influence of bone fragments size/type and their location within chicken breast samples on the time and the frequency domain ultrasound parameters, a multifactor analysis of variance (ANOVA) was considered. Multifactor ANOVA models were adjusted independently for each computed-parameter. The mean comparisons were performed by using Fisher's Least Significant Difference (LSD) test with a 95% confidence interval. Further, an ANOVA test based on the L values was also performed to examine whether the inserted bone fragments affected the thickness of measured samples.

2.7 Mathematical modeling

In order to assess the feasibility of using the ultrasound images computed in the time-domain and frequency-domain to detect the BFs within chicken breast samples, three different approaches, namely time-domain approach (TDA), frequency-domain approach (FDA) and the integration of TDA and FDA: time-frequency domain approach (TFDA) were proposed (Fig. 4). TDA approach used all the parameters computed in the time-domain (PP, ENG, INT, V_e, VAR_t, SKE_t, KUR_t and ENT_t; Fig. 4C), while FDA used the ones in the frequency-domain (M_0, F_r, MP, VAR_{sp}, SKE_{sp}, KUR_{sp} and ENT_{sp}; Fig. 4C). Finally, the TFDA integrates all features extracted of both time and frequency domain in the same dataset (PP, ENG, INT, V_e, VAR_t, SKE_t, KUR_t and ENT_t, M_0, F_r, MP, VAR_{sp}, SKE_{sp}, KUR_{sp} and ENT_{sp}).

To improve the speed of analysis and to facilitate the modelling procedures, each image was unfolded as a feature vector (Achata et al., 2018). For this, each image was reshaped as a vector of i×j×k (Fig. 4.F-G). As an example, one image in TDA (i = 5 × j = 5 × k = 8) was rearranged from a 3D-matrix to a 1D-row vector of dimension 200 (Fig. 4F). Thus, each combination of approaches has matrices of unfolded images with different dimensions: TDA ([162-all data both control and OC] × i×j×k = 200), FDA ([162-all data both control and OC] × i×j×k = 175) and TFDA ([162-all data both control and OC] × i×j×k = 375).

As stated in the "Introduction section", addressing the challenge of detecting BF in the poultry meat industry is of critical importance. To this end, various data-driven modeling strategies should be explored to identify the most effective solution. In this study, two main strategies were proposed. The first involved an unsupervised method, applying a PCA model within the framework of MIA-based Multivariate Statistical Process Control (MSPC). The second approach aimed to assess whether detection performance could be improved through Latent Variable-based Machine Learning (LV-ML) techniques. The statistical modeling procedures for both the unsupervised and supervised strategies are illustrated in Fig. 5 and described in detail in sections 2.7.1 and 2.7.2.

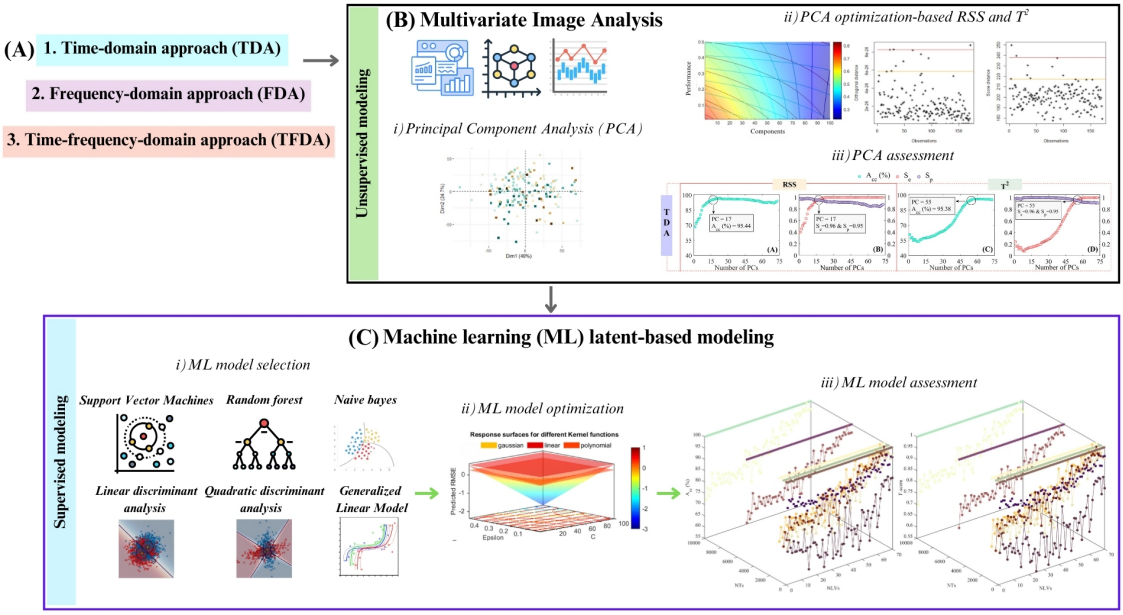


Fig. 5. Statistical modeling procedure used for both unsupervised and supervised strategies in the detection of bone fragments in poultry meat. Analysis of different approaches time-domain, frequency-domain and time-frequency-domain (A), unsupervised modeling using principal component analysis (PCA) and statistical optimization via Multivariate Statistical Process Control (MSPC) based Residual Sum Squares (RSS) and Hotelling's T-square (T^2) statistics (B) and supervised modeling and optimization based on Latent Variable-Machine Learning (LV-ML).

2.7.1 Unsupervised modeling

2.7.1.1 MIA based PCA-MSPC

The MIA procedure was followed according to reported by Colucci et al. (2019) and Verdú et al., (2025). The PCA model was employed to extract the latent eigenspace of unfolded control images (without BF). For this purpose, control data sets (81 control images) were randomly split into a segment of 90% of experimental data for model calibration (C_{cal}; 73 samples). The remaining samples, not included in model training, comprised 10 % of the control data (C_{EV}; 8 samples) and all OC images (81 samples), which were reserved for external validation. This validation aimed to assess the feasibility of the calibrated PCA model in detecting BF (Reis, 2015).

Firstly, the segment of data for PCA calibration (Eq. 19) was mean-centered and scaled to have unit variance. The PCA model used the Singular Value Decomposition (SVD) algorithm to extract the orthogonal latent eigenspace by compressing the image information into a LVs (Kruse et al., 2014). During the scaling process, both the mean and standard deviation vectors obtained from scaling process were saved as PCA control coordinates. Furthermore, the external validation dataset (C_{EV} + OC images) was scaled (C_{EVsc} + OC_{sc}) using the control coordinates and then projected onto the latent space by using the loadings (P^T_{cal}; Eq. 19-21)

from the control model. The residual sum of squares (RSS; Eq. 22-24) and the Hotelling's T-squared (T^2; Eq. 25-27) multivariate control statistics were computed.

To assess the detection capability of MIA-MSPC, a design of experiment (DoE) was formulated. This DoE consisted in analyze the influence of computing the control limit (CL) of both RSS_{cal} (Eq. 22) and T^2_{cal} (Eq. 25) from control images, using four different levels (90%, 95%, 97.5% and 99%) by percentile method (Vitale et al., 2016). Additionally, a limit augmentation (LA) of 0%, 50%, 75% and 100% was used to increase the decision boundary of computed CLs (Sinisterra-Solís et al., 2024).

$$\hat{C}_{cal} = t_{cal} P_{cal}^{T} \tag{19}$$

$$t_{EV} = C_{EVsc} P_{cal} \tag{20}$$

$$t_{OC} = OC_{sc} P_{cal} \tag{21}$$

$$RSS_{cal} = \sum_{a=1}^{A} (C_{cal} - t_{cal} P_{cal}^{T})_a^2 \tag{22}$$

$$RSS_{EV} = \sum_{a=1}^{A} (C_{EVsc} - t_{EV} P_{cal}^{T})_a^2 \tag{23}$$

$$RSS_{OC} = \sum_{a=1}^{A} (OC_{sc} - t_{OC} P_{cal}^{T})_a^2 \tag{24}$$

$$T^2{}_{cal} = \sum_{a=1}^{A} \frac{t^2_{cal_a}}{\sigma_{cal_a}^2} \tag{25}$$

$$T^2{}_{EV} = \sum_{a=1}^{A} \frac{t^2_{EV_a}}{\sigma_{cal_a}^2} \tag{26}$$

$$T^2{}_{OC} = \sum_{a=1}^{A} \frac{t^2_{OC_a}}{\sigma_{cal_a}^2} \tag{27}$$

where $\hat{C}_{cal}$ is the predicted control images based on the calibrated PCA model, t_{cal}, t_{EV} and t_{OC} are the scores of projected C_{cal}, C_{EV} and OC images in the PCA space and $\sigma^2_{cal,a}$ is the variance of each *a* computed LV. Statistical modeling and computing procedure was performed using MATLAB programming language.

2.7.1.2 Analysis of sample size in the detection of BF using MIA-MSPC

In order to elucidate the influence of the number of USI (both control and OC) used in the detection of BF within poultry samples, four different datasets varying in their number of images were used in the mathematical modeling. The total of experimental dataset (81 control images and 81 OC images, i.e. 162 images) was split into further three different ratios with 75% (61 control images and 60 OC images, equal to 121 images), 50% (41 control images

and 40 OC images, equal to 81 images) and 25% (21 control images and 21 OC images, equal to 42 images). Thus, TDA using the total of experimental data set (100%) consisted in 81 control images and 81 OC images, both with dimensions of i = 5 (number of points on the X axis)× j = 5 (number of points on the Y axis) × k = 8 (number of parameters computed in the time-domain). FDA was composed of 81 control images and 81 OC images, both with dimensions of i = 5 × j = 5 × k = 7 and TFDA integrating both TDA and FDA, as result 81 control images and 81 OC images, both with dimensions of i = 5 × j = 5 × k = 15. Furthermore, the other datasets (75%, 50% and 25%) were also considered for TDA, FDA and TFDA. These datasets were also unfolded as a vector of i×j×k. As a result each combination of approaches has matrices with different dimensions: TDA ([162 unfolded images; UNI-100%, 121 UNI-75%, 81 UNI-50% , 42 UNI-25%] × i×j×k = 200), FDA ([162 UNI-100%, 121 UNI-75%, 81 UNI-50% , 42 UNI-25%] × i×j×k = 175) and TFDA ([162 UNI-100%, 121 UNI-75%, 81 UNI-50% , 42 UNI-25%] × i×j×k = 375).

2.7.1.3 Analysis of first order statistics on ultrasound energy-magnitude-distribution parameters

Alternatively to all strategies described, a new approach was proposed. Feature extraction approach using first order statistics applied on ultrasound energy-magnitude-distribution parameters. This approach consisted in compute from each channel of the images (Fig. 4D-4E), the mean, standard deviation, range, skewness and kurtosis. Then, new approaches named feature-extraction time-domain approach (feTDA), feature-extraction frequency-domain approach (feFDA) and feature-extraction time-frequency domain approach (feTFDA) were also evaluated. As an example, in TDA, which used the PP, ENG, INT, V_e, VAR_t, SKE_t, KUR_t and ENT_t (81 control images and 81 OC images, both with dimensions of i = 5 × j = 5 × k = 8), for each channel, five first order statistics such as mean, standard deviation, range, skewness and kurtosis were calculated, as result feTDA considering all data has dimensions of 162 images × [5 first order statistics × 8 channels =40]. The same procedure was applied for FDA (162 × 35) and TFDA (162 × 75). This approach was proposed to reduce the dimensionality of the original high-dimensional data space. The rationale behind this method lies in the assumption that the presence of BF alters the statistical distribution of the ultrasound energy-magnitude-distribution parameters. Specifically, changes in first-order statistical descriptors (mean, standard deviation, range, skewness, and kurtosis) extracted from each channel are hypothesized to reflect these distributional shifts.

2.7.1.4 Classification performance and statistical validation of unsupervised modeling strategy

The capability of PCA in the detection of BFs was assessed varying from 1 LV to the maximal number of LVs for each approach (section 2.7.1.1, section 2.7.1.2 and section 2.7.1.3) and sample ratio. Both RSS and T^2 multivariate statistical control charts where employed to quantify the classification performance of the models, joint to the use of two corresponding

confusion matrices (CFM, Eq. 28; computed using "confusionmat" MATLAB function). In this way, control images from calibration and internal validation dataset with values of RSS and T^2 below the LA-control limits indicate true negatives (TN), whereas if they exceed the LA-control limits, they indicate false positives (FP). Regarding the OC images, true positive (TP) means OC images exceeding the LA-control limit, hence being correctly detected as samples with FBs, whereas OC images not exceeding the LA-control limit are considered as false negative (FN). The goodness of classification of each multivariate statistic was assessed by computing figures of merit such as the overall accuracy (A_{cc}, Eq. 29), sensibility (S_e, Eq. 30) and specificity (S_p, Eq. 31) (Craig et al., 2018).

$$\mathrm{CFM} = \begin{matrix} \\ \text{Predicted} \end{matrix} \begin{matrix} \begin{matrix} \text{Real} \\ \text{OC} \quad \text{Control} \end{matrix} \\ \begin{bmatrix} \text{TP} & \text{FP} \\ \text{FN} & \text{TN} \end{bmatrix} \end{matrix} \begin{matrix} \\ \text{OC} \\ \text{Control} \end{matrix} \tag{28}$$

$$A_{cc}(\%) = \frac{TP+TN}{TP+TN+FP+FN} * 100 \tag{29}$$

$$S_e = \frac{TP}{TP+FN} \tag{30}$$

$$S_p = \frac{TN}{TN+FP} \tag{31}$$

In order to optimize the PCA model based on RSS and T^2, a multi-objective optimization problem was formulated. The objective function was finding the number of LVs (optimal number of principal components, OPCs) of the PCA model which simultaneously maximize both the S_e and S_p. Thus, the response surface methodology (RSM) and desirability function (D_e) was performed (Kumar et al., 2019; Yolmeh and Jafari, 2017). In this context, the same desirability value was assigned to both Se and Sp, assuming equal importance for correctly identifying positive and negative cases. This balanced approach ensures that the optimization process does not favor one metric over the other, leading to a more robust and generalizable model (Costa and Lourenço, 2023).

Optimization processes were carried out using the "fmincon" MATLAB function. All calculations were repeated 100 times to estimate the influence of considering different randomly partitions of the control matrices to calibrate the PCA model and further its influence on the detection of BFs. For the selection of the number of LVs of the optimized PCA model in each approach maximizing the classification performance, a multifactor ANOVA model considering the random data partition (as a blocking factor), the control limits and their LA and the goodness of classification metrics as responses (A_{cc}, S_e, and S_p) was carried out. All multifactor ANOVA models (sections 2.6 and 2.8) were subjected to residual validation (Marques et al., 2020). This process involved conducting different tests on the residuals to assess normality (Shapiro-Wilk's test and normal probability plot), independence (Ljung-Box's test), and homoscedasticity (multiple linear regression-MLR on square residuals). Hypothesis tests and fulfillment of statistical assumptions were assessed at a

confidence level of 95%. The statistical analysis was conducted using STATGRAPHICS Centurion XVIII (Manugistics, Inc., Rockville, MD, USA).

2.7.2 Supervised machine learning latent-based classification models

As already mentioned in Section 2.7, the second approach aimed to explore whether applying machine learning (ML) to the MIA-PCA framework could improve the detection of BF in poultry meat. To this end, an expanded and more in-depth modeling procedure was employed, including analysis, multi-objective optimization, and variable selection strategies. Based on the previously extracted latent eigenspace from the PCA model, the images projected onto the control model for calibration (t_{cal}), control images for external validation (t_{EV}), and OC images (t_{OC}) were combined into a single matrix for each approach (TDA, FDA and TFDA). These matrices served as the basis for calibrating various ML techniques using the extracted LVs. The LV-ML techniques included: SVM, RT, RF, NB, LDA, QDA, and GLM. These techniques were applied in binary classification mode to mathematically distinguish between control (t_{cal}+t_{EV}) and OC (t_{OC}) USIs, based on the extracted latent eigenspace (see section 2.7.1). For this purpose, the scores (both from control and OC images) were used as input (regressors) of these ML classifiers.

An additional approach emerge since TFDA is the combination of two different approaches with different number channels (Fig. 4D) and these channels belong to two different domains, the temporal domain and frequency domain. In the previous section 2.7.1, TDA and FDA were integrated into same matrix without any preprocessing tool, wherein the naturally related these descriptors and the number of in each group were not considered. In order to elucidate whether this combination led to an any improvement in the detection of BF, two types of block scaling; "Hard" and "Soft" were applied to TFDA (Eriksson et al., 2016). Thus, TFDA block-scale hard (TFDABH, Eq. 32) and TFDA block-scale soft (TFDABS, Eq. 33) were also considered as additional approaches.

$$\mathrm{TFDABH} = \left[\frac{\frac{x_{TDA} - \hat{x}_{TDA}}{\sigma_{TDA}}}{\sqrt{nVTDA}} \quad \frac{\frac{x_{FDA} - \hat{x}_{FDA}}{\sigma_{FDA}}}{\sqrt{nVFDA}} \right] \tag{32}$$

$$\mathrm{TFDABS} = \left[\frac{\frac{x_{TDA} - \hat{x}_{TDA}}{\sigma_{TDA}}}{\sqrt[4]{nVTDA}} \quad \frac{\frac{x_{FDA} - \hat{x}_{FDA}}{\sigma_{FDA}}}{\sqrt[4]{nVFDA}} \right] \tag{33}$$

Where X_{TDA} and X_{FDA} are the matrices formed by the UNI in TDA and FDA respectively, $\hat{X}_{TDA}$, σ_{TDA}, and nVTDA, and $\hat{X}_{FDA}$, σ_{FDA}, and nVFDA are the mean, standard deviation and number of variables in X_{TDA} and X_{FDA}, respectively.

As the detection of BF was primarily an experimental data-driven task, exploring various configuration of each ML technique became imperative in order to properly address the experimental variability of hyperparameters in the detection of BF (Collazos-Escobar et al.,

2023b). Thus, in order to optimize the hyperparameters of each LV-ML technique, different multilevel factorial DoEs were formulated.

In the case of SVM, a DoE ($5^1 6^1 2^1 3^1 71^1$) using all approaches (TDA, FDA, TFDA, TFDABH and TFDABS), 6 kernel functions (KF; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), two types (C-svc and nu-svc), regularization parameter (C; 100, 500.5, and 1000) and NLVs (1, 2, 3, 4,…71) was formulated (Table 1). Computational procedure using all ML techniques was carried out using the statistical software R (R Core Team, 2025). Further, model fitting via SVM was performed using the R-package developed by Karatzoglou et al. (2024).

Table 1. Summary of the Support Vector Machine (SVM) configuration used for supervised modeling of bone fragment detection in poultry meat. The table reports the R function, associated R packages, selected hyperparameters, tuning ranges, and the design of experiments (DoE) applied in the analysis.

R-function	R-package	Approach	HyP	Tuning range	DoE
ksvm	kernlab	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	KF	i) rbfdot ii) polydot iii) laplacedot iv) vanilladot v) besseldot vi) anovadot	($5^1 6^1 2^1 3^1 71^1$): 12780 models
			Type	i) C-svc ii) nu-svc	
			C	i) 100 ii) 500.5 iii) 1000	
			NLVs	i) 1 ii) 2 iii) 3 …… lxxi) 71	

TDA (time-domain approach) frequency-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), HyP (hyperparameter), KF (kernel function), C (regularization parameters) and NLVs (number of latent variables).

For RF a DoE ($5^1 6^1 71^1$) was set up considering all approaches (TDA, FDA, TFDA, TFDABH and TFDABS), different number of trees (NTs ;1= DTe, 50, 500, 1000, 5000 and 10000) and number of latent variables (NLVs; 1, 2, 3, 4,…71; Table 2) (Liaw and Wiener, 2002), whereas NB (Table 3) considered a DoE ($5^1 2^1 71^1$) which employed all approaches (TDA, FDA, TFDA, TFDABH and TFDABS), application of Laplace Smoothing (LS; 0-No application and 1-application) and NLVs (1, 2, 3, 4,…71; Table 3) (Meyer et al., 2024).

Table 2. Summary of the Random Forest (RF) configuration used for supervised modeling of bone fragment detection in poultry meat. The table reports the R function, associated R packages, selected hyperparameters, tuning ranges, and the design of experiments (DoE) applied in the analysis.

R-function	R-package	Approach	HyP	Tuning range	DoE
randomForest	randomForest	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	NTs	i) 1 = DTe ii) 50 iii) 500 iv) 1000 v) 5000 vi) 10000	$(5^1 6^1 71^1)$: 2130 models
			NLVs	i) 1 ii) 2 iii) 3 …… lxxi) 71	

TDA (time-domain approach) frequency-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), HyP (hyperparameter), NTs (number of trees), DTe (Decision tree) and NLVs (number of latent variables).

Table 3. Summary of the Naïve Bayes (NB) configuration used for supervised modeling of bone fragment detection in poultry meat. The table reports the R function, associated R packages, selected hyperparameters, tuning ranges, and the design of experiments (DoE) applied in the analysis.

R-function	R-package	Approach	HyP	Tuning range	DoE
naiveBayes	e1071	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	LS	i) 0 ii) 1	$(5^1 2^1 71^1)$: 710 models
			NLVs	i) 1 ii) 2 iii) 3 …… lxxi) 71	

TDA (time-domain approach) frequency-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), HyP (hyperparameter), LS (Laplace Smoothing) and NLVs (number of latent variables).

Regarding LDA (DoE= $5^1 71^1$), QDA (DoE=$5^1 55^1$) and GLM (DoE= $5^1 71^1$), all of these DoEs considered all approaches (TDA, FDA, TFDA, TFDABH and TFDABS) and for LDA and GLM a NLVs (1, 2, 3, 4,…71; Table 4). In the case of QDA, a reduced NLVs (from 1 to 55) was employed because the independent estimation and inversion of the covariance matrix for each class (control and OC) became ill-conditioned when using a higher number of latent variables. This is due to the fact that, in QDA, each class requires the computation of its own covariance matrix, which must be invertible. When the number of variables approaches or

exceeds the number of observations in a class, the covariance matrix becomes singular or nearly singular, preventing model computation (Siqueira et al., 2017). Thus, in order to avoid this issue, a lower number of NLVs (up to 55) was selected. Although up to 60 NLVs could theoretically be used given that the class size was 60 (60 control and 60 OC in training 75% dataset) a more conservative cutoff was applied to ensure numerical stability and avoid potential problems related to near-singular covariance matrices during QDA computation. Computing modeling of LDA, QDA and GLM was performed using different R-function-packages (Table 4). LDA and QDA were fitted using the method developed by Venables & Ripley. (2002) while GLM was calculated using the *stats* package R Core Team (2025).

Table 4. Summary of the Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Generalized Linear Model (GLM) configuration used for supervised modeling of bone fragment detection in poultry meat. The table reports the R function, associated R packages, selected hyperparameters, tuning ranges, and the design of experiments (DoE) applied in the analysis.

R-function	R-package	Approach	HyP	Tuning range	DoE
LDA: *lda*	MASS	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	NLVs	i) 1 ii) 2 iii) 3 …… lxxi) 71	$(5^1 71^1)$: 355 models
QDA: *qda*	MASS	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	NLVs	i) 1 ii) 2 iii) 3 …… lv) 55	$(5^1 55^1)$: 275 models
GLM: *glm*	stats	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	NLVs	i) 1 ii) 2 iii) 3 …… lxxi) 71	$(5^1 71^1)$: 355 models

TDA (time-domain approach) frequency-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), HyP (hyperparameter) and NLVs (number of latent variables).

2.7.2.1 Variable selection based RF model

In order to explore strategies for selecting latent variables (LVs) that maximize the predictive performance of the LV-machine learning (LV-ML) models in the detection of BF in chicken breast samples, the feasibility of using the Mean Decrease Accuracy (MDA) criterion for variable selection (VS) within a Random Forest (RF) model was evaluated (Collazos-Escobar et al., 2024). Consequently, the Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) approach was also considered (Fig. 6). To achieve this, during the model tuning process of SVM, RF, NB, LDA, QDA and GLM techniques across all

dataset (TDA, FDA, TFDA, TFDABH and TFDABS), a PCA model was first used to extract the latent eigenspace (considering in this case the maximal number of LVs = 71; Tables 1 to 4). Subsequently, a RF model using 10000 NTs was calibrated on the extracted latent space. When this RF model was trained on the training dataset (see Section 2.7.2.2), it achieved an A_{cc} of 100% in classifying control and OC samples, which is indicative of an overfitting phenomenon in ML models (see "Results and discussion" section). While overfitting is a known limitation in ML models (de Andrade et al., 2020), it can be strategically leveraged during the VS process to identify features that strongly contribute to model performance. Nevertheless, the effectiveness of the selected variables must be rigorously tested using an external validation dataset to ensure generalizability (Collazos-Escobar et al., 2023a). Thus, using this calibrated RF, the MDA criterion was used to rank the most important first 30 LVs to better differentiation of images with and without BFs. Subsequently, these ranked LVs were used in the computer modeling of RF-VS-LV-SVM, RF-VS-LV-RF, RF-VS-LV-NB, RF-VS-LV-LDA, RF-VS-LV-QDA and RF-VS-LV-GLM (Table 5).

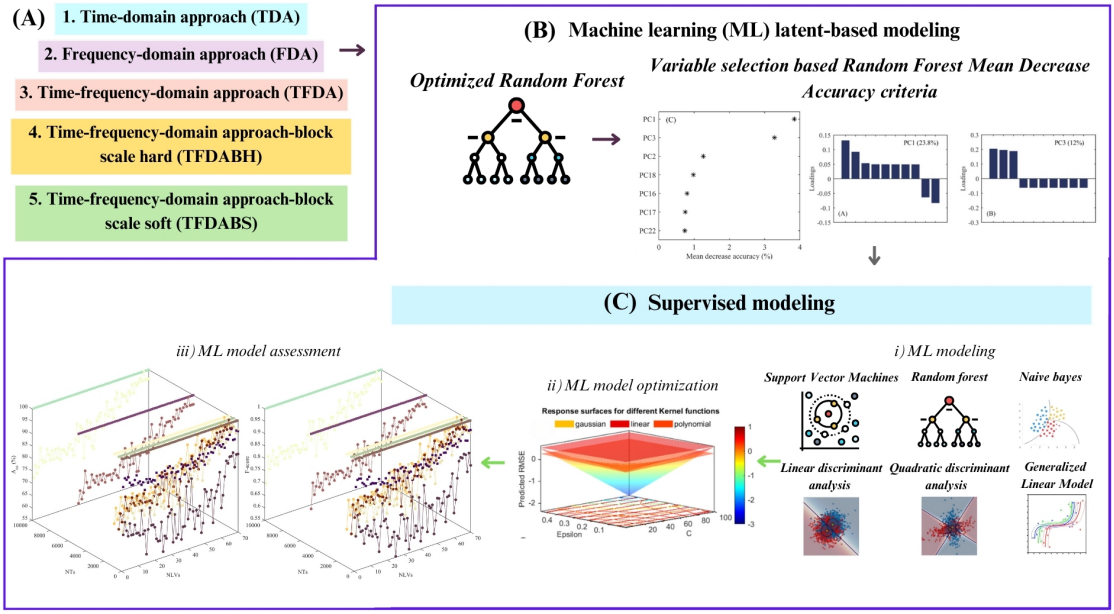


Fig. 6. Statistical modeling procedure used to evaluate the feasibility of variable selection (VS) based on mean decrease in accuracy from a Random Forest (RF) model for tuning and optimizing supervised machine learning (ML) techniques. The analysis includes: time-frequency-domain approaches (A), RF-based variable selection (RF-VS; B) and supervised modeling and optimization based on selected latent variables using machine learning (RF-VS-LV-ML; C).

Table 5. Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) approach. Design of Experiments (DoE) for RF-VS-LV-Support Vector Machine (RF-VS-LV-SVM), RF-VS-LV-Random Forest (RF-VS-LV-RF), RF-VS-LV-Naïve Bayes (RF-VS-LV-NB), RF-VS-LV-Linear Discriminant Analysis (RF-VS-LV-LDA), RF-VS-LV-Quadratic Discriminant Analysis (RF-VS-LV-QDA) and RF-VS-LV-Generalized Linear Model (RF-VS-LV -GLM).

Technique	Approach	HyP	Tuning range	DoE
RF-VS-LV-SVM	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	KF	i) rbfdot ii) polydot iii) laplacedot iv) vanilladot v) besseldot vi) anovadot	$(5^1 6^1 2^1 3^1)$: 180 models
		Type	i) C-svc ii) nu-svc	
		C	i) 1 ii) 500.5 iii) 1000	
RF-VS-LV-RF	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	NTs	i) 1 = DTe ii) 50 iii) 500 iv) 1000 v) 5000 vi) 10000	$(5^1 6^1)$: 30 models
RF-VS-LV-NB	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS	LS	i) 0 ii) 1	$(5^1 2^1)$: 10 models
RF-VS-LV-LDA	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS		—	(5^1): 5 models
RF-VS-LV-QDA	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS		—	(5^1): 5 models
RF-VS-LV-GLM	i) TDA ii) FDA iii) TFDA iv) TFDABH v) TFDABS		—	(5^1): 5 models

TDA (time-domain approach) frequency-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), HyP (hyperparameter), KF (kernel function), C (regularization parameters), NTs (number of trees), DTe (Decision tree) and LS (Laplace Smoothing).

2.7.2.2 Classification performance and statistical validation of supervised modeling strategy

The classification performance of all LV-ML (section 2.7.2) and RF-VS-LV-ML (section 2.7.2.1) was also assessed via the CFM (Eq. 28) and all of the figures of merit derived from the CFM (A_{cc}, S_e and S_e; section 2.7.1.4). Additionally, to gain a deeper understanding of model behavior and robustness, complementary classification performance metrics were computed, including precision (P_r; Eq. 34), recall (R_e; Eq. 35), F-score (F_s; Eq. 36) (Galdón-Navarro et al., 2018), the area under the Receiver Operating Characteristic curve (AUC_{ROC}) (Debón and García-Díaz, 2012), and the Matthews correlation coefficient (MCC) (El Zein et al., 2025). These metrics provide a more comprehensive assessment of the models' classification capabilities, which is particularly relevant in the context of BF detection within chicken breast samples.

$$P_r = \frac{TP}{TP+FP} \tag{34}$$

$$R_e = \frac{TP}{TP+FN} \tag{35}$$

$$F_s = 2 \times \frac{P_r \times R_e}{P_r + R_e} \tag{36}$$

$$AUC_{ROC} = \text{trapz}(ROC) \tag{37}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)\,(TP+FN)\,(TN+FP)(TN+FN)}} \tag{38}$$

Where "trapz" is the area-under-curve of ROC curve. AUC_{ROC} was computed using the *performance-"auc"* R-function from the *performance* R-package Lüdecke et al. (2021). In addition, the *confusionMatrix* R-function from the *Caret* R-package (Kuhn, 2008) was used to compute the CFM in this section.

In the case of both LV-ML and RF-VS-LV-ML, a modeling strategy based on using a segment of 75% of the experimental dataset for models' training and the remaining 25% to calculate their predictive power was used (Debón and García-Díaz, 2012). In this sense, all approaches (TDA, FDA, TFDA, TFDABH and TFDABS), the experimental images (section 2.72) were randomly split 100 times (repeated hold-out strategy) by (Ruiz de Miras et al., 2024) in two data sets, for training dataset the 121 images (60 ± 1 of control and 60 ± 1 of OC) and 41 images (20 ± 1 of control and 20 ± 1 of OC) for validation purposes. Additionally, during the computational procedure all models were trained and validated using identical datasets. This ensured that all models underwent training and validation with the same data split in each partition, eliminating any bias linked to the use of different training and validation data (Collazos-Escobar et al., 2023b). The goodness of fit of LV-ML and RF-VS-LV-ML was calculated for both training and validation datasets. Additionally, the DoEs purposed in Tables 1 to 4 were replicate 100 times using in each iteration the same partition

for training and the remaining for validation across all LV-ML and RF-VS-LV-ML, As a result, for LV-SVM a total of 1'278.000 runs (12780 models from DoE; Table 1 × 100 times), LV-RF a total of 213,000 runs (2130 models from DoE; Table 2 × 100 times), LV-NB a total of 71,000 runs (710 models from DoE; Table 3 × 100 times), LV-LDA a total of 35,500 runs (355 models from DoE; Table 4 × 100 times), LV-QDA a total of 27,500 runs (275 models from DoE; Table 4 × 100 times) and LV-GLM a total of 35,500 runs (355 models from DoE; Table 4 × 100 times). Meanwhile, in the RF-VS-LV-ML the number of runs were: 18000 runs for RF-VS-LV-SVM (180 models from DoE; Table 5 × 100 times), 3000 runs for RF-VS-LV-RF (30 models from DoE; Table 5 × 100 times), 1000 runs for RF-VS-LV-NB (10 models from DoE; Table 5 × 100 times) and 500 runs for each of RF-VS-LV-LDA, RF-VS-LV-QDA and RF-VS-LV-GLM (5 models from DoE; Table 5 × 100 times). Additionally, computation times (CT, s) were recorded using the system time R-function, both to measure the duration required to train the ML techniques and to evaluate their computational cost. All computations were performed on an Intel Core i7 processor running at 2.2 GHz with 16 GB of RAM.

2.7.2.3 Multi-objective hyperparameters optimization of LV-ML and RF-VS-LV-ML models

In order to optimize the hyperparameters of each ML techniques in both LV-ML and RF-VS-LV-ML, a multi-objective strategy based on PLS modeling was applied. Partial Least Square Regression (PLSR) technique was used to simultaneously assess the influence of hyperparameters combination based-DoEs (Tables 1 to 5) of each technique on all figures of merit in a multivariate way. The application of PLSR on this task depicts an important advance in ML model's optimization compared to the conventional univariate way (optimization strategy based-MLR/ANOVA on one goodness of fit metric such as mean-square error; MSE or coefficient of determination; R^2), since PLSR allowed to model the latent relationship between both regressors and responses simultaneously and to find a latent eigenspace wherein to maximize the covariance of projected input and response variables (Duma et al., 2024). Thus, the calibrated/validated PLSR model can be further used within an optimization framework (Paris et al., 2024).

The first step in this approach consisted in organizing the statistical results obtained from each ML into a datasets structured to be further used in the modeling procedure via PLSR. In this sense, in the LV-SVM, the data approaches (TDA, FDA, TFDA, TFDABH and TFDABS), type (C-svc and nu-svc), KF (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), C (1, 500.5, and 1000) factors from DoE (Table 1) were transformed into dummy variables (0-ausence and 1-presence of each condition) to facilitate the mathematical PLSR modeling since these factors are naturally categorical. Additionally, replication of runs (100 random partition of the experimental datasets) was also considered as categorical input. Regarding the NLVs (1 to 71; Table 1) was considered as continuous numerical variable. All of these variables were considered as regressors the model regressors in the X matrix space. The Y response variable space was formed by the A_{cc}, S_e, S_p, P_r, R_e, F_s, AUC_{ROC} and MCC

for both training (A_{ccT}, S_{eT}, S_{pT}, P_{rT}, R_{eT}, F_{sT}, AUC_{ROCT} and MCC_T) and validation (A_{ccV}, S_{eV}, S_{pV}, P_{rV}, R_{eV}, F_{sV}, AUC_{ROCV} and MCC_V). The idea behind organize the response matrix composed of classification performance metrics for training and validation datasets was to elucidate the effect of regressors on both dataset and in the further optimization process, use the model to find the best LV-SVM model (from those 12780 models) to simultaneously maximize the classification performance of control and OC images from training and validation datasets.

For all LV-RF, LV-NB, LV-LDA, LV-QDA and LV-GLM, the dataset approaches (TDA, FDA, TFDA, TFDABH and TFDABS) and replication of runs (100 random partition of the experimental datasets) were set as dummy variables and NLVs (1 to 71; Table 1) was considered as numerical variable. In the case of LV-RF and LV-NB, the NTs and LS (Table 2 and Table 3) were also converted into dummy variables. Therefore, the DTe, NTs=50, NTs=500, NTs=1000, NTs=5000 and NTs=10000 and LS= 0 and LS=1 were considered in RF and NB as model regressors, respectively. In the case of RF-VS-LV-ML, the data approaches (TDA, FDA, TFDA, TFDABH and TFDABS), hyperparameters belong to each ML technique and the replication runs were set as dummy variables. The main difference between LV-ML and RF-VS-LV-ML is that, in the RF-VS-LV-ML approach, the tunned RF-NTs=10000 (as detailed in section 2.7.2.1) determined the most important LVs to maximize the classification performance of control and OC images for the training dataset. For that reason, the NLVs is not a factor to be considered in the PLSR modeling and it is not a factor to be further optimized. Once the datasets were organized, a PCA model was performed prior to the PLSR model calibration to verify that all response variables were correlated and to ensure the successful calibration of the PLSR model. Otherwise, if the response variables were not correlated, it would be necessary to calibrate a separate PLSR model for each group of correlated variables. Subsequently, the datasets were independently modeled using a PLSR approach. To achieve this, the X (regressors, Eq. 39) and Y (response variables, Eq. 40) matrices were mathematically decomposed by finding the maximum covariance and linear relationships between the scores T and U (Eq. 39 to 40) (Brendel et al., 2020). The objective function in a PLSR model considered maximizing the covariance between X and Y spaces via computation of latent variables based on both data spaces. The decision variables obtained were the W and C regression coefficients (weights) and their regressors (scores T and U) (Barrera Jiménez et al., 2023).

$$X_{pred} = TW^T + E \tag{39}$$

$$Y_{pred} = UC^T + F \tag{40}$$

Where X_{pred} and Y_{pred} are the predicted X and Y matrices, T and U are the score matrices, W and C are the weight matrices, E and F are the residual matrices of X and Y data matrices, respectively. The selection of the optimal number of PLSR components (ONPLSR) was carried out by using a K-fold-Cross-Validation (K=5) in order to avoid the overtraining of the PLSR model and the Variable Importance for the Projection (VIPs) metric were used to

quantify the global importance of the TDA, FDA, TFDA, TFDABH and TFDABS, hyperparameters of each ML technique, NLVs and random partition (RS) of experimental data set (Craig et al., 2018). X-variables with VIP values below 0.5 were excluded from the model, subsequently a screened PLSR model was obtained (Kahrıman and Liland, 2021). Additionally, both the RSS and T^2 were used to detect and remove outlier observations from the experimental data set (Sánchez-Jiménez et al., 2023). The goodness of fit of PLSR models was assessed using the coefficient of determination for training (R^2; Eq. 41) and K-Fold cross validation (Q^2; Eq. 42) datasets and the root mean square error in both training ($RMSE_{TR}$; Eq. 43) and for K-Fold cross validation set ($RMSE_{CV}$; Eq. 44).

$$R^2(\%) = 100 - \frac{\sum_{i=1}^{n}(Y - Y_{pred})^2}{\sum_{i=1}^{n}(\overline{Y} - Y_{pred})^2} \tag{41}$$

$$Q^2(\%) = 100 - \frac{\sum_{i=1}^{n}(Y_{CV} - Y_{CVpred})^2}{\sum_{i=1}^{n}(\overline{Y_{CV}} - Y_{CVpred})^2} \tag{42}$$

$$RMSE_{TR} = \sqrt{\frac{\sum_{i=1}^{n}(Y - Y_{pred})^2}{n}} \tag{43}$$

$$RMSE_{CV} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{CV} - Y_{CVpred})^2}{n_{CV}}} \tag{44}$$

Where Y_{CV} and Y_{CVpred} are the experimental and predicted data matrices for the K-fold cross validation dataset used in the PLSR model calibration and n and n_{CV} are the number of observations in both training and K-fold cross validation datasets. All parameter estimation in the PLSR modeling was conducted using the Non-Iterative Partial Least Square (NIPALS) algorithm implemented in the *opls* R-function of the OPLS R-package (Thévenot et al., 2015). Once the PLSR models were independently calibrated and validated in each LV-ML and RF-VS-LV-ML approaches (Fig. 7), they were used as digital model in an optimization problem.

