

Adaptación de inteligencia artificial por el modelo de regresión múltiple estocástica para determinar la calidad de la fibra de alpaca (*Lama pacos*)

Artificial intelligence adaptation by the stochastic multiple regression model to determine the fibre quality of alpaca (*Lama pacos*)

Abdel Alejandro Portocarrero Banda², Eric Vilca Cayllahua², Briguit Stefany Ortiz Quispe², Lilia Mary Miranda Ramos², Hugo Guillermo Jiménez Pacheco^{1,2*}

RESUMEN

Se describe la aplicación de inteligencia artificial basada en el modelo de regresión lineal múltiple con gradiente descendiente estocástica con la finalidad de determinar la calidad de la fibra de alpaca Huacaya de color blanco. Se analizaron 1200 fibras correspondientes a seis muestras de alpaca. Las fibras se caracterizaron mediante microscopía óptica y con el equipo analizador óptico de diámetro de fibra (OFDA 100). Se consideraron como variables independientes al diámetro de fibra, diámetro de médula, porcentaje de medulación por volumen, factor de confort, fibras objetables y como variable de respuesta al factor «Soft». Esta última variable resultante de la diferencia del factor de confort y fibras objetables sirvió para darle un ordenamiento lógico a la matriz de datos y obtener un modelo de predicción preciso. Los valores promedio fueron 26.80 ± 6.95 para el diámetro de fibra, 14.10 ± 5.92 en diámetro de médula, 24.75 ± 13.20 μm para el porcentaje de medulación por volumen y $71.56 \pm 13.04\%$ para el factor de confort. El modelamiento de regresión lineal múltiple de *machine learning* se adaptó con gran precisión a un tamaño muestral pequeño, mostrando errores mínimos, y optimizado con el algoritmo de gradiente

¹ Universidad Católica de Santa María, Arequipa, Perú

² Instituto de Investigación e Innovación en Energías Renovables y Medio Ambiente, Arequipa, Perú

* E-mail: hjimenez@ucsm.edu.pe

Recibido: 5 de julio de 2022

Aceptado para publicación: 18 de febrero de 2023

Publicado: 28 de abril de 2023

©Los autores. Este artículo es publicado por la Rev Inv Vet Perú de la Facultad de Medicina Veterinaria, Universidad Nacional Mayor de San Marcos. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0) [<https://creativecommons.org/licenses/by/4.0/deed.es>] que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada de su fuente original

descendiente estocástico predijo un factor Soft muy cercano al factor Soft observado. Se concluye que la técnica de regresión lineal múltiple con el enfoque estocástico satisface la predicción del nuevo factor denominado «soft» y que representa el modelamiento adecuado para la predicción de calidad de fibras en la industria textil.

Palabras clave: fibra de alpaca, inteligencia artificial, factor Soft, regresión múltiple estocástica

ABSTRACT

The application of artificial intelligence based on the multiple linear regression model with stochastic descending gradient is described in order to determine the quality of the white Huacaya alpaca fibre. In total, 1200 fibres corresponding to six alpaca samples were analysed. The fibres were characterized by optical microscopy and with the optical fibre diameter analyser (OFDA 100) equipment. Fibre diameter, medulla diameter, percentage of medullation by volume, comfort factor, and objectionable fibres were considered as independent variables, and the «Soft» factor was considered as a response variable. This last variable resulting from the difference in the comfort factor and objectionable fibres served to give a logical order to the data matrix and obtain an accurate prediction model. The average values were 26.80 ± 6.95 for the fibre diameter, 14.10 ± 5.92 for the medulla diameter, $24.75 \pm 13.20 \mu\text{m}$ for the percentage of medullation by volume and $71.56 \pm 13.04\%$ for the comfort factor. The machine learning multiple linear regression modelling fitted a small sample size with high precision, showing minimal errors, and optimized with the stochastic gradient descent algorithm predicted a Soft factor very close to the observed Soft factor. It is concluded that the multiple linear regression technique with the stochastic approach satisfies the prediction of the new factor called «soft» and that it represents the appropriate modelling for the prediction of fibre quality in the textile industry.