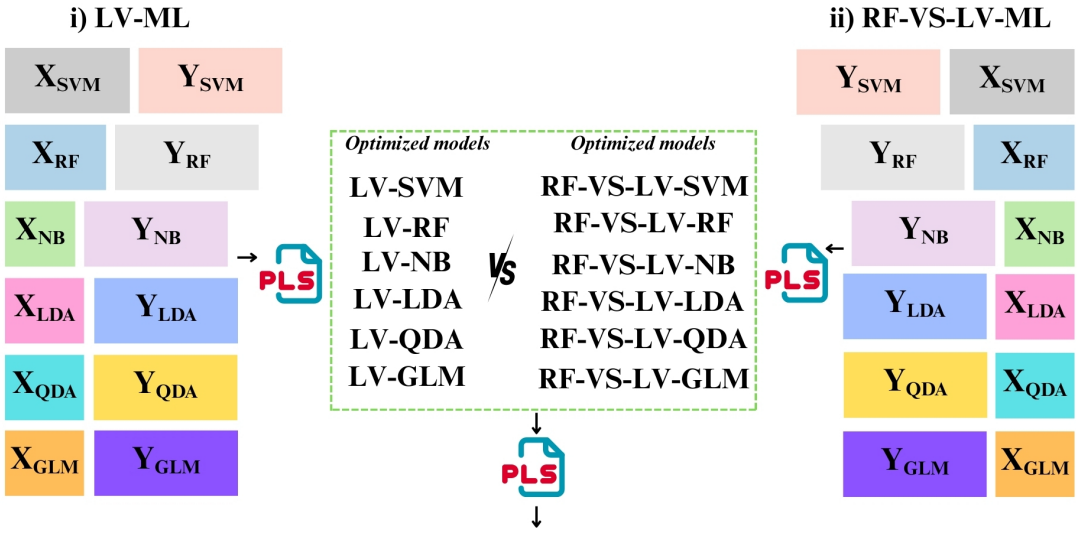


Fig. 7. Statistical modeling procedure based Partial Least Squares Regression (PLSR) used in the multi-objective optimization of Latent Variable-Machine Learning (LV-ML) and Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) models.

The multi-objective optimization process using the PLSR models was carried out in two consecutive steps. The first step consisted of independently determining the best hyperparameter configuration for each ML technique (LV-ML and RF-VS-LV-ML). In this phase, a total of 12 PLSR models (Fig. 7) were calibrated. These models were then used to identify the best dataset approach (TDA, FDA, TFDA, TFDABH, and TFDABS), hyperparameter configuration (KF, type, C, NTs, and LS), and NLVs for each technique in order to maximize the classification performance for both control and OC images in the training and validation datasets.

The second step involved calibrating a new PLSR model to compare all optimized LV-ML and RF-VS-LV-ML models and to determine which one achieved the highest classification performance for both control and OC images. In this step, the dataset summarized the previously optimized models, encoded as dummy variables in the X-space, with all figures of merit as responses (Y-space). Additionally, the CT was included as a new response variable. The goal using this PLSR model was therefore to identify the best optimized model that maximized classification performance in both the training and validation datasets while minimizing CT during the training process. Further, to select the optimized model that maximized classification performance while minimizing CT, two different ANOVA models were performed on A_{cc} and CT as complementary tools to the PLSR-based multi-objective decision optimization.

In the multi-objective optimization process using the calibrated PLSR models, equal weights were assigned to the sixteen response variables considered in the Y-space (A_{ccT}, S_{eT}, S_{pT}, P_{rT}, R_{eT}, F_{sT}, AUC_{ROCT}, MCC_T, A_{ccV}, S_{eV}, S_{pV}, P_{rV}, R_{eV}, F_{sV}, AUC_{ROCV} and MCC_V), to guarantee that all classification performance metrics contributed equally to the overall optimization criterion. The search space of the better conditions (dataset approach, hyperparameter configuration of ML techniques and NLVs) in the optimization process was bounded by

predefined lower and upper limits of X-matrices, containing the dummy variables associated with the dataset approaches (TDA, FDA, TFDA, TFDABH and TFDABS) and ML hyperparameters (KF, type, C, NTs and LS) and a continuous variable related with NLVs (ranged between 1 to 71 LVs). For each candidate X-vector input in its raw (unscaled form), standardization was performed using the mean and standard deviation parameters extracted from the PLSR model calibration. Then, the response Y-vector corresponding to this standardized X-vector was computed using the coefficients (W/C) matrices obtained from the calibrated PLSR models.

Since the PLSR models allowed to a multivariate prediction of all response variables (A_{ccT}, S_{eT}, S_{pT}, P_{rT}, R_{eT}, F_{sT}, AUC_{ROCT}, MCC_T, A_{ccV}, S_{eV}, S_{pV}, P_{rV}, R_{eV}, F_{sV}, AUC_{ROCV} and MCC_V) per each X-vector candidate solution, the objective function (FO) used in the optimization process required a scalarized objective function to compare alternatives. This was defined as the weighted sum of the predicted responses (Eq. 45).

$$FO(X) = \sum_{j=1}^{m} w_j * Y_{pred}(X) \tag{45}$$

Where w_j is the relative importance assigned to the *j-th* Y-response. As we previously mentioned, equal weights were used to ensure that each classification performance metric contributed identically to the FO. This fact was carried out to avoid the potential bias caused by the differences in scale among response variables and encoded the intended trade-off structure, such that improvements in any metric would have a proportional impact on FO. Nevertheless, the weights assigned to each response variable can be adjusted across different multi-objective scenarios to reflect alternative assessment of the contribution of these response variables to FO. Thus, a systematic sensitivity analysis constitutes a relevant matter for further research.

When CT is included alongside performance metrics to be maximized, Eq. (45) is generalized by introducing a direction coefficient (S_j; Eq. 46).

$$FO_{CT}(X) = \sum_{j=1}^{m} w_j * S_j * Y_{pred}(X) \tag{46}$$

Where $S_j \in \{-1, 1\}$, with $S_j=1$ for responses to be maximized and $S_j= -1$ for responses to be minimized.

The discrete structure in this multi-objective problem required strict combinatorial constraints, within the dataset block (TDA, FDA, TFDA, TFDABH, TFDABS) exactly one variable could be active (value equal to one after rounding), and similarly for specific hyperparameter blocks such as KF, type, C, NTs, LS. Any violation of these one-hot conditions (for instance activating multiple datasets or multiple KFs, C, NTs and/or LS simultaneously) was considered infeasible solution. To enforce these restrictions within a continuous multi-objective optimization framework, a large penalty term was added whenever a one-hot constraint was violated.

Additionally, in order to avoid this fact, a custom initialization function (auxiliary function) generated feasible seed vectors that inherently satisfied all structural constraints by randomly selecting one active position in (TDA, FDA, TFDA, TFDABH, TFDABS), randomly selecting one active position in hyperparameters, and assigning a random integer to the NLVs between 1 and 71 was programed. For instance, in the case of LV-SVM (used here as an illustrative X-vector example), the auxiliar function allowed us to randomly generate a vector which contains a selected data approach (TDA = 0, FDA =1, TFDA=0, TFDABH=0 and TFDABS=0), KF (rbfdot=1, polydot=0, laplacedot=0, vanilladot=0, besseldot=0 and anovadot=0), C parameter (C_{100}=0, $C_{500.5}$=1 and C_{1000}=0) and NLVs=8. Thus, a condition of LV-SVM using the FDA approach, using a rbfdot-KF, C parameter of 500.5 and 8 NVLs was predicted.

Optimization process was performed using the Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization method, which efficiently handles bound-constrained optimization problems with continuous and discrete variables. A total of 100 independent optimization runs were conducted, each initialized from a distinct feasible seed (generated using the auxiliar function). Upon completion, the solution with the highest score (defined as the weighted sum of predicted responses with no penalties) was selected as the optimal configuration, representing the best attainable balance across all classification performance metrics under the imposed structural and operational constraints. All multi-objective computational assessments were conducted using the *optim* R-function from the *stats* R-package.

3. Results and discussion

3.1 Influence of BF on the ultrasound signals in the time and frequency domains

The influence of BFs on contact ultrasound signals in the time-domain is illustrated in Figs. 8 and 9, while the results for the frequency-domain are depicted in Fig. 10. Figs. 8 and 9 show the contact ultrasound signals obtained from the center point (point 13, Fig. 1G), along with their corresponding OC ultrasound images (using PP image as example) wherein the different BFs were inserted into chicken breast pieces. Meanwhile, Fig. 10 depicts the phs of control and OC also obtained from the center point of control and OC ultrasound images, respectively.

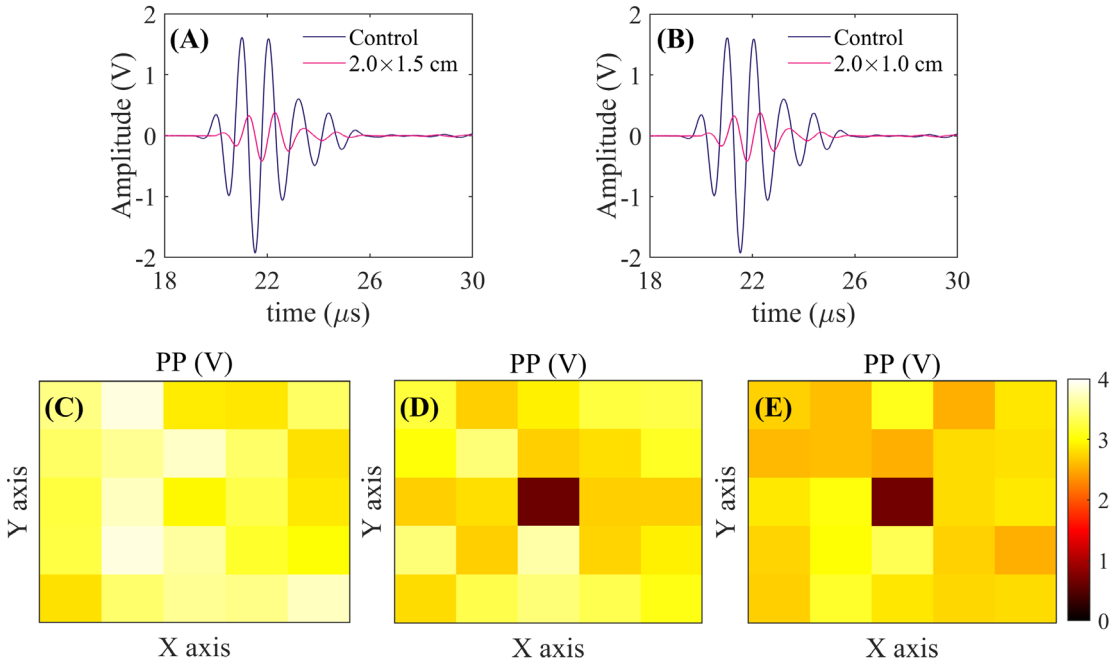


Fig. 8. Ultrasound signals and example of PP (peak-to-peak) image for detection of bone fragments within the center of chicken breast samples. Bone fragments of size 2.0 × 1.5 cm (A, D) and size of 2.0 × 1.0 cm (B, E), control sample image (C).

A consistent trend was observed in all cases: the BF presence disturbed the time-domain control ultrasound signals and the control frequency-domain spectra. In the case of temporal domain, the BFs of 2.0 × 1.5 cm (Fig. 8A) and 2.0 × 1.0 cm (Fig. 8B) promoted an important decrease in signal amplitude. Moreover, the presence of these BFs led to a pronounced reduction in the maximum peak of the phs (for 2.0 × 1.5 cm, Fig. 10A, and 2.0 × 1.0 cm, Fig. 10B). Conversely, for the BFs of 1.5 × 0.3 cm (Fig. 9A and Fig. 10C), 1.0 × 0.3 cm (Fig. 9B and Fig. 10D), and 0.5 × 0.3 cm (Fig. 9C and Fig. 10E), the influence of the bone fragments led to less pronounced drops in the maximum amplitude of temporal ultrasound signals and the maximum peak of the phs, compared to the samples containing BFs of larger size. Additionally, the control ultrasound images (Fig. 8C) and the pixels around point 13 of OC images (control pixels) evidenced PP values between 2.8 to 4 V (color bar from orange to yellow), meanwhile in the OC ultrasound images (Fig. 8 and Fig. 9): the higher the size of BFs the lower the PP values (ranging from 0.5 V, clear-red to 2 V, dark-red).

Ultrasound waves are partially reflected and transmitted when they are passing through materials with different acoustical impedance (defined by density and velocity). As regards the detection of FBs such as BFs, glass or metal fragments within foods, a strong energy reflection and scattering effects between the food material and the FBs interfaces are expected (Cho and Irudayaraj, 2003). Therefore, the physical properties, structure and nature of FBs determine the energy attenuation of ultrasound wave and its velocity (Fariñas et al., 2021a).

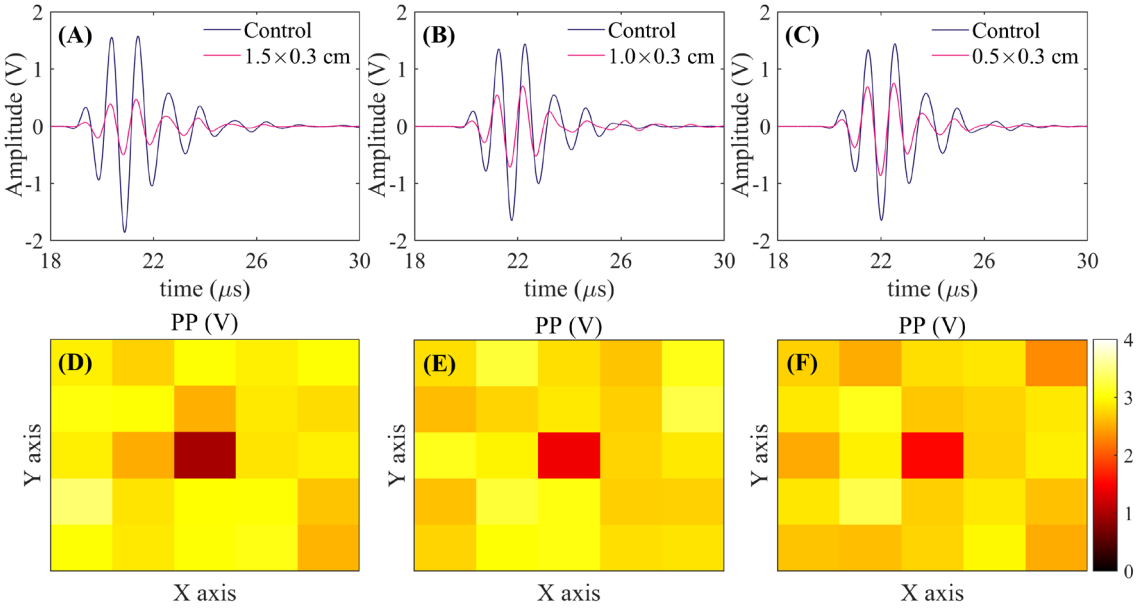


Fig. 9. Ultrasound signals and example of PP (peak-to-peak) image for detection of bone fragments within the center of chicken breast samples. Bone fragments of size 1.5 × 0.3 cm (A, D), size of 1.0 × 0.3 cm (B, E) and size of 0.5 × 0.3 cm (C, F).

In this sense, the results observed from time-domain signals (Fig. 8 and 9) and the phs (Fig. 10) can be analyzed by considering the BF within chicken breast as a heterogeneous system, wherein the disparities in their acoustical impedance resulted in increased reflection of the ultrasound waves on sample's surface as well as scattering of ultrasound energy when it passes through the chicken breast, consequently amplifying the attenuation. Similarly, the presence of a gas-filled bone structure within the BFs potentially contributes to the energy attenuation (Fariñas, et al., 2021). Furthermore, the size of the BF within chicken breast was related to the attenuation of ultrasound waves, the larger sizes of BF reflected and/or absorbed (scattering inside the bone) more energy of the wave compared to the smaller sizes (Correia et al., 2008).

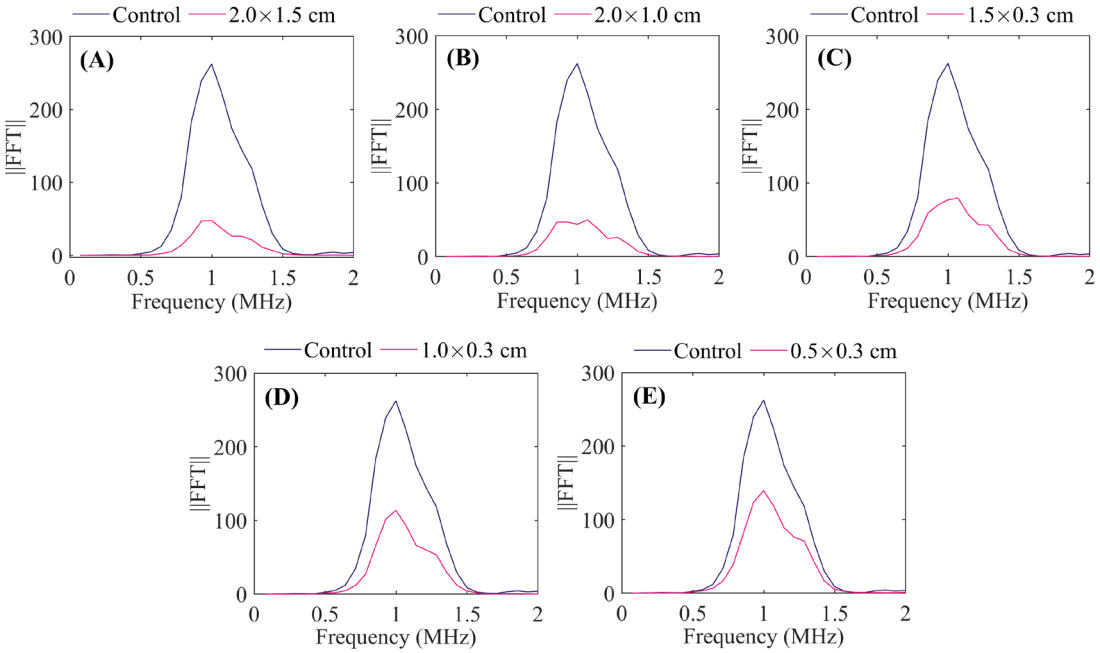


Fig. 10. Example of the frequency spectrum of chicken breast samples with and without bone fragments. Bone fragments of size 2.0 × 1.5 cm (A), size of 2.0 × 1.0 cm (B), size 1.5 × 0.3 cm (C), size of 1.0 × 0.3 cm (D) and size of 0.5 × 0.3 cm (E).

In order to quantify the abovementioned attenuation effects of BFs within chicken breast samples, a multifactor ANOVA examining the influence of BF size and their location inside the samples, on the energy-related ultrasound parameters computed in the time-frequency domain and ultrasound velocity (section 2.5), was carried out (Tables 6 to 9). Further, the results of the multifactor ANOVA assessed on the L are also shown in Table 6.

Table 6. Ultrasound parameters computed in the time-domain (energy-magnitude related and velocity) and thickness for each bone size and location within the sample. Multifactor ANOVA homogeneous groups.

Position: center (point 13 of the matrix)					
Type	PP (V)	ENG (V^2)	INT (V μs)	V_e (m s^{-1})	L (m)
Control	3.8 ± 0.1^{aA}	424.5 ± 3.9^{aA}	204.9 ± 3.5^{aA}	1547.8 ± 11.6^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
2.0×1.5 cm	0.7 ± 0.1^{bA}	11.5 ± 2.3^{bA}	31.0 ± 2.8^{bA}	1433.8 ± 13.0^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
2.0×1.0 cm	0.8 ± 0.1^{bA}	6.9 ± 2.8^{bA}	30.6 ± 1.9^{bA}	1444.9 ± 13.2^{bA}	$0.01 \pm 1 \times 10^{-3aA}$
1.5×0.3 cm	1.1 ± 0.2^{cA}	60.8 ± 2.4^{cA}	66.7 ± 3.6^{cA}	1551.2 ± 7.2^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
1.0×0.3 cm	1.5 ± 0.1^{dA}	59.7 ± 1.8^{cA}	84.4 ± 3.8^{dA}	1552.6 ± 9.1^{aA}	$0.02 \pm 2 \times 10^{-3aA}$
0.5×0.3 cm	1.8 ± 0.2^{eA}	123.2 ± 2.2^{dA}	110.1 ± 2.3^{eA}	1545.9 ± 11.1^{aA}	$0.01 \pm 2 \times 10^{-3aA}$
Position: top-left (point 7 of the matrix)					
Control	3.8 ± 0.1^{aA}	425.9 ± 2.6^{aA}	202.9 ± 1.9^{aA}	1555.2 ± 11.3^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
2.0×1.5 cm	0.6 ± 0.1^{bA}	8.0 ± 3.8^{bA}	31.9 ± 2.9^{bA}	1444.2 ± 11.7^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
2.0×1.0 cm	0.7 ± 0.1^{bA}	9.6 ± 0.5^{bA}	31.0 ± 2.1^{bA}	1445.5 ± 17.0^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
1.5×0.3 cm	0.9 ± 0.1^{cA}	59.9 ± 3.2^{cA}	67.6 ± 7.1^{cA}	1548.4 ± 12.0^{aA}	$0.02 \pm 1 \times 10^{-3aA}$
1.0×0.3 cm	1.5 ± 0.2^{dA}	58.9 ± 3.7^{cA}	84.4 ± 2.6^{dA}	1549.0 ± 12.9^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
0.5×0.3 cm	1.8 ± 0.2^{eA}	120.5 ± 3.4^{dA}	112.5 ± 3.0^{eA}	1536.2 ± 11.1^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
Position: bottom-left (point 17 of the matrix)					
Control	3.8 ± 0.2^{aA}	424.2 ± 2.8^{aA}	205.6 ± 2.1^{aA}	1546.1 ± 9.5^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
2.0×1.5 cm	0.6 ± 0.2^{bA}	11.3 ± 1.3^{bA}	31.1 ± 4.0^{bA}	1450.4 ± 10.3^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
2.0×1.0 cm	0.7 ± 0.1^{bA}	10.9 ± 1.5^{bA}	30.3 ± 2.8^{bA}	1459.0 ± 9.5^{bA}	$0.01 \pm 1 \times 10^{-3aA}$
1.5×0.3 cm	1.1 ± 0.1^{cA}	59.4 ± 3.3^{cA}	68.1 ± 3.0^{cA}	1534.2 ± 9.0^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
1.0×0.3 cm	1.6 ± 0.1^{dA}	58.3 ± 4.1^{cA}	82.3 ± 2.5^{dA}	1538.1 ± 13.5^{aA}	$0.01 \pm 2 \times 10^{-3aA}$
0.5×0.3 cm	1.9 ± 0.2^{eA}	117.1 ± 2.2^{dA}	113.4 ± 2.2^{eA}	1549.9 ± 11.7^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
Position: top-right (point 9 of the matrix)					
Control	3.7 ± 0.2^{aA}	427.6 ± 2.9^{aA}	203.3 ± 3.3^{aA}	1551.6 ± 11.2^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
2.0×1.5 cm	0.7 ± 0.1^{bA}	12.5 ± 2.3^{bA}	31.9 ± 4.2^{bA}	1446.5 ± 7.0^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
2.0×1.0 cm	0.7 ± 0.1^{bA}	11.4 ± 1.6^{bA}	28.1 ± 4.7^{bA}	1456.4 ± 7.6^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
1.5×0.3 cm	1.0 ± 0.1^{cA}	57.9 ± 2.5^{cA}	68.0 ± 5.2^{cA}	1543.4 ± 8.6^{aA}	$0.02 \pm 1 \times 10^{-3aA}$
1.0×0.3 cm	1.5 ± 0.1^{dA}	59.6 ± 4.0^{cA}	85.3 ± 3.1^{dA}	1544.8 ± 10.5^{aA}	$0.02 \pm 1 \times 10^{-3aA}$
0.5×0.3 cm	1.8 ± 0.2^{eA}	117.2 ± 1.9^{dA}	114.3 ± 3.1^{eA}	1547.1 ± 8.0^{aA}	$0.02 \pm 1 \times 10^{-3aA}$
Position: bottom-right (point 19 of the matrix)					
Control	3.7 ± 0.2^{aA}	429.2 ± 2.2^{aA}	202.3 ± 2.1^{aA}	1548.2 ± 8.1^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
2.0×1.5 cm	0.6 ± 0.1^{bA}	11.5 ± 2.4^{bA}	29.8 ± 2.4^{bA}	1434.0 ± 10.6^{bA}	$0.02 \pm 1 \times 10^{-3aA}$
2.0×1.0 cm	0.7 ± 0.2^{bA}	11.5 ± 1.3^{bA}	30.2 ± 2.6^{bA}	1437.4 ± 11.0^{bA}	$0.01 \pm 1 \times 10^{-3aA}$
1.5×0.3 cm	1.0 ± 0.1^{cA}	57.1 ± 3.4^{cA}	67.9 ± 4.6^{cA}	1540.1 ± 11.3^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
1.0×0.3 cm	1.5 ± 0.2^{dA}	57.8 ± 2.8^{cA}	86.4 ± 4.6^{dA}	1540.7 ± 12.1^{aA}	$0.01 \pm 1 \times 10^{-3aA}$
0.5×0.3 cm	1.9 ± 0.1^{eA}	123.3 ± 2.3^{dA}	114.6 ± 2.6^{eA}	1549.2 ± 8.0^{aA}	$0.02 \pm 1 \times 10^{-3aA}$

PP (peak-to-peak), ENG (energy), INT (integral), Ve (ultrasound velocity), and L (thickness). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) for each ultrasound parameter as a function of the size of bone fragments. Uppercase letters indicate statistically significant differences (95%) for the location of these bone pieces within the chicken breast.

Table 7. Ultrasound parameters computed in the time-domain (energy-distribution) for each bone size and location within the sample. Multifactor ANOVA homogeneous groups.

Type	VAR_t (V^2)	SKE_t	KUR_t	ENT_t
Position: center (point 13 of the matrix)				
Control	$0.34 \pm 6 \times 10^{-3aA}$	$-0.11 \pm 9 \times 10^{-3aA}$	8.92 ± 0.06^{aA}	3.3 ± 0.1^{aA}
2.0×1.5 cm	$7 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 7 \times 10^{-3bA}$	8.03 ± 0.07^{bA}	2.1 ± 0.1^{bA}
2.0×1.0 cm	$8 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.11 \pm 8 \times 10^{-3bA}$	8.12 ± 0.06^{bA}	2.1 ± 0.1^{bA}
1.5×0.3 cm	$0.04 \pm 6 \times 10^{-3cA}$	$-0.03 \pm 8 \times 10^{-3cA}$	8.21 ± 0.07^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	$0.07 \pm 3 \times 10^{-3dA}$	$-0.03 \pm 8 \times 10^{-3cA}$	8.30 ± 0.07^{dA}	2.6 ± 0.2^{dA}
0.5×0.3 cm	$0.10 \pm 6 \times 10^{-3eA}$	$-0.06 \pm 6 \times 10^{-3dA}$	8.52 ± 0.07^{eA}	2.7 ± 0.1^{eA}
Position: top-left (point 7 of the matrix)				
Control	$0.34 \pm 8 \times 10^{-3aA}$	$-0.12 \pm 6 \times 10^{-3aA}$	8.84 ± 0.06^{aA}	3.3 ± 0.1^{aA}
2.0×1.5 cm	$7 \times 10^{-3} \pm 2 \times 10^{-3bA}$	$0.10 \pm 8 \times 10^{-3bA}$	8.06 ± 0.06^{bA}	2.1 ± 0.1^{bA}
2.0×1.0 cm	$9 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 7 \times 10^{-3bA}$	8.06 ± 0.06^{bA}	2.1 ± 0.1^{bA}
1.5×0.3 cm	$0.04 \pm 8 \times 10^{-3cA}$	$-0.03 \pm 5 \times 10^{-3cA}$	8.19 ± 0.05^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	$0.08 \pm 2 \times 10^{-3dA}$	$-0.03 \pm 6 \times 10^{-3cA}$	8.29 ± 0.08^{dA}	2.6 ± 0.1^{dA}
0.5×0.3 cm	$0.10 \pm 7 \times 10^{-3eA}$	$-0.07 \pm 8 \times 10^{-3dA}$	8.55 ± 0.07^{eA}	2.7 ± 0.1^{eA}
Position: bottom-left (point 17 of the matrix)				
Control	$0.34 \pm 7 \times 10^{-3aA}$	$-0.11 \pm 6 \times 10^{-3aA}$	8.87 ± 0.05^{aA}	3.2 ± 0.1^{aA}
2.0×1.5 cm	$7 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 7 \times 10^{-3bA}$	8.08 ± 0.06^{bA}	2.1 ± 0.1^{bA}
2.0×1.0 cm	$8 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 6 \times 10^{-3bA}$	8.05 ± 0.05^{bA}	2.1 ± 0.1^{bA}
1.5×0.3 cm	$0.04 \pm 6 \times 10^{-3cA}$	$-0.04 \pm 7 \times 10^{-3cA}$	8.24 ± 0.07^{cA}	2.5 ± 0.1^{cA}
1.0×0.3 cm	$0.08 \pm 3 \times 10^{-3dA}$	$-0.05 \pm 6 \times 10^{-3cA}$	8.33 ± 0.06^{dA}	2.6 ± 0.1^{dA}
0.5×0.3 cm	$0.10 \pm 7 \times 10^{-3eA}$	$-0.07 \pm 6 \times 10^{-3dA}$	8.56 ± 0.06^{eA}	2.7 ± 0.1^{eA}
Position: top-right (point 9 of the matrix)				
Control	$0.33 \pm 5 \times 10^{-3aA}$	$-0.11 \pm 8 \times 10^{-3aA}$	8.92 ± 0.08^{aA}	3.2 ± 0.1^{aA}
2.0×1.5 cm	$7 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 7 \times 10^{-3bA}$	8.06 ± 0.07^{bA}	2.1 ± 0.1^{bA}
2.0×1.0 cm	$8 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 7 \times 10^{-3bA}$	8.14 ± 0.08^{bA}	2.2 ± 0.1^{bA}
1.5×0.3 cm	$0.04 \pm 5 \times 10^{-3cA}$	$-0.03 \pm 5 \times 10^{-3cA}$	8.23 ± 0.07^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	$0.08 \pm 4 \times 10^{-3dA}$	$-0.03 \pm 5 \times 10^{-3cA}$	8.32 ± 0.07^{dA}	2.6 ± 0.1^{dA}
0.5×0.3 cm	$0.10 \pm 8 \times 10^{-3eA}$	$-0.06 \pm 5 \times 10^{-3dA}$	8.54 ± 0.05^{eA}	2.7 ± 0.1^{eA}
Position: bottom-right (point 19 of the matrix)				
Control	$0.34 \pm 8 \times 10^{-3aA}$	$-0.11 \pm 8 \times 10^{-3aA}$	8.91 ± 0.05^{aA}	3.2 ± 0.1^{aA}
2.0×1.5 cm	$7 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.10 \pm 8 \times 10^{-3bA}$	8.07 ± 0.07^{bA}	2.1 ± 0.1^{bA}
2.0×1.0 cm	$8 \times 10^{-3} \pm 1 \times 10^{-3bA}$	$0.11 \pm 6 \times 10^{-3bA}$	8.16 ± 0.05^{bA}	2.1 ± 0.1^{bA}
1.5×0.3 cm	$0.04 \pm 4 \times 10^{-3cA}$	$-0.03 \pm 5 \times 10^{-3cA}$	8.21 ± 0.06^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	$0.07 \pm 3 \times 10^{-3dA}$	$-0.03 \pm 6 \times 10^{-3cA}$	8.33 ± 0.06^{dA}	2.5 ± 0.1^{dA}
0.5×0.3 cm	$0.10 \pm 7 \times 10^{-3eA}$	$-0.07 \pm 5 \times 10^{-3dA}$	8.54 ± 0.06^{eA}	2.7 ± 0.1^{eA}

VARt (variance in time-domain), SKEt (skewness in time-domain), KURt (kurtosis in time-domain) and ENTt (entropy in time-domain). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) for each ultrasound parameter as a function of the size of bone fragments. Uppercase letters indicate statistically significant differences (95%) for the location of these bone pieces within the chicken breast.

Table 8. Ultrasound parameters computed in the frequency-domain (energy-magnitude related) for each bone size and location within the sample. Multifactor ANOVA homogeneous groups.

Position: center (point 13 of the matrix)			
Type	M_0 (MHz)	F_r	MP
Control	2177.8 ± 64.6^{aA}	1.00 ± 0.01^{aA}	296.2 ± 16.9^{aA}
2.0×1.5 cm	396.1 ± 64.6^{bA}	1.02 ± 0.01^{bA}	42.7 ± 3.7^{bA}
2.0×1.0 cm	333.7 ± 59.4^{bA}	1.03 ± 0.01^{bcA}	43.1 ± 5.7^{bA}
1.5×0.3 cm	770.1 ± 62.9^{cA}	1.03 ± 0.02^{bcA}	89.8 ± 6.7^{cA}
1.0×0.3 cm	787.5 ± 73.5^{cA}	1.02 ± 0.01^{bcA}	98.0 ± 6.6^{cA}
0.5×0.3 cm	827.7 ± 62.4^{cA}	1.03 ± 0.01^{cA}	103.4 ± 11.3^{cA}
Position: top-left (point 7 of the matrix)			
Control	2094.8 ± 65.9^{aA}	1.00 ± 0.01^{aA}	297.0 ± 8.3^{aA}
2.0×1.5 cm	354.3 ± 56.3^{bA}	1.02 ± 0.01^{bA}	41.4 ± 5.5^{bA}
2.0×1.0 cm	383.0 ± 67.7^{bA}	1.02 ± 0.00^{bcA}	46.7 ± 10.1^{bA}
1.5×0.3 cm	791.6 ± 63.4^{cA}	1.02 ± 0.01^{bcA}	89.1 ± 7.5^{cA}
1.0×0.3 cm	831.5 ± 65.3^{cA}	1.03 ± 0.01^{bcA}	89.4 ± 10.6^{cA}
0.5×0.3 cm	862.9 ± 71.2^{cA}	1.03 ± 0.01^{cA}	94.5 ± 14.5^{cA}
Position: bottom-left (point 17 of the matrix)			
Control	2164.4 ± 62.7^{aA}	1.01 ± 0.01^{aA}	294.1 ± 14.0^{aA}
2.0×1.5 cm	393.6 ± 49.2^{bA}	1.02 ± 0.01^{bA}	35.5 ± 8.4^{bA}
2.0×1.0 cm	395.0 ± 67.6^{bA}	1.02 ± 0.01^{bcA}	46.8 ± 8.1^{bA}
1.5×0.3 cm	805.2 ± 68.6^{cA}	1.03 ± 0.01^{bcA}	87.2 ± 14.0^{cA}
1.0×0.3 cm	785.8 ± 57.8^{cA}	1.02 ± 0.02^{bcA}	95.2 ± 14.3^{cA}
0.5×0.3 cm	851.9 ± 61.2^{cA}	1.03 ± 0.00^{cA}	103.7 ± 11.3^{cA}
Position: top-right (point 9 of the matrix)			
Control	2116.5 ± 61.9^{aA}	1.00 ± 0.01^{aA}	295.7 ± 8.1^{aA}
2.0×1.5 cm	376.6 ± 54.8^{bA}	1.02 ± 0.01^{bA}	38.9 ± 8.1^{bA}
2.0×1.0 cm	379.8 ± 49.2^{bA}	1.03 ± 0.02^{bcA}	47.0 ± 8.4^{bA}
1.5×0.3 cm	757.9 ± 67.2^{cA}	1.02 ± 0.02^{bcA}	93.6 ± 8.1^{cA}
1.0×0.3 cm	816.1 ± 55.1^{cA}	1.03 ± 0.01^{bcA}	97.9 ± 8.1^{cA}
0.5×0.3 cm	835.1 ± 60.7^{cA}	1.03 ± 0.02^{cA}	108.4 ± 8.4^{cA}
Position: bottom-right (point 19 of the matrix)			
Control	2057.9 ± 71.7^{aA}	1.00 ± 0.01^{aA}	298.3 ± 2.5^{aA}
2.0×1.5 cm	305.9 ± 58.8^{bA}	1.02 ± 0.00^{bA}	33.5 ± 6.4^{bA}
2.0×1.0 cm	314.9 ± 58.9^{bA}	1.04 ± 0.01^{bcA}	40.4 ± 4.1^{bA}
1.5×0.3 cm	733.9 ± 57.6^{cA}	1.02 ± 0.01^{bcA}	90.7 ± 4.8^{cA}
1.0×0.3 cm	782.2 ± 67.7^{cA}	1.03 ± 0.01^{bcA}	92.8 ± 7.3^{cA}
0.5×0.3 cm	798.5 ± 67.4^{cA}	1.02 ± 0.01^{cA}	96.5 ± 5.8^{cA}

M0 (zero-order moment), Fr (center frequency), and MP (maximum peak of the frequency spectrum). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) for each ultrasound parameter as a function of the size of bone fragments. Uppercase letters indicate statistically significant differences (95%) for the location of these bone pieces within the chicken breast.

Table 9. Ultrasound parameters computed in the frequency-domain (energy-distribution) for each bone size and location within the sample. Multifactor ANOVA homogeneous groups.

Position: center (point 13 of the matrix)				
Type	VAR_{sp}	SKE_{sp}	KUR_{sp}	ENT_{sp}
Control	6782.2 ± 68.2^{aA}	1.45 ± 0.02^{aA}	3.7 ± 0.1^{aA}	0.5 ± 0.1^{aA}
2.0×1.5 cm	160.4 ± 42.8^{bA}	1.29 ± 0.07^{bcA}	3.3 ± 0.3^{bcA}	2.8 ± 0.2^{bA}
2.0×1.0 cm	166.4 ± 56.8^{bA}	1.36 ± 0.05^{bA}	3.2 ± 0.1^{bA}	2.7 ± 0.2^{bA}
1.5×0.3 cm	954.8 ± 75.3^{cA}	1.35 ± 0.17^{cA}	3.5 ± 0.1^{cA}	2.3 ± 0.1^{cA}
1.0×0.3 cm	1776.8 ± 88.7^{dA}	1.35 ± 0.06^{bcA}	3.3 ± 0.2^{bcA}	2.3 ± 0.1^{cA}
0.5×0.3 cm	1884.7 ± 75.5^{dA}	1.31 ± 0.03^{bcA}	3.2 ± 0.1^{bcA}	2.2 ± 0.1^{cA}
Position: top-left (point 7 of the matrix)				
Control	6769.8 ± 69.2^{aA}	1.45 ± 0.03^{aA}	3.7 ± 0.1^{aA}	0.3 ± 0.1^{aA}
2.0×1.5 cm	107.1 ± 53.9^{bA}	1.32 ± 0.21^{bcA}	3.4 ± 0.2^{bcA}	2.9 ± 0.1^{bA}
2.0×1.0 cm	120.6 ± 42.6^{bA}	1.25 ± 0.2^{bA}	3.1 ± 0.3^{bA}	2.9 ± 0.2^{bA}
1.5×0.3 cm	955.3 ± 76.2^{cA}	1.38 ± 0.09^{cA}	3.5 ± 0.3^{cA}	2.4 ± 0.2^{cA}
1.0×0.3 cm	1715.3 ± 64.0^{dA}	1.32 ± 0.15^{bcA}	3.3 ± 0.4^{bcA}	2.4 ± 0.1^{cA}
0.5×0.3 cm	1895.1 ± 77.7^{dA}	1.38 ± 0.03^{bcA}	3.5 ± 0.1^{bcA}	2.5 ± 0.1^{cA}
Position: bottom-left (point 17 of the matrix)				
Control	6776.9 ± 62.2^{aA}	1.43 ± 0.03^{aA}	3.6 ± 0.1^{aA}	0.4 ± 0.1^{aA}
2.0×1.5 cm	100.3 ± 45.3^{bA}	1.35 ± 0.11^{bcA}	3.3 ± 0.2^{bcA}	2.9 ± 0.1^{bA}
2.0×1.0 cm	129.1 ± 47.1^{bA}	1.36 ± 0.10^{bA}	3.2 ± 0.4^{bA}	2.8 ± 0.1^{bA}
1.5×0.3 cm	898.9 ± 69.2^{cA}	1.42 ± 0.02^{cA}	3.3 ± 0.1^{cA}	2.5 ± 0.1^{cA}
1.0×0.3 cm	1726.4 ± 58.9^{dA}	1.37 ± 0.12^{bcA}	3.2 ± 0.2^{bcA}	2.4 ± 0.2^{cA}
0.5×0.3 cm	1886.2 ± 54.6^{dA}	1.38 ± 0.03^{bcA}	3.3 ± 0.1^{bcA}	2.5 ± 0.2^{cA}
Position: top-right (point 9 of the matrix)				
Control	6746.8 ± 77.7^{aA}	1.45 ± 0.03^{aA}	3.7 ± 0.1^{aA}	0.6 ± 0.2^{aA}
2.0×1.5 cm	122.6 ± 42.6^{bA}	1.30 ± 0.12^{bcA}	3.2 ± 0.2^{bcA}	3.0 ± 0.2^{bA}
2.0×1.0 cm	138.7 ± 56.8^{bA}	1.25 ± 0.12^{bA}	3.1 ± 0.2^{bA}	2.9 ± 0.1^{bA}
1.5×0.3 cm	931.0 ± 65.3^{cA}	1.36 ± 0.09^{cA}	3.2 ± 0.1^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	1712.3 ± 78.7^{dA}	1.32 ± 0.03^{bcA}	3.2 ± 0.2^{bcA}	2.5 ± 0.1^{cA}
0.5×0.3 cm	1890.3 ± 51.9^{dA}	1.34 ± 0.06^{bcA}	3.0 ± 0.2^{bcA}	2.5 ± 0.2^{cA}
Position: bottom-right (point 19 of the matrix)				
Control	6795.3 ± 67.6^{aA}	1.45 ± 0.02^{aA}	3.7 ± 0.1^{aA}	0.4 ± 0.2^{aA}
2.0×1.5 cm	128.0 ± 38.8^{bA}	1.33 ± 0.08^{bcA}	3.2 ± 0.3^{bcA}	2.9 ± 0.1^{bA}
2.0×1.0 cm	141.6 ± 36.8^{bA}	1.28 ± 0.05^{bA}	3.1 ± 0.2^{bA}	3.0 ± 0.1^{bA}
1.5×0.3 cm	888.5 ± 40.6^{cA}	1.32 ± 0.05^{cA}	3.2 ± 0.2^{cA}	2.4 ± 0.1^{cA}
1.0×0.3 cm	1722.1 ± 59.0^{dA}	1.37 ± 0.10^{bcA}	3.3 ± 0.2^{bcA}	2.4 ± 0.2^{cA}
0.5×0.3 cm	1840.2 ± 47.6^{dA}	1.39 ± 0.02^{bcA}	3.7 ± 0.1^{bcA}	2.5 ± 0.2^{cA}

VARsp (spectral-variance), SKEsp (spectral-skewness), KURsp (spectral-kurtosis) and ENTsp (spectral-entropy). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) for each ultrasound parameter as a function of the size of bone fragments. Uppercase letters indicate statistically significant differences (95%) for the location of these bone pieces within the chicken breast.

A statistically significant ($p<0.05$) effect of the BF size was found on the time and frequency domain ultrasound parameters (Tables 6 to 9). The presence of BFs within chicken breast samples disrupted the energy level of control signals (reduction in signal amplitude) in both temporal (Table 6 and Fig. 8) and frequency (Table 8 and Fig. 10). BF led to a statistically significant ($p<0.05$) reduction in the energy-related (Table 6) and energy distribution (Table 7) ultrasound parameters. Conversely, the location of the BF did not significantly ($p>0.05$) affect the ultrasonic parameters, which shows the robustness of the technique to measure the presence of bones of different sizes, regardless of their spatial location (Tables 6-9). Moreover, non-statistically significant ($p>0.05$) effect of BF size or location was found on L (Table 6), which demonstrates that the incorporation of BFs of varying sizes, in different locations, had no impact on the final thickness of the sample, which could have altered the ultrasonic measurements.

By using the average values of PP, ENG and INT (Table 6), it was possible to sort the samples from the lowest to the highest energy attenuation in five homogeneous groups (LSD intervals ($p<0.05$) from ANOVA), as follows: control (3.8 V, 424.5 V^2 and 204.9 V µs) > 0.5 × 0.3 cm (1.8 V, 123.2 V^2 and 110.1 V µs) > 1.0 × 0.3 cm (1.5 V, 59.7 V^2 and 84.4 V µs) > 1.5 × 0.3 cm (1.1 V, 60.8 V^2 and 66.7 V µs), 2.0 × 1.0 cm (0.8 V, 6.9 V^2 and 30.6 V µs) > 2.0 × 1.5 cm (0.7 V, 11.5 V^2 and 31 V µs). As can be seen in Table 6, the larger sizes of BFs exhibited a higher attenuation, compared to the smaller ones. These results were consistent with the experimental signals and images depicted in Figs. 8 and 9. In addition, the statistical results for V_e indicated that there were no statistically significant differences ($p>0.05$) in the speed of ultrasound waves between the control group and BFs of sizes 1.5 × 0.3 cm, 1.0 × 0.3 cm, and 0.5 × 0.3 cm. However, a noticeable decrease in the ultrasound velocity was found for the larger BF sizes (2.0 × 1.5 cm and 2.0 × 1.0 cm). These results suggest that for small bones the wave front (used to calculate velocity) travels only through the meat flesh (where velocity is higher) and therefore velocity is not altered, compared to the control sample. However, when the bone size is larger, the wave front has traveled through the bone (with lower ultrasound velocity than flesh) and therefore ultrasonic velocity decreases.

Similar results were reported by Correia et al. (2008) in the detection of BFs of different sizes (large size = 15.75 mm^2, medium size = 9.92 mm^2 and small size = 6.18 mm^2) inserted in skinless chicken breasts by using a single point-measure pulse-eco ultrasound technology. These authors quantified the influence of these BFs on both the amplitude ratio and V_e . They found that the presence of BFs led to a statistically significant ($p<0.05$) increase in ultrasonic attenuation, while non statistically significant differences ($p>0.05$) were found in V_e. This study reported V_e values of 1564 ± 2 m/s for chicken breast samples, similar to the values of the present work (Table 6). Although Correia et al. (2008) claimed that the V_e could not be used to detect the presence of BFs, our results showed that the largest BFs, which were obtained from other parts of the chicken skeleton such as vertebra (2.0 × 1.5 cm, Fig 1E i) and chest (2.0 × 1.0 cm, Fig 1E ii), were detected by using V_e.

The greater energy attenuation and delay of the ultrasound wave generated in larger bone fragments (Table 6), can be attributed to the presence of air within the bone structure. When

ultrasound waves propagate through the bone internal structure, they disperse due to the presence of internal air voids, phenomenon that will be more pronounced as the size of the bone increases.

The presence of BFs also influenced the energy-distribution parameters (VARt, SKEt, KURt and ENTt, Table 7). As for the energy magnitude parameters, five homogeneous groups were found by using the LSD intervals. The control, 0.5 × 0.3 cm BF, 1.0 × 0.3 cm BF, 1.5 × 0.3 cm BF, while the last group was integrated by sizes of 2.0 × 1.5 cm and 2.0 × 1.0 cm. The higher BF sizes led to a decrease in the dispersion of the ultrasound waves (VAR_t), left-skewed the time-domain signals (positive SKE_t values), reduced the tailedness (KUR_t) and the randomness (ENT_t). Thus, a noticeable trend was observed from the energy time-domain distribution: the larger the BFs, the most pronounced changes in the energy-distribution of the ultrasound signals waves.

These results can be explained since the increase in BF size represents a barrier for the propagation of the ultrasonic wave causing different effects related to wave amplitude and energy distribution (Collazos-Escobar et al., 2025d). As a result, the presence of BF led to an important modifications in the statistical first-order parameters of the ultrasonic time-domain signals. The reduction in VAR_t values of samples with BF indicated a lower heterogeneity in wave propagation (Table 7), suggesting that larger BFs act as reflector/scatterer of energy within the propagation medium. Positive SKE_t values in an ultrasound signal, when measured in the presence of BF compared to control signals, indicate changes in waveform asymmetry consistent with phase shifts or energy reflection patterns resulting from BF-induced interference at the propagation medium interfaces between the sample and the BF. Furthermore, the reduction in KUR_t and ENT_t (Table 7) values signifies a flattening of the amplitude distribution, indicative of fewer extreme deviations in signal intensity and a shift toward a more deterministic signal (less entropic behavior) profile. From an applied perspective, these features provide a measurable acoustic fingerprint that can be used for the non-destructive detection and size characterization of BFs within biological or food matrices.

As regard of the energy-magnitude related variables computed in the frequency-domain (M_0, F_r and MP, Table 8), the BF presence within chicken breast samples significantly ($p<0.05$) reduced M_0 and MP (Fig. 5) and also modified the center-frequency of phs (Fr). Multifactor ANOVA of M_0 and MP showed three independent groups were clustered by LSD intervals. Control, a group integrated by the BF of 1.5 × 0.3 cm, 1.0 × 0.3 cm and 0.5 × 0.3 cm, and another group for sizes of 2.0 × 1.5 cm and 2.0 × 1.0 cm. While the results of F_r led to also discriminate three groups (Table 8).

These changes can be explained by the fact that the central energy distribution in the phs is altered when the propagating ultrasound wave encounters BFs (Fig. 10), which act as acoustic impedance mismatches. Such mismatches induce partial reflection, scattering, and frequency-dependent attenuation. As a result, the spectral energy shifts toward lower frequencies, reducing the overall magnitude parameters (M_0 and MP) and modifying F_r. This behavior is consistent with the combined effects of diffraction and absorption within

heterogeneous medium, where the size and composition of BFs control the degree of spectral distortion and energy distribution across the frequency domain.

Finally, the BFs presence also influenced the phs energy-distribution (Table 9). The higher BFs produced an important decrease in the dispersion of phs (VAR_{sp}), right-skewed (positive SKE_{sp} values but smaller than control; Fig. 10), reduced the tailedness and increased the disorder in the distribution of energy of phs (ENT_{sp}) compared to the control phs. Thus, the frequency-domain energy-magnitude and distribution parameters were also adequate to detect the presence of BFs in the breasts.

3.2 BF detection using USI and latent-based multivariate statistical process control

The statistical results considering the experimental ultrasound images for the detection of BFs by using the RSS and T^2 and both the TDA, FDA and TFDA approaches (section 2.7.1), are summarized in Tables 10 to 13. Additionally, the results of the feature-extraction approach (section 2.7.1.3) are represented as Supplementary material (Table S1-S3).

The modeling results (Tables 10, 11 and 12) reported that the average values of A_{cc} ranged between 88.2 to 96.07%, S_e were between 0.88 to 0.96 and S_p varied from 0.88 to 0.96, for TDA, FDA and TFDA in all control limits of both RSS and T^2 and their LA. Conversely, the statistical performance of the RSS and T^2 considering the feature extraction approach (Supplementary material, Table S1, S2 and S3) exhibited values of A_{cc} varied from 80 to 93%, while both S_e values were between 0.63 to 0.94 and S_p were between 0.86 to 0.94 for feTDA, feFDA and feTFDA in all control limits and LA. To complement this, the ANOVA results revealed that both A_{cc}, S_e and S_p of TDA, FDA and TFDA were significantly ($p<0.05$) higher than those of feTDA, feFDA and feTFDA. This demonstrates the noteworthy performance of the former approach over feature extraction for detecting BFs using both multivariate control charts (RSS and T^2) computed in the MIA-based MSPC strategy.

The statistical results of multifactor ANOVA models computed from TDA, FDA and TFDA manifested significant differences ($p<0.05$) in the average A_{cc} S_e and S_p values of the optimized PCA models by using different control limits and LA. In all cases, the optimized PCA models obtained from the multi-objective optimization process in which the goal was to simultaneously maximize both S_e and S_p with the lower number of LVs (Fig. 11) were considered the best classifiers in each approach and in each multivariate statistic used. Moreover, it can be seen in Fig 11 the typical plateau of the classification metrics via RSS employing TDA (A_{cc}, Fig. 11A, S_e vs S_p, Fig. 11B), FDA (A_{cc}, Fig. 11E, S_e vs S_p, Fig. 11F) and TFDA (A_{cc}, Fig. 11I, S_e vs S_p, Fig. 11J) and for T^2 using TDA (A_{cc}, Fig. 11C, S_e vs S_p, Fig. 11D), FDA (A_{cc}, Fig. 11G, S_e vs S_p, Fig. 11H) and TFDA (A_{cc}, Fig. 11K, S_e vs S_p, Fig. 11L) wherein the multi-objective optimization problem found the OPCs in all cases in the point of crossing of S_e and S_p.

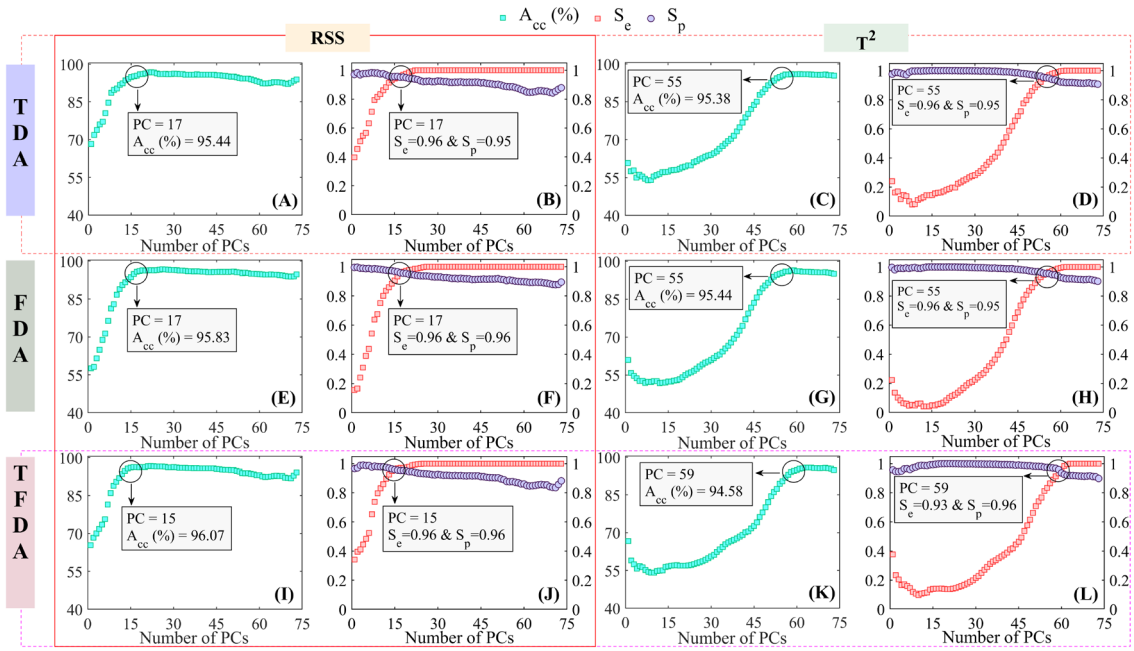


Fig. 11. Classification performance of the multivariate control statistics used for detection of bone fragments in chicken breast. Average A_{cc} for both RSS and T^2 considering TDA (A, C), FDA (E, G) and TFDA (I, K) approaches. Average Se and Sp for both RSS and T2 considering TDA (B, D), FDA (F, H) and TFDA (J, L) approaches. TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency domain approach), A_{cc} (overall accuracy), S_e (sensibility), S_p (specificity), RSS (Residual Sum Squares) and T^2 (Hotelling's T-squared).

Table 10. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using time-domain approach (TDA).

TDA-RSS					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	6	89.00 ± 0.84^{aA}	0.90 ± 0.02^{aA}	0.88 ± 0.02^{aA}
	95	8	91.61 ± 0.81^{aB}	0.92 ± 0.01^{aB}	0.91 ± 0.02^{aB}
	97.5	12	93.90 ± 0.85^{aC}	0.94 ± 0.02^{aC}	0.94 ± 0.02^{aC}
	99	14	94.25 ± 1.04^{aC}	0.94 ± 0.03^{aC}	0.95 ± 0.02^{aC}
50	**90**	**17**	$\mathbf{95.44 \pm 0.90^{bA}}$	$\mathbf{0.96 \pm 0.01^{bA}}$	$\mathbf{0.95 \pm 0.02^{bA}}$
	95	19	95.49 ± 0.82^{bA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	97.5	21	95.75 ± 0.82^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	99	23	95.67 ± 0.95^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
75	90	20	95.61 ± 0.84^{bA}	0.95 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	23	95.60 ± 0.83^{bA}	0.96 ± 0.01^{bA}	0.95 ± 0.02^{bA}
	97.5	25	95.78 ± 0.81^{bA}	0.96 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	99	27	95.90 ± 0.92^{bB}	0.96 ± 0.02^{bA}	0.96 ± 0.02^{bA}
100	90	23	95.72 ± 0.79^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	25	95.53 ± 0.80^{bA}	0.96 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	97.5	26	95.19 ± 0.74^{bA}	0.95 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	99	28	95.55 ± 1.05^{bA}	0.95 ± 0.02^{bA}	0.96 ± 0.02^{bA}
TDA-T^2					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	42	88.20 ± 1.10^{aA}	0.88 ± 0.02^{aA}	0.88 ± 0.02^{aA}
	95	45	91.40 ± 1.04^{aB}	0.92 ± 0.02^{aB}	0.91 ± 0.02^{aB}
	97.5	47	93.38 ± 0.91^{aC}	0.94 ± 0.02^{aC}	0.93 ± 0.02^{aC}
	99	48	94.48 ± 1.01^{aD}	0.95 ± 0.02^{aD}	0.94 ± 0.02^{aC}
50	90	54	94.89 ± 0.97^{bA}	0.94 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	**55**	$\mathbf{95.38 \pm 1.01^{bB}}$	$\mathbf{0.96 \pm 0.02^{bA}}$	$\mathbf{0.95 \pm 0.02^{bA}}$
	97.5	55	95.40 ± 0.99^{bB}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	55	95.38 ± 1.00^{bB}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
75	90	57	95.20 ± 1.02^{cA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	57	95.18 ± 1.01^{cA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	97.5	57	95.17 ± 1.00^{cA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	57	95.14 ± 0.99^{cA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
100	90	58	94.94 ± 1.00^{bcA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	58	94.93 ± 1.00^{bcA}	0.94 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	97.5	58	94.93 ± 0.99^{bcA}	0.94 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	58	94.92 ± 0.98^{bcA}	0.94 ± 0.02^{bA}	0.95 ± 0.02^{bA}

TDA (time-domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), LA (limit augmentation), CL (control limit), OPCs (optimal number of principal components), Acc (overall accuracy), Se (sensibility) and Sp (specificity). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) of each goodness of classification metric (Acc, Se and Sp) as a function of the LA. Uppercase letters indicate statistically significant differences (95%) of Acc, Se and Sp as a function of the computed CL.

Table 11. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using frequency-domain approach (FDA).

FDA-RSS					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	6	89.31 ± 1.23^{aA}	0.90 ± 0.04^{aA}	0.88 ± 0.02^{aA}
	95	8	91.75 ± 0.84^{aB}	0.92 ± 0.01^{aA}	0.91 ± 0.02^{aB}
	97.5	10	93.35 ± 1.13^{aB}	0.93 ± 0.03^{aA}	0.94 ± 0.02^{aC}
	99	12	95.40 ± 1.10^{aC}	0.96 ± 0.02^{aB}	0.95 ± 0.02^{aC}
50	90	15	95.44 ± 0.96^{bA}	0.95 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	95	**17**	$\mathbf{95.83 \pm 0.98^{bA}}$	$\mathbf{0.96 \pm 0.02^{bAB}}$	$\mathbf{0.96 \pm 0.02^{bA}}$
	97.5	19	95.85 ± 0.95^{bA}	0.96 ± 0.01^{bB}	0.96 ± 0.02^{bA}
	99	20	95.63 ± 0.99^{bA}	0.96 ± 0.02^{bB}	0.96 ± 0.02^{bA}
75	90	18	95.84 ± 0.91^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	20	95.80 ± 0.84^{bA}	0.95 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	97.5	21	95.44 ± 0.92^{bA}	0.95 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	99	23	95.77 ± 0.81^{bA}	0.96 ± 0.01^{bA}	0.95 ± 0.02^{bA}
100	90	21	95.72 ± 0.84^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	23	95.93 ± 0.73^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	97.5	24	95.74 ± 0.90^{bA}	0.96 ± 0.02^{bA}	0.96 ± 0.02^{bA}
	99	25	95.61 ± 0.81^{bA}	0.96 ± 0.02^{bA}	0.96 ± 0.02^{bA}
FDA-T^2					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	41	88.07 ± 1.32^{aA}	0.88 ± 0.03^{aA}	0.88 ± 0.02^{aA}
	95	43	91.35 ± 1.36^{aB}	0.91 ± 0.03^{aB}	0.92 ± 0.02^{aB}
	97.5	45	93.67 ± 1.17^{aC}	0.94 ± 0.02^{aC}	0.94 ± 0.02^{aC}
	99	46	94.94 ± 0.92^{aD}	0.95 ± 0.02^{aC}	0.95 ± 0.02^{aC}
50	90	55	95.49 ± 0.88^{bA}	0.96 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	**55**	$\mathbf{95.44 \pm 0.85^{bA}}$	$\mathbf{0.95 \pm 0.02^{bA}}$	$\mathbf{0.95 \pm 0.02^{bA}}$
	97.5	55	95.44 ± 0.84^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	55	95.41 ± 0.83^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
75	90	57	95.30 ± 0.90^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	57	95.25 ± 0.91^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	97.5	57	95.25 ± 0.91^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	57	95.24 ± 0.91^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
100	90	59	95.29 ± 0.96^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	95	59	95.30 ± 0.96^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	97.5	59	95.30 ± 0.96^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}
	99	59	95.30 ± 0.96^{bA}	0.95 ± 0.02^{bA}	0.95 ± 0.02^{bA}

FDA (frequency-domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), LA (limit augmentation), CL (control limit), OPCs (optimal number of principal components), Acc (overall accuracy), Se (sensibility) and Sp (specificity). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) of each goodness of classification metric (Acc, Se and Sp) as a function of the LA. Uppercase letters indicate statistically significant differences (95%) of Acc, Se and Sp as a function of the computed CL.

Table 12. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using time-frequency domain approach (TFDA).

TFDA-RSS					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	6	88.24 ± 1.20^{aA}	0.88 ± 0.03^{aA}	0.89 ± 0.02^{aA}
	95	8	93.02 ± 0.65^{aB}	0.94 ± 0.01^{aB}	0.92 ± 0.02^{aB}
	97.5	9	94.46 ± 0.86^{aC}	0.95 ± 0.02^{aC}	0.94 ± 0.02^{aC}
	99	11	95.42 ± 0.96^{aC}	0.95 ± 0.02^{aC}	0.95 ± 0.02^{aC}
50	**90**	**15**	$\mathbf{96.07 \pm 0.86^{bA}}$	$\mathbf{0.96 \pm 0.01^{bA}}$	$\mathbf{0.96 \pm 0.02^{bA}}$
	95	16	96.03 ± 0.85^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	97.5	19	95.88 ± 0.81^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	99	20	95.81 ± 0.81^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
75	90	19	95.81 ± 0.78^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	21	95.80 ± 0.87^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	97.5	22	95.77 ± 0.91^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	99	23	95.75 ± 0.84^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
100	90	22	95.85 ± 0.80^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	95	24	95.70 ± 0.83^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	97.5	26	95.70 ± 0.86^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
	99	28	95.89 ± 0.81^{bA}	0.96 ± 0.01^{bA}	0.96 ± 0.02^{bA}
TFDA-T^2					
LA (%)	CL (%)	OPCs	A_{cc} (%)	S_e*	S_p*
0	90	47	88.20 ± 1.19^{aA}	0.88 ± 0.03	0.88 ± 0.02
	95	50	91.99 ± 0.87^{aB}	0.92 ± 0.02	0.92 ± 0.02
	97.5	51	93.57 ± 1.03^{aC}	0.93 ± 0.02	0.94 ± 0.02
	99	52	94.23 ± 0.85^{aC}	0.94 ± 0.01	0.94 ± 0.02
50	**90**	**59**	$\mathbf{94.58 \pm 1.01^{bA}}$	$\mathbf{0.93 \pm 0.03}$	$\mathbf{0.96 \pm 0.03}$
	95	59	94.54 ± 1.03^{bA}	0.93 ± 0.03	0.96 ± 0.03
	97.5	59	94.54 ± 1.03^{bA}	0.93 ± 0.03	0.96 ± 0.03
	99	59	94.54 ± 1.03^{bA}	0.93 ± 0.03	0.96 ± 0.03
75	90	60	93.09 ± 1.44^{cC}	0.92 ± 0.05	0.94 ± 0.03
	95	60	93.09 ± 1.43^{cC}	0.92 ± 0.05	0.94 ± 0.03
	97.5	60	93.09 ± 1.43^{cC}	0.92 ± 0.05	0.94 ± 0.03
	99	60	93.09 ± 1.43^{cC}	0.92 ± 0.05	0.94 ± 0.03
100	90	61	92.91 ± 2.63^{cC}	0.93 ± 0.08	0.93 ± 0.03
	95	61	92.91 ± 2.63^{cC}	0.93 ± 0.08	0.93 ± 0.03
	97.5	61	92.91 ± 2.63^{cC}	0.93 ± 0.08	0.93 ± 0.03
	99	61	92.91 ± 2.63^{cC}	0.93 ± 0.08	0.93 ± 0.03

TFDA (time-frequency domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), LA (limit augmentation), CL (control limit), OPCs (optimal number of principal components), Acc (overall accuracy), Se (sensibility) and Sp (specificity). Results are expressed as mean ± standard error. Different lowercase letters indicate statistically significant differences (95%) of Acc as a function of the LA. Uppercase letters indicate statistically significant differences (95%) of Acc as a function of the computed CL. *The residuals from multifactor analysis of variance (ANOVA) models failed to meet the assumptions of normality and homoscedasticity, thus rendering both models unsuitable for practical inference.

The results of TDA (Table 10) showed a great classification performance of control and OC images. Based on the LSD intervals, the optimized PCA model using RSS with 17 OPCs, employing 50% augmented the 90% control limit exhibited high classification performance (A_{cc} = 95.44, S_e = 0.96 and S_p = 0.95) with the minimum number of LVs. In contrast, the optimized PCA model using T^2 required more LVs (OPCs= 55), maintaining the augmentation of 50% of the limit 95% to achieved similar classification performance to RSS (A_{cc} = 95.38, S_e = 0.96 and S_p = 0.95).

Closely, the statistical results of FDA (Table 11) exhibited quite similar behavior of TDA, the optimized PCA using RSS with 17 LVs, augmented 50% control limit at 95% showed an A_{cc} = 95.85, S_e = 0.96 and S_p = 0.96 and T^2 control statistic reached an A_{cc} = 95.44, S_e = 0.95 and S_p = 0.95 using 55 LVs and the same LA and control limit of RSS. As expected, the classification results for TDA and FDA were closely aligned by using both RSS and T^2 (Table 10 and 11). This result could be attributed to the fact that the presence of BFs produced an important attenuation of the ultrasound energy, which is closely related to both temporal and frequency domains (Suen et al., 2016). Therefore, the PCA model based RSS and T^2 were able to satisfactorily detect the BFs by using both energy-magnitude and energy distribution parameters from the time and frequency domains (as explained in section 3.1). The detection via RSS suggested that the presence of BFs led to a detectable breakage in the correlation structure of the control model and T^2 indicated extreme values (lower energy related and magnitude values of ultrasound parameters) in these images compared to the control ones (Kruse et al., 2014). Nevertheless, the RSS statistic was the most robust classifier to maximize the goodness of classification of control and OC images due to its simplicity in the use of lower number of LVs than the T^2 statistic.

The statistical results of the last approach, which integrated TDA and FDA (TFDA, Table 12), revealed a slight but statistically non-significant ($p>0.05$) enhancement in RSS classification performance and did not evidence an improvement using T^2 ($p>0.05$). In this regard, when model input variables potentially contribute to describe the response, selecting specific input variables can improve model results. Conversely, adding more variables could worsen the model's accuracy (Zhang, 2014). Therefore, in the case of T^2 the combination of TDA and FDA in the same framework to feed the PCA model caused redundance (features which have explained the same extreme values) and promote the use of more LVs (Fig. 11K and 11L).

Nonetheless, the use of TFDA-RSS contributed to reduce 2 LVs (15 OPCs, Table 12, Fig. 11I and 11J) maintaining the 50% augmented control limit at 90% and non-significant differences ($p>0.05$) in the figures of merits (A_{cc}=96.07%, S_e=0.96 and S_p=0.96) compared to TDA and FDA (A_{cc}=95.83%, S_e=0.96 and S_p=0.96). This result suggests that the combination of both energy-related and energy-magnitude ultrasound parameters computed in the time and frequency domains made the PCA model more robust for the detection of any disturbance in the correlation structure not only between the variables referred with the energy and distribution in time and frequency domains but also the relationship between both spaces, thus, less LVs were needed to maximize the classification of control and OC images.

In order to analyze in detail the classification performance of the LSD based selected optimized PCA models influenced by the approaches considered, the statistical results in Table 12 illustrate together the CFMs, the number of images correctly classified and misclassified by using both multivariate control statistics.

As can be seen, the RSS and T^2 using TDA, FDA and TFDA were able to detect all of OC images within the larger BFs (2.0 × 1.5 cm and 2.0 × 1.0 cm), while in the detection of the smaller BFs sizes (1.5 × 0.3 cm, 1.0 × 0.3 cm and 0.5 × 0.3 cm), there were almost 1 ± 1 (from the 100 times randomly partition of data sets) of these OC images incorrectly classified. Additionally, these optimized models did not correctly classify at least 3 ± 1 control images in all approaches (Table 8). This result can be attributed to the natural variability of poultry meat and their respective images (Garrido-Novell et al., 2018). The inherent variability such as compositional variations (lean meat and fat components) and intricate structural arrangement (tendons and connective tissues of poultry samples) could contribute to increase variability in control USIs (Fariñas et al., 2021). As already explained, the use of TFDA-RSS did not improve the goodness of classification metrics compared to TDA-RSS and FDA-RSS but also reduced the LVs (Table 8). Thus, the use of TFDA-RSS and TDA-T^2/FDA-T^2 could be considered as the best options for practical industrial implementation in the quality inspection of the presence of BF within poultry meat products.

Results obtained in the present work were similar to those obtained by Zhao et al. (2006) in the detection of glass fragments within beverages packaged in glass containers by the integration of CUS and Artificial Neural Networks (ANN). These authors considered a training (n = 60 signals) and validation (n = 20 signals) datasets, to develop a detection of glass fragment based CUS-ANN. They obtained a successful classification rate (A_{cc}) of 95% and claimed that the combination of CUS and ML proved to be feasible for non-invasive and real-time quality inspection of foodstuffs. Nonetheless, the differential aspect of this work lies in the capability of our ultrasonic system to inspect the entire product, rather than being confined to single-point measurements for detecting FBs, further, the use of a high number of experimental images for the calibration and external validation of mathematical models. This capability provides a significant advantage in analyzing the presence of FBs, regardless of their location within food products.

Table 13. Classification performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics in the detection of varying-size bone fragments using the time-domain (TDA), frequency-domain (FDA) and time-frequency domain (TFDA) approaches.

TDA		Number of samples (predicted)	
Type	Number of samples (real)	RSS	T^2
2.0×1.5 cm	15	15	15
2.0×1.0 cm	17	17	17
1.5×0.3 cm	**17**	**16**	**16**
1.0×0.3 cm	**15**	**14**	**14**
0.5×0.3 cm	**17**	**15**	**15**
OC (all types)	81	TP = 78 ± 1 FN = 3 ± 1	TP = 78 ± 1 FN = 3 ± 1
Control (C_{cal+} C_{EV})	81	TN = 77 ± 2 FP = 4 ± 2	TN = 77 ± 2 FP = 4 ± 2
C_{cal}	**73**	TN_{cal} = **71**	TN_{cal} = **71**
C_{EV}	**8**	TN_{IV} = **6**	TN_{IV} = **6**
FDA		Number of samples (predicted)	
Type	Number of samples (real)	RSS	T^2
2.0×1.5 cm	15	15	15
2.0×1.0 cm	17	17	17
1.5×0.3 cm	**17**	17	**16**
1.0×0.3 cm	**15**	**14**	**14**
0.5×0.3 cm	**17**	**15**	**15**
OC (all types)	81	TP = 78 ± 1 FN = 3 ± 1	TP = 77 ± 1 FN = 4 ± 1
Control (C_{cal+} C_{EV})	81	TN = 78 ± 2 FP = 3 ± 2	TN = 77 ± 2 FP = 4 ± 2
C_{cal}	**73**	TN_{cal} = **72**	TN_{cal} = **71**
C_{EV}	**8**	TN_{IV} = **6**	TN_{IV} = **6**
TFDA		Number of samples (predicted)	
Type	Number of samples (real)	RSS	T^2
2.0×1.5 cm	15	15	15
2.0×1.0 cm	17	17	17
1.5×0.3 cm	**17**	**16**	**15**
1.0×0.3 cm	**15**	**14**	**14**
0.5×0.3 cm	**17**	**16**	**15**
OC (all types)	81	TP = 78 ± 1 FN = 3 ± 1	TP = 76 ± 2 FN = 5 ± 2
Control (C_{cal+} C_{EV})	81	TN = 78 ± 1 FP = 3 ± 1	TN = 78 ± 2 FP = 3 ± 2
C_{cal}	**73**	TN_{cal} = **72**	TN_{cal} = **72**
C_{EV}	**8**	TN_{IV} = **6**	TN_{IV} = **6**

TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), OC (out-of-control), Ccal (control images for PCA calibration), CEV (control images for PCA external validation), TP (true positive), TN (true negative), TNcal (true negative for calibration images), TNIV (true negative for external validation images), FP (false positive) and FN (false negative). Results are expressed as mean ± standard error.

3.2.1 Influence of training and validation dataset size on the BF detection using the latent-based statistical process control approach

The statistical results for the detection of BFs by using both RSS and T^2 which considered the TDA, FDA and TFDA approaches and four different ratios of datasets (100%, 75%, 50% and 25%) alongside with statistical comparison using an ANOVA model, are depicted in Fig. 12.

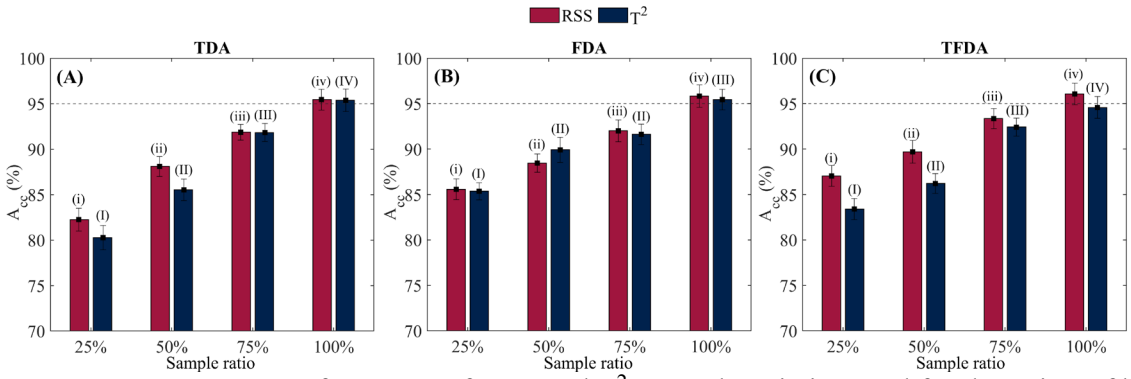


Fig. 12. Average A_{cc} performance of RSS and T^2 control statistics used for detection of bone fragments in chicken breast using different number of ultrasound images. Results for TDA (A), FDA (B) and TFDA (C). TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency domain approach), A_{cc} (overall accuracy), RSS (Residual Sum Squares) and T^2 (Hotelling's T-squared).

As can be observed (Fig. 12) the higher the number of ultrasound images in the analysis, the better performance (progressively and statistically significant increase of A_{cc}, ranging between 80% to 96%) of both statistics for all the approaches. The increase in the number of ultrasound images led the model become more robust with more images for model calibration (Hu et al., 2018). Thus, as can be seen for the RSS, which has provided the best classification results in all of approaches, the difference between considering the entire batch of samples or 75% of the total samples, is statistically ($p<0.05$) lower (< 4% for TDA and < 3% for FDA and TFDA, respectively, Figs. 12A, 12B and 12C). Furthermore, the use of T^2 also provided great detection results in all approaches. However, the statistical results were better by using RSS than T^2 with lower number of LVs. This fact (RSS behavior) indicates that the typical plateau value that appears when plotting A_{cc} vs the number of samples (Hu et al., 2018) is already being reached. This behavior (Fig. 12A, 12B and 12C), in addition to the high value in the percentage of correctly classified samples (>95%), indicates that the number of samples tested in the present study (162) suffices for the selected approach (TDA, FDA and or TFDA) and RSS multivariate control statistic.

3.3 BF detection using USI and latent-based machine learning classifiers

The second part of this work was focused to assessing the performance of the supervised Latent Variable-based Machine Learning (LV-ML) strategies, with and without the RF-based variable selection stage (RF-VS-LV-ML), for the detection of BF in chicken breast samples. These models included the use of Latent Variable-Support Vector Machine (LV-SVM), Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA), Latent Variable-Generalized Linear Model (LV-GLM), Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM), Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF), Random Forest-Variable Selection-Latent Variable-Naïve Bayes (RF-VS-LV-NB), Random Forest-Variable Selection-Latent Variable-Linear Discriminant Analysis (RF-VS-LV-LDA), Random Forest-Variable Selection-Latent Variable-Quadratic Discriminant Analysis (RF-VS-LV-QDA) and Random Forest-Variable Selection-Latent Variable-Generalized Linear Model (RF-VS-LV-GLM).

In contrast to the unsupervised MIA approach, these methods aimed to leverage the extracted latent eigenspace in combination with different classification ML techniques to enhance classification performance of chicken breast samples with and without BF. Particular emphasis was placed on assessing how the choice of hyperparameter configurations and the inclusion of variable selection affected model robustness and predictive accuracy. To provide a comprehensive view of model behavior, the performance metrics were systematically analyzed and reported separately for the training (75%) and validation (25%) datasets, allowing for a direct comparison of fitting ability and generalization power. Accordingly, the following section presents the statistical results obtained with the LV-ML and RF-VS-LV-ML models, highlighting the main differences in accuracy, computational time, and complementary figures of merit.