Key words: alpaca fiber, artificial intelligence, Soft factor, stochastic multiple regression

INTRODUCCIÓN

El Perú concentra el 72% de la población mundial alpacas, con 4.3 millones de cabezas, siendo Arequipa la tercera región con mayor porcentaje de alpacas representado por el 12.7% de la población total (MIDAGRI 2021). La comercialización de las fibras de alpaca como materia prima para la industria textil en las regiones altoandinas es una de las actividades de mayor importancia, lo que brinda oportunidades de desarrollo mediante la exportación de fibras procesadas o como productos textiles incrementando su valor agregado (ONUUDI, 2006; Guillen *et*

al., 2023). La industria textil de Arequipa es la principal demandante de fibras de alpaca, abarcando el 85% de la producción nacional, la cual es mayormente adquirida por empresas que acopian fibras de pequeños y medianos productores (66.5%) (Infoalpaca, 2013), que, en base a sus características textiles, establecen el precio (Frank, 2008). Algunas empresas privadas dan mejores precios por vellones más finos (González *et al.*, 2008). La producción y comercialización de fibras desde el punto de vista socioeconómico es de gran relevancia, puesto que gran parte de pequeños, medianos productores y empresas agropecuarias dependen económicamente de su explotación (Carpio, 2017).

La finura de la fibra está relacionada directamente con la calidad de la fibra, pues diámetros de fibra más finos son más apreciados para la comercialización y transformación en prendas (Cordero *et al.*, 2011). El grado de finura se expresa en el factor de confort (Bustinza, 2001), el cual se relaciona a la suavidad del tejido, y se define como el porcentaje de diámetro de fibra debajo de los 30 μm (Naylor y Hansford 1999; Frank *et al.*, 2014; Tadesse *et al.*, 2021) ya que la presencia de fibras gruesas le confiere una apariencia heterogénea a los tejidos o hilos (Wang *et al.*, 2005) y aumenta el efecto de picazón en la piel (Professor, 2021). También se ha encontrado que una alta incidencia de medulación y tipo de medulación en las fibras causan el mismo efecto (Frank *et al.*, 2014, 2017). Las fibras meduladas que tienen efecto sobre el factor de confort son consideradas como objetables, y según IWTO (1998) son las que tienen una opacidad superior al 94% y un diámetro superior a 25 μm .

Por lo tanto, para la determinación de la calidad de fibra y su grado de confort se debería de abordar variables adicionales que relacionen las principales características textiles, tales como el diámetro de fibra, diámetro de médula, porcentaje de medulación por volumen y porcentaje de fibras objetables. Por otro lado, se dispone de métodos para la determinación de las principales características textiles como el Analizador de Diámetro de Fibra Óptica (OFDA), que es el más aplicado por las industrias textiles (Brims *et al.*, 1999), pero no muestra una relación entre características, de allí que se hace necesario realizar un estudio que evalúe estas interacciones en un solo factor con el cual se pueda identificar y dar una predicción precisa y eficaz; lo cual debería incluir modelos predictivos aplicando inteligencia artificial.

El desarrollo de inteligencia artificial (IA) ha permitido que las computadoras puedan procesar automáticamente datos y obtener una predicción a partir de ellos, mediante el cual es posible obtener valores muy cercanos a la realidad (Krizhevsky *et al.*, 2012).

Diversos estudios han demostrado que la aplicación de un modelo basado en IA, como el modelo convencional de regresión lineal múltiple (MLR), puede proporcionar datos precisos, incluso cuando la cantidad de datos proporcionados es baja (Uyak *et al.*, 2005; Kulkarni y Chellam, 2010; Peleato *et al.*, 2018). Por lo tanto, el presente estudio tuvo como objetivo el uso de inteligencia artificial basada en el modelo de MLR evaluando un nuevo factor para determinar la calidad en fibras de alpaca Huacaya.