The statistical results of the LV-ML and RF-VS-LV-ML models in the detection of BF within chicken breast samples (section 2.7.2, considering the 100% of experimental dataset, 81 control and 81 OC images, since it was the best data size defined in section 3.2.1) exhibited significant differences in the goodness of fit metrics depending on the hyperparameters configurations belonging to each ML model and the NVLs used in modeling procedure (Fig. 13-36). Figs. 13 to 36 illustrate together the average values of A_{cc} (%) and CT (s) of LV-ML and RF-VS-LV-ML for both training and validation datasets, while the remaining figures of merit (S_{eT}, S_{eV}, S_{pT}, S_{pV}, P_{rT}, P_{rV}, R_{eT}, R_{eV}, F_{sT}, F_{sV}, AUC_{ROCT}, AUC_{ROCV}, MCC_T and MCC_V) alongside with their variability are compiled in "Supplementary material" (section 9). Additionally, the variability associated with the A_{ccT} (75% for training) and A_{ccV} (25% for validation) are also presented as supplementary material.

The statistical results were depicted in average values since the huge number of runs, combination between techniques, data approaches, hyperparameters, NLVs and goodness of fit metric tested from DoE (section 2.7.2; Tables 1-5) did not facilitate the data representation.

Thus, in order to improve the data visualization, the A_{cc} and CT were selected as main figures of merit metrics of all evaluated LV-ML and RF-VS-LV-ML models.

As previously stated in "Material and Methods" section, the task of exploring different data-driven modeling strategies to address the challenge of detecting BF in the poultry meat industry is of critical importance. Thereby, the use of LV-ML and RF-VS-LV-ML aimed to assess whether detection performance based on the MIA-PCA strategy could be improved through ML techniques. For this reason and to facilitate the analysis of how ML influenced the classification performance of contaminated/uncontaminated chicken breast samples, all figures summarizing the A_{cc} results included a line/grid reference to compare the statistical results of MIA-MSPC (Tables 10-12 and Fig. 11) with those obtained from LV-ML and RF-VS-LV-ML approaches. The first red line/grid colored was include to represent the maximum A_{cc} results using the MIA-MSPC-RSS-TFDA (Table 12). Furthermore, an additional green line/grid colored was included in all figures as a reference of A_{cc} = 99% with the aim to rapidly identify if any of LV-ML and RF-VS-LV-ML models exceeded the A_{cc} = 96.07% (Table 12) and reached values of 96.07% ≤ A_{cc} ≤ 99% and/or A_{cc} ≥ 99%.

As a general trend, the statistical results of A_{cc} in all LV-SVM, LV-RF, LV-NB, LV-LDA, LV-QDA, LV-GLM, RF-VS-LV-SVM, RF-VS-LV-RF, RF-VS-LV-NB, RF-VS-LV-LDA, RF-VS-LV-QDA and RF-VS-LV-GLM models employing different hyperparameter configurations and NLVs and using the different data approaches such as TDA, FDA, TFDA, TFDABH and TFDABS showed a significant drop between training and validation datasets.

In the case of LV-SVM, the statistical classification performance as a function of the five data approaches tested (Figs. 13 to 17) revealed both consistent patterns and important differences related to KFs SVM's hyperparameter, and the NLVs. A pronounced differences in A_{cc} between training and validation was observed, particularly with non-linear KFs such as rbfdot, laplacedot, polydot, vanilladot, and besseldot, regardless of whether the representation was time domain (TDA; Fig. 1S), frequency domain (FDA; Fig. 2S), time-frequency domain (TFDA; Fig. 3S), and/or block-scaled time-frequency domain (TFDABH; Fig. 4S and TFDABS; Fig. 5S). This indicates that non-linear mappings capture complex structures in training data but often fail to generalize, especially at high NLVs. Overfitting also depended on the SVM configuration. For instance, using the same KF, C-svc and nu-svc frequently displayed different A_{cc} trends, showing that the each model configuration classify both control and OC images with difference A_{cc} as a function of the NVLs used.

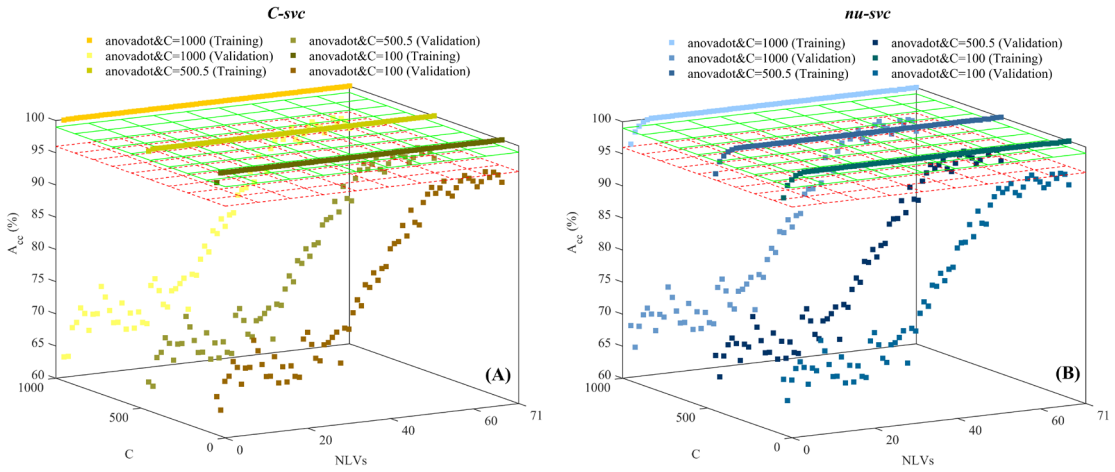


Fig. 13. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

Increasing NLVs generally improved model performance, with A_{cc} frequently exceeding 96% in both training and validation under TDA, FDA, TFDA, and TFDABH. However, in TFDABS, performance decrease severely using KF-besseldot, training A_{cc} remained below 60% and validation A_{cc} did not increase, even with large NLVs. This result reflects that this SVM combination of hyperparameters was not compatible, as a result a low goodness of fit was evidenced. Respect to the influence of SVM's C parameter, it was not significantly. Varying in C values between 100 and 1000 did not alter A_{cc} trends, indicating that LV preprocessing stabilized the optimization problem, thereby reducing the influence of margin regularization. Among KF, anovadot showed the most robust behavior (Fig. 13A and 13B to 15A and 15B). In TDA, FDA, and TFDA, it yielded high training and validation A_{cc} for both C-svc and nu-svc, with narrow train/validation gaps. This suggests that anovadot effectively exploits orthogonal latent projections by modeling additive and interaction effects while avoiding the severe overfitting observed with other non-linear kernels. Both C-svc and nu-svc exhibited similar performance, with nu-svc showing slightly smoother validation trends at larger NLVs.

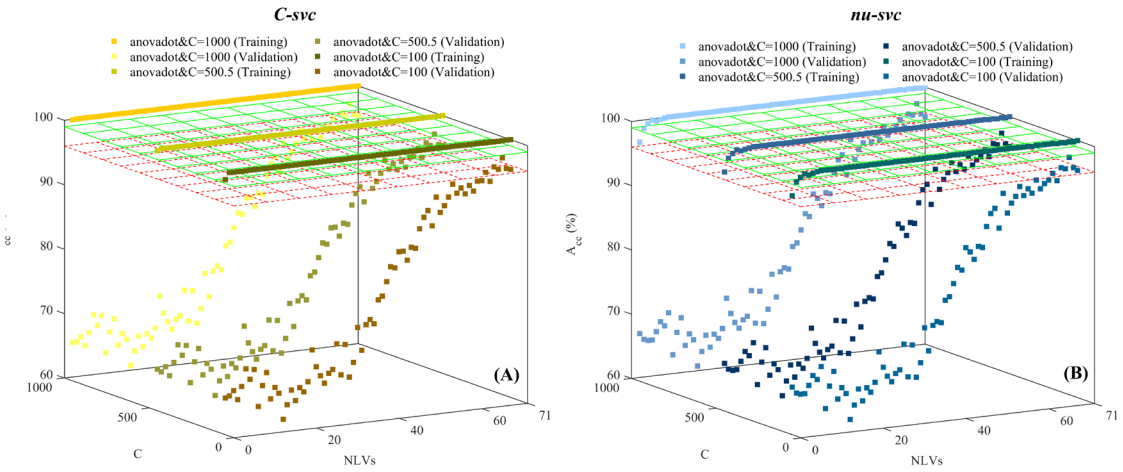


Fig. 14. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

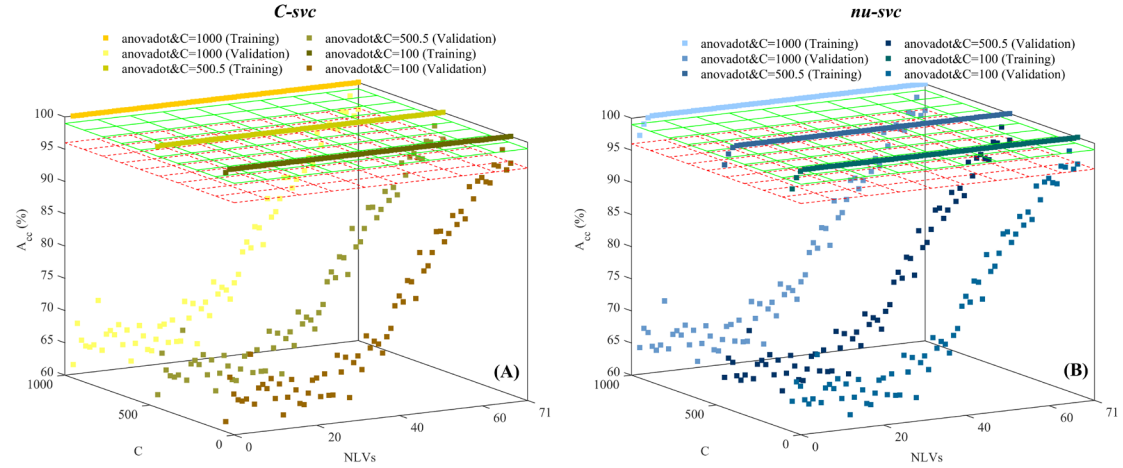


Fig. 15. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

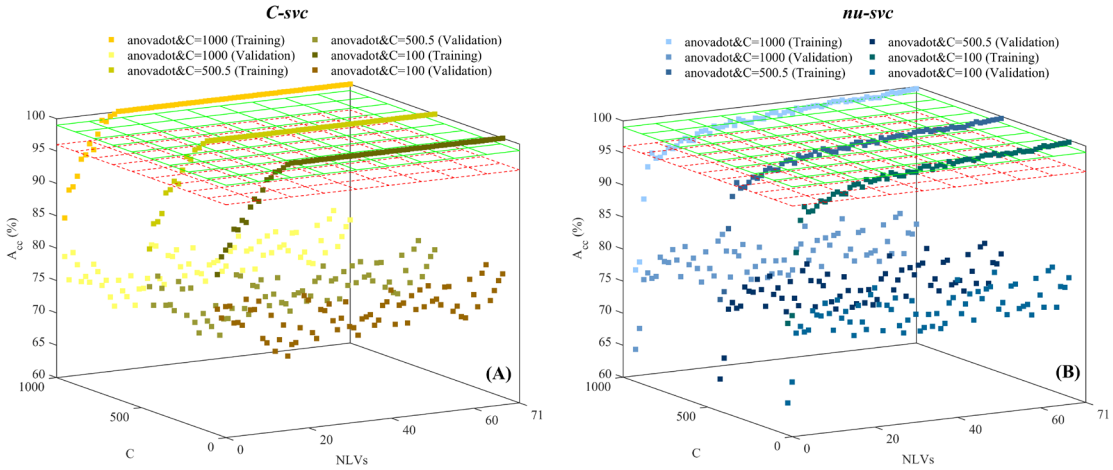


Fig. 16. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

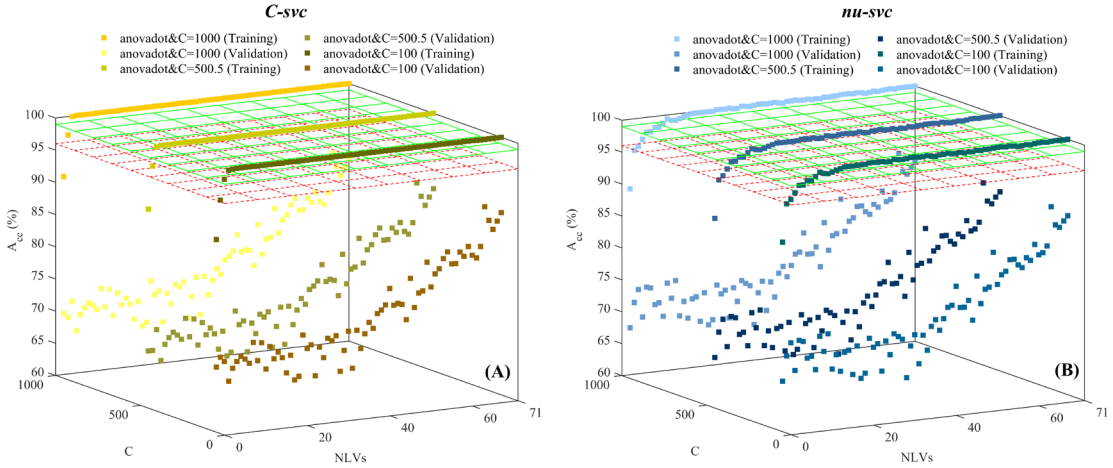


Fig. 17. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) obtained with the "anovadot" kernel function are reported separately for the training (75%) and the validation (25%) datasets.

The analysis of CT in Figs. 18 to 22 characterizes the behavior of the LV-SVM framework under different hyperparameter and NVLs assessed. CT was measured during training sing 75% of the dataset, with variations determined by KF, Type (C-svc vs. nu-svc), the C parameter and the NLVs considered during model tuning. The figures quantify both the computational cost imposed by SVM hyperparameters and the additional cost introduced by increasing model dimensionality through larger NLVs.

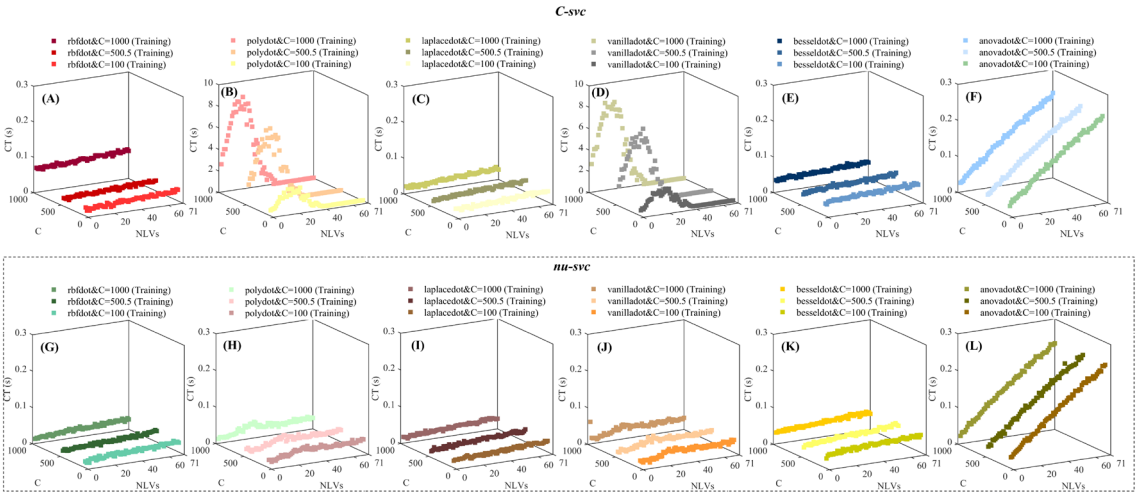


Fig. 18. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

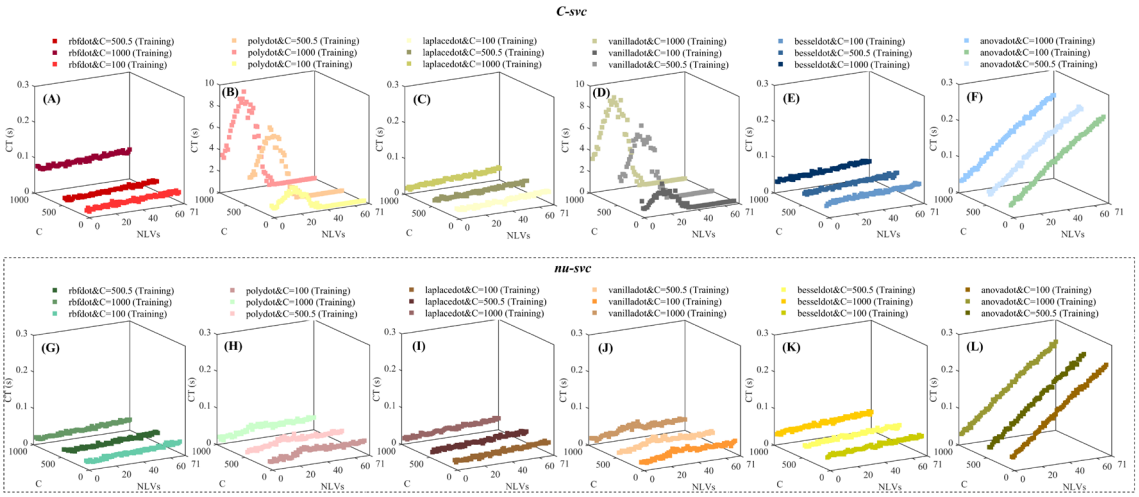


Fig. 19. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Regarding to the TDA (Fig. 18), a clear trend was observed. Linear and polynomial functions (vanilladot, polydot) consistently show the lowest CT, while more complex kernels (anovadot and besseldot) are markedly more expensive, with rbfdot and laplacedot occupying an intermediate position. Increasing C from 100 to 1000 only modestly increases CT, indicating that solver convergence was only mildly sensitive to tighter margin constraints in this domain. Differences between C-svc and nu-svc remain small overall, with nu-svc tending to be slightly slower under non-linear kernels at high C values.

Considering the FDA (Fig. 19), the relative kernel ordering persists, but absolute CTs for the simpler kernels are generally lower than in TDA. This suggests that frequency representations

compact relevant information, thereby reducing the effective per-component computational load. In FDA, CT grows more gently with NLVs, especially for vanilladot and polydot, whereas non-linear kernels still show steep scaling with dimensionality. The role of C and SVM type remains secondary, except under the heaviest kernels. In practical terms, FDA achieves a favorable computational profile for low-to-moderate NLVs but does not alleviate the cost of highly non-linear kernels.

The results of CT using the TFDA (Fig. 20) raises the computational demand more substantially. By concatenating time and frequency features, the dimensionality increases, shifting CT upward across all kernels. The relative differences between kernels become more pronounced: while vanilladot and polydot remain relatively efficient, rbfdot, laplacedot, and especially anovadot and besseldot exhibit steep increases in CT as NLVs grow. Although the effect of C remains moderate, high values exacerbate training times when combined with non-linear kernels. Once again, the CT-NLVs relationship is nearly linear, but with a steeper slope than in TDA and FDA due to the higher effective dimensionality of TFDA representations. Additionally, by using TFDABH (Fig. 21) the CT represents the most computationally demanding scenario. The hard block-scale of TFDA introduced heterogeneous features that dramatically inflate kernel evaluations. As a result, CT is substantially higher than in all previous representations, particularly for anovadot and besseldot. The effect of C parameter becomes non-negligible in this setting, amplifying runtimes significantly when large margins are enforced. Similarly, nu-svc tends to incur additional overhead under these demanding conditions. Most critically, the slope of CT versus NLVs is steepest in TFDABH: each additional latent variable imposes a disproportionately large computational resources, reflecting the compounded cost of block-scale interactions in high dimensions.

By contrast, the TFDABS (Fig. 22) provides a more balanced profile. Although CT remains higher than in TFDA, it is consistently lower than in TFDABH. The KF ranking remains the same, but the escalation of CT with NLVs is considerably more moderate than in the hard variant. Similarly, the influence of C and SVM type on runtime is less dramatic, returning to a pattern closer to TFDA. This indicates that soft scaling mitigates some of the inefficiencies introduced by hard block partitioning, allowing richer feature representations without incurring prohibitive computational costs. Thereby, the statistical CT results revealed that the changes in training time were primarily influenced by kernel choice and the number of latent variables, while C plays an amplifying but secondary role, and SVM type contributes only marginal differences except in extreme conditions. Across representations, a clear runtime ordering can be established: FDA is typically the most efficient, followed by TDA, then TFDA, with TFDABS moderately heavier and TFDABH the most computationally expensive.

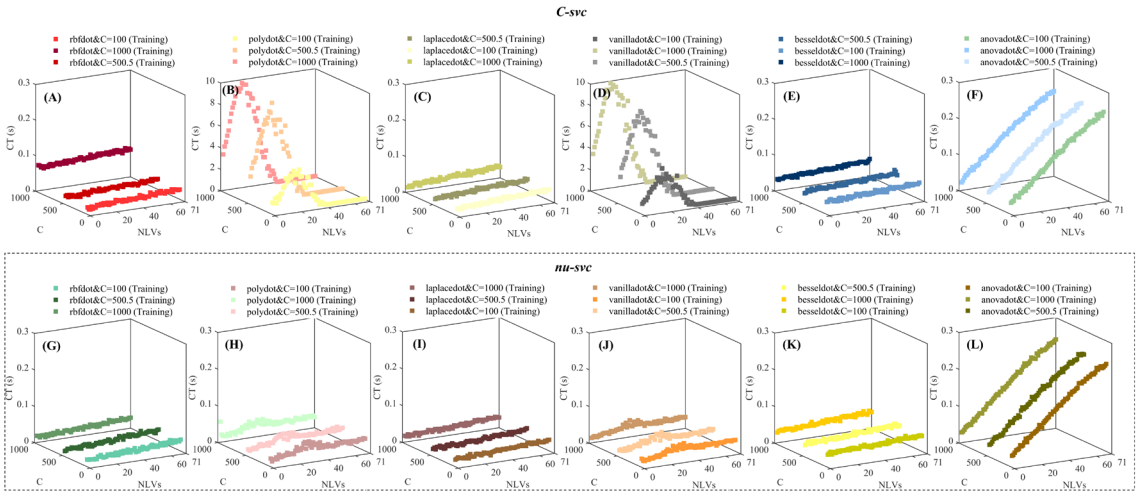


Fig. 20. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

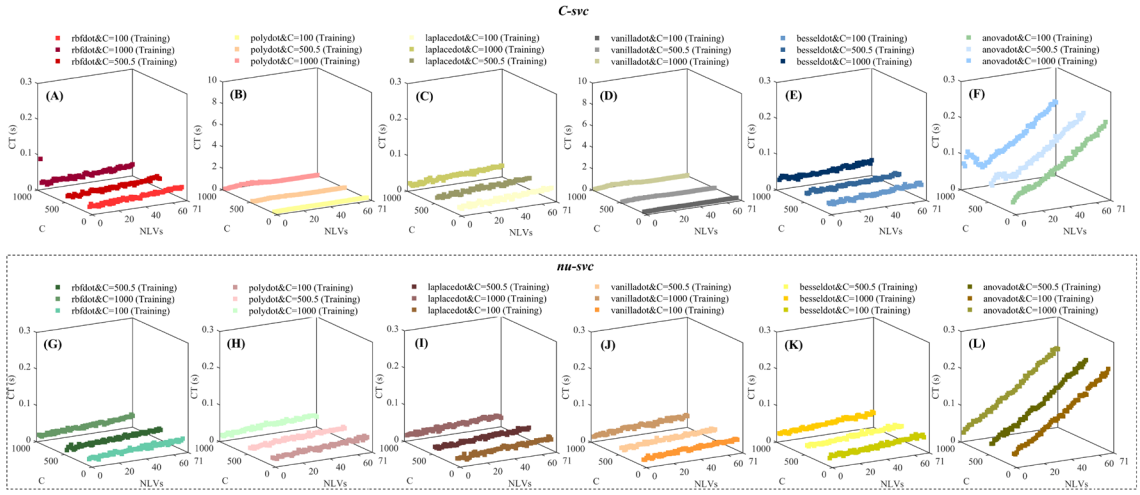


Fig. 21. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

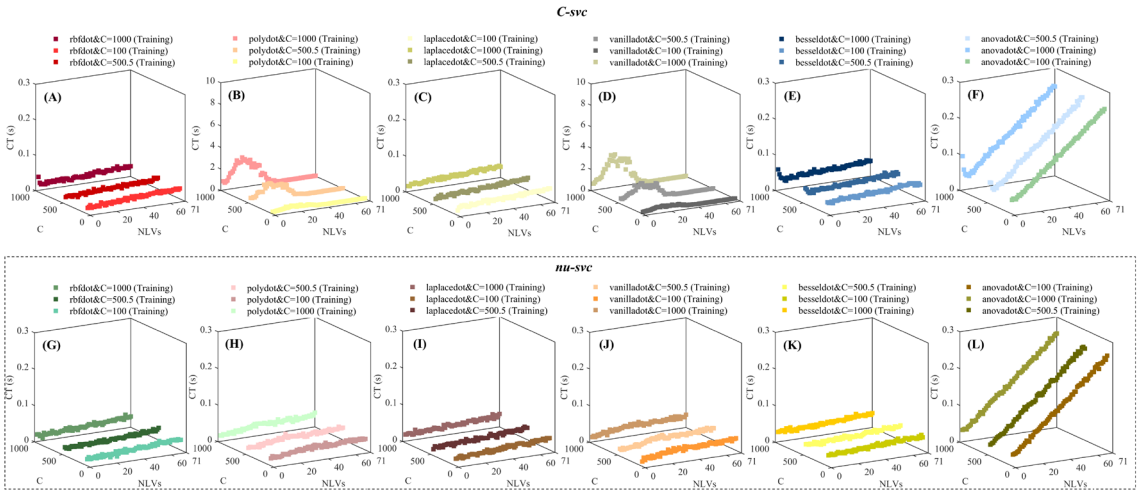


Fig. 22. Computational time (CT) of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of CT are presented for the training (75%) dataset.

Regarding the statistical results of the other LV-ML models. Considering the TDA (Fig. 23), the performance of LV-RF, LV-NB, LV-LDA, LV-QDA, and LV-GLM exhibits pronounced sensitivity to both the NLVs and hyperparameter tuning. In this setting, training accuracies consistently exceed those obtained on the validation set, indicating that model fitting is strongly influenced by the dimensionality of the latent representation. This divergence, which becomes more pronounced at higher NLVs, reveals clear overfitting tendencies, particularly in models such as LV-RF and LV-QDA, where decision boundaries adapt excessively to training-specific characteristics. Simpler models, such as LV-NB and LV-GLM, yield more stable but comparatively lower accuracies, reflecting their reduced flexibility in capturing the intrinsic variability of TDA features.

Considering the FDA (Fig. 24), the dependence on NLVs and hyperparameters remains evident, but notable distinctions emerge. Models such as LV-RF and LV-LDA achieve improved generalization compared with their time-domain counterparts, suggesting that spectral features encode discriminative structure more effectively. Validation accuracies under FDA are, in several configurations, closer to training values, especially at intermediate NLVs, thereby narrowing the gap observed in TDA. This improvement suggests that frequency-domain features act as a form of natural regularization, filtering redundant or noisy components inherent in raw temporal signals. Nonetheless, the persistence of measurable discrepancies between training and validation accuracies across all models indicates that overfitting risks are not fully mitigated by the transition from temporal to spectral representations.

Using the TFDA (Fig. 25) data approach, most models achieve higher accuracies than under TDA or FDA, confirming the advantage of leveraging complementary information domains. For instance, LV-RF consistently attains superior training performance, whereas LV-LDA and LV-GLM exhibit notable gains in validation accuracy, underscoring the utility of hybrid

features for both flexible and linear models. However, the training/validation gap remains a dominant trend. Models with high representational capacity, such as LV-RF and LV-QDA, are particularly prone to overfitting, while simpler models display less variance but lower peak accuracies. This trade-off highlights the challenge of balancing feature richness with model robustness.

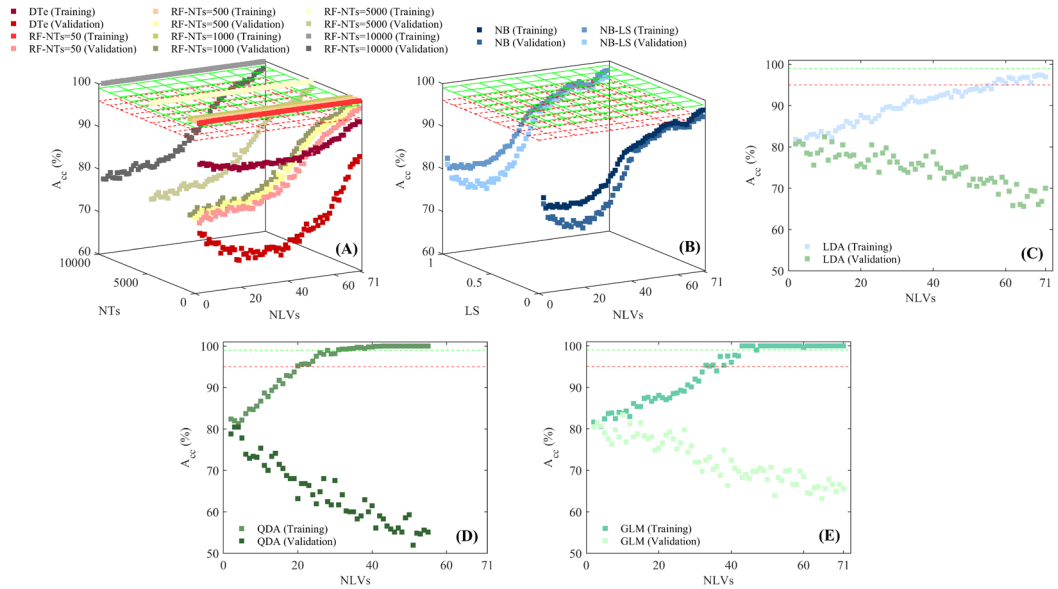


Fig. 23. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-domain approach (TDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

The imposition of block-scale constraints in the TFDA domain adds another dimension to the analysis. Under the TFDABH (Fig. 26), training accuracies for several models, especially LV-RF and LV-QDA, increase sharply, reflecting the adaptability of these classifiers to highly localized time-frequency structures. However, validation performance does not consistently follow this trend, and in many cases the training–validation gap widens substantially. This outcome underscores the risk that overly rigid block constraints, while effective in capturing local dependencies, may exacerbate overfitting in the absence of suitable regularization. Conversely, the TFDABS (Fig. 27) yields a more favorable balance between training and validation performance. Here, the gains in training accuracy are more moderate than under TFDABH, but validation accuracies are generally higher and more stable across NLVs. This suggests that the TFDABS design provides a more flexible mechanism for integrating local time-frequency dependencies without enforcing rigid structural constraints. As a result, models trained under TFDABS exhibit stronger generalization capability, positioning this formulation as a potentially more robust strategy for real-world scenarios where unseen variability must be accommodated.

Beyond classification accuracy, the assessment of CT (Fig. 28) constitutes a critical dimension of model feasibility. The results reveal a clear trade-off between predictive performance and computational efficiency. Frameworks incorporating enriched representations, such as TFDA, TFDABH, and TFDABS, consistently demand longer CTs compared with TDA and FDA. This computational cost was especially pronounced for complex models such as LV-RF, where both training and inference phases are computationally intensive. By contrast, simpler models like LV-NB and LV-GLM maintain relatively low computational demands across all approaches, albeit at the cost of reduced accuracy. This trade-off emphasizes the practical consideration that while enriched feature domains can maximize predictive performance, their elevated computational cost may constrain applicability in real-time or resource-limited contexts. Simpler TDA or FDA strategies, though less powerful in absolute accuracy, remain viable alternatives when efficiency is prioritized.

These results demonstrate that model performance was jointly conditioned by feature-domain representation, classifier complexity, and the NLVs. The persistent training/validation discrepancy in A_{cc} across all approaches highlights the central role of regularization and model selection strategies in mitigating overfitting. Furthermore, the computational analysis shows that the most accurate approaches are not necessarily the most practical, reinforcing the importance of balancing predictive reliability with computational efficiency in application-specific settings.

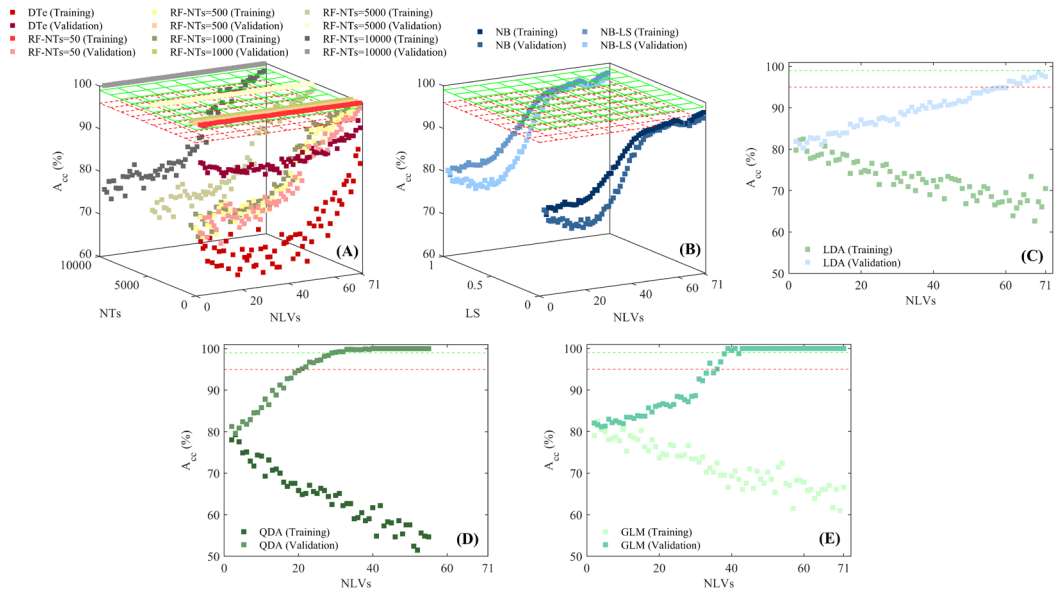


Fig. 24. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the frequency-domain approach (FDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

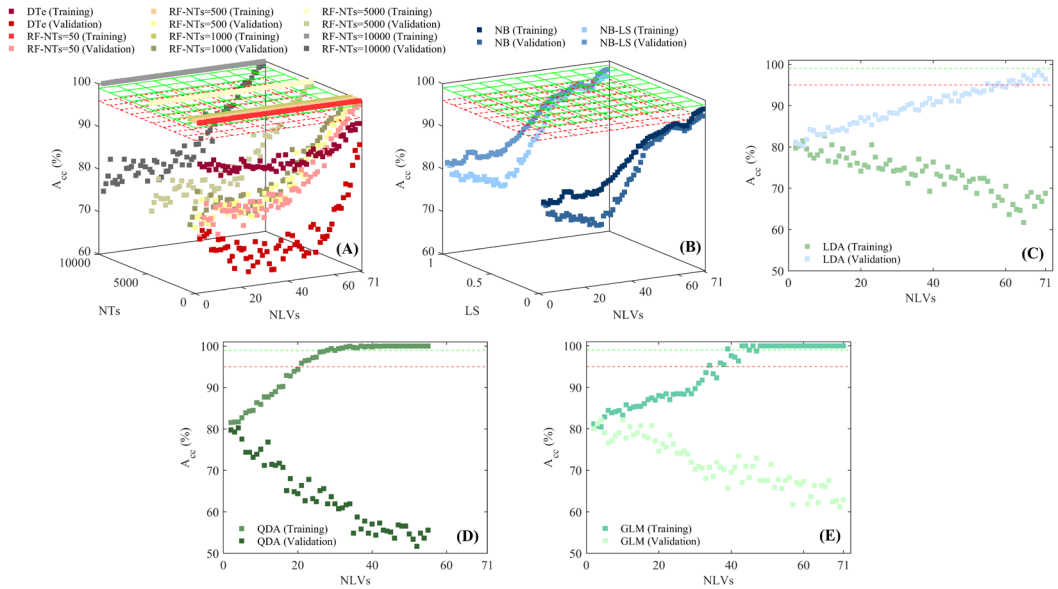


Fig. 25. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach (TFDA), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

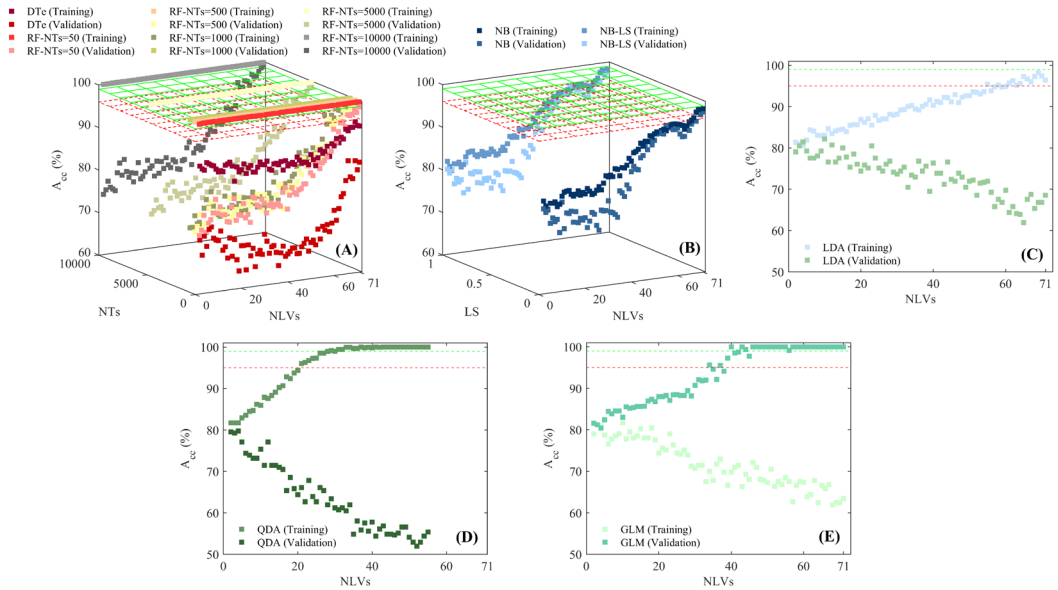


Fig. 26. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

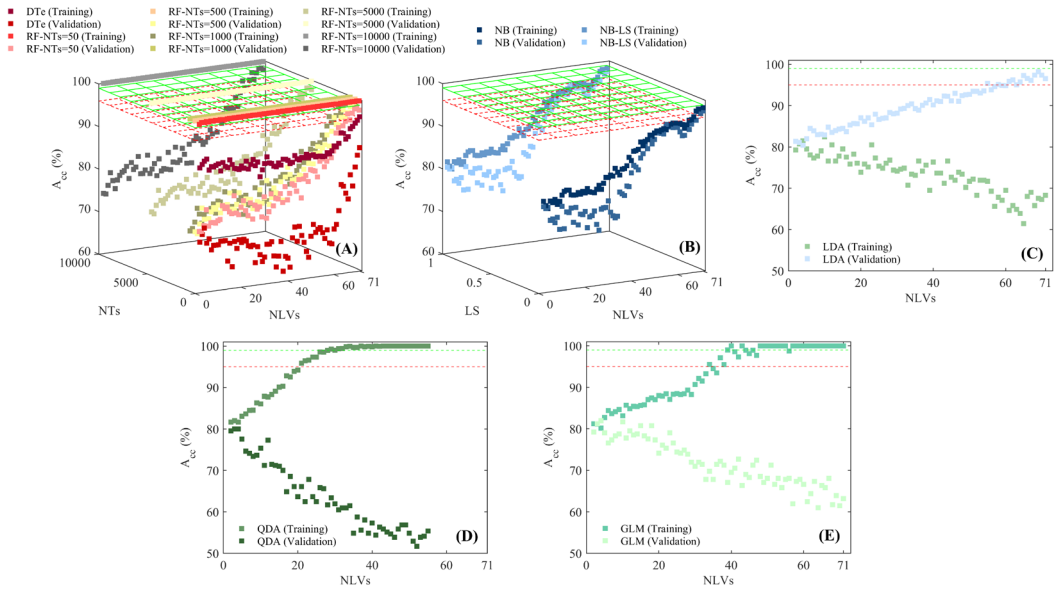


Fig. 27. Statistical classification performance of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

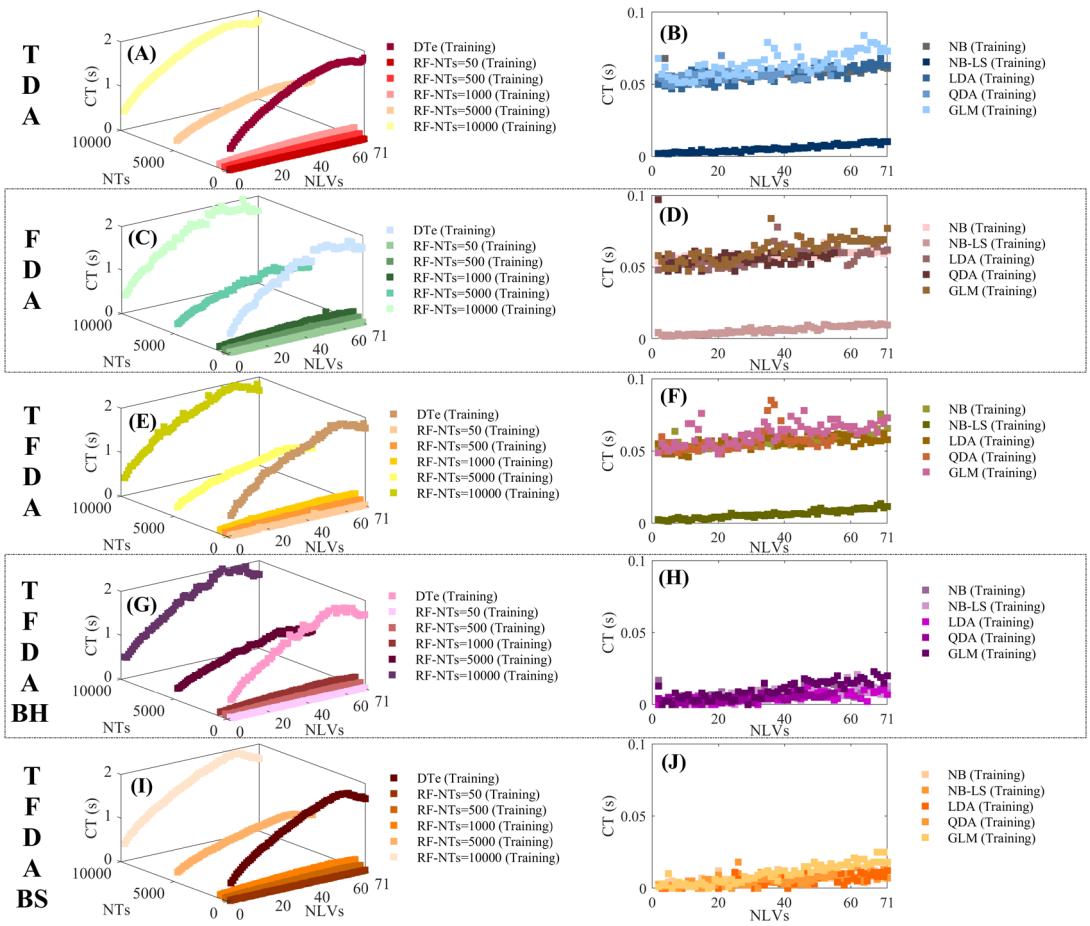


Fig. 28. Computational time (CT) of the Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the ML hyperparameters and the number of latent variables (NLVs) tested. Results of CT are reported for the training (75%) dataset.

The assessment of RF-VS-LV-SVM models under the TDA (Fig. 29) highlights the sensitivity of classification A_{cc} to the choice of kernel function, SVM type, and the regularization parameter. Training accuracies tend to remain consistently higher than validation accuracies, a pattern that becomes accentuated for larger values of the regularization parameter (C = 1000). Kernels such as rbfdot and laplacedot generally provide superior training performance, whereas validation performance is more stable for polydot and vanilladot. The differences between training and validation results suggests that highly flexible kernels, while effective in optimizing the decision boundary during training, increase the risk of overfitting. In terms of CT, training with more complex kernels, such as anovadot and besseldot, requires significantly longer runtimes, indicating that kernel choice directly impacts computational feasibility.

When the same model is applied considering the FDA (Fig. 30), the overall performance patterns remain consistent, but the validation accuracies improve relative to TDA. This improvement indicates that spectral representations enhance the discriminative structure of the data, particularly when combined with regularization. For example, the rbfdot kernel achieves balanced results with reduced gaps between training and validation accuracies, especially for intermediate C values (C = 500.5). However, the CT under FDA is generally higher than in TDA, reflecting the added complexity of frequency-domain feature extraction. The improvement in generalization combined with the computational penalty emphasizes the importance of balancing accuracy gains with processing cost.

The integration of temporal and spectral features through the TFDA (Fig. 31) further amplifies these observations. The highest training accuracies are consistently achieved under this configuration, particularly for rbfdot and laplacedot KF. Validation accuracies also improve compared with both TDA and FDA, indicating that joint time-frequency representations provide richer discriminatory information. Nevertheless, the gap between training and validation remains significant in models using high-capacity kernels, especially under the C-svc type with large regularization constants. CT reach their maximum under TFDA, confirming that while enriched representations boost predictive performance, they impose a heavy computational load during training.

The introduction of block-scale constraints modifies these trend. By using the TFDABH (Fig. 32), training performance exhibits sharp increases across most KFs, but validation performance does not scale proportionally, thereby widening the discrepancy. Kernels with strong non-linear mapping capacity, such as rbfdot, display pronounced overfitting, while simpler kernels maintain lower but more stable validation accuracies. Computational times under TFDABH are consistently higher than in TFDA, highlighting the additional cost introduced by block-scale partitioning. Conversely, in the TFDABS (Fig. 33) yields more balanced outcomes. Training A_{cc} values were slightly lower than under TFDABH, but validation A_{cc} are consistently higher and less variable across KF and C values. The reduced gap between training and validation suggests that the soft-block constraint provides an effective balance between model flexibility and generalization. Computational costs under TFDABS remain elevated but are somewhat reduced compared with TFDABH, which reinforces the potential of this approach as a more computationally efficient compromise.

A comparison with the RF-VS-LV-RF (Fig. 34) model reveals complementary trends. The RF-based framework benefits less from kernel flexibility but is highly sensitive to the NTs and DTe depth. Training A_{cc} tend to saturate rapidly with increasing NTs, whereas validation accuracies plateau at lower levels, highlighting the diminishing returns of increasing ensemble size. Computational time rises nearly linearly with NTs, reflecting the inherent cost of ensemble growth. Across all domain approaches, RF models are computationally heavier than Naïve Bayes or GLM counterparts but lighter than SVMs with complex KF.

The RF-VS-LV-NB (Fig. 35) model exhibits markedly different behavior. Performance is less sensitive to hyperparameter tuning, with LS exerting only marginal influence on both

training and validation A_{cc}. Compared with SVM and RF models, NB shows more modest peak accuracies but narrower gaps between training and validation sets, suggesting reduced overfitting. CT remain minimal across all approaches, establishing NB as a computationally efficient though less competitive classifier in terms of absolute performance.

Finally, the performance of RF-VS-LV-LDA, RF-VS-LV-QDA, and RF-VS-LV-GLM (Fig. 36) underscores the influence of classifier complexity as a function of the different feature data approaches tested. LDA and GLM deliver consistently low accuracies, with a small difference in A_{cc} between training and validation datasets, reflecting limited predictive capability despite their robustness against overfitting. In contrast, QDA attains higher training A_{cc} but exhibits pronounced overfitting, particularly under TFDA and TFDABH, where validation performance decreased sharply. Although these models achieve the shortest CT (Figs. 36B, 36D, 36F, 36H and 36J), their limited classification power positions them as suitable only for resource-constrained settings where efficiency is prioritized over accuracy. The comparative analysis confirms that predictive performance was highly influenced by the interaction between data approach, classifier complexity, and hyperparameter configuration of these models.

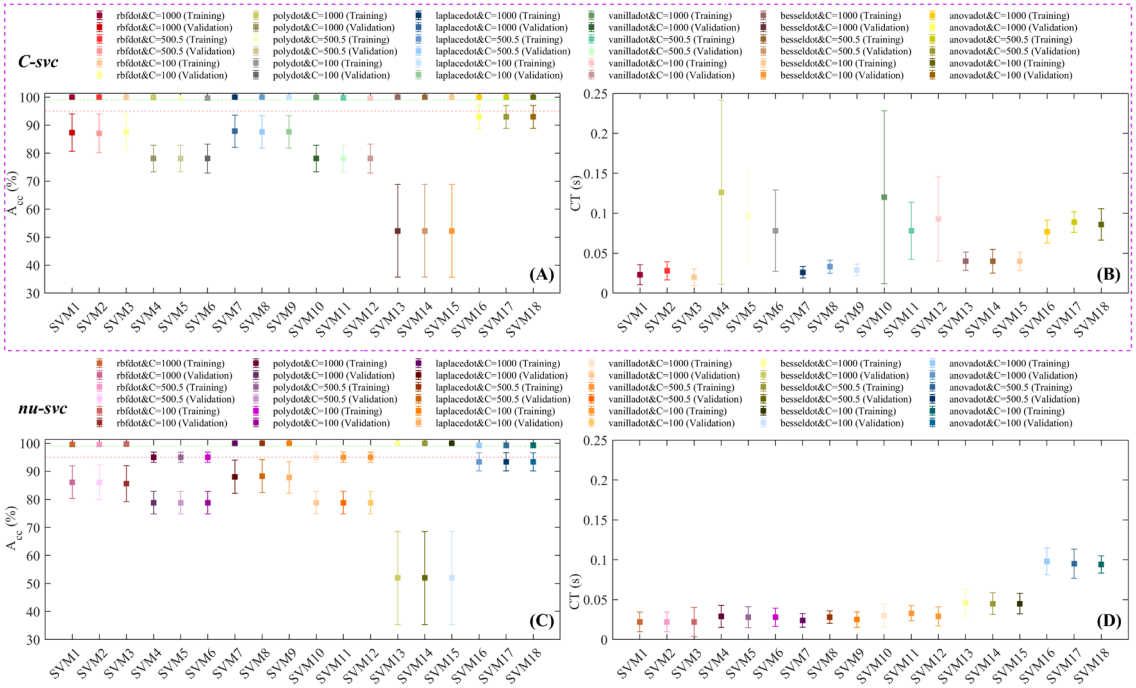


Fig. 29. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

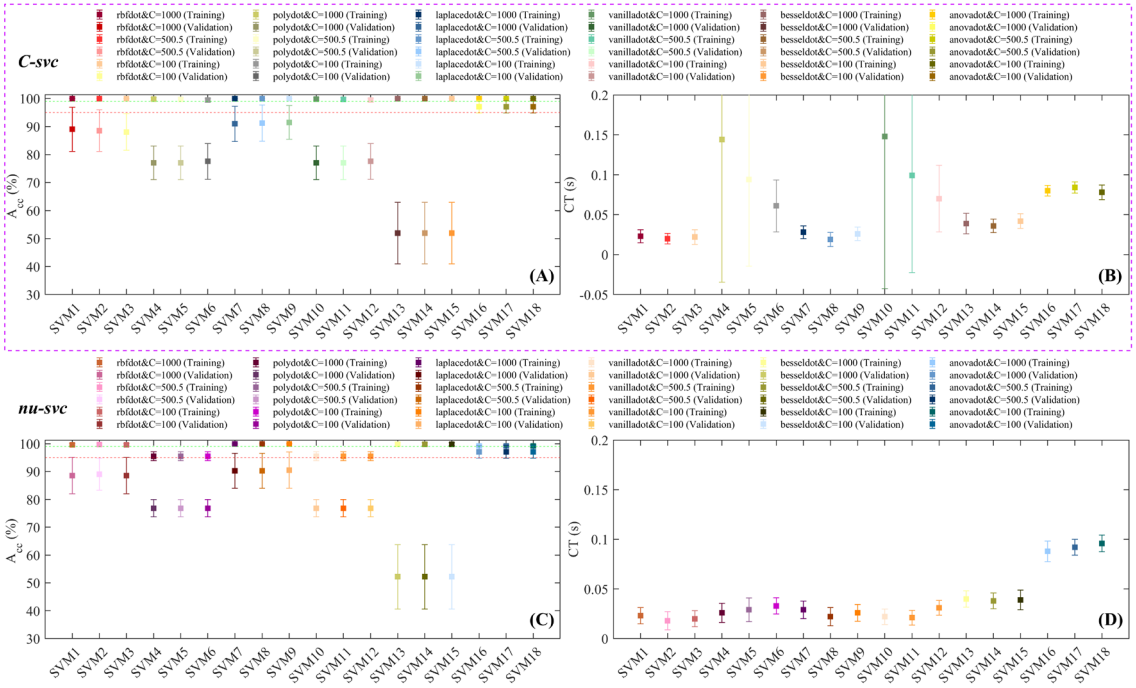


Fig. 30. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

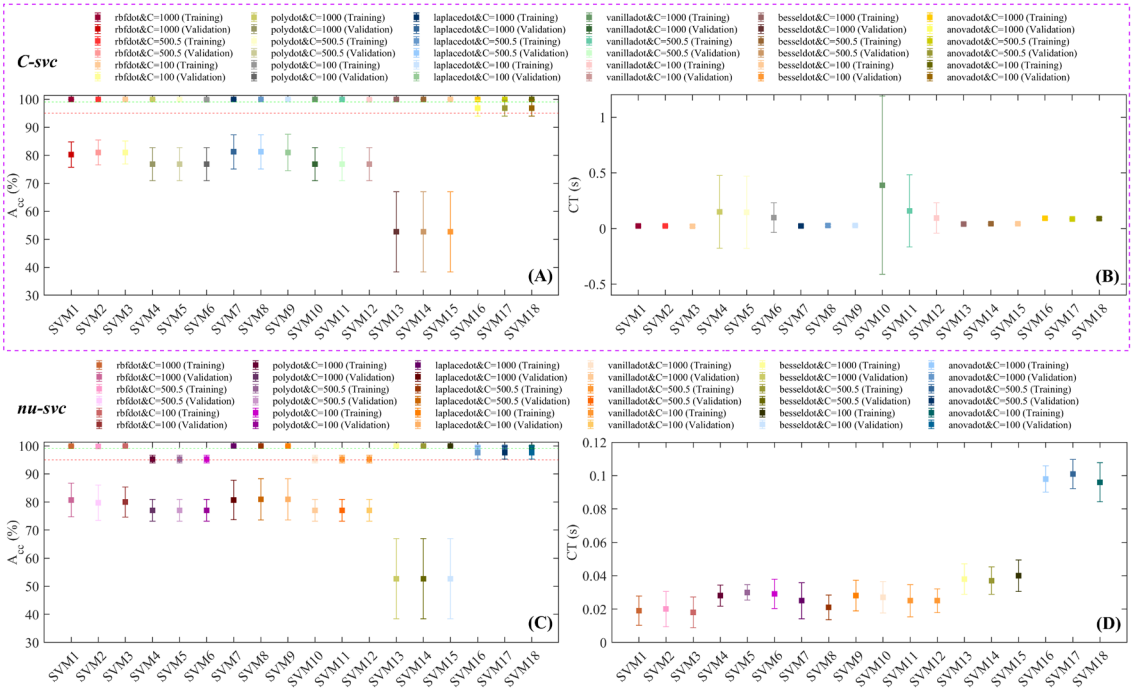


Fig. 31. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

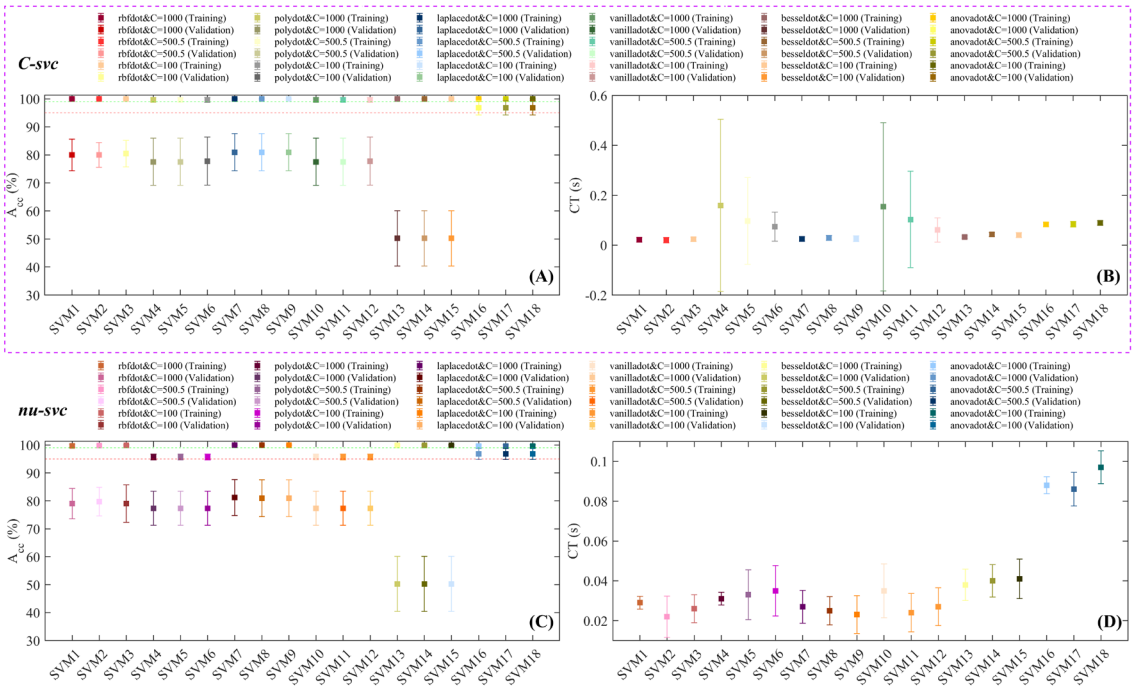


Fig. 32. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

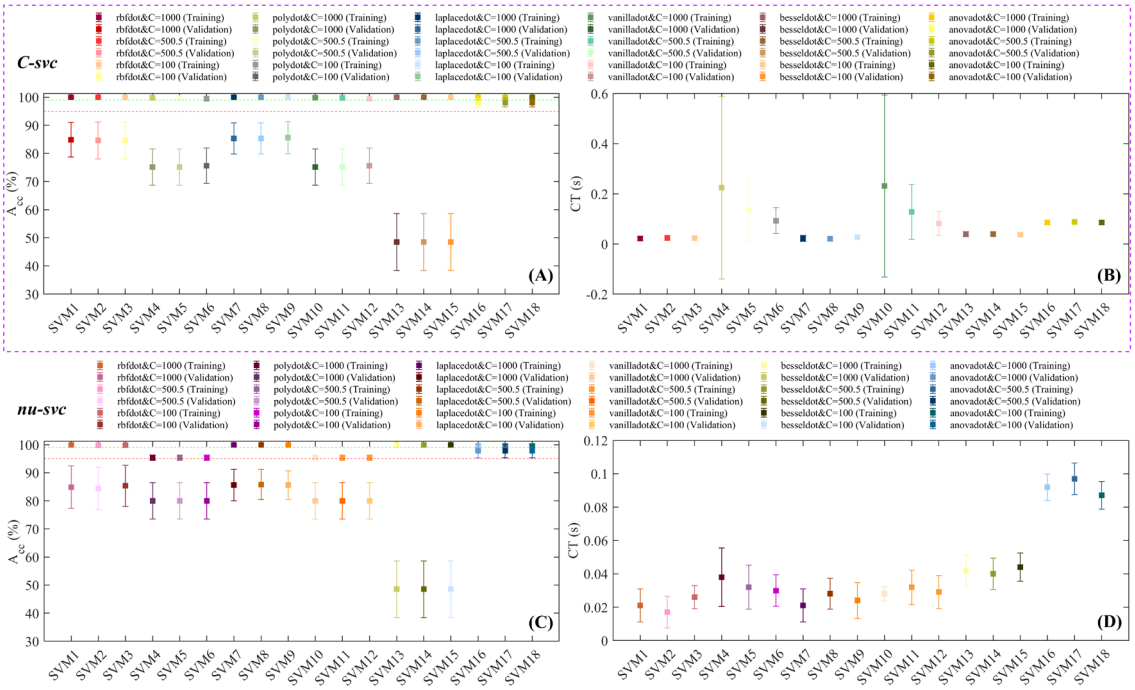


Fig. 33. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Furthermore, computational time (CT) is also presented a mean ± standard deviation for the training process. Kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), SVM' type (C-svc and nu-svc) and regularization parameter (C; 100, 500.5, and 1000).

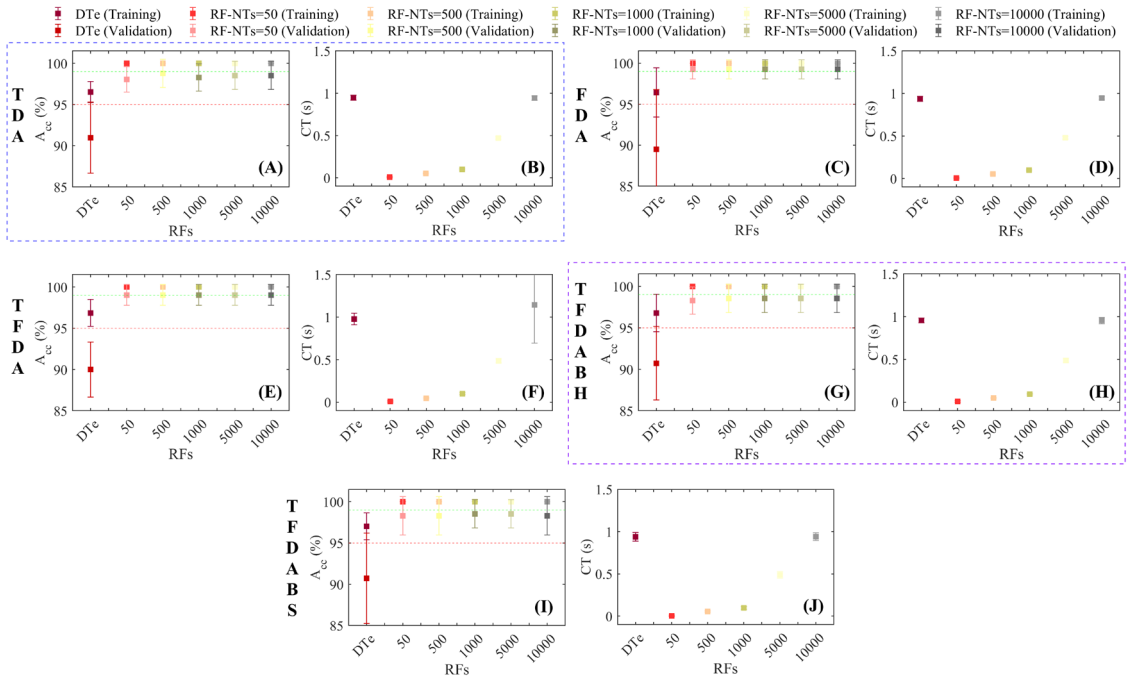


Fig. 34. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the RF hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process. DTe (decision tree), NTs (number of trees).

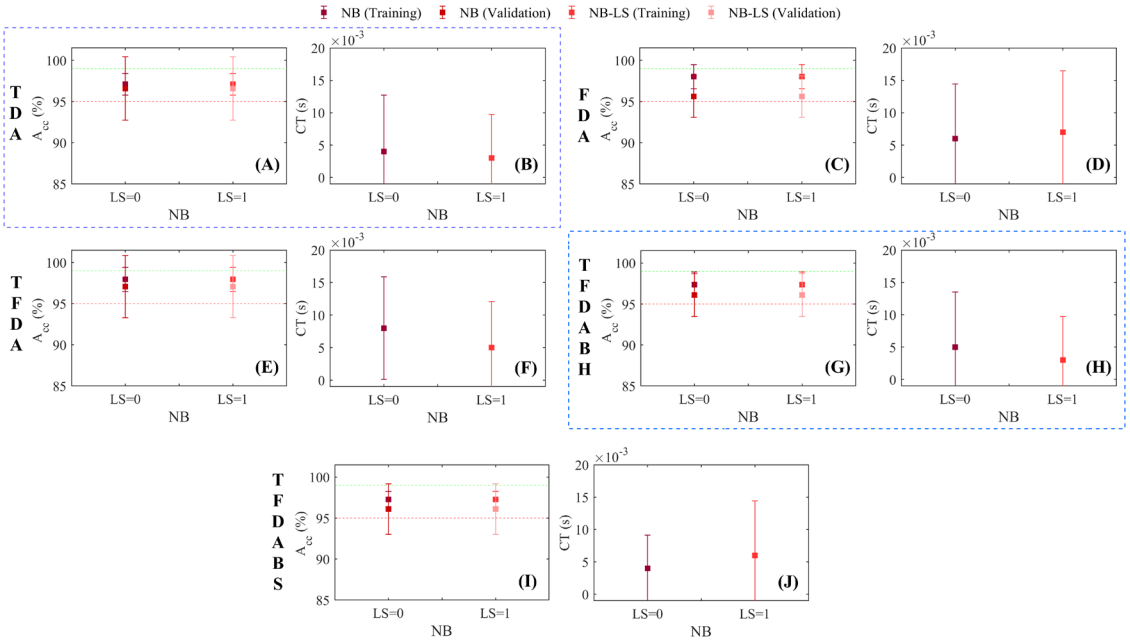


Fig. 35. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Naïve Bayes (RF-VS-LV-NB) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the NB hyperparameters. Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process. LS (Laplace Smoothing).

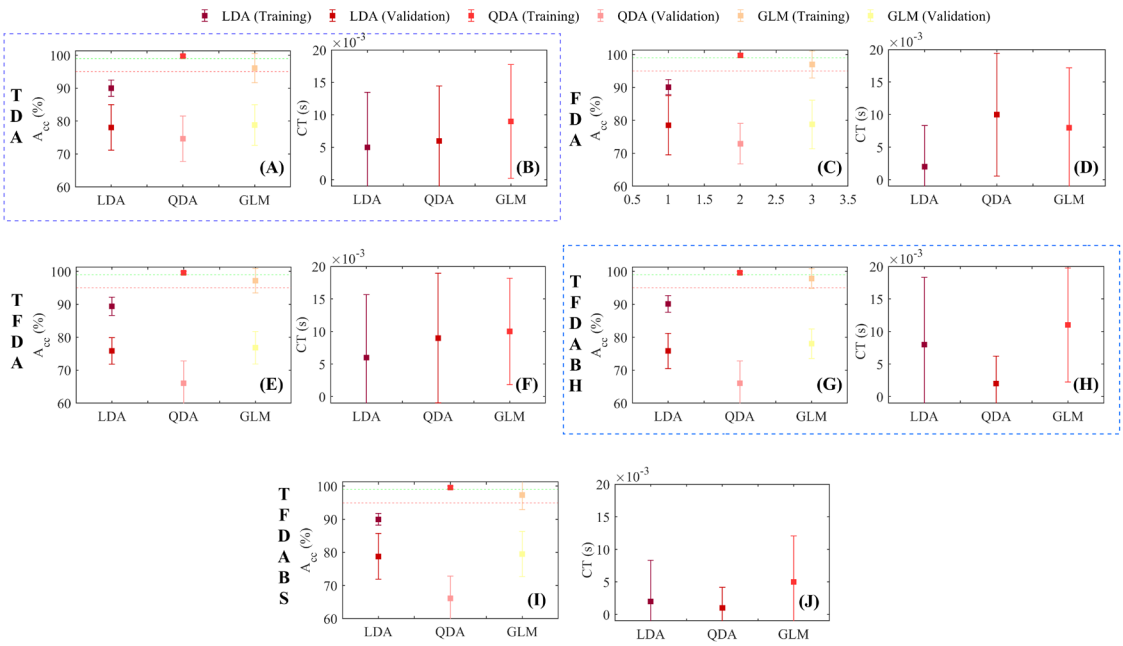


Fig. 36. Statistical classification performance of the Random Forest-Variable Selection-Latent Variable-Linear Discriminant Analysis (RF-VS-LV-LDA), Random Forest-Variable Selection-Latent Variable-Quadratic Discriminant Analysis (RF-VS-LV-QDA) and Random Forest-Variable Selection-Latent Variable-Generalized Linear Model Analysis (RF-VS-LV-GLM) using the time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS). Results of overall accuracy (A_{cc}) are reported as a mean ± standard deviation separately for the training (75%) and the validation (25%) datasets. Additionally, computational time (CT) is also presented a mean ± standard deviation for the training process.

The application of PLSR modeling strategy provides an advanced framework to assess the statistical classification performance of LV-ML and RF-VS-LV-ML models while simultaneously accounting for hyperparameter configuration, NLVs, and domain-specific data representations (TDA, FDA, TFDA, TFDABH and TFDABS). In the case of LV-SVM, the PLSR model (Fig. 37) reveals that KF, SVM's type (C-svc vs. nu-svc), and the C parameters were among the most influential factors driving predictive performance accuracy. Models employing non-linear KF such as rbfdot and laplacedot achieve higher A_{ccT}, but these gains are often accompanied by wider discrepancies with A_{ccV}, suggesting model overfitting. Performance indicators such as S_e, S_p, P_r, R_e, and F_s follow similar trends, with higher values consistently observed in the training dataset compared with validation. The $RMSE_{TR}$ and $RMSE_{CV}$, R^2 and Q^2 values confirmed that there was a slightly difference between model fitting for training and cross-validation datasets. Variables with higher VIP scores include TFDA, TFDABH, TFDABS, Type, all of KF and NVLs, highlighting the importance of these variables in influence of the statistical performance of LV-SVM model. While variables such as C and RS did not significantly influence the responses.

When the screened PLSR model is considered (Fig. 38), where variables with VIP<0.5 (TDA, FDA, C=100, C=500, C=1000 and RS) are removed, robustness is improved and spurious correlations are reduced. The results show a more balanced relationship between training and validation performance. This result indicated that variable screening enhances the stability of the PLSR interpretation and reduces noise introduced by low-influence regressors.

The analysis of the LV-RF using the PLSR (Fig. 39) demonstrates a distinct dependency on ensemble hyperparameters, specifically the NTs and DTe. Training accuracies approach saturation rapidly as NTs increase, whereas validation accuracies plateau earlier, emphasizing diminishing returns in ensemble expansion. The VIP analysis highlights NTs as one of the most critical predictors, with large ensembles (NTs ≥ 5000) leading to high training fitting results but limited improvement in validation performance. Performance metrics such as MCC and AUC reinforce this pattern, showing consistently higher training values compared with validation. Removal of low-VIP regressors improves the interpretability of the model but does not fully mitigate the overfitting tendency of large ensembles.

The PLSR modeling fitting on the LV-NB (Fig. 40) results provides a contrasting perspective. In this model, the influence of hyperparameters was less pronounced, with LS contributing only marginally to performance. The model yields moderate training accuracies but relatively stable validation accuracies, with smaller gaps between R^2 and Q^2 compared with SVM or RF. P_r, R_e, and F_s values were lower than in LV-SVM and LV-RF but balanced across training and validation datasets, indicating limited overfitting. VIP analysis places greater weight on the NVLs, suggesting that classification performance was highly dependent of the NVLs used in the model training irrespective of the data approach used and the applications of LS.

In the case of LV-LDA (Fig. 41), PLSR results confirm the stability of this linear classifier. Validation accuracies remain close to training values, and R^2 and Q^2 show smaller discrepancies relative to more complex models. Although overall accuracy is lower than that of LV-SVM or LV-RF, performance measures such as sensitivity and specificity are consistent across training and validation, reflecting good generalization. Low RMSECV values and modest RSS confirm the model's robustness. Variables with higher VIP scores include NLVs and data representation, again pointing to the feature extraction stage as the most critical determinant of performance in LDA.

Conversely, the LV-QDA (Fig. 42) displays higher training accuracies but reduced generalization capability. Discrepancies between R^2 and Q^2 are more pronounced than in LDA, and $RMSE_{CV}$ values increase, indicating susceptibility to overfitting. Performance indicators such as recall and F-score are elevated for training but drop considerably for validation datasets. VIP analysis highlights the strong dependence on NLVs, which, when set too high, introduce variance increases. The LV-GLM (Fig. 43) exhibits intermediate behavior, with performance metrics indicating modest but reliable classification capacity. R^2 and Q^2 values are generally aligned, and residual errors remain moderate. S_e and P_r show less fluctuation between training and validation compared with QDA, highlighting GLM as a more stable alternative compared to LV-LDA and LV-QDA .

When PLSR is applied to models incorporating RF-VS, additional patterns emerge. For RF-VS-LV-SVM (Fig. 44), Type and KF continue to influence the statistical figures of merit of models. The RF-VS-LV-RF (Fig. 45) shows similar trends, with ensemble hyperparameters remaining critical, but with reduced overfitting compared with LV-RF, confirming that variable screening moderates excessive variance introduced in the model.