MATERIALES Y MÉTODOS

Lugar de Estudio, Animales y Muestras

El trabajo de investigación se realizó con muestras de vellón obtenidas de alpacas del centro poblado Chalhuanca, distrito de Yanque, provincia de Caylloma, Región Arequipa, Perú, ubicada a 4349 msnm. Las muestras de fibra (20 g) se tomaron del costillar medio sobre la antepenúltima costilla de alpacas Huacaya (Aylan-Parker y McGregor 2002). El tamaño de muestra corresponde a una mínima cantidad requerida por el modelo de inteligencia artificial que comprende seis alpacas blancas, de sexo macho y con edades de entre 1 y 3 años.

Análisis de Laboratorio

Las muestras de fibra de alpaca fueron analizadas en el Laboratorio de Investigación e Innovación en Energías Renovables y Medio Ambiente (INNOVERGY) de la Universidad Católica de Santa María, en Arequipa, Perú.

Las muestras fueron lavadas repetidas veces con agua a 50 °C con detergente industrial (50 g/l), luego se sumergieron en alcohol etílico al 96% (PA) y se secaron a condiciones de medioambiente. Con ayuda de pinzas y un estereoscopio (Stemi 508, Zeiss), se colocaron las fibras de manera aleatoria

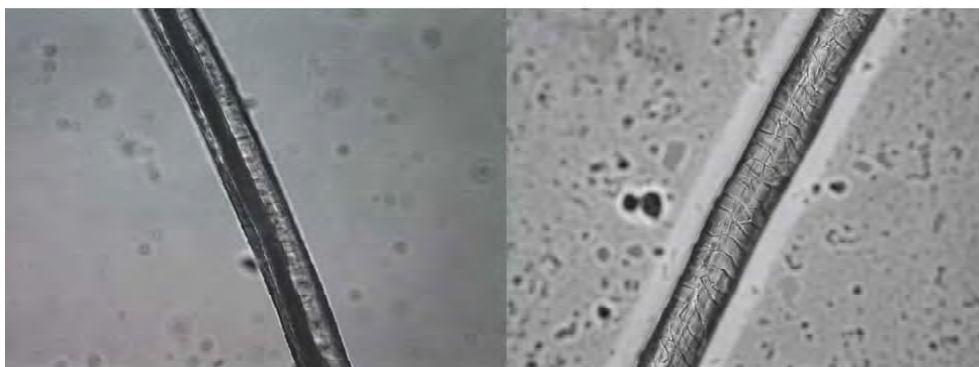


Figura 1. Muestra de fibra bajo microscopio óptico con objetivo de 10X y aumento de 40X para caracterizar fibras.

en el portaobjetos en forma vertical para luego colocar el cubreobjetos y proceder a la observación por microscopía óptica (MO) (Zhong *et al.* 2016; Xing *et al.*, 2020; Zang *et al.*, 2021).

Se analizaron 1200 fibras en total, 200 fibras por cada paquete de vellón colectado de las 6 alpacas, nombrándolos como M1, M2, M3, M4, M5 y M6. Seguidamente, se realizó el análisis con MO con objetivo de 10x y 40x con la finalidad de analizar fibras independientes, descartando así las fibras sobrepuestas (Figura 1). Asimismo, se tomaron fotografías mediante el microscopio trinocular con una cámara CMEX-10 PRO 10 mp de Euromex acoplada en conjunto con el software de visualización, captura de imágenes y medición Zen Blue v. 2.3.

En cada fibra se determinó el diámetro total y diámetro de médula. Se determinó el diámetro medio de fibra (DMF), la desviación estándar de diámetro fibra (DEF), el coeficiente de variación de diámetro de fibra (CVF), el factor de confort (FC) que considera el porcentaje de fibras menores de 30 μm (McGregor 1997; Frank *et al.*, 2006), el diámetro medio de médula (DMM), la desviación estándar de diámetro de médula (DEM), el porcentaje de fibras objetables

(FO) que considera el porcentaje de fibras con medulas mayores a 25 μm (IWTO-8-2011 2017) y el porcentaje de modulación por volumen (MV) que se calculó con la fórmula de Bray (% Med. Vol. = $([Nm/Nw]*100*[d2m + o2m/d2w + o2w])$) [Ecuación 1], donde Nw = número total de fibras medidas, $d2w$ = promedio de diámetro de fibra, $o2w$ = desviación estándar de diámetro de fibra, Nm = número de fibras con médula, $d2m$ = promedio de diámetro de médula y $o2m$ = desviación estándar de diámetro de médula (Merrick, 1998).

Desarrollo del Modelo Usando Inteligencia Artificial

Se utilizó un modelo de *machine learning*, regresión lineal múltiple (RLM), usando el método de optimización de descenso de gradientes estocástico (Pillaud-Vivien *et al.*, 2018), que fue implementado para la predicción de un nuevo factor denominado Soft (FS) de fibras de alpaca. La implementación del modelo de IA se realizó en el computador HP Pavilion Intel Core i5 de décima generación de 2.50 GHz hasta 4.50 GHz con TurboBoost de Intel. Para el desarrollo del modelo se usó el lenguaje de programación Python 3.9.2 en conjunto con las librerías *pandas* y *numpy*.

Toma de muestra y configuración del sistema

Las muestras se tomaron por conjuntos de datos, cada uno de estos contenía los parámetros DMF, DMM, MV, FO y el FC de 30 fibras individuales para generar una muestra (dato). Para generar un *dataset* se tomaron 1200 fibras individuales obteniendo 40 datos. Se incluyen también seis (6) datos adicionales al *dataset* provenientes del equipo OFDA.

Hold out set

A partir de la información tomada de las muestras se construyó la matriz de datos. Un dato contiene cinco (5) variables independientes (DMF, DMM, MV, FO y FC) y un dato dependiente (% Soft). De los 46 datos se realizó una división aleatoria con el propósito de generalizar el modelo de forma optimizada seleccionando datos aleatorios, 75% del *dataset* para datos de entrenamiento y verificación de modelo, y 25% para datos de prueba y la validación de modelo.

Descenso de gradientes estocástico

La aplicación de la regresión lineal múltiple (RLM) mediante el método del descenso de gradientes estocástico (SGD) es computacionalmente más eficiente que otros algoritmos, adaptativo a *datasets* pequeños permitiendo realizar la generalización de predicciones, gestionando anomalías de manera correcta y evitando la sobreestimación de la correlación de la variable dependiente y las predicciones. Se formula el modelo (Ecuación 2) que define la estructura de las variables independientes y la variable dependiente, en la que $n=5$, almacenando los pesos procesados $(\{\theta_i, x_i\}_{i=1}^n)$ permitiendo actualizar los pesos. Se aplica regularización para mejorar la generalización (Ecuación 3) para mejorar la estimación de datos nuevos que ingresen al modelo, y permitir un modelo escalable y preparado para retroalimentación y mantener un modelo estable (Zhang, 2004).

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n = \theta^T x \quad (2)$$

$$J(\theta) = \frac{1}{2m} \left[\sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^n \theta_j^2 \right] \quad (3)$$

Estimación de muestras foráneas

El cálculo de la estimación mediante mínimos cuadrados permite la transformación de datos numéricos continuos en forma de matrices a datos enteros continuos, de esta manera utilizar los pesos actualizados y operarlos con los datos entregados, y poder calcular la estimación por un conjunto de datos, o de manera individual.