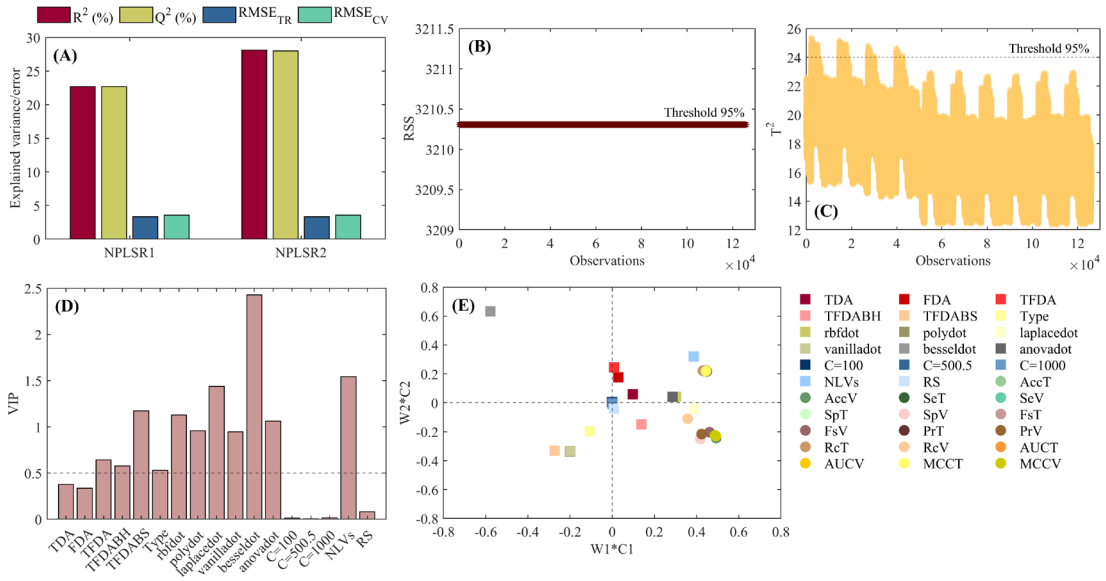


Fig. 37. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Support Vector Machines (LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the SVM's hyperparameters and number of latent variables (NLVs). NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

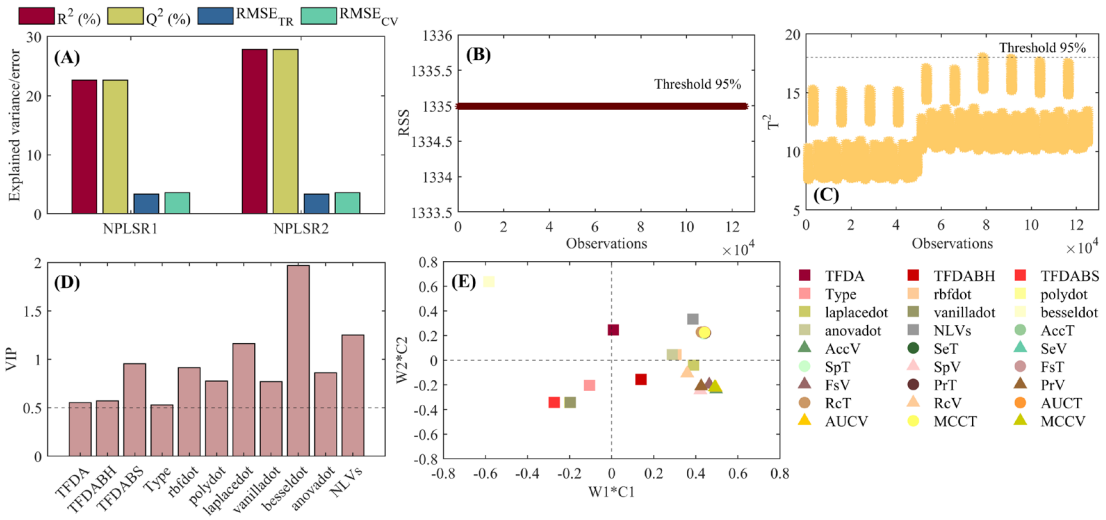


Fig. 38. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Support Vector Machines (LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the SVM's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

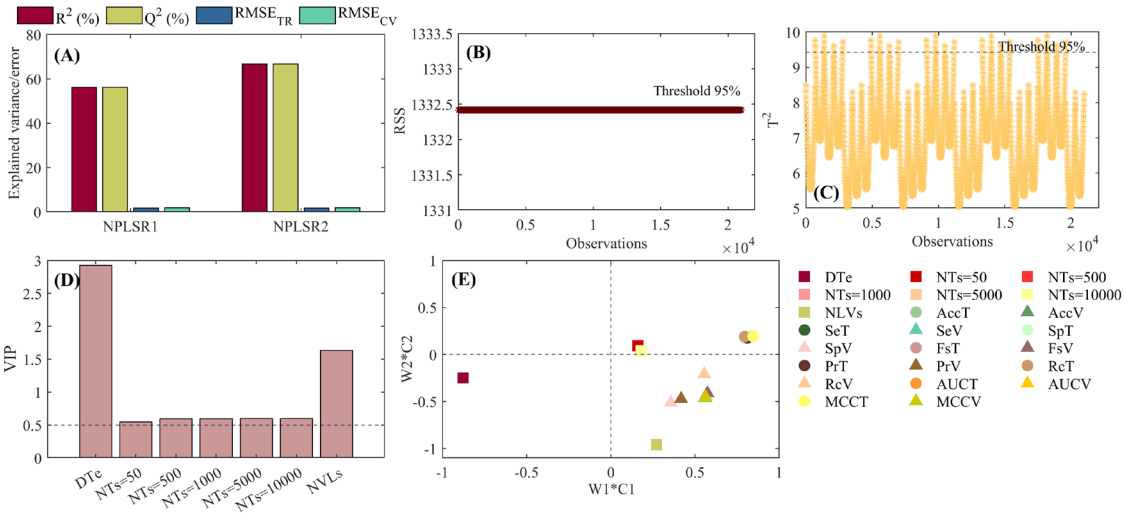


Fig. 39. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Random Forest (LV-RF) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the RF's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), DTe (decision tree), NTs (number of trees; 50, 500 1000 5000 10000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

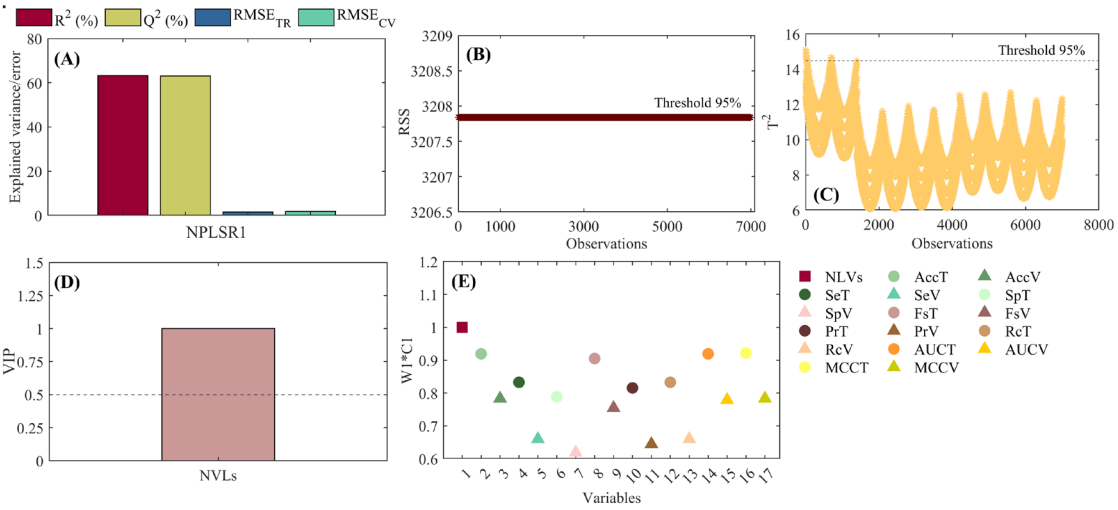


Fig. 40. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Naïve Bayes (LV-NB) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and the NB's hyperparameters and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

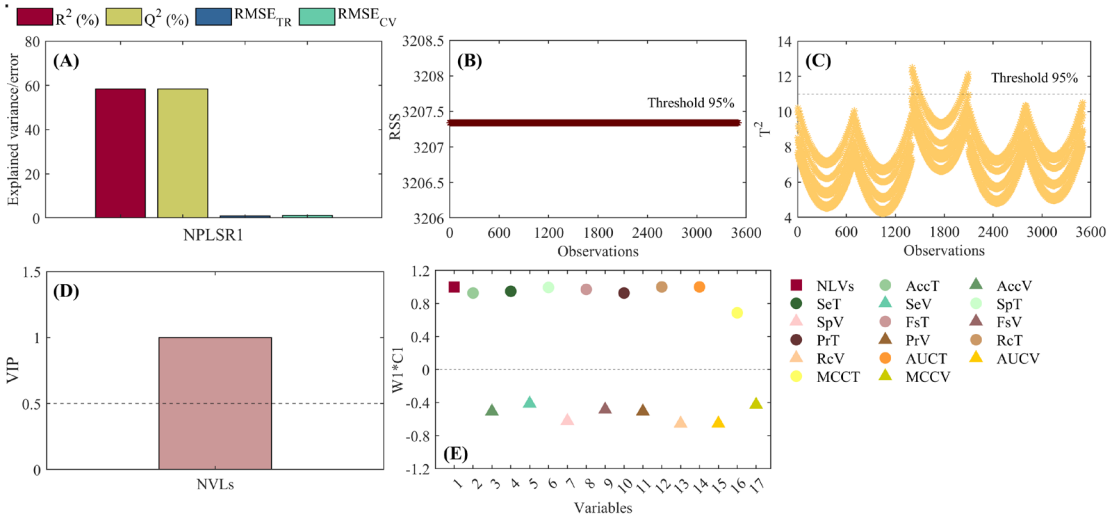


Fig. 41. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Linear Discriminant Analysis (LV-LDA) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

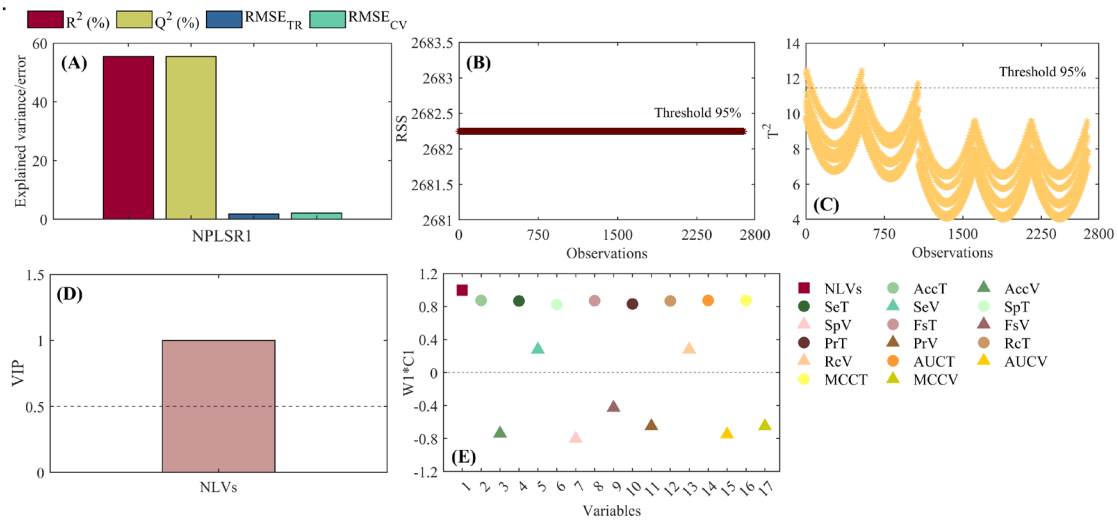


Fig. 42. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Quadratic Discriminant Analysis (LV-QDA) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

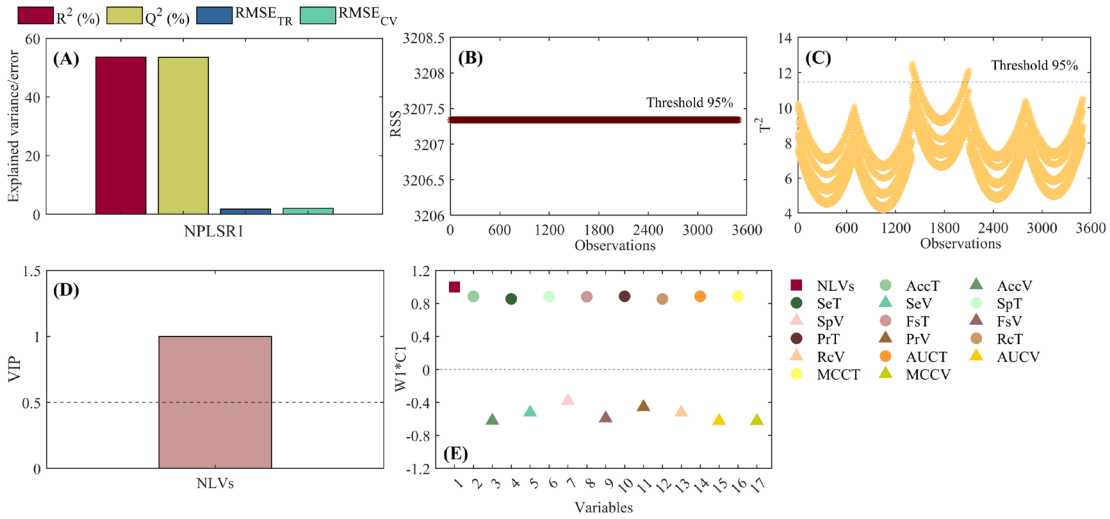


Fig. 43. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Latent Variable-Generalized Linear Model (LV-GLM) considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS), and number of latent variables (NLVs). Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

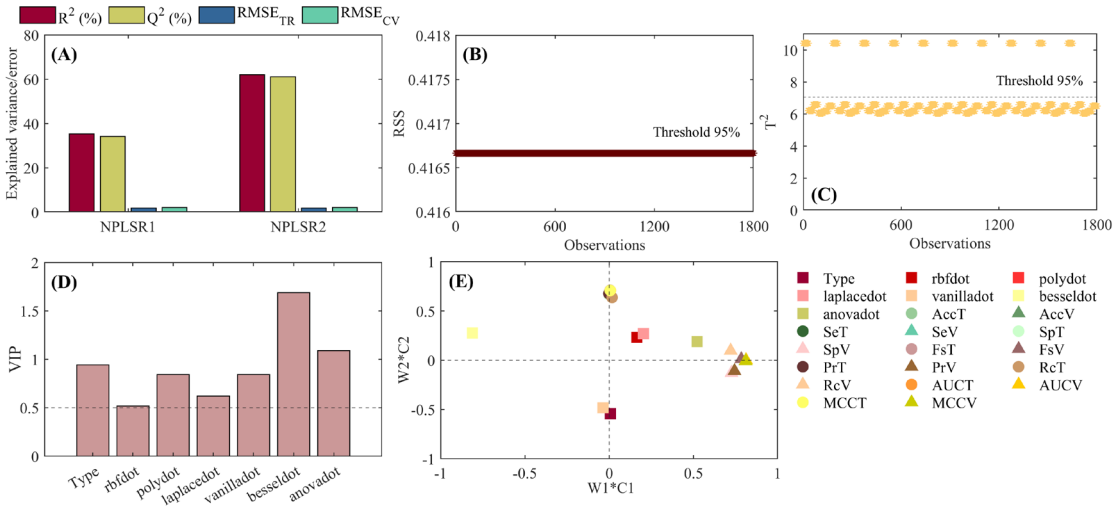


Fig. 44. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Random Forest-Variable Selection-Latent Variable-Support Vector Machines (RF-VS-LV-SVM) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS) and the SVM's hyperparameters. Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), KF (kernel functions; rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), type (C-svc and nu-svc), C (regularization parameter; 100, 500.5, and 1000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

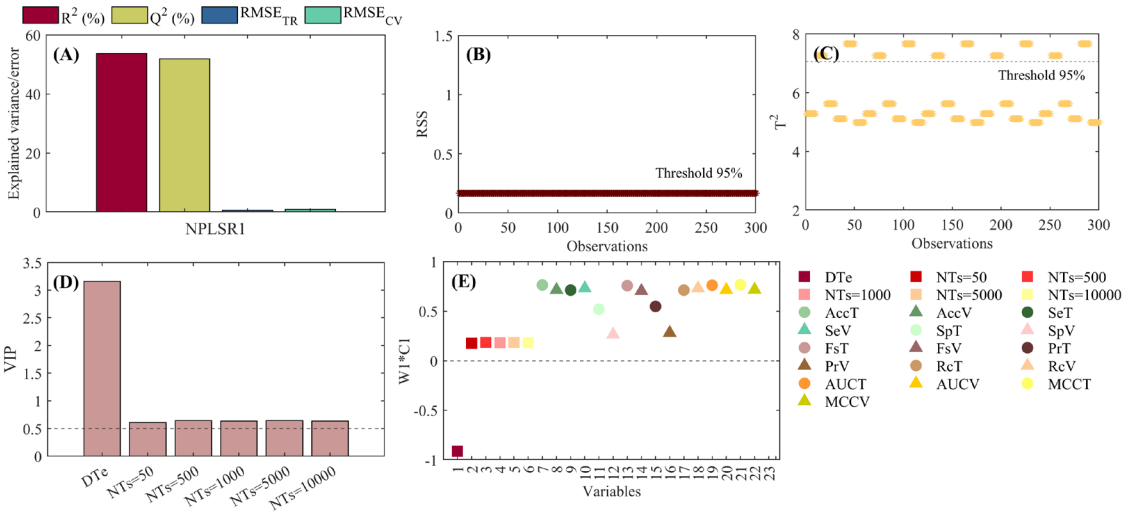


Fig. 45. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) model considering simultaneously the data approach: time-domain approach (TDA), frequency-domain approach (FDA), time-frequency-domain approach (TFDA), time-frequency-domain approach-block-scale hard (TFDABH) and time-frequency-domain approach-block-scale soft (TFDABS) and the RF's hyperparameters. Results of the PLSR model are presented for the screened model; regressor variables with VIP values lower than 0.5 were removed to improve model robustness. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), DTe (decision tree), NTs (number of trees; 50, 500, 1000, 5000 and 10000), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

The statistical results in the use of PLSR for modeling and multi-objective optimizing the LV-ML and RF-VS-LV-ML models are summarized in Table 14 to 20.

Table 14. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Support Vector Machine (LV-SVM) and Latent Variable-Random Forest (LV-RF) and the number of latent variables (NLVs) used in model tunning.

LV-SVM				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.375		ns	
FDA	0.336		ns	
TFDA	0.443		0.554	
TFDABH	0.579		0.572	
TFDABS	1.174		0.953	
Type	0.529		0.530	
rbfdot	1.127		0.912	
polydot	0.958	ONPLSR= 2	0.777	ONPLSR= 2
laplacedot	1.437	R^2= 28.1%	1.163	R^2= 27.8%
vanilladot	0.948	Q^2= 28%	0.769	Q^2= 27.7%
besseldot	2.428	$RMSE_{TR}$= 3.34	1.969	$RMSE_{TR}$= 3.35
anovadot	1.063	$RMSE_{CV}$= 3.59	0.860	$RMSE_{CV}$= 3.59
C=100	0.012		ns	
C=500.5	0.004		ns	
C=1000	0.015		ns	
NLVs	1.542		1.250	
RS	0.080		ns	
LV-RF				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.206		ns	
FDA	0.009		ns	
TFDA	0.064		ns	
TFDABH	0.074		ns	
TFDABS	0.076		ns	
DTe	2.924	ONPLSR=2	2.924	ONPLSR= 2
NTs=50	0.547	R^2= 66.9%	0.547	R^2= 66.8%
NTs=500	0.592	Q^2= 66.8%	0.592	Q^2= 66.7%
NTs=1000	0.595	$RMSE_{TR}$= 1.66	0.595	$RMSE_{TR}$= 1.67
NTs=5000	0.597	RMSEcv= 1.92	0.597	RMSEcv= 1.92
NTs=10000	0.597		0.597	
NLVs	1.635		1.635	
RS	0.029		ns	

TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), VIP (variable importance for the projection), ONPLSR (optimal number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RS (replication of runs), kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), C (regularization parameter), DTe (decision tree) and NTs (number of trees).

Table 15. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Naïve Bayes (LV-NB) and Latent Variable-Linear Discriminant Analysis (LV-LDA) and the number of latent variables (NLVs) used in model tunning.

LV-NB				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.032		ns	
FDA	0.211		ns	
TFDA	0.079		ns	
TFDABH	0.043	ONPLSR= 1 R^2= 63.5% Q^2= 63.3% $RMSE_{TR}$= 1.61 $RMSE_{CV}$= 1.82	ns	ONPLSR= 1 R^2= 63.2% Q^2= 63.1% $RMSE_{TR}$= 1.62 $RMSE_{CV}$= 1.87
TFDABS	0.057		ns	
LS=0	0.000		ns	
LS=1	0.000		ns	
NLVs	2.990		1.000	
RS	0.035		ns	
LV-LDA				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.029		ns	
FDA	0.006		ns	
TFDA	0.008	ONPLSR= 1 R^2= 52.1% Q^2= 52.0% $RMSE_{TR}$= 1.61 $RMSE_{CV}$= 1.72	ns	ONPLSR= 1 R^2= 58.4% Q^2= 58.3% $RMSE_{TR}$= 1.01 $RMSE_{CV}$= 1.24
TFDABH	0.005		ns	
TFDABS	0.010		ns	
NLVs	2.645		1.000	
RS	0.033		ns	

TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), VIP (variable importance for the projection), ONPLSR (optimal number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RS (replication of runs) and LS (Laplace Smoothing).

Table 16. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach, hyperparameters belonging to the Latent Variable-Quadratic Discriminant Analysis (LV-QDA) and Latent Variable-Generalized Linear Model (LV-GLM) and the number of latent variables (NLVs) used in model tunning.

LV-QDA				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.022	ONPLSR= 1 R^2= 55.4% Q^2= 55.3% $RMSE_{TR}$= 1.83 $RMSE_{CV}$= 2.05	ns	ONPLSR= 1 R^2= 55.4% Q^2= 55.3% $RMSE_{TR}$= 1.83 $RMSE_{CV}$= 2.10
FDA	0.027		ns	
TFDA	0.022		ns	
TFDABH	0.035		ns	
TFDABS	0.034		ns	
NLVs	2.644		1.000	
RS	0.075		ns	
LV-GLM				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.013	ONPLSR= 1 R^2= 53.6% Q^2= 53.5% $RMSE_{TR}$= 1.83 $RMSE_{CV}$= 2.01	ns	ONPLSR= 1 R^2= 53.6% Q^2= 53.5% $RMSE_{TR}$= 1.83 $RMSE_{CV}$= 2.01
FDA	0.030		ns	
TFDA	0.011		ns	
TFDABH	0.016		ns	
TFDABS	0.016		ns	
NLVs	2.645		1.000	
RS	0.048		ns	

TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), VIP (variable importance for the projection), ONPLSR (optimal number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset) and RS (replication of runs).

Table 17. Statistical results of Partial Least Square Regression (PLSR) model fitting to assess the influence of data approach and hyperparameters belonging to each Machine Learning model.

Random Forest-Variable Selection-Latent Variable-Support Vector Machine (RF-VS-LV-SVM)				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.099		ns	
FDA	0.149		ns	
TFDA	0.091		ns	
TFDABH	0.180		ns	
TFDABS	0.035		ns	
Type	1.420		0.943	
rbfdot	0.780	ONPLSR= 2 R^2= 62.5% Q^2= 61.2% $RMSE_{TR}$= 1.79 $RMSE_{CV}$= 2.10	0.518	ONPLSR= 2 R^2= 62.0% Q^2= 61.1% $RMSE_{TR}$= 1.79 $RMSE_{CV}$= 2.10
polydot	1.271		0.844	
laplacedot	0.933		0.619	
vanilladot	1.271		0.844	
besseldot	2.542		1.688	
anovadot	1.642		1.090	
C=100	0.019		ns	
C=500.5	0.004		ns	
C=1000	0.023		ns	
RS	0.203		ns	
Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF)				
Variable	Raw model		Screened model	
	VIP	Goodness of fit	VIP	Goodness of fit
TDA	0.047		ns	
FDA	0.038		ns	
TFDA	0.039		ns	
TFDABH	0.028		ns	
TFDABS	0.001	ONPLSR= 1 R^2= 53.7% Q^2= 50.8% $RMSE_{TR}$= 0.65 $RMSE_{CV}$= 0.92	ns	ONPLSR= 1 R^2= 53.7% Q^2= 51.9% $RMSE_{TR}$= 0.65 $RMSE_{CV}$= 0.89
DTe	3.161		3.161	
NTs=50	0.606		0.606	
NTs=500	0.643		0.643	
NTs=1000	0.635		0.635	
NTs=5000	0.644		0.644	
NTs=10000	0.634		0.634	
RS	0.043		ns	
Models	RF-VS-LV-NB*		RF-VS-LV-LDA*	
	RF-VS-LV-QDA*		RF-VS-LV-GLM*	

TDA (time-domain approach), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), TFDABS (time-frequency-domain approach-block-scale soft), VIP (variable importance for the projection), ONPLSR (optimal number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RS (replication of runs), kernel functions (rbfdot, polydot, laplacedot, vanilladot, besseldot, and anovadot), C (regularization parameter), DTe (decision tree) and NTs (number of trees). *No model was built because the first predictive component was already not significant.

Table 18. Multi-objective optimized Latent Variable-Machine Learning (LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Goodness of fit	Optimized models			
	LV-SVM	LV-RF	LV-NB	LV-LDA
	Configuration i) TFDA ii) Type: C-svc iii) KF: anovadot iv) C=100 v) NLVs=69	Configuration i) TFDA ii) NTs=50 iii) NLVs=71	Configuration i) TFDABH ii) LS=1 iii) NVLs=69	Configuration i) FDA ii) NLVs=3
A_{cc} (%)	Training: 100.00 ± 0.00^{aA} Validation: 98.05 ± 2.24^{aA}	Training: 100.00 ± 0.00^{aA} Validation: 99.02 ± 1.26^{aA}	Training: 98.26 ± 0.91^{aA} Validation: 97.56 ± 2.30^{aA}	Training: 81.24 ± 2.31^{bA} Validation: 82.20 ± 4.75^{bA}
S_e	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00	Training: 0.84 ± 0.03 Validation: 0.83 ± 0.08
S_p	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.04	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.03	Training: 0.96 ± 0.02 Validation:0.95 ± 0.04	Training: 0.78 ± 0.02 Validation: 0.82 ± 0.09
F_s	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.02	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.01	Training: 0.98 ± 0.01 Validation:0.97 ± 0.02	Training:0.82 ± 0.03 Validation: 0.82 ± 0.05
P_r	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.05	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.03	Training: 0.97 ± 0.02 Validation: 0.95 ± 0.04	Training: 0.80 ± 0.03 Validation:0.81 ± 0.10
R_e	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00	Training: 0.84 ± 0.03 Validation: 0.83 ± 0.08
AUC	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.02	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.01	Training: 0.98 ± 0.01 Validation: 0.98 ± 0.02	Training: 0.81 ± 0.02 Validation: 0.83 ± 0.05
MCC	Training: 1.00 ± 0.00 Validation: 0.96 ± 0.04	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.03	Training: 0.97 ± 0.02 Validation: 0.95 ± 0.04	Training: 0.62 ± 0.05 Validation: 0.65 ± 0.10
TP	Training: 60 ± 2 Validation: 20 ± 1	Training: 60 ± 3 Validation: 21 ± 3	Training: 62 ± 3 Validation: 19 ± 3	Training: 52 ± 4 Validation: 16 ± 2
TN	Training: 61 ± 2 Validation: 20 ± 1	Training: 61 ± 3 Validation: 20 ± 3	Training: 57 ± 3 Validation: 21 ± 3	Training: 47 ± 3 Validation:17 ± 2
FP	Training: 0 ± 0 Validation: 0 ± 0	Training: 0 ± 0 Validation: 0 ± 0	Training:0 ± 0 Validation: 0 ± 0	Training: 10 ± 2 Validation: 3 ± 2
FN	Training: 0 ± 0 Validation: 0 ± 1	Training: 0 ± 0 Validation:0 ± 1	Training: 2 ± 1 Validation: 1 ± 1	Training: 13 ± 2 Validation:4 ± 2
CT (s)	Training: 0.237 ± 0.009^{a}	Training: 0.013 ± 0.007^{b}	Training: 0.007 ± 0.008^{c}	Training:0.056 ± 0.011^{d}

A_{cc} (overall accuracy), S_e (sensibility), S_p (specificity), P_r (precision), R_e (recall), F_s (F-score), AUC (area under the Receiver Operating Characteristic curve), MCC (Matthews correlation coefficient), CT (computation time), FDA (frequency-domain approach), TFDA (time-frequency-domain approach), TFDABH (time-frequency-domain approach-block-scale hard), KF (kernel function), C (regularization parameter), NTs (number of trees), LS (Laplace Smoothing), and NLVs (number of latent variables). Lowercase letters indicate statistically significant differences ($p<0.05$) between models, while uppercase letters indicate statistically significant differences ($p<0.05$) between the goodness-of-fit performance of training and validation datasets.

Table 19. Multi-objective optimized Latent Variable-Machine Learning (LV-ML) and Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Goodness of fit	Optimized models			
	LV-QDA	LV-GLM	RF-VS-LV-SVM	RF-VS-LV-RF
	Configuration i) TDA ii) NVLs=3	Configuration i) TDA ii) NVLs=10	Configuration i) TFDABS ii) Type: C-svc, iii) KF: anovadot iv) C=100	Configuration i) FDA ii) NTs=50
A_{cc} (%)	Training: 81.98 ± 2.67^{bA} Validation: 80.49 ± 5.39^{bA}	Training: 83.72 ± 1.56^{bA} Validation: 83.41 ± 4.42^{bA}	Training: 100.00 ± 0.00^{aA} Validation: 98.05 ± 1.54^{aA}	Training: 100.00 ± 0.00^{aA} Validation: 99.27 ± 1.18^{aA}
S_e	Training: 0.84 ± 0.04 Validation: 0.80 ± 0.07	Training: 0.86 ± 0.02 Validation: 0.83 ± 0.07	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00
S_p	Training: 0.80 ± 0.03 Validation: 0.82 ± 0.09	Training: 0.81 ± 0.02 Validation: 0.84 ± 0.09	Training: 1.00 ± 0.00 Validation: 0.97 ± 0.03	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02
F_s	Training: 0.82 ± 0.03 Validation: 0.80 ± 0.06	Training: 0.84 ± 0.01 Validation: 0.83 ± 0.04	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.02	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.01
P_r	Training: 0.81 ± 0.03 Validation: 0.80 ± 0.10	Training: 0.82 ± 0.01 Validation: 0.84 ± 0.08	Training: 1.00 ± 0.00 Validation: 0.97 ± 0.03	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02
R_e	Training: 0.84 ± 0.04 Validation: 0.80 ± 0.07	Training: 0.86 ± 0.02 Validation: 0.83 ± 0.07	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02	Training: 1.00 ± 0.00 Validation: 1.00 ± 0.00
AUC	Training: 0.82 ± 0.03 Validation: 0.81 ± 0.06	Training: 0.84 ± 0.02 Validation: 0.83 ± 0.05	Training: 1.00 ± 0.00 Validation: 0.98 ± 0.01	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.01
MCC	Training: 0.64 ± 0.05 Validation: 0.61 ± 0.11	Training: 0.67 ± 0.03 Validation: 0.67 ± 0.09	Training: 1.00 ± 0.00 Validation: 0.96 ± 0.03	Training: 1.00 ± 0.00 Validation: 0.99 ± 0.02
TP	Training: 51 ± 4 Validation: 16 ± 2	Training: 53 ± 2 Validation: 16 ± 2	Training: 60 ± 3 Validation: 20 ± 3	Training: 60 ± 2 Validation: 21 ± 2
TN	Training: 48 ± 4 Validation: 17 ± 2	Training: 48 ± 3 Validation: 18 ± 3	Training: 61 ± 3 Validation: 20 ± 3	Training: 61 ± 2 Validation: 20 ± 2
FP	Training: 10 ± 2 Validation: 4 ± 2	Training: 8 ± 1 Validation: 3 ± 1	Training: 0 ± 0 Validation: 0 ± 0	Training: 0 ± 0 Validation: 0 ± 0
FN	Training: 12 ± 2 Validation: 4 ± 2	Training: 11 ± 1 Validation: 3 ± 2	Training: 0 ± 0 Validation: 1 ± 1	Training: 0 ± 0 Validation: 0 ± 0
CT (s)	Training: 0.054 ± 0.010^{d}	Training: 0.052 ± 0.010^{d}	Training: 0.086 ± 0.009^{e}	Training: 0.006 ± 0.008^{c}

A_{cc} (overall accuracy), S_e (sensibility), S_p (specificity), P_r (precision), R_e (recall), F_s (F-score), AUC (area under the Receiver Operating Characteristic curve), MCC (Matthews correlation coefficient), CT (computation time), TDA (time-domain approach), FDA (frequency-domain approach), TFDABS (time-frequency-domain approach-block-scale soft), KF (kernel function), C (regularization parameter), NTs (number of trees) and NLVs (number of latent variables). Lowercase letters indicate statistically significant differences ($p<0.05$) between models, while uppercase letters indicate statistically significant differences ($p<0.05$) between the goodness-of-fit performance of training and validation datasets.

Table 20. Multi-objective optimized Random Forest-Variable Selection-Latent Variable-Machine Learning (RF-VS-LV-ML) models for maximizing bone fragment detection in chicken breast samples. Results are expressed as mean ± standard error and presented separately for training (75%) and validation (25%) datasets.

Goodness of fit	Optimized models			
	RF-VS-LV-NB	RF-VS-LV-LDA	RF-VS-LV-QDA	RF-VS-LV-GLM
	Configuration i) TFDA ii) LS=1	Configuration i) TFDABS	Configuration i) TDA	Configuration i) TFDABS
A_{cc} (%)	Training: 97.93 ± 1.47^{aA} Validation: 97.07 ± 3.78^{aA}	Training: 90.00 ± 1.76^{cA} Validation: 78.78 ± 6.90^{bB}	Training: 99.75 ± 0.40^{aA} Validation: 74.63 ± 6.92^{cB}	Training: 97.36 ± 4.44^{aA} Validation: 79.51 ± 6.82^{bB}
S_e	Training: 1.00 ± 0.01 Validation: 1.00 ± 0.00	Training: 0.89 ± 0.03 Validation: 0.73 ± 0.17	Training: 0.99 ± 0.01 Validation: 0.98 ± 0.03	Training: 0.98 ± 0.04 Validation: 0.74 ± 0.16
S_p	Training: 0.96 ± 0.03 Validation: 0.95 ± 0.07	Training: 0.91 ± 0.04 Validation: 0.85 ± 0.07	Training: 1.00 ± 0.00 Validation: 0.49 ± 0.13	Training: 0.97 ± 0.05 Validation: 0.84 ± 0.09
F_s	Training: 0.98 ± 0.01 Validation: 0.97 ± 0.04	Training: 0.90 ± 0.02 Validation: 0.76 ± 0.12	Training: 1.00 ± 0.00 Validation: 0.80 ± 0.06	Training: 0.97 ± 0.04 Validation: 0.77 ± 0.12
P_r	Training: 0.96 ± 0.02 Validation: 0.94 ± 0.07	Training: 0.91 ± 0.02 Validation: 0.82 ± 0.08	Training: 1.00 ± 0.00 Validation: 0.68 ± 0.09	Training: 0.97 ± 0.05 Validation: 0.83 ± 0.07
R_e	Training: 1.00 ± 0.01 Validation: 1.00 ± 0.00	Training: 0.89 ± 0.03 Validation: 0.73 ± 0.17	Training: 0.99 ± 0.01 Validation: 0.98 ± 0.03	Training: 0.98 ± 0.04 Validation: 0.74 ± 0.16
AUC	Training: 0.98 ± 0.01 Validation: 0.97 ± 0.03	Training: 0.90 ± 0.02 Validation: 0.79 ± 0.07	Training: 1.00 ± 0.00 Validation: 0.74 ± 0.06	Training: 0.97 ± 0.04 Validation: 0.79 ± 0.07
MCC	Training: 0.96 ± 0.03 Validation: 0.94 ± 0.07	Training: 0.80 ± 0.04 Validation: 0.59 ± 0.13	Training: 1.00 ± 0.00 Validation: 0.55 ± 0.11	Training: 0.95 ± 0.09 Validation: 0.60 ± 0.13
TP	Training: 60 ± 4 Validation: 20 ± 4	Training: 54 ± 4 Validation: 15 ± 4	Training: 59 ± 3 Validation: 21 ± 3	Training: 59 ± 3 Validation: 15 ± 4
TN	Training: 58 ± 4 Validation: 19 ± 3	Training: 55 ± 5 Validation: 18 ± 2	Training: 61 ± 3 Validation: 10 ± 3	Training: 58 ± 5 Validation: 18 ± 3
FP	Training: 0 ± 0 Validation: 0 ± 0	Training: 6 ± 2 Validation: 6 ± 4	Training: 0 ± 0 Validation: 0 ± 1	Training: 1 ± 3 Validation: 5 ± 3
FN	Training: 2 ± 2 Validation: 1 ± 2	Training: 6 ± 2 Validation: 3 ± 2	Training: 0 ± 0 Validation: 10 ± 3	Training: 2 ± 3 Validation: 3 ± 2
CT (s)	Training: 0.005 ± 0.007^{c}	Training: 0.002 ± 0.006^{c}	Training: 0.006 ± 0.008^{c}	Training: 0.005 ± 0.007^{c}

A_{cc} (overall accuracy), S_e (sensibility), S_p (specificity), P_r (precision), R_e (recall), F_s (F-score), AUC (area under the Receiver Operating Characteristic curve), MCC (Matthews correlation coefficient), CT (computation time), TDA (time-domain approach), TFDA (time-frequency-domain approach), TFDABS (time-frequency-domain approach-block-scale soft) and LS (Laplace Smoothing). Lowercase letters indicate statistically significant differences ($p<0.05$) between models, while uppercase letters indicate statistically significant differences ($p<0.05$) between the goodness-of-fit performance of training and validation datasets.

The statistical PLSR results on LV-SVM model (Table 14), showed that a the PLSR fit with ONPLSR= 2 yields modest explanatory power (R^2= 28.1%, Q^2= 28.0%, $RMSE_{TR}$= 3.34 and $RMSE_{CV}$= 3.59), and the screened specification preserves essentially the same generalization (R^2= 27.8%, Q^2 = 27.7%, $RMSE_{TR}$= 3.35 and $RMSE_{CV}$ = 3.59), indicating that pruning low-influence regressors enhances parsimony without changing performance. VIP values revealed that the Type and KF mostly influence the statistical performance results: laplacedot (VIP 1.437 to 1.163), besseldot (2.428 to 1.969), rbfdot (1.127 to 0.912), anovadot (1.063 to 0.860), and polydot/vanilladot (0.95 to 0.77) retained important influence after model screening, whereas the SVM regularization parameter C was not statistically significant in the tested range (VIP= 0.01-0.02; non-significant post-screening). The SVM "type" (C-svc vs nu-svc) remains moderately informative (VIP= 0.53 in both stages). The NVLs exhibited a high VIP (1.542 to 1.250), confirming that latent dimensionality is a primary lever for bias-variance control in this classifier. Data domain approach (TDA, FDA, TFDA, TFDABH and TFDABS) carries mixed raw VIPs, TFDABS is relatively large in the raw model (1.174) but attenuates post-screening (0.953), which signals redundancy between domain features and kernel-latent interactions. RS was negligible (VIP= 0.08; non-significant), manifesting that the data partition of the experimental dataset did not affect the goodness of fit of model in both training and validation sets.

In the LV-RF model (Table 14), ONPLSR= 2 with high goodness of fit (R^2= 66.9%, Q^2= 66.8%, $RMSE_{TR}$= 1.66 and $RMSE_{CV}$ = 1.92) that remains stable after screening (R^2= 66.8% and Q^2= 66.7%). The results showed that the DTe mostly influence the goodness of fit metrics (VIP= 2.924, unchanged after screening) and the ensemble size NTs were consistently influential (VIP= 0.55–0.60 for NTs= 50, 500, 1000, 5000, 10000), while NLVs also ranks high (VIP= 1.635). Data-domain indicators show very low VIPs and drop to non-significance upon screening, which implies that the statistical performance in RF is driven by ensemble architecture rather than by the data approach used.

Regarding to eh LV-NB (Table 15), the results indicated that this model was largely controlled by latent dimensionality. With ONPLSR= 1, the raw model attains R^2= 63.5% and Q^2= 63.3% ($RMSE_{TR}$= 1.61 and $RMSE_{CV}$= 1.82) and the screened model maintains similar generalization (R^2 = 63.2% and Q^2= 63.1%; $RMSE_{CV}$ = 1.87). NLVs shows a very high VIP that collapses to unity after screening (2.990 to 1.000), while LS was not significant (VIP= 0 for both LS levels). Domain variables have small VIPs and become non-significant under screening, supporting the view that the NB decision surface benefits mainly from how the latent space was parameterized. For LV-LDA, ONPLSR= 1 with a notable improvement due to screening: R^2/Q^2 rise from 52.1 to 52.0% and 58.4 to 58.3%, and errors drop sharply ($RMSE_{TR}$= 1.61 to 1.01 and $RMSE_{CV}$ 1.72 to 1.24). Additionally, NLVs concentrates the explanatory power (VIP= 2.645 to 1.000), while data approaches were not statistically significant.

Similarly to LV-LDA, the LV-QDA and LV-GLM (Table 16), the LV-QDA, used an ONPLSR= 1 with R^2 = 55.4% and Q^2= 55.3% in both raw and screened models; $RMSE_{CV}$ increases slightly after screening (2.05 to 2.10), a tolerable bias increment exchanged for interpretability. NLVs drives the performance (VIP 2.644 to 1.000), whereas domain

variables remain weak. The GLM fit is similarly stable (ONPLSR= 1; R^2/Q^2= 53.6/53.5% unchanged by screening; $RMSE_{CV}$ = 2.01), with NLVs again showed the highest importance (VIP 2.645 to 1.000) and data approaches remained non-significance.

Regarding the RF-VS based models (Table 17) which quantify how this strategy of variable selection interacts with downstream classifiers. For RF-VS-LV-SVM, ONPLSR= 2 with R^2= 62.5% and Q^2= 61.2% ($RMSE_{TR}$= 1.79 and $RMSE_{CV}$ = 2.10) and a screened specification that preserves Q^2 (62.0% to 61.1%). The SVM type remains highly informative after screening (VIP 1.420 to 0.943), and the KF continues to be the principal determinant of variability: besseldot (2.542 to 1.688) and anovadot (1.642 to 1.090) lead, followed by polydot and vanilladot (1.27 to 0.844) and then laplacedot and rbfdot (0.93 and 0.78 raw; 0.62 and 0.52 screened). The C parameter was also negligible (VIP= 0.02; non-significant), and data approach become non-significant after screening. PLSR model fitting on the RF-VS-LV-RF results (Table 17) used ONPLSR= 1 with moderate fit and generalization (R^2= 53.7%; Q^2= 51-52%; $RMSE_{TR}$= 0.65 and $RMSE_{CV}$= 0.89 to 0.92) that is again stable after screening. DTe remains the dominating factor (VIP= 3.161), NTs holds substantial influence (VIP= 0.60-0.64), and domain indicators are weak and non-significant. Notably, no PLSR was retained for RF-VS-LV-NB, RF-VS-LV-LDA, RF-VS-LV-QDA and RF-VS-LV-GLM because the first predictive component was already not significant, which is consistent with the earlier observation that these models were mainly influenced by NLVs and simple structural choices that RF-VS has already filtered.

The multi-objective optimized LV-ML (Table 18) established the best configuration of SVM model to maximize the BF detection capability. The best performing optimized LV-SVM model was defined by employing TFDA data approach, Type: C-svc, KF: anovadot, C=100 (or any of the other two, C=500.5 or C=1000) and NLVs= 69. Using this configuration, the model reached A_{ccT}= 100.00 ± 0.00% and A_{ccV}= 98.05 ± 2.24%,, with AUC_T=1.00 and AUC_V=0.98 and MCC_T= 1.00 and MCC_V=0.96; CT was moderate (0.237 s). LV-RF model using TFDA, NTs= 50 and NLVs= 71, achieved A_{ccT}= 100.00 ± 0.00% and A_{ccV}= 99.02 ± 1.26% with AUC_T=1.00 and AUC_V=0.99 and MCC_T= 1.00 and MCC_V=0.98, but with a significantly lower CT (0.013 s), indicating a superior accuracy-efficiency balance within the LV-ML. LV-NB by employing TFDABH with LS= 1 and NLVs= 69 sustains high validation (97.56 ± 2.30%) and strong MCC (0.95) at minimal computational cost (0.007 s), positioning NB as an ultra-efficient alternative with only a small performance gap relative to SVM/RF. LV-LDA under FDA with NLVs= 3 provided stable but lower ceilings (A_{ccV}= 82.20% and MCC_V= 0.65) with moderate time (0.056 s).