Análisis de robustez y sensibilidad

El análisis de sensibilidad se llevó a cabo utilizando cuatro métodos estadísticos. El error medio absoluto (MAE), el error medio cuadrado (MSE), la raíz cuadrada del error medio cuadrado (RMSE) y el coeficiente de determinación (R^2) (Amor *et al.*, 2022). Se utilizó el MAE para poder analizar las diferencias individuales de las estimaciones y los datos reales (Ecuación 4). Se empleó el MSE para poder observar la diferencia generalizada de valores atípicos o anomalías (Ecuación 5). Se utilizó el RMSE para calcular la desviación estándar de la varianza y poder analizar el ajuste de estimación (Ecuación 6). Se aplica por último el coeficiente de determinación R^2 cuadrado para determinar el ajuste de regresión (Ecuación 7).

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (4)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (5)$$

$$RMSE = \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (6)$$

$$R^2 = 1 - \frac{ss_{res}}{ss_{tot}} \quad (7)$$

Validación de la inteligencia artificial

El proceso de validación consistió en comparar el método MLR para determinar el FS. Por un lado, inicialmente se evaluaron seis muestras de fibras de alpaca Huacaya mediante MO, donde se recopilaban datos correspondientes a las 1200 fibras. Por otro lado, se utilizaron seis ejemplares de alpaca Huacaya blanca, cada muestra preparada de acuerdo con la International Wool Textile Organization (IWTO, 2017a) y se evaluaron siguiendo el procedimiento descrito por dicha institución (IWTO, 2017b) mediante el equipo OFDA 100 en el laboratorio de control de calidad de la empresa Inca Tops S.A., Arequipa. Para concluir la validación se evaluaron las mismas fibras mediante la IA basada en MLR donde se obtuvo el FS de cada muestra.

RESULTADOS Y DISCUSIÓN

Características de las Fibras

Diámetro de fibra

En el Cuadro 1 se muestra la distribución estadística descriptiva del diámetro de la fibra de las muestras trabajadas.

El diámetro de fibra promedio fue $26.80 \pm 6.95 \mu\text{m}$ (Rango: 23.37 - 31.63 μm), los cuales mostraron diferencias significativas entre medias ($p < 0.05$), lo que indica una constante por efecto de la edad, y se corrobora con otros estudios que reportan un impacto significativo por esta variable en el DMF, ya que a menor edad se pueden observar en la alpaca fibras más finas (Gutierrez, 2011). Machaca *et al.* (2017) encontraron diámetros de fibra en el rango de 21.61 y 23.87 μm en alpacas de 1 y 3 años, respectivamente. Asimismo, Lupton *et al.* (2006) reportaron que en alpacas criadas en EEUU se obtuvo 24.3, 26.5 y 30.1 μm para animales de 1, 2 y >2 , respectivamente. El diámetro de fibra se considera como la característica más importante, ya que de esta depende la calidad, el rendimiento de la hilabilidad, la vellosidad del hilo, el valor y el uso final (Lupton *et al.*, 2006).

Diámetro de médula

En el Cuadro 2 se muestra que todas las muestras presentaron algún porcentaje de fibras meduladas. Esto confirma que la presencia de medulación está presente en la mayoría de fibras muestreadas, pudiendo determinarse el nivel mediante marcadores genéticos (Allain y Renieri, 2010). Por otro lado, el análisis de medulación es más sencillo por microscopía si las fibras son blancas o claras (Moore *et al.*, 2011).

Cuadro 1. Promedios del diámetro de fibra de alpaca del centro poblado Chalhuanca, provincia de Caylloma, Arequipa (Perú), según la muestra

Muestra	Media (μm)	DE	CVDF	IC (95%)
M1	23.37	6.10	0.26	22.58 - 24.16
M2	26.65	6.29	0.24	25.87 - 27.44
M3	25.05	6.15	0.25	23.93 - 26.16
M4	31.63	9.56	0.30	30.52 - 32.75
M5	28.08	7.60	0.27	26.97 - 29.20
M6	26.03	6.79	0.26	24.9 - 27.14
Media general	26.80	6.95	0.260	--

DE: Desviación estándar; CVDF: Coeficiente de variación de diámetro de fibra

Cuadro 2. Promedios del diámetro de médula (DMM) y porcentaje de medulación (%M) de la fibra de alpaca Huacaya del centro poblado Chalhuanca, provincia de Caylloma, Arequipa (Perú), según muestra y presencia de medulación

Medulación	DMM (μm)	DE (μm)	CVF	IC de 95%	%M
M1					
Medulada	20.59	7.38	0.36	19.75 - 21.433	68
No medulada					32
M2					
Medulada	11.63	5.59	0.48	11.00 - 12.27	40
No medulada					60
M3					
Medulada	12.21	4.69	0.38	11.45 - 12.96	49
No medulada					51
M4					
Medulada	14.69	6.31	0.43	13.67 - 15.71	79
No medulada					21
M5					
Medulada	12.10	5.64	0.47	11.19 - 13.01	64
No medulada					36
M6					
Medulada	13.35	5.90	0.44	12.40 - 14.30	57
No medulada					43
Media general	14.10	5.92	0.43		

El diámetro de médula de las muestras se presentó en un rango entre 2.22 μm como valor mínimo y 48.33 μm como valor máximo, siendo valores semejantes a los reportados por Cruz *et al.* (2019) en alpacas Huacaya (14.00 - 38.80 μm), indicando que la medición realizada por microscopía se encuentra dentro del rango de los diámetros de médula analizados por equipamiento especializado como el OFDA.

Las fibras de alpaca se diferenciaron en meduladas y no meduladas (Cuadro 2), obteniendo un rango promedio de DMM de 11.63 a 20.59 μm . La muestra 1 presentó el mayor valor de DMM con un valor de 20.59 μm , así como el segundo mayor porcentaje de fibras meduladas (68%), antecedido por la muestra 4 (79% de fibras meduladas). Por otra parte, la muestra 2 presentó el menor DMM (11.63 μm y 40% de fibras meduladas,

lo cual indica que las fibras de alpaca presentan varios grados de medulación, lo cual afecta la heterogeneidad de los tejidos y el factor de confort (Pinares *et al.*, 2018; Bonilla *et al.*, 2022). Sin embargo, las fibras meduladas también pueden ser utilizadas en la industria para artículos que requieran ser más livianos como ropa deportiva, o para ropa de invierno y sacos de dormir (Shishoo, 1995; Tao, 2001). Además, las fibras meduladas proporcionan un mayor aislamiento térmico natural en comparación con las fibras no meduladas, lo que proporciona a las prendas una alta calidad aislante y térmica (Tao 2001; Wang *et al.*, 2005).

Se determinaron diferencias significativas entre las medias del diámetro de médula ($p < 0.05$), además de una variabilidad media-baja de sus datos.

Cuadro 3. Medias del porcentaje de medulación por volumen de fibra de alpaca Huacaya de comunidades de la provincia de Caylloma, Arequipa (Perú)

Muestra	Media %	DE	IC 95%
M1	14.48	20.41	6.00 - 22.97
M2	15.77	8.56	7.29 - 24.26
M3	21.14	13.51	9.14 - 33.14
M4	42.32	5.54	30.32 - 54.32
M5	28.20	5.54	16.20 - 40.20
M6	26.60	11.67	14.60 - 38.60
Media	24.75	13.20	---

Porcentaje de medulación por volumen

La cantidad de fibras meduladas presentes en una muestra puede ser representada en porcentaje de acuerdo con el número, el volumen o por ambos, según el método de medición utilizado (Merrick, 1998). Para fines específicos de este estudio se consideró que la medulación por volumen porcentual era la más relevante, debido a que considera más variables en su desarrollo (porcentaje de medulación por número, y la media y la desviación estándar del diámetro de fibra y del diámetro de médula) (Merrick, 1998).