The results of LV-ML and RF-VS-LV-ML (Table 19) comparison in reveals the decisive effect of RF-VS when coupled with high-capacity learners. RF-VS-LV-SVM with TFDABS, C-svc, anovadot and C = 100 achieved A_{ccT}=100.00 ± 0.00% and A_{ccV}= 98.05 ± 1.54%, with high S_e, S_p, P_r, R_e, AUC and MCC at a higher cost than RF (0.086 s). RF-VS-LV-RF with FDA and NTs= 50 delivered the strongest validated performance among all models (A_{ccV}= 99.27 ± 1.18% and S_{eV}, S_{pV} and MCC_V >0.99) with the lowest CT reported in the table (0.006 s). LV-QDA and LV-GLM models exhibited a low goodness of fit with A_{cc} <80% in both training and validation datasets.

The statistical results of RF-VS-LV-NB using TFDA and LS= 1 exhibited A_{ccV}= 97.07 ± 3.78%, AUC_V 0.97 and MCC_V 0.94 and an extremely low CT (0.005 s), making it a compelling choice when throughput dominates. RF-VS-LV-QDA trained on TDA exhibits marked overfitting, A_{ccT}= 99.75 ± 0.40% compared to A_{ccV}= 74.63 ± 6.92%, which indicates that quadratic boundaries inflate variance in this latent regime unless covariance shrinkage or stricter dimensionality control is imposed. RF-VS-LV-GLM under TFDABS shows intermediate behavior (A_{ccV}=79.51%) and RF-VS-LV-LDA under TFDABS remains the lowest among the hybrids in validation (A_{ccV}= 78.78%).

The global PLSR modeling integrating all optimized models (Fig. 46) highlights systematic relationships across all classifiers. Models with higher flexibility, such as LV-SVM and LV-RF, achieve the highest training accuracies and associated R^2 values, but also exhibit the largest gaps in Q^2 and validation metrics. By contrast, simpler models such as LV-NB, LV-LDA, and LV-GLM achieve lower peak performance but present more balanced training-validation behavior, as reflected in their MCC and AUC values. Hybrid approaches involving RF-based variable selection consistently show improved generalization relative to their non-screened counterparts, particularly for RF-VS-LV-SVM and RF-VS-LV-RF models, confirming the importance of dimensionality reduction in mitigating overfitting. The overall PLSR results emphasize that model selection cannot rely solely on training accuracy but must be informed by validation-oriented metrics such as Q^2, $RMSE_{CV}$, AUC_V and MCC_V as well as by VIP-based screening to identify the most influential hyperparameters and data approaches. Subsequently, the use of PLSR model in the multi-objective optimization process allowed to determine that the RF-VS-LV-RF was the most accurate with the minimal CT model in the detection of BF within chicken breast samples. Thus, this model was considered as the best performing model in all frameworks and strategies tested. Additionally, the ANOVA results for A_{cc} and CT (Section 2.7.2.3) revealed statistically significant differences ($p<0.05$) across all optimized models. The homogeneous groups defined by LSD intervals further indicated that the RF-VS-LV-RF model was the most effective for detecting BF in chicken breast samples, as it achieved significantly higher A_{cc} and lower CT ($p<0.05$).

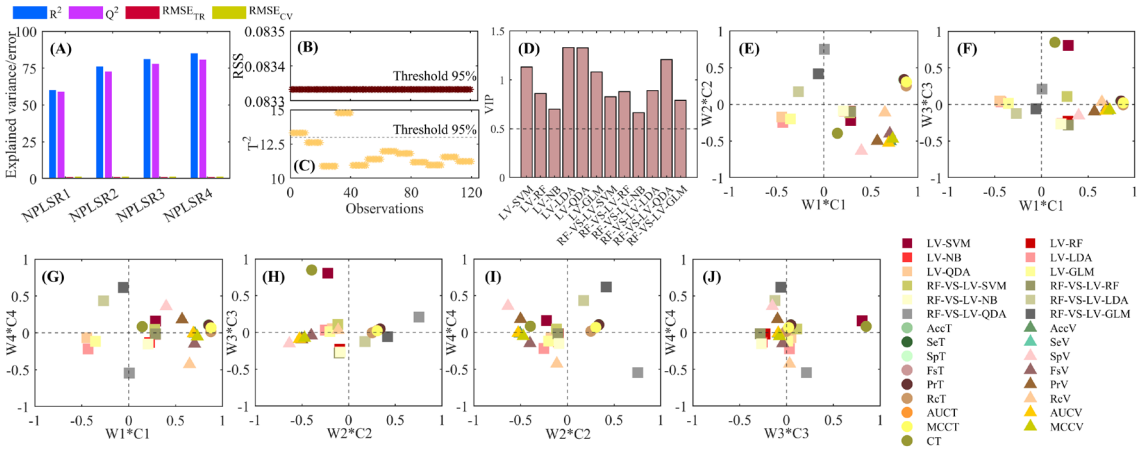


Fig. 46. Partial Least Square Regression (PLSR) modeling to assess the statistical classification performance results of the optimized Latent Variable-Support Vector Machines (LV-SVM), Latent Variable-Random Forest (LV-RF), Latent Variable-Naïve Bayes (LV-NB), Latent Variable-Linear Discriminant Analysis (LV-LDA), Latent Variable-Quadratic Discriminant Analysis (LV-QDA), Latent Variable-Generalized Linear Model (LV-GLM), Random Forest-Variable Selection-Latent Variable-Support Vector Machines (RF-VS-LV-SVM), Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF), Random Forest-Variable Selection-Latent Variable-Naïve Bayes (RF-VS-LV-NB), Random Forest-Variable Selection-Latent Variable- Linear Discriminant Analysis (RF-VS-LV-LDA), Random Forest-Variable Selection-Latent Variable- Quadratic Discriminant Analysis (RF-VS-LV-QDA) and Random Forest-Variable Selection-Latent Variable- Generalized Linear Model (RF-VS-LV-GLM) models. NPLSR (computed number of PLSR components), R^2 (coefficient of determination for training dataset), Q^2 (coefficient of determination for K-Fold cross validation dataset), $RMSE_{TR}$ (root mean square error for training dataset), $RMSE_{CV}$ (root mean square error for K-Fold cross validation dataset), RSS (residual sum squares), T^2 (Hotelling's T-squared), VIP (variable Importance for the projection), A_{ccT} (overall accuracy for training dataset), A_{ccV} (overall accuracy for validation dataset), S_{eT} (sensibility for training dataset), S_{eV} (sensibility for validation dataset), S_{pT} (specificity for training dataset), S_{pV} (specificity for validation dataset), P_{rT} (precision for training dataset), P_{rV} (precision for validation dataset), R_{eT} (recall for training dataset), R_{eV} (recall for validation dataset), F_{sT} (F-score for training dataset), F_{sV} (F-score for validation dataset), AUC_T (area under the Receiver Operating Characteristic curve for training dataset), AUC_T (area under the Receiver Operating Characteristic curve for validation dataset), MCC_T (Matthews correlation coefficient for training dataset) and MCC_V (Matthews correlation coefficient for validation dataset).

The optimization of LVs through RF-VS-LV-RF revealed critical insights into feature selection and model calibration. Using the MDA criterion, the RF model identified a subset of LVs with maximal contribution to classification accuracy (Fig. 47A). This eigenspace was subsequently employed within the RF-VS-LV-RF framework, resulting in the best-performing optimized model when using frequency-domain features (RF-VS-LV-RF-FDA). The MDA-based ranking highlights that spectral descriptors extracted from the ultrasound frequency distribution provide the strongest differentiation power between control samples and those containing BF. Among the most relevant variables, energy-magnitude parameters such as M_0, spectral shape VAR_{sp}, SKE_{sp}), KUR_{sp}, ENT_{sp}, as well as frequency-domain peaks such as F_r and M_P, consistently appear in the top 30 FDA features (Figs. 47B to 47H). Their elevated importance scores confirm that perturbations in spectral energy distribution, caused by acoustic scattering from BF, constitute a robust biomarker for automated classification procedure.

The eigenspace defined by the nine most important LVs (as determined by MDA) was further analyzed through PC projections (Fig. 48). The three-dimensional score plots reveal a clear clustering structure, with separation between control tissues and those containing BF of different sizes. Larger defects (2.0 × 1.5 cm, 2.0 × 1.0 cm) form distinct clusters distant from control samples, while smaller fragments (1.0 × 0.3 cm and 0.5 × 0.3 cm) remain closer but still separable in the reduced eigenspace. The two-dimensional projection of PC70 vs. PC1 further illustrates the differentiation capacity of the selected eigenspace, with minimal overlap across classes. This visualization confirms that the reduced LV space retains the essential information required for robust classification, validating the RF-VS feature selection strategy.

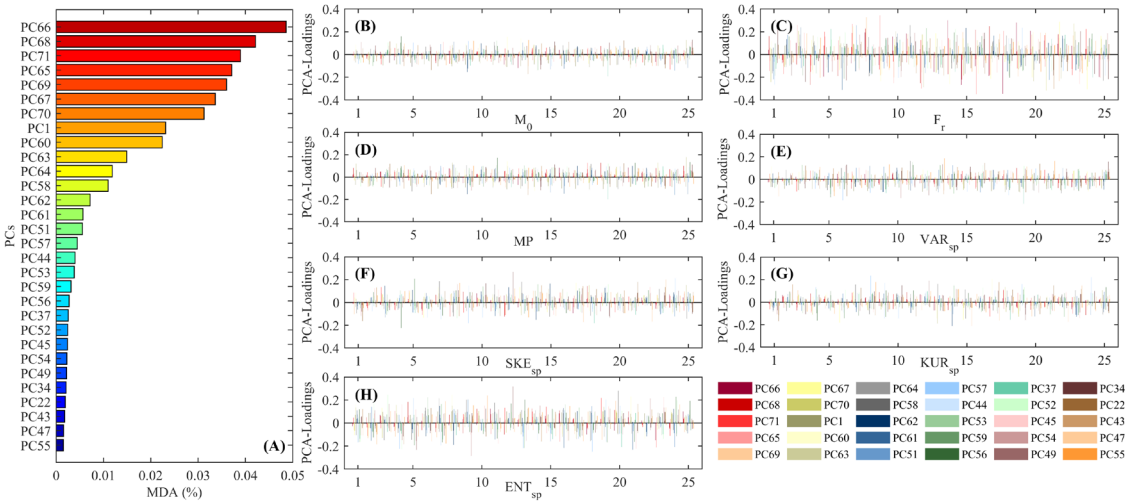


Fig. 47. Optimal Latent Variables (LVs) selected using the Mean Decrease Accuracy (MDA, %) criterion from the Random Forest (RF) model. This eigenspace was employed for calibration and validation of the Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) framework. Using frequency-domain (FDA) features to feed the RF-VS-LV-RF yielded the best-performing optimized model (RF-VS-LV-RF-FDA). Variable importance of each LV in maximizing sample classification (with vs. without bone fragments) according to RF accuracy (A). Loading plots of the 30 most important FDA energy-magnitude-distribution ultrasound parameters ranked by MDA (B to H). Parameters include M_0 (zero-order moment), F_r (center frequency of the phase spectrum), MP (maximum peak of the frequency spectrum), VAR_{sp} (spectral variance of the phase spectrum), SKE_{sp} (spectral skewness of the phase spectrum), KUR_{sp} (spectral kurtosis of the phase spectrum), and ENT_{sp} (spectral entropy of the phase spectrum).

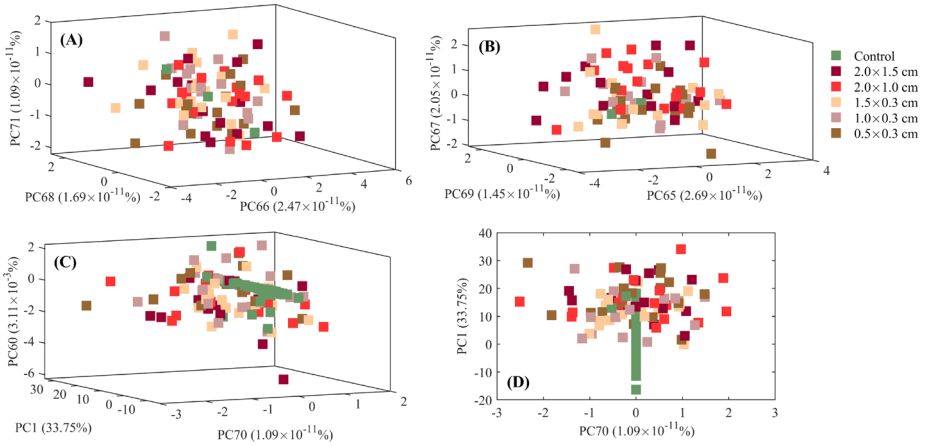


Fig. 48. Three-dimensional score plots of the nine most important Latent Variables (LVs; Fig. 47) selected by the Mean Decrease Accuracy (MDA) criterion from the Random Forest-Variable selection (RF-VS) strategy. Panels (A to C) show representative combinations of principal components (PCs) derived from the frequency-domain (FDA) features, while panel (D) presents the two-dimensional score projection of PC70 vs PC1 for clustering comparison. Sample groups correspond to Control and different bone fragment defect sizes (2.0×1.5 cm, 2.0×1.0 cm, 1.5×0.3 cm, 1.0×0.3 cm, and 0.5 × 0.3 cm). This selected eigenspace was used to feed the RF-VS-LV-RF yielded the best-performing optimized model (RF-VS-LV-RF-FDA).

A deeper validation of the optimized RF-VS-LV-RF-FDA model (NTs= 50) is provided (Table 21), where classification performance is evaluated against real BF sizes in chicken breast samples. The model achieved perfect detection across all fragment categories, from large fragments (2.0×1.5 cm) to the smallest (0.5×0.3 cm), without a single misclassification across bone-containing samples (total OC = 81/81 correctly predicted). For control samples, the classification was correct in nearly all cases (80/81), with only a single instance of misclassification across 100 RS of the experimental dataset. This outcome highlights both the robustness and the extremely high sensitivity of the RF-VS-LV-RF-FDA framework, capable of identifying even the smallest embedded bone fragments with near-perfect accuracy and representing an improvement in the previous detection results based on MIA-MSPC (Table 13).

The statistical results defined the RF-VS-LV-RF-FDA as the most effective strategy for BF detection in chicken breast samples. The integration of FDA features with RF-driven variable selection and latent variable compression achieves an optimal balance between accuracy, generalization, and computational efficiency. Compared with conventional LV-based classifiers, the hybrid framework demonstrates a decisive performance advantage, achieving statistical metrics that approach the theoretical upper limit of classification reliability. This confirms that the combination of eigenspace optimization, FDA ultrasound feature analysis, and ensemble ML strategies constitutes a powerful tool for non-invasive food safety monitoring using ultrasound spectroscopy.

Table 21. Classification performance of the optimized Random Forest-Variable Selection-Latent Variable-Random Forest (RF-VS-LV-RF) model using frequency-domain (FDA) features, yielding the best-performing configuration (RF-VS-LV-RF-FDA) for detecting bone fragments (BF) of varying sizes in chicken breast samples.

RF-VS-LV-RF-FDA-NTs=50			
	Type of BF	Number of samples (real)	Number of samples (predicted)
OC (all types)	2.0×1.5 cm	15	15
	2.0×1.0 cm	17	17
	1.5×0.3 cm	17	17
	1.0×0.3 cm	15	15
	0.5×0.3 cm	17	17
Total OC		**81**	**81**
Control		**81**	**80***

NTs (number of random trees), OC (Out-of-Control). *Classification performance of all control samples using the RF-VS-LV-RF-FDA model with NTs = 50, failed in at least one sample across the 100 random partitions of the experimental dataset.

4. Conclusions

This work provides a comprehensive demonstration of the potential of ultrasound imaging, combined with advanced multivariate image analysis and machine learning modeling-strategies, to achieve robust, accurate, and computationally efficient detection of bone fragments in chicken breast based products.

Ultrasound contact imaging has proven to be an effective and valuable technology for detecting bone fragments of varying sizes, regardless of their location within the chicken breast. The difference in acoustic impedance between the chicken breast and the bone fragments, mainly due to the presence of air gaps within bone pieces, significantly contributed to increase energy attenuation and changed ultrasound velocity. The detection of bone fragments was influenced by their size, although good classification results were found for all the different sizes considered.

The energy-magnitude and energy-distribution ultrasound parameters, computed in the time-frequency domains, effectively detected the bone fragments within chicken breast fillets. Both temporal and frequency-based approaches quantified similar information regarding ultrasound signal attenuation and alterations in the wave distribution caused by the presence of varying sizes of bone fragments.

The Residual Sum Squares multivariate control statistic has proven to be the most robust unsupervised model for detecting bone fragments (overall accuracy >95%) within chicken breasts, irrespective of the ultrasound parameters (time-frequency) used during the model's tuning. This approach has emerged as a valuable tool for integration into a monitoring system, facilitating the classification of contact ultrasound images of control chicken breasts and those containing bone fragments.

The integration of Multivariate Image Analysis with Latent Variable-based Machine Learning models provides a reliable strategy to enhance the detection of bone fragments in chicken breast samples. Compared with purely unsupervised approaches, the proposed framework significantly improved overall classification accuracy and robustness across fragment sizes (improvement of 3%>). The incorporation of Random Forest-based variable selection further reduced computational complexity while increasing model performance, confirming the value of combining feature-space dimensionality reduction with Machine Learning classification.

The integration of Random Forest-based Variable Selection with Latent Variable-based Machine Learning models significantly enhanced the detection of bone fragments while reducing model complexity. By combining feature-space dimensionality reduction, eigenspace optimization, and Machine Learning classification, this strategy achieved higher statistical performance in both training and validation datasets. Further, Mean Decrease Accuracy-based variable importance analysis improved detection by ranking latent variables derived from the unsupervised approach according to their contribution to Random Forest

accuracy. The pruning strategy of the eigenspace retained only the most relevant latent variables, thereby maximizing the classification performance for contaminated and uncontaminated chicken breast samples.

Partial Least Squares Regression model proved to be a reliable tool for advancing the multi-objective optimization of Machine Learning models. It enabled the exploration of underlying relationships between model hyperparameters, the number of latent variables obtained from the unsupervised approach, and data approaches (time, frequency, and time-frequency domains). Variable screening using the Variable Importance for the Projection criterion computed from the Partial Least Squares Regression model reduced noise and improved robustness, while also identifying which hyperparameters and feature types most strongly influenced classification performance. This information was then used to optimize models in a multi-objective manner, maximizing performance across both training and validation datasets. Based on this approach, the optimized Random Forest-Variable Selection-Latent Variable-Random Forest model achieved highly accurate detection of bone fragments of different sizes (overall accuracy: training=100%, validation=99.27%), enabling real-time quality monitoring of bone presence in chicken breast products. The model also combined high accuracy with very short training times (0.006 s), clearly outperforming computationally heavier alternatives tested.

The multi-objective optimized Random Forest-Variable Selection-Latent Variable-Random Forest model revealed that the frequency-domain approach provides the most robust fingerprints for bone fragment detection. Particularly, the energy-magnitude and energy-distribution parameters such as the zero-order moment, maximum peak of phase spectra, phase spectrum center frequencies, spectral variance, spectral skewness, spectral kurtosis, and spectral entropy proved to be the most relevant ultrasound parameters for maximizing the differentiation between control and out-of-control samples.

This research demonstrates that ultrasound imaging, when combined with advanced multivariate image analysis, machine learning, and optimization strategies, provides an effective approach to food safety monitoring. The proposed multi-objective framework ensured that the resulting model was not only accurate, but also generalizable, interpretable, and computationally efficient. These findings represent an important step toward the development of intelligent, non-invasive monitoring systems for the poultry industry.

Future work should be conducted in order to assess the detection limits of contact ultrasound technology for detecting small-sized bone fragments and to detect foreign bodies of different nature, such as plastics, glass, and metal pieces, which could also contaminate the chicken breast during the manufacturing process. This will be essential to develop a robust industrial prototype that can be used for real-time quality monitoring of the entire poultry meat production.

5. Declaration of competing interest

None.

6. Formal acknowledgments

The authors express their gratitude for the funding provided by the ULTRADIGITAL project (AGROALNEXT/2022/045) as part of the AGROALNEXT program. This program is supported by the MCIN, with funding from the European Union NextGenerationEU (PRTR-C17.I1), as well as the Generalitat Valenciana. Additionally, Gentil A. Collazos-Escobar's doctoral scholarship (PRE2020-092255) is acknowledged, which was granted through the State Training Subprogram of the State Plan for Scientific and Technical Research and Innovation 2017-2020, in conjunction with the European Social Fund.

I would like to express my deepest gratitude to my M.Sc. advisor, Dr. José Manuel Prats Montalbán, for his unwavering support and guidance throughout the entire process of this master's degree. He has always been a constant voice of encouragement and motivation to move forward, always willing to explain statistical models and multivariate image analysis with patience and clarity. His generosity, kindness, and tireless dedication, together with his invaluable support and mentorship, have been fundamental pillars during these years. I feel truly fortunate to have had him as a mentor and guide in this journey. It has been a true honor to work and collaborate with him, and I sincerely hope to continue our collaboration in the future.

I would like to thank Dr. José Vicente García Pérez for his support and for his dedication in securing and managing the funding that made this work possible.

I would like to thank Dr. José Javier Benedito Fort for his guidance, dedication, and the many hours he devoted to teaching me the analysis of ultrasonic signals and spectra. I am deeply grateful for the knowledge he shared with me and for transmitting his critical vision and insightful perspective on ultrasonic analysis, which have greatly enriched my work.

I would like to deeply thank M.Sc. Elisa Lincetti for her hard work and dedication in the experimental acquisition of ultrasound images, the development of experimental procedures, and the formulation and analysis of models. Her commitment, precision, and valuable contributions have been fundamental to the progress of this work.

I would like to sincerely thank M.Sc. Maider Aguerralde Martin for her trustworthy support in the formulation of PLS models for the multi-objective optimization process of machine learning algorithms. Her kind voice of encouragement, generosity, and willingness to share her knowledge have been invaluable to me, and her guidance has greatly enriched my work.

I would like to thank M.Sc. Miguel Ángel Vecina for his valuable guidance and support in the implementation and understanding of Convolutional Neural Networks. His technical

expertise, insightful advice, and willingness to help have been of great importance in advancing this work.

I would like to thank Dr. Nelson Sinisterra for his valuable contribution to the analysis and modeling of different statistical approaches and techniques applied to various parts of this work. His expertise and guidance have been essential in strengthening the methodological foundation of this research.

Finally, I extend my deepest gratitude to Dr. Eva Vallada Regalado for her unwavering support and guidance throughout all the administrative procedures of the master's program. Her kindness, efficiency, and invaluable assistance in managing enrollments, withdrawals, extensions, and every administrative matter over these years have made this journey considerably smoother.

7. Informal acknowledgments

To my parents, I dedicate this stage that now comes to an end. Their unconditional love, sacrifice, and guidance have been the foundation of my growth, and without their constant encouragement and support, reaching this milestone would not have been possible.

To my dear fiancée, Martha Patricia Medina Casas, for her unwavering support, patience, and love. Her understanding, encouragement, and constant presence have been a source of strength and inspiration throughout this journey.

I would like to express my heartfelt gratitude to my dear friends Walbersy Navarro, Patricia Mas Monsonís, Miguel Ángel Vecina García, Maider Aguerralde Martin and Nelson Sinisterra for their unwavering friendship, constant support, and for being a guiding and encouraging voice throughout this journey.

8. References

Achata, E.M., Esquerre, C., Gowen, A.A., O'Donnell, C.P., 2018. Feasibility of near infrared and Raman hyperspectral imaging combined with multivariate analysis to assess binary mixtures of food powders. Powder Technol 336, 555–566. https://doi.org/10.1016/j.powtec.2018.06.025

Algamal, Z.Y., 2018. A new method for choosing the biasing parameter in ridge estimator for generalized linear model. Chemometrics and Intelligent Laboratory Systems 183, 96–101. https://doi.org/10.1016/j.chemolab.2018.10.014

Ali, M.M., Hashim, N., 2021. Non-destructive methods for detection of food quality, in: Future Foods: Global Trends, Opportunities, and Sustainability Challenges. Elsevier, pp. 645–667. https://doi.org/10.1016/B978-0-323-91001-9.00003-7

An, J., Zhang, Z., 2025. Optimizing soft sensor costs through feature selection: A comparative study of sensory and chemical parameters in wine grade prediction. Chemometrics and Intelligent Laboratory Systems 262. https://doi.org/10.1016/j.chemolab.2025.105404

Babamoradi, H., Van den Berg, F., Rinnan, Å., 2013. Comparison of bootstrap and asymptotic confidence limits for control charts in batch MSPC strategies. Chemometrics and Intelligent Laboratory Systems 127, 102–111. https://doi.org/10.1016/j.chemolab.2013.06.005

Barrera Jiménez, A. A., Matsunami, K., Van Hauwermeiren, D., Peeters, M., Stauffer, F., dos Santos Schultz, E., Kumar, A., De Beer, T., Nopens, I., 2023. Partial least squares regression to calculate population balance model parameters from material properties in continuous twin-screw wet granulation. Int J Pharm 123040. https://doi.org/10.1016/j.ijpharm.2023.123040

Barrios-Rodriguez, Y., Cordoba-Salazar, G. A., Bahamón-Monje, A. F., Gutiérrez-Guzmán, N., 2021. Effect of roast degree, preparation method, and variety in the sensory and chemical characteristics of coffee (coffea arabica): A mid-infrared spectrum analysis. Coffee Sci 16, 1–12. https://doi.org/10.25186/.v16i.1964

Belaud, J.P., Prioux, N., Vialle, C., Sablayrolles, C., 2019. Big data for agri-food 4.0: Application to sustainability management for by-products supply chain. Comput Ind 111, 41–50. https://doi.org/10.1016/j.compind.2019.06.006

Bowler, A., Ozturk, S., di Bari, V., Glover, Z.J., Watson, N.J., 2023. Machine learning and domain adaptation to monitor yoghurt fermentation using ultrasonic measurements. Food Control 147, 109622. https://doi.org/10.1016/j.foodcont.2023.109622

Bowler, A.L., Bakalis, S., Watson, N.J., 2020. Monitoring mixing processes using ultrasonic sensors and machine learning. Sensors (Switzerland) 20. https://doi.org/10.3390/s20071813

Brendel, R., Schwolow, S., Rohn, S., Weller, P., 2020. Comparison of PLSR, MCR-ALS and Kernel-PLSR for the quantification of allergenic fragrance compounds in

complex cosmetic products based on nonlinear 2D GC-IMS data. Chemometrics and Intelligent Laboratory Systems 205. https://doi.org/10.1016/j.chemolab.2020.104128

Cabana, E., Lillo, R.E., 2022. Robust adjusted discriminant analysis based on shrinkage with application to geochemical and environmental fields. Chemometrics and Intelligent Laboratory Systems 221. https://doi.org/10.1016/j.chemolab.2021.104488

Caesarendra, W., Tjahjowidodo, T., 2017. A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing. Machines. https://doi.org/10.3390/machines5040021

Caladcad, J.A., Cabahug, S., Catamco, M.R., Villaceran, P.E., Cosgafa, L., Cabizares, K.N., Hermosilla, M., Piedad, E.J., 2020. Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal. Comput Electron Agric 172. https://doi.org/10.1016/j.compag.2020.105327

Cavaglia, J., Schorn-García, D., Giussani, B., Ferré, J., Busto, O., Aceña, L., Mestres, M., Boqué, R., 2020. Monitoring wine fermentation deviations using an ATR-MIR spectrometer and MSPC charts. Chemometrics and Intelligent Laboratory Systems 201. https://doi.org/10.1016/j.chemolab.2020.104011

Chauchard, F., Cogdill, R., Roussel, S., Roger, J.M., Bellon-Maurel, V., 2004. Application of LS-SVM to non-linear phenomena in NIR spectroscopy: Development of a robust and portable sensor for acidity prediction in grapes. Chemometrics and Intelligent Laboratory Systems 71, 141–150. https://doi.org/10.1016/j.chemolab.2004.01.003

Chen, H., Liu, X., Jia, Z., Liu, Z., Shi, K., Cai, K., 2018. A combination strategy of random forest and back propagation network for variable selection in spectral calibration. Chemometrics and Intelligent Laboratory Systems 182, 101–108. https://doi.org/10.1016/j.chemolab.2018.09.002

Chen, Y., Dan, C., He, Y., Zheng, X., 2025. Generalized continuum regression (GCR): An advanced multivariate method for precise dimensionality reduction and efficient regression modeling. Chemometrics and Intelligent Laboratory Systems 262. https://doi.org/10.1016/j.chemolab.2025.105407

Cho, B.-K., Irudayaraj, J.M.K., 2003. Foreign Object and Internal Disorder Detection in Food Materials Using Noncontact Ultrasound Imaging, JOURNAL OF FOOD SCIENCE.

Collazos-Escobar, G.A., Barrios-Rodriguez, Y.F., Bahamón-Monje, A.F., Gutiérrez-Guzmán, N., 2023a. Uses of mid-infrared spectroscopy and chemometric models for differentiating between dried cocoa bean varieties. Revista Brasileira de Engenharia Agricola e Ambiental 27, 803–810. https://doi.org/10.1590/1807-1929/agriambi.v27n10p803-810

Collazos-Escobar, G.A., Barrios-Rodríguez, Y.F., Bahamón-Monje, A.F., Gutiérrez-Guzmán, N., 2024. Mid-infrared spectroscopy and machine learning as a complementary tool for sensory quality assessment of roasted cocoa-based products. Infrared Phys Technol 141. https://doi.org/10.1016/j.infrared.2024.105482

Collazos-Escobar, G.A., Gutiérrez-Guzmán, N., Váquiro, H.A., García-Pérez, J. V., Cárcel, J.A., 2025a. Analysis of Machine Learning Algorithms for the Computer Simulation of Moisture Sorption Isotherms of Coffee Beans. Food Bioproc Tech. https://doi.org/10.1007/s11947-025-03785-x

Collazos-Escobar, G.A., Gutiérrez-Guzmán, N., Váquiro-Herrera, H.A., Bon, J., Cárcel, J.A., García-Pérez, J. V., 2023b. Model-based investigation of water adsorption in Achira (Canna edulis K.) biscuits. LWT 189. https://doi.org/10.1016/j.lwt.2023.115472

Collazos-Escobar, G.A., Hurtado-Cortés, V., Bahamón-Monje, A.F., Gutiérrez-Guzmán, N., 2025b. Mathematical modeling of water sorption isotherms in specialty coffee beans processed by wet and semidry postharvest methods. Sci Rep 15, 3898. https://doi.org/10.1038/s41598-024-83702-y

Collazos-Escobar, G.A., Lincetti, E., Spilimbergo, S., Prats-Montalbán, J.M., García-Pérez, J. V., Benedito, J., 2025c. Integrated use of ultrasound imaging and multivariate image analysis for detecting bone fragments in poultry meat. Food Research International 206. https://doi.org/10.1016/j.foodres.2025.116047

Collazos-Escobar, G.A., Prats-Montalbán, J.M., Giacomozzi, A.S., Benedito, J., Gómez Álvarez-Arenas, T.E., García-Pérez, J. V., 2025d. Non-invasive detection of internal foreign bodies in foods by using air-coupled ultrasound: case studies in beef burger patties and jelly plates. J Food Eng 112777. https://doi.org/10.1016/j.jfoodeng.2025.112777

Colucci, D., Prats-Montalbán, J.M., Fissore, D., Ferrer, A., 2019. Application of multivariate image analysis for on-line monitoring of a freeze-drying process for pharmaceutical products in vials. Chemometrics and Intelligent Laboratory Systems 187, 19–27. https://doi.org/10.1016/j.chemolab.2019.02.004

Conde, T., Mulet, A., Clemente, G., Benedito, J., 2008. Detection of internal cracks in manchego cheese using the acoustic impulse-response technique and ultrasounds. J Dairy Sci 91, 918–927. https://doi.org/10.3168/jds.2007-0661

Corona, E., Garcia-Perez, J. V., Gomez Alvarez-Arenas, T.E., Watson, N., Povey, M.J.W., Benedito, J., 2013. Advances in the ultrasound characterization of dry-cured meat products. J Food Eng 119, 464–470. https://doi.org/10.1016/j.jfoodeng.2013.06.023

Correia, L.R., Mittal, G.S., Basir, O.A., 2008. Ultrasonic detection of bone fragment in mechanically deboned chicken breasts. Innovative Food Science and Emerging Technologies 9, 109–115. https://doi.org/10.1016/j.ifset.2007.06.004

Costa, N., Lourenço, J., 2023. Assessing the resilience of optimal solutions in multiobjective problems. Chemometrics and Intelligent Laboratory Systems 239. https://doi.org/10.1016/j.chemolab.2023.104850

Craig, A.P., Botelho, B.G., Oliveira, L.S., Franca, A.S., 2018. Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. Food Chem 245, 1052–1061. https://doi.org/10.1016/j.foodchem.2017.11.066

Cristina Duarte Marques, R., Resende Oliveira, É., Silva Mendes Coutinho, G., Emannuele Chaves Ribeiro, A., Souza Teixeira, C., Soares Soares Júnior, M., Caliari, M., 2020. Modeling sorption properties of maize by-products obtained using the Dynamic Dewpoint Isotherm (DDI) method. Food Biosci 38. https://doi.org/10.1016/j.fbio.2020.100738

de Andrade, B.M., de Gois, J.S., Xavier, V.L., Luna, A.S., 2020. Comparison of the performance of multiclass classifiers in chemical data: Addressing the problem of overfitting with the permutation test. Chemometrics and Intelligent Laboratory Systems 201. https://doi.org/10.1016/j.chemolab.2020.104013

Debón, A., Carlos Garcia-Díaz, J., 2012. Fault diagnosis and comparing risk for the steel coil manufacturing process using statistical models for binary data. Reliab Eng Syst Saf 100, 102–114. https://doi.org/10.1016/j.ress.2011.12.022

Dixon, S.J., Brereton, R.G., 2009. Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure. Chemometrics and Intelligent Laboratory Systems 95, 1–17. https://doi.org/10.1016/j.chemolab.2008.07.010

Djekic, I., Jankovic, D., Rajkovic, A., 2017. Analysis of foreign bodies present in European food using data from Rapid Alert System for Food and Feed (RASFF). Food Control 79, 143–149. https://doi.org/10.1016/j.foodcont.2017.03.047

Duchesne, C., Liu, J.J., MacGregor, J.F., 2012. Multivariate image analysis in the process industries: A review. Chemometrics and Intelligent Laboratory Systems. https://doi.org/10.1016/j.chemolab.2012.04.003

Duma, Z.S., Susiluoto, J., Lamminpää, O., Sihvonen, T., Reinikainen, S.P., Haario, H., 2024. KF-PLS: Optimizing Kernel Partial Least-Squares (K-PLS) with Kernel Flows. Chemometrics and Intelligent Laboratory Systems 254. https://doi.org/10.1016/j.chemolab.2024.105238

Edward, M. C., Stringer, M. F. 2007. Observations on patterns in foreign material investigations. Food Control, 18, 773-782. https://doi.org/10.1016/j.foodcont.2006.01.007

El Zein, A.K., Cobre, A. de F., Luna Lazo, R.E., Antunes, K.A., Manfron, J., Ferreira, L.M., Pontarolo, R., 2025. Identification of Monteverdia ilicifolia by fourier-transform mid-infrared spectroscopy associated with chemometrics and machine

learning. Chemometrics and Intelligent Laboratory Systems 263. https://doi.org/10.1016/j.chemolab.2025.105420

Eriksson, L., Johansson, E., Antti, H., & Holmes, E. 2016. Multi-and Megavariate Data Analysis: Finding and Using Regularities in Metabonomics Data. Metabolomics in Toxicity Assesment. Chapter 8. DOI: https://doi.org/10.1201/b14117

Eskelinen, J.J., Alavuotunki, A.P., Hæggström, E., Alatossava, T., 2007. Preliminary study of ultrasonic structural quality control of swiss-type cheese. J Dairy Sci 90, 4071–4077. https://doi.org/10.3168/jds.2007-0105

Fang, X., Ye, H., Zhang, S., Guo, L., Xu, Y., Zhang, D., Nie, Q., 2023. Investigation of potential genetic factors for growth traits in yellow-feather broilers using weighted single-step genome-wide association study. Poult Sci 102. https://doi.org/10.1016/j.psj.2023.103034

Fariñas, L., Contreras, M., Sanchez-Jimenez, V., Benedito, J., Garcia-Perez, J. V., 2021a. Use of air-coupled ultrasound for the non-invasive characterization of the textural properties of pork burger patties. J Food Eng 297. https://doi.org/10.1016/j.jfoodeng.2021.110481

Fariñas, L., Sanchez-Torres, E.A., Sanchez-Jimenez, V., Diaz, R., Benedito, J., Garcia-Perez, J. V., 2021b. Assessment of avocado textural changes during ripening by using contactless air-coupled ultrasound. J Food Eng 289. https://doi.org/10.1016/j.jfoodeng.2020.110266

Fariñas, M.D., Sanchez-Jimenez, V., Benedito, J., Garcia-Perez, J. V., 2023. Monitoring physicochemical modifications in beef steaks during dry salting using contact and non-contact ultrasonic techniques. Meat Sci 204, 109275. https://doi.org/10.1016/j.meatsci.2023.109275

Fernández Pierna, J.A., Vincke, D., Baeten, V., Grelet, C., Dehareng, F., Dardenne, P., 2016. Use of a multivariate moving window PCA for the untargeted detection of contaminants in agro-food products, as exemplified by the detection of melamine levels in milk using vibrational spectroscopy. Chemometrics and Intelligent Laboratory Systems 152, 157–162. https://doi.org/10.1016/j.chemolab.2015.10.016

Fink, F., Stawski, T.M., Emmerling, F., Falkenhagen, J., 2025. A novel machine-learning approach to unlock technical lignin classification by NIR spectroscopy - bench to handheld. Chemometrics and Intelligent Laboratory Systems 264. https://doi.org/10.1016/j.chemolab.2025.105467

Fu, L., Xie, H.L., Xu, X.R., Yang, H.J., Nie, X. Du, 2014. Combining random forest with multi-amino acid features to identify protein palmitoylation sites. Chemometrics and Intelligent Laboratory Systems 135, 208–212. https://doi.org/10.1016/j.chemolab.2014.04.009

Galdón-Navarro, B., Prats-Montalbán, J.M., Cubero, S., Blasco, J., Ferrer, A., 2018. Comparison of latent variable-based and artificial intelligence methods for impurity detection in PET recycling from NIR hyperspectral images. J Chemom 32. https://doi.org/10.1002/cem.2980

Gan, W.S., 2020. Signal processing and image processing for acoustical imaging, Signal Processing and Image Processing for Acoustical Imaging. Springer Singapore. https://doi.org/10.1007/978-981-10-5550-8

Garcia-Perez, J. V., de Prados, M., Martinez, G., Gomez Alvarez-Arenas, T.E., Benedito, J., 2019. Ultrasonic online monitoring of the ham salting process. Methods for signal analysis: Time of flight calculation. J Food Eng 263, 87–95. https://doi.org/10.1016/j.jfoodeng.2019.05.032

Garrido-Novell, C., Garrido-Varo, A., Pérez-Marín, D., Guerrero, J.E., 2018. Using spectral and textural data extracted from hyperspectral near infrared spectroscopy imaging to discriminate between processed pork, poultry and fish proteins. Chemometrics and Intelligent Laboratory Systems 172, 90–99. https://doi.org/10.1016/j.chemolab.2017.11.011

Gholizadeh, M., Jamei, M., Ahmadianfar, I., Pourrajab, R., 2020. Prediction of nanofluids viscosity using random forest (RF) approach. Chemometrics and Intelligent Laboratory Systems 201. https://doi.org/10.1016/j.chemolab.2020.104010

Godoy, J.L., Vega, J.R., Marchetti, J.L., 2014. Relationships between PCA and PLS-regression. Chemometrics and Intelligent Laboratory Systems 130, 182–191. https://doi.org/10.1016/j.chemolab.2013.11.008

Guisset, S., Martin, M., Govaerts, B., 2019. Comparison of PARAFASCA, AComDim, and AMOPLS approaches in the multivariate GLM modelling of multi-factorial designs. Chemometrics and Intelligent Laboratory Systems 184, 44–63. https://doi.org/10.1016/j.chemolab.2018.11.006

He, T., Lai, W., Li, M., Feng, Y., Liu, Y., Yu, T., Tang, H., Zhang, T., Li, H., 2021. The detonation heat prediction of nitrogen-containing compounds based on quantitative structure-activity relationship (QSAR) combined with random forest (RF). Chemometrics and Intelligent Laboratory Systems 213. https://doi.org/10.1016/j.chemolab.2021.104249

Hu, M.H., Zhao, Y., Zhai, G.T., 2018. Active learning algorithm can establish classifier of blueberry damage with very small training dataset using hyperspectral transmittance data. Chemometrics and Intelligent Laboratory Systems 172, 52–57. https://doi.org/10.1016/j.chemolab.2017.11.012

Jiang, H., Yoon, S.C., Zhuang, H., Wang, W., Lawrence, K.C., Yang, Y., 2018. Tenderness classification of fresh broiler breast fillets using visible and near-infrared hyperspectral imaging. Meat Sci 139, 82–90. https://doi.org/10.1016/j.meatsci.2018.01.013

Jiménez-Carvelo, A.M., González-Casado, A., Bagur-González, M.G., Cuadros-Rodríguez, L., 2019. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity – A review. Food Research International. https://doi.org/10.1016/j.foodres.2019.03.063

Jin, C., Bouzembrak, Y., Zhou, J., Liang, Q., van den Bulk, L.M., Gavai, A., Liu, N., van den Heuvel, L.J., Hoenderdaal, W., Marvin, H.J.P., 2020. Big Data in food safety-A review. Curr Opin Food Sci. https://doi.org/10.1016/j.cofs.2020.11.006

Kahrıman, F., Liland, K.H., 2021. SelectWave: A graphical user interface for wavelength selection and spectral data analysis. Chemometrics and Intelligent Laboratory Systems 212. https://doi.org/10.1016/j.chemolab.2021.104275

Karatzoglou, A., Smola, A., & Hornik, K. 2024. kernlab-An S4 Package for Kernel Methods in R. Journal of Statistical Software, 11 (9), 1-20. DOI: https://doi.org/10.18637/jss.v011.i09.