El porcentaje promedio de MV fue $24.75 \pm 13.20 \mu\text{m}$, con un rango entre 14.48 a $42.32 \mu\text{m}$ ($p < 0.05$). La muestra 4 presenta el mayor MV, con lo que se puede afirmar que casi el 50% del volumen de algunas de sus fibras está ocupado por la médula, siendo considerada como una muestra altamente medulada y de médula gruesa. Este resultado indica el alto grado de medulación de algunas fibras de alpaca, lo cual tendría gran influencia en la apariencia y la capacidad de teñido (Lupton *et al.*, 2006).

Factor de confort

Los valores de confort presentaron un promedio de $71.56 \pm 13.04\%$ (rango: 46.0 - 85.33%) (Cuadro 4), valores promedio inferiores a los reportados por Lupton *et al.* (2006), quienes obtuvieron 82.7, 74.1 y 58.6% en alpacas de 1, 2 y >2 años, respectivamente. Por otro lado, Machaca *et al.* (2017) reportaron para alpacas Huacaya de Cotaure valores entre 92.38 y 86.45%, además de indicar que el color, el sitio de muestreo y el sexo tienen influencia sobre el factor de confort. En el presente estudio, el sexo, color, y sitio de muestreo fueron fijos por lo que estas diferencias entre las medias de las muestras ($p < 0.05$) puede estar muy relacionado a la edad.

Cuadro 4. Medias del factor de confort de fibra de alpaca Huacaya comunidades de la provincia de Caylloma, Arequipa (Perú)

Muestras	Media (%)	DE	IC
M1	85.33	8.19	76.95 - 93.71
M2	74.67	8.04	66.28 - 83.05
M3	80.67	16.57	68.81 - 92.52
M4	46.00	23.9	34.1 - 57.9
M5	64.67	13.86	52.81 - 76.52
M6	78.00	10.70	66.14 - 89.86
Media	71.56	13.04	---

Fibras objetables

Las fibras objetables se caracterizan por ser fuertemente meduladas (Hunter *et al.* 2013), siendo consideradas como fibras que parecen tener una opacidad superior al 94% y un diámetro superior a $25 \mu\text{m}$ (IWTO-57, 1998). Se obtuvieron valores de fibras objetables entre 3.33 y 26.0%, resultados similares a los obtenidos por Lupton *et al.*

Cuadro 5. Matriz de correlación de Pearson de las variables de fibra de alpaca Huacaya para la predicción factor denominado Soft (FS)

	DMF	DMM	MV	FC	FO	FS
DMF	1	0.775***	0.688***	-0.941***	0.361**	-0.943***
DMM		1	0.975***	-0.692***	0.667***	-0.791***
MV			1	-0.595***	0.72***	-0.715***
FC				1	-0.261*	0.972***
FO					1	-0.479**

DMF: diámetro medio de fibra; DMM: diámetro medio de médula; MV: porcentaje de modulación por volumen; FC: factor de confort; FO: porcentaje de fibras objetables

*** Correlación fuerte; ** Correlación moderada; * Correlación débil

(2006), quienes reportaron valores entre 0.09 % y 22.37%. Por otro lado, no hubo diferencia entre las medias de las muestras para este factor.

Análisis de Correlación

Efecto del diámetro de médula y el porcentaje de modulación por volumen

La relación entre FS con las variables DMF, DMM, MV y FO presentaron correlaciones negativas, lo que indica que son inversamente proporcionales. Por otro lado, la correlación entre FS y FC fue positiva y fuerte (Cuadro 5), lo que indica que ante un mayor valor de confort resultará en un mayor valor de FS. Asimismo, se obtuvo una correlación alta entre DMM y MV (Cuadro 5), lo que indica que a mayores diámetros de médula se obtiene un mayor espacio ocupado por la médula (Merrick, 1998).