Kruse, O.M.O., Prats-Montalbán, J.M., Indahl, U.G., Kvaal, K., Ferrer, A., Futsaether, C.M., 2014. Pixel classification methods for identifying and quantifying leaf surface injury from digital images. Comput Electron Agric 108, 155–165. https://doi.org/10.1016/j.compag.2014.07.010

Kuhn, M. 2008) Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1–26. DOI: https://doi.org/10.18637/jss.v028.i05

Kumar, M., Dahuja, A., Sachdev, A., Kaur, C., Varghese, E., Saha, S., Sairam, K.V.S.S., 2019. Valorisation of black carrot pomace: microwave assisted extraction of bioactive phytoceuticals and antioxidant activity using Box–Behnken design. J Food Sci Technol 56, 995–1007. https://doi.org/10.1007/s13197-018-03566-9

Lemaigre, S., Adam, G., Goux, X., Noo, A., De Vos, B., Gerin, P.A., Delfosse, P., 2016. Transfer of a static PCA-MSPC model from a steady-state anaerobic reactor to an independent anaerobic reactor exposed to organic overload. Chemometrics and Intelligent Laboratory Systems 159, 20–30. https://doi.org/10.1016/j.chemolab.2016.09.010

Li, Y., Bian, X., Sheng, J., Yang, S., 2025. Macroscopic properties and air pores of tailings concrete under dry-wet cycles of chloride attack based on principal component analysis (PCA). Constr Build Mater 489. https://doi.org/10.1016/j.conbuildmat.2025.142233

Liaw, A., & Wiener, M. 2002. Classification and Regression by randomForest. R News 2(3), 18-22. https://CRAN.R-project.org/doc/Rnews/

Liberda, D., Kosowska, K., Koziol, P., Wrobel, T.P., 2021. Spatial sampling effect on data structure and random forest classification of tissue types in High Definition and Standard Definition FT-IR imaging. Chemometrics and Intelligent Laboratory Systems 217. https://doi.org/10.1016/j.chemolab.2021.104407

Lim, H., Lee, J., Lee, S., Cho, H., Lee, H., Jeon, D., 2022. Low-density foreign body detection in food products using single-shot grid-based dark-field X-ray imaging. J Food Eng 335. https://doi.org/10.1016/j.jfoodeng.2022.111189

López, F., Prats, J.M., Ferrer, A., Valiente, J.M., 2006. Defect detection in random colour textures using the MIA T2 defect maps. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4142 LNCS, 752–763. https://doi.org/10.1007/11867661_68

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., Makowski, D. 2021. An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software, 6(60), 3139. DOI: https://doi.org/10.21105/joss.03139

Macgregor, J.F., Kourtl, T., 1995. STATISTICAL PROCESS CONTROL OF MULTIVARIATE PROCESSES, Control Fag. Practice.

Malaslı, M.Z., Akkoyunlu, M.C., Pekel, E., Taşova, M., Dursun, S.K., Akkoyunlu, M.T., 2025. Prediction of drying kinetics and energy consumption values of purple carrots dried in a temperature-controlled microwave dryer by decision tree, random forest and ada boost approaches. Chemometrics and Intelligent Laboratory Systems 260. https://doi.org/10.1016/j.chemolab.2025.105352

McFarlane, N.J.B., Speller, R.D., Bull, C.R., Tillett, R.D., 2003. Detection of bone fragments in chicken meat using X-ray backscatter. Biosyst Eng 85, 185–199. https://doi.org/10.1016/S1537-5110(03)00036-9

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. 2024. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien_. R package version 1.7-16, https://CRAN.R-project.org/package=e1071.

Mohd Khairi, M.T., Ibrahim, S., Md Yunus, M.A., Faramarzi, M., 2018. Noninvasive techniques for detection of foreign bodies in food: A review. J Food Process Eng. https://doi.org/10.1111/jfpe.12808

Mohd Khairi, M.T., Ibrahim, S., Md Yunus, M.A., Faramarzi, M., 2015. Contact and non-contact ultrasonic measurement in the food industry: A review. Meas Sci Technol 27. https://doi.org/10.1088/0957-0233/27/1/012001

Ni, D., Xiao, Z., Lim, M.K., 2020. A systematic review of the research trends of machine learning in supply chain management. International Journal of Machine Learning and Cybernetics 11, 1463–1482. https://doi.org/10.1007/s13042-019-01050-0

Nibouche, O., Asharindavida, F., Wang, H., Vincent, J., Liu, J., van Ruth, S., Maguire, P., Rahman, E., 2024. A new sub-class linear discriminant for miniature spectrometer based food analysis. Chemometrics and Intelligent Laboratory Systems 250. https://doi.org/10.1016/j.chemolab.2024.105136

Nielsen, M.S., Lauridsen, T., Christensen, L.B., Feidenhans'l, R., 2013. X-ray dark-field imaging for detection of foreign bodies in food. Food Control 30, 531–535. https://doi.org/10.1016/j.foodcont.2012.08.007

Otchere, D.A., 2023. Fundamental error in tree-based machine learning model selection for reservoir characterisation. Energy Geoscience. https://doi.org/10.1016/j.engeos.2023.100229

Ozturk, S., Bowler, A., Rady, A., Watson, N.J., 2023. Near-infrared spectroscopy and machine learning for classification of food powders during a continuous process. J Food Eng 341. https://doi.org/10.1016/j.jfoodeng.2022.111339

Paris, A., Duchesne, C., Poulin, É., 2024. Adjusting plant operating conditions to widen multivariate specification regions for incoming raw materials–An optimization framework. Chemometrics and Intelligent Laboratory Systems 244. https://doi.org/10.1016/j.chemolab.2023.104991

Pérez-Santaescolástica, C., Fraeye, I., Barba, F.J., Gómez, B., Tomasevic, I., Romero, A., Moreno, A., Toldrá, F., Lorenzo, J.M., 2019. Application of non-invasive technologies in dry-cured ham: An overview. Trends Food Sci Technol. https://doi.org/10.1016/j.tifs.2019.02.011

Prats-Montalbán, J.M., de Juan, A., Ferrer, A., 2011. Multivariate image analysis: A review with applications. Chemometrics and Intelligent Laboratory Systems. https://doi.org/10.1016/j.chemolab.2011.03.002

Prats-Montalbán, J.M., Ferrer, A., Bro, R., Hancewicz, T., 2009. Prediction of skin quality properties by different Multivariate Image Analysis methodologies. Chemometrics and Intelligent Laboratory Systems 96, 6–13. https://doi.org/10.1016/j.chemolab.2008.10.012

Prats-Montalbán, J.M., Jerez-Rozo, J.I., Romañach, R.J., Ferrer, A., 2012. MIA and NIR Chemical Imaging for pharmaceutical product characterization. Chemometrics and Intelligent Laboratory Systems 117, 240–249. https://doi.org/10.1016/j.chemolab.2012.04.002

Preys, S., Vigneau, E., Mazerolles, G., Cheynier, V., Bertrand, D., 2007. Multivariate prototype approach for authentication of food products. Chemometrics and Intelligent Laboratory Systems 87, 200–207. https://doi.org/10.1016/j.chemolab.2007.01.003

Ramtanon, I., Lacoue-Nègre, M., Berlioz-Barbier, A., Le Masle, A., Renault, J.H., 2025. A selective genetic algorithm - PLS-DA approach based on untargeted LC-HRMS: Application to complex biomass samples. Chemometrics and Intelligent Laboratory Systems 261. https://doi.org/10.1016/j.chemolab.2025.105381

Reis, M.S., 2015. An integrated multiscale and multivariate image analysis framework for process monitoring of colour random textures: MSMIA. Chemometrics and Intelligent Laboratory Systems 142, 36–48. https://doi.org/10.1016/j.chemolab.2015.01.008

Ruiz de Miras, J., Gacto, M.J., Blanc, M.R., Arroyo, G., López, L., Torres, J.C., Martín, D., 2024. Machine learning regression algorithms for generating chemical element maps from X-ray fluorescence data of paintings. Chemometrics and Intelligent Laboratory Systems 248. https://doi.org/10.1016/j.chemolab.2024.105116

Sánchez-Jiménez, V., Collazos-Escobar, G.A., González-Mohino, A., Alvarez-Arenas, T.E., Benedito, J., García-Pérez, J. V., 2023. Non-invasive monitoring of potato drying by means of air-coupled ultrasound. Food Control 109653. https://doi.org/10.1016/j.foodcont.2023.109653

Scatigno, C., Festa, G., 2022. FTIR coupled with machine learning to unveil spectroscopic benchmarks in the Italian EVOO. Int J Food Sci Technol 57, 4156–4162. https://doi.org/10.1111/ijfs.15735

Sinisterra-Solís, N., Sanjuán, N., Ribal, J., Estruch, V., Clemente, G., Rozakis, S., 2024. Developing a composite indicator to assess agricultural sustainability: Influence of some critical choices. Ecol Indic 161. https://doi.org/10.1016/j.ecolind.2024.111934

Siqueira, L.F.S., Araújo Júnior, R.F., de Araújo, A.A., Morais, C.L.M., Lima, K.M.G., 2017. LDA vs. QDA for FT-MIR prostate cancer tissue classification. Chemometrics and Intelligent Laboratory Systems 162, 123–129. https://doi.org/10.1016/j.chemolab.2017.01.021

Suen, Y., Xiao, S., Hao, S., Zhao, X., Xiong, Y., Liu, S., 2016. Time-frequency representation measurement based on temporal Fourier transformation. Opt Commun 376, 86–91. https://doi.org/10.1016/j.optcom.2016.05.017

Sun, H., Hu, X., 2017. Attribute selection for decision tree learning with class constraint. Chemometrics and Intelligent Laboratory Systems 163, 16–23. https://doi.org/10.1016/j.chemolab.2017.02.004

Thévenot, E.A., Roux, A., Xu, Y., Ezan, E., Junot, C., 2015. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. J Proteome Res 14, 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

van Herwerden, D., O'Brien, J.W., Choi, P.M., Thomas, K. V., Schoenmakers, P.J., Samanipour, S., 2022. Naive Bayes classification model for isotopologue detection in LC-HRMS data. Chemometrics and Intelligent Laboratory Systems 223. https://doi.org/10.1016/j.chemolab.2022.104515

Velásquez, S., Franco, A.P., Peña, N., Bohórquez, J.C., Gutierrez, N., 2021a. Effect of coffee cherry maturity on the performance of the drying process of the bean: Sorption isotherms and dielectric spectroscopy. Food Control 123. https://doi.org/10.1016/j.foodcont.2020.107692

Velásquez, S., Franco, A.P., Peña, N., Bohórquez, J.C., Gutiérrez, N., 2021b. Classification of the maturity stage of coffee cherries using comparative feature and machine learning. Coffee Sci 16, 1. https://doi.org/10.25186/.v16i.1710

Velásquez, S., Peña, N., Bohórquez, J.C., Gutierrez, N., Sacks, G.L., 2019. Volatile and sensory characterization of roast coffees – Effects of cherry maturity. Food Chem 274, 137–145. https://doi.org/10.1016/j.foodchem.2018.08.127

Verdú, S., García, I., Roda, C., Barat, J.M., Grau, R., Ferrer, A., Prats-Montalbán, J.M., 2025. Multivariate image analysis for assessment of textural attributes in transglutaminase-reconstituted meat. Chemometrics and Intelligent Laboratory Systems 256. https://doi.org/10.1016/j.chemolab.2024.105280

Villalba, P., Sanchis, J., Ferrer, A., 2019. A graphical user interface for PCA-based MSPC: A benchmark software for multivariate statistical process control in

MATLAB. Chemometrics and Intelligent Laboratory Systems 185, 135–152. https://doi.org/10.1016/j.chemolab.2018.12.004

Vitale, R., Prats-Montalbán, J.M., López-García, F., Blasco, J., Ferrer, A., 2016. Segmentation techniques in image analysis: A comparative study. J Chemom 30, 749–758. https://doi.org/10.1002/cem.2854

Vranckx, I., Raymaekers, J., De Ketelaere, B., Rousseeuw, P.J., Hubert, M., 2021. Real-time discriminant analysis in the presence of label and measurement noise. Chemometrics and Intelligent Laboratory Systems 208. https://doi.org/10.1016/j.chemolab.2020.104197

Wang, F., Ma, S., Yan, G., 2023. A PLS-based random forest for NOx emission measurement of power plant. Chemometrics and Intelligent Laboratory Systems 240. https://doi.org/10.1016/j.chemolab.2023.104926

Wang, T., Jiao, L., Yan, C., He, Y., Li, M., Zhang, T., Li, H., 2019. Simultaneous quantitative analysis of four metal elements in oily sludge by laser induced breakdown spectroscopy coupled with wavelet transform-random forest (WT-RF). Chemometrics and Intelligent Laboratory Systems 194. https://doi.org/10.1016/j.chemolab.2019.103854

Xu, X., Guan, L., Wang, Z., Yao, R., Guan, X., 2025. A double-layer forecasting model for PV power forecasting based on GRU-Informer-SVR and Blending ensemble learning framework. Appl Soft Comput 172. https://doi.org/10.1016/j.asoc.2025.112768

Yaqoob, M., Sharma, S., Aggarwal, P., 2021. Imaging techniques in Agro-industry and their applications, a review. Journal of Food Measurement and Characterization. https://doi.org/10.1007/s11694-021-00809-w

Yeap, D., McCartney, M.M., Rajapakse, M.Y., Fung, A.G., Kenyon, N.J., Davis, C.E., 2020. Peak detection and random forests classification software for gas chromatography/differential mobility spectrometry (GC/DMS) data. Chemometrics and Intelligent Laboratory Systems 203. https://doi.org/10.1016/j.chemolab.2020.104085

Yolmeh, M., Jafari, S.M., 2017. Applications of Response Surface Methodology in the Food Industry Processes. Food Bioproc Tech. https://doi.org/10.1007/s11947-016-1855-2

Yoon, S. C., Lawrence, K. C., Smith, D. P., Park, B., & Windham, W. R. 2007. Bone Fragment Detection in Chicken Breast Fillets Using Transmittance Image Enhancement. Trans ASABE 51, 331–339. DOI: https://doi.org/10.13031/2013.24209

Zhang, Y., 2014. An improved QSPR method based on support vector machine applying rational sample data selection and genetic algorithm-controlled training parameters optimization. Chemometrics and Intelligent Laboratory Systems 134, 34–46. https://doi.org/10.1016/j.chemolab.2014.03.004

Zhao, B., Yang, P., Basir, O.A., Mittal, G.S., 2006. Ultrasound based glass fragments detection in glass containers filled with beverages using neural networks and short time Fourier transform. Food Research International 39, 686–695. https://doi.org/10.1016/j.foodres.2006.01.008

Zhu, J., Deng, J., Zhao, X., Xu, L., Jiang, H., 2024. Accurate identification of cadmium pollution in peanut oil using microwave technology combined with SVM-RFE. Sens Actuators A Phys 368. https://doi.org/10.1016/j.sna.2024.115085

9. Supplementary material

The results of the unsupervised PCA-MSPC detection of BF, obtained through the feature extraction approach based on first-order statistics applied to ultrasound energy-magnitude-distribution parameters (Section 2.7.1.3), are summarized in Tables 1S to 3S. In addition, the results regarding the improvement of BF detection using latent-variable-based machine learning techniques (Section 2.7.2) have been systematically compiled into 120 Excel files, provided as supplementary material. These files include the complete datasets, detailed outcomes, and computational analyses that support and validate the findings of this study. Specifically, they summarize the results of multivariate statistical modeling using machine learning techniques within the latent variables framework (SVM, DTe, RF, NB, LDA, QDA, and GLM), as well as those obtained from the Random Forest-Variable Selection strategy (RF-RV-LV-ML) applied to SVM, DTe, RF, NB, LDA, QDA, and GLM. Furthermore, the statistical results of all hyperparameters configurations tested for LV-SVM are separately represented considering TDA (Fig. 1S), FDA (Fig. 2S), TFDA (Fig. 3S), TFDABH (Fig. 4S) and TFDABS (Fig. 5S) data approaches.

Table 1S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction time-domain approach (feTDA).

feTDA-RSS					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	29	85.67 ± 1.48	0.86 ± 0.03	0.86 ± 0.02
0	95	26	87.25 ± 1.05	0.86 ± 0.02	0.89 ± 0.02
0	97.5	26	88.13 ± 1.11	0.84 ± 0.02	0.92 ± 0.02
0	99	26	87.96 ± 1.41	0.81 ± 0.03	0.95 ± 0.02
50	90	25	89.10 ± 1.29	0.85 ± 0.02	0.93 ± 0.02
50	95	26	88.36 ± 1.26	0.81 ± 0.03	0.96 ± 0.02
50	97.5	26	87.09 ± 1.59	0.77 ± 0.03	0.97 ± 0.02
50	99	26	84.52 ± 2.00	0.71 ± 0.04	0.98 ± 0.02
75	90	25	88.98 ± 1.18	0.82 ± 0.02	0.96 ± 0.02
75	95	26	87.10 ± 1.60	0.77 ± 0.03	0.97 ± 0.02
75	97.5	26	85.19 ± 1.71	0.72 ± 0.04	0.98 ± 0.02
75	99	25	82.30 ± 2.10	0.66 ± 0.05	0.99 ± 0.01
100	90	26	87.64 ± 1.52	0.79 ± 0.03	0.96 ± 0.02
100	95	26	85.99 ± 1.81	0.75 ± 0.04	0.97 ± 0.01
100	97.5	26	83.67 ± 2.26	0.69 ± 0.05	0.98 ± 0.01
100	99	25	80.86 ± 1.68	0.63 ± 0.04	0.99 ± 0.01
feTDA-T^2					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	31	86.22 ± 1.00	0.86 ± 0.01	0.86 ± 0.02
0	95	37	88.56 ± 0.70	0.89 ± 0.01	0.88 ± 0.02
0	97.5	40	89.80 ± 0.76	0.89 ± 0.01	0.90 ± 0.01
0	99	40	90.20 ± 0.90	0.88 ± 0.01	0.92 ± 0.02
50	90	40	89.68 ± 1.04	0.85 ± 0.02	0.95 ± 0.02
50	95	40	89.33 ± 1.10	0.84 ± 0.02	0.95 ± 0.02
50	97.5	40	88.98 ± 1.10	0.83 ± 0.02	0.95 ± 0.02
50	99	40	88.49 ± 1.00	0.81 ± 0.02	0.96 ± 0.02
75	90	40	88.59 ± 0.83	0.82 ± 0.02	0.96 ± 0.02
75	95	40	88.20 ± 0.86	0.81 ± 0.02	0.96 ± 0.02
75	97.5	40	87.59 ± 0.88	0.79 ± 0.02	0.96 ± 0.02
75	99	40	86.94 ± 1.07	0.78 ± 0.03	0.96 ± 0.02
100	90	40	87.43 ± 0.93	0.78 ± 0.02	0.96 ± 0.02
100	95	40	86.77 ± 1.05	0.77 ± 0.02	0.97 ± 0.02
100	97.5	40	85.90 ± 1.18	0.75 ± 0.03	0.97 ± 0.02
100	99	40	85.02 ± 1.20	0.73 ± 0.03	0.97 ± 0.02

feTDA (feature-extraction time-domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), limit augmentation (LA), OPCs (optimal number of principal components), A_{cc} (overall accuracy), S_e (sensibility) and S_p (specificity). Results are expressed as mean ± standard error.

Table 2S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction frequency-domain approach (feFDA).

feFDA-RSS					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	21	85.67 ± 1.18	0.86 ± 0.03	0.86 ± 0.02
0	95	27	88.94 ± 1.36	0.88 ± 0.03	0.89 ± 0.02
0	97.5	27	88.87 ± 1.66	0.85 ± 0.04	0.93 ± 0.02
0	99	28	87.69 ± 1.89	0.81 ± 0.04	0.95 ± 0.02
50	90	30	88.41 ± 1.99	0.85 ± 0.04	0.91 ± 0.02
50	95	27	87.69 ± 1.56	0.80 ± 0.03	0.95 ± 0.02
50	97.5	27	85.85 ± 2.01	0.76 ± 0.05	0.96 ± 0.02
50	99	27	83.25 ± 2.61	0.69 ± 0.06	0.97 ± 0.02
75	90	31	87.96 ± 2.22	0.82 ± 0.05	0.94 ± 0.02
75	95	31	85.68 ± 2.99	0.77 ± 0.06	0.95 ± 0.02
75	97.5	27	83.63 ± 2.16	0.70 ± 0.05	0.97 ± 0.02
75	99	28	81.83 ± 2.49	0.66 ± 0.06	0.98 ± 0.02
100	90	31	86.67 ± 2.40	0.79 ± 0.05	0.94 ± 0.02
100	95	31	83.83 ± 3.27	0.72 ± 0.07	0.95 ± 0.02
100	97.5	27	82.54 ± 2.19	0.67 ± 0.05	0.98 ± 0.01
100	99	28	80.07 ± 2.54	0.62 ± 0.05	0.98 ± 0.01
feFDA-T^2					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	32	85.86 ± 1.03	0.86 ± 0.02	0.86 ± 0.02
0	95	35	90.20 ± 0.87	0.91 ± 0.02	0.89 ± 0.02
0	97.5	35	91.20 ± 0.91	0.90 ± 0.02	0.92 ± 0.02
0	99	35	91.60 ± 0.88	0.90 ± 0.02	0.93 ± 0.02
50	90	35	89.64 ± 1.03	0.83 ± 0.03	0.96 ± 0.02
50	95	35	88.56 ± 1.29	0.81 ± 0.03	0.96 ± 0.02
50	97.5	35	87.35 ± 1.59	0.78 ± 0.04	0.97 ± 0.02
50	99	35	86.70 ± 1.57	0.77 ± 0.04	0.97 ± 0.02
75	90	35	86.28 ± 1.34	0.76 ± 0.03	0.97 ± 0.02
75	95	35	84.94 ± 1.47	0.73 ± 0.03	0.97 ± 0.02
75	97.5	35	83.59 ± 1.60	0.70 ± 0.04	0.97 ± 0.02
75	99	35	82.80 ± 1.54	0.68 ± 0.04	0.97 ± 0.02
100	90	35	83.06 ± 1.68	0.69 ± 0.04	0.98 ± 0.02
100	95	35	81.84 ± 1.77	0.66 ± 0.04	0.98 ± 0.02
100	97.5	35	81.05 ± 1.65	0.64 ± 0.04	0.98 ± 0.02
100	99	35	80.64 ± 1.44	0.63 ± 0.04	0.98 ± 0.02

feFDA (feature-extraction frequency-domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), limit augmentation (LA), OPCs (optimal number of principal components), A_{cc} (overall accuracy), S_e (sensibility) and S_p (specificity). Results are expressed as mean ± standard error.

Table 3S. Optimized Principal Component models (PCA) and statistical performance of the Residual Sum Squares (RSS) and Hotelling's T-squared (T^2) multivariate control statistics for detection of bone fragments using feature-extraction time-frequency domain approach (feTFDA).

feTFDA-RSS					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	14	86.12 ± 1.09	0.86 ± 0.01	0.86 ± 0.02
0	95	21	88.83 ± 1.12	0.89 ± 0.02	0.89 ± 0.02
0	97.5	30	90.43 ± 1.55	0.90 ± 0.04	0.90 ± 0.02
0	99	33	91.17 ± 1.61	0.90 ± 0.04	0.92 ± 0.02
50	90	30	90.99 ± 1.19	0.91 ± 0.02	0.91 ± 0.02
50	95	34	92.05 ± 1.10	0.91 ± 0.02	0.93 ± 0.02
50	97.5	36	92.83 ± 1.15	0.93 ± 0.03	0.93 ± 0.02
50	99	37	92.61 ± 1.25	0.92 ± 0.03	0.93 ± 0.02
75	90	35	92.85 ± 1.05	0.93 ± 0.02	0.93 ± 0.02
75	95	37	93.04 ± 1.00	0.94 ± 0.02	0.93 ± 0.02
75	97.5	38	93.04 ± 1.02	0.93 ± 0.02	0.93 ± 0.02
75	99	40	92.78 ± 1.18	0.93 ± 0.02	0.93 ± 0.02
100	90	36	92.39 ± 1.15	0.92 ± 0.02	0.93 ± 0.02
100	95	38	92.74 ± 0.89	0.93 ± 0.02	0.93 ± 0.01
100	97.5	41	92.85 ± 1.07	0.93 ± 0.02	0.93 ± 0.02
100	99	43	92.78 ± 1.08	0.92 ± 0.02	0.93 ± 0.02
feTFDA-T^2					
LA (%)	Control limit (%)	OPCs	A_{cc} (%)	S_e	S_p
0	90	37	86.14 ± 0.87	0.86 ± 0.02	0.86 ± 0.02
0	95	41	88.23 ± 1.00	0.88 ± 0.01	0.89 ± 0.02
0	97.5	44	90.28 ± 0.99	0.90 ± 0.02	0.90 ± 0.01
0	99	45	91.01 ± 0.87	0.90 ± 0.02	0.92 ± 0.02
50	90	50	92.37 ± 1.04	0.92 ± 0.02	0.93 ± 0.02
50	95	51	93.12 ± 0.89	0.94 ± 0.01	0.93 ± 0.02
50	97.5	51	92.99 ± 0.90	0.93 ± 0.01	0.93 ± 0.02
50	99	51	92.85 ± 0.99	0.93 ± 0.02	0.93 ± 0.02
75	90	52	92.79 ± 0.85	0.93 ± 0.02	0.93 ± 0.02
75	95	52	92.53 ± 0.93	0.92 ± 0.02	0.93 ± 0.02
75	97.5	53	93.30 ± 0.86	0.94 ± 0.02	0.93 ± 0.02
75	99	53	93.22 ± 0.91	0.93 ± 0.02	0.93 ± 0.02
100	90	54	93.44 ± 0.98	0.93 ± 0.02	0.93 ± 0.02
100	95	54	93.22 ± 0.88	0.93 ± 0.02	0.94 ± 0.02
100	97.5	54	93.08 ± 0.93	0.92 ± 0.02	0.94 ± 0.02
100	99	55	93.63 ± 0.94	0.94 ± 0.01	0.93 ± 0.02

feTFDA (feature-extraction time-frequency domain approach), RSS (Residual Sum Squares), T^2 (Hotelling's T-squared), limit augmentation (LA), OPCs (optimal number of principal components), A_{cc} (overall accuracy), S_e (sensibility) and S_p (specificity). Results are expressed as mean ± standard error.

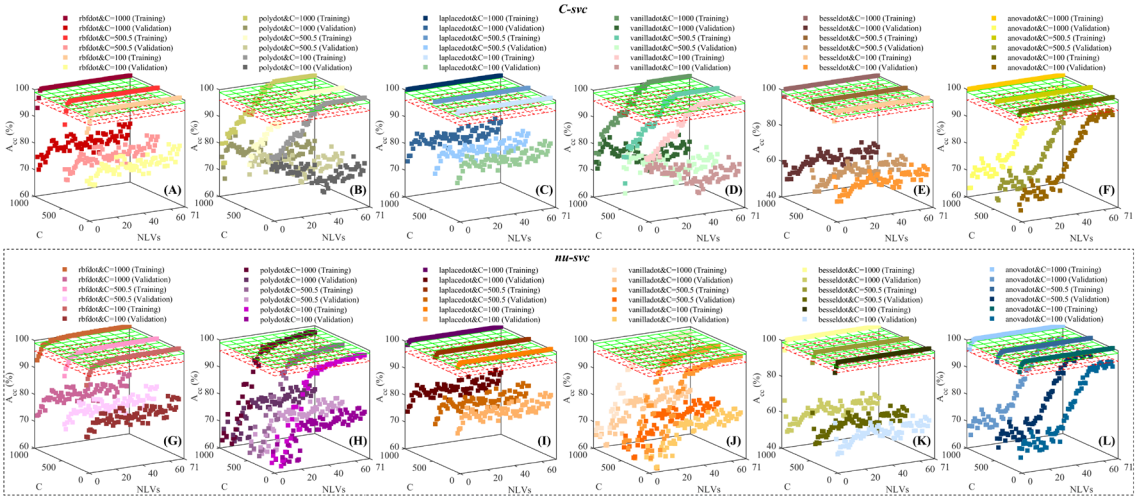


Fig. 1S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-domain approach (TDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

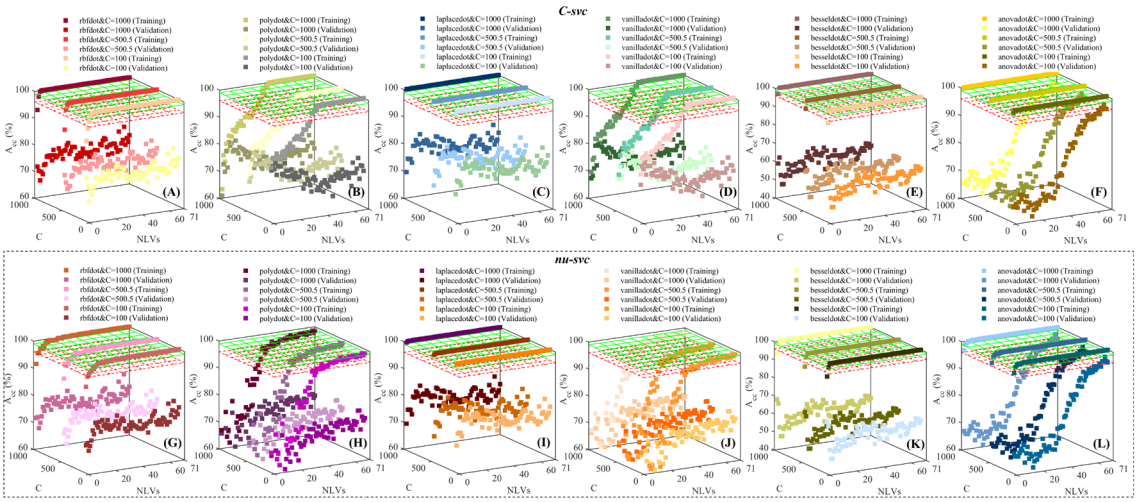


Fig. 2S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the frequency-domain approach (FDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

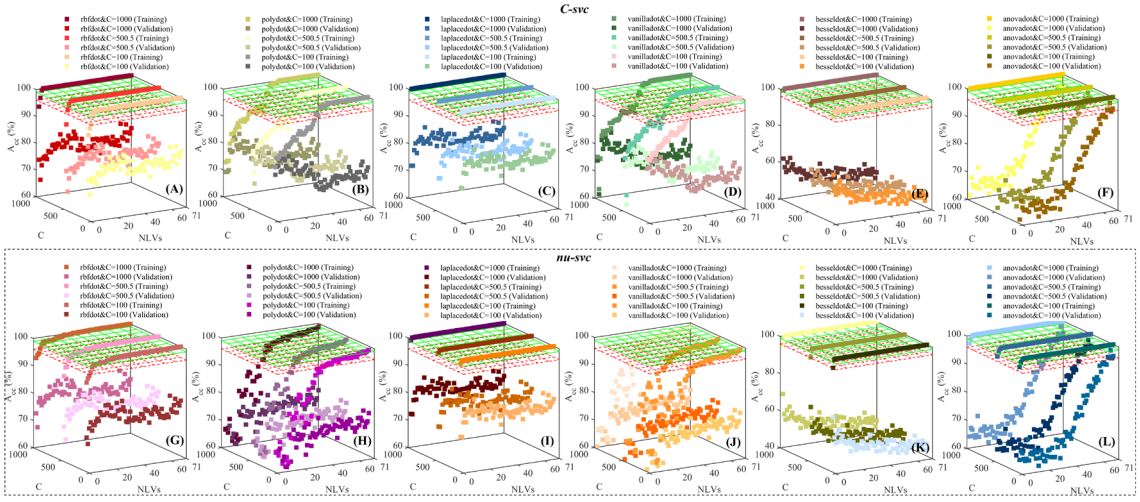


Fig. 3S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach (TFDA), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

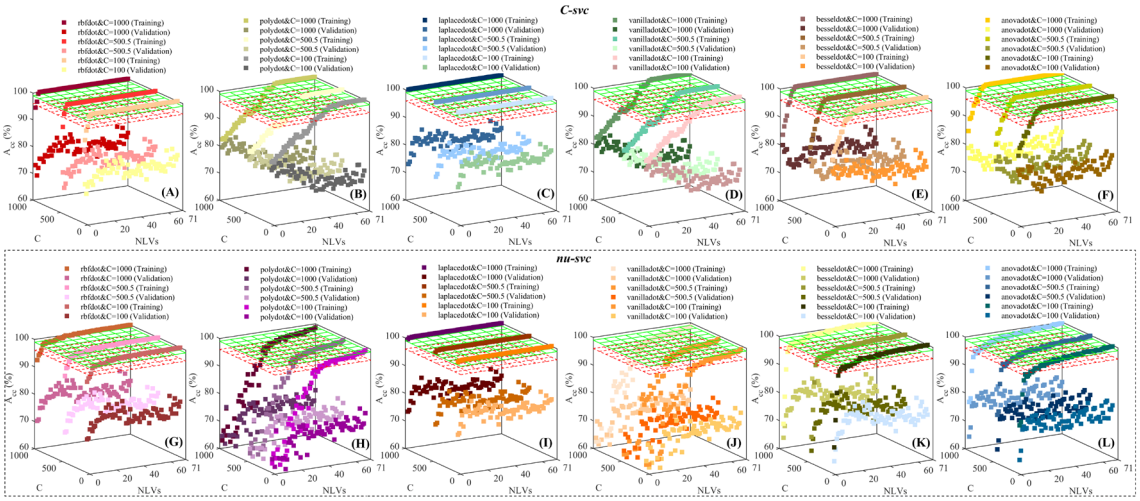


Fig. 4S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale hard (TFDABH), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

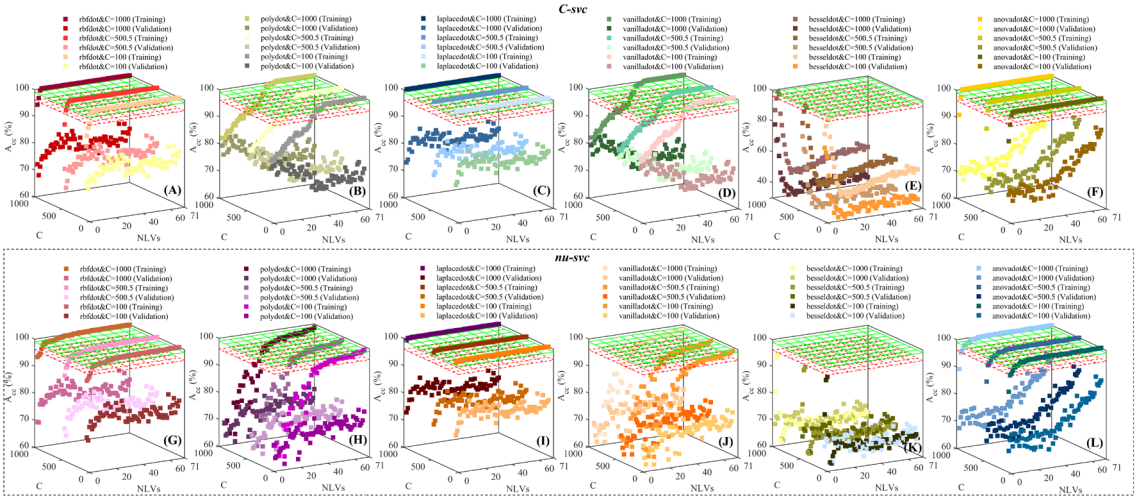


Fig. 5S. Statistical classification performance of the Latent Variable-Support Vector Machine (LV-SVM) using the time-frequency-domain approach-block-scale soft (TFDABS), shown as a function of the SVM hyperparameters and the number of latent variables (NLVs) tested. Results of overall accuracy (A_{cc}) are reported separately for the training (75%) and the validation (25%) datasets.

10. Annexed

Sustainable development goals

This master's thesis, entitled: Non-invasive detection of bone fragments in poultry meat by the integration of ultrasound imaging, multivariate image analysis and machine learning, aligns closely with the triple impact goals of the United Nations Sustainable Development Goals (SDGs). Firstly, by ensuring the safety and quality of poultry and food products, this research supports SDG 3: Good Health and Well-being, by reducing the risk of injuries caused by bone fragments. In addition, this research supports SDG 2: Zero Hunger, as improving food safety, minimizing waste, and increasing the reliability of animal protein production directly contribute to global food security and the availability of safe, nutritious food for all.

Secondly, the adoption of non-invasive, real-time detection methods enhances operational efficiency and reduces waste, contributing to SDG 12: Responsible Consumption and Production. Lastly, the implementation of advanced technologies such as machine learning and ultrasound based-imaging promotes innovation and sustainable industrial practices, aligning with SDG 9: Industry, Innovation, and Infrastructure. The integration of these cutting-edge technologies not only addresses food safety concerns but also fosters sustainable practices within the food industry, thereby creating a comprehensive impact on health, production efficiency, and technological advancement.