Al relacionar el DMF y DMM se obtuvo una correlación moderada positiva con el diámetro de fibra, relación que ha sido muy estudiada, y se corrobora con en el estudio de Berolatti *et al.* (2021), quienes obtuvieron una correlación positiva moderada alta en alpacas adultas entre modulación y diámetro de fibra. Así mismo, Onal *et al.* (2007) y Cordero *et al.* (2011) obtuvieron un compor-

tamiento similar al correlacionar la modulación con el diámetro de fibra en fibras de alpaca y otros, obteniendo un valor positivo con una correlación moderada alta y moderada, respectivamente.

Las correlaciones entre DMM y MV con FC fueron negativas y moderadas, indicando relaciones inversas, ya que con una mayor incidencia de fibras meduladas gruesas se obtiene porcentajes de confort inferiores. Bonilla *et al.* (2022) indican, asimismo, que al aumentar la incidencia de modulación, la sensación de picor y la heterogeneidad en los tejidos son más intensos. La fibra medular y el tipo de modulación son responsables de la disminución de la sensación de confort (Frank *et al.*, 2014, 2017).

Se obtuvo una correlación moderada entre las fibras objetables con el DMM y el MV, por lo que una mayor cantidad de fibras meduladas gruesas provoca un aumento en el porcentaje de fibras objetables, ya que las fibras objetables se caracterizan por ser meduladas y de mayor diámetro (IWTO-57, 1998). En este sentido, tanto el porcentaje de modulación como el diámetro de médula son variables de importancia para las industrias textiles, con las que puedan obtener productos textiles de calidad, sin que interfieran con el proceso de teñido.

Efecto del factor de confort

Los resultados de la correlación de Pearson muestran una fuerte correlación entre el FC con el FS. La fuerte correlación de DMF y el FC, entendiéndose que a mayor finura de fibra se brinda mayor comodidad (McGregor y Butler, 2004). Por lo contrario, a fibras más gruesas la sensación de picazón aumenta, la cual se ajusta a una distribución de ley potencia (Mamani *et al.*, 2022). Las correlaciones entre DMM y MV con FS fueron significativas y negativas, en tanto que la correlación entre FO y FS mostró una significancia débil, pero igualmente negativa. Es así que el aumento de DMM, MV y FO disminuye la FS. La relación de estas tres variables afecta directamente a la calidad de fibra, ya que a mayor DMM y MV se aumenta la incidencia de FO. McGregor (2006) reportó que la incidencia de medulación se incrementa entre 10 a 60% en fibras de 22 y 40 μm .

Efecto del porcentaje de fibras objetables

Se encontró una correlación moderada negativa entre FO y FS, por lo que ante un aumento de las fibras objetables se tienen menores valores del factor Soft. Asimismo, se obtuvo una fuerte correlación entre FO con FC y DMF; y una correlación moderada positiva con DMM y MV. Por lo tanto, un mayor porcentaje de FO se obtendrá el existir más fibras meduladas y que posean un diámetro de médula mayor. Esta misma tendencia se observó en el estudio de Lupton *et al.* (2006).

Ajuste de Datos para el Modelamiento de IA

Para la predicción de FS en fibras de alpaca se trató de extraer datos que estuvieran relacionados entre ellos para ser parte del *dataset* del modelo de RLM, toda vez que la precisión y confiabilidad del modelo se basa principalmente en este aspecto (Fong *et al.*, 2020). Inicialmente se seleccionaron variables relacionadas (diámetro de fibra, diá-

metro de médula, porcentaje de fibras meduladas, porcentaje de medulación por volumen y factor de confort), pero el error era muy grande (MAE training = 5.1028×10^{-28} ; MAE test = 4.0192×10^{-28}), afectando al modelo. Para solucionarlo, se retiró el porcentaje de fibras meduladas y se agregó la variable de porcentaje de fibras objetables. Este tipo de fibras está más relacionado al disconfort por fibras que causan comezón o picazón (Naebe *et al.*, 2018) La vinculación entre confort y disconfort está fuertemente asociada a la medulación y al diámetro de fibra (Pinares *et al.*, 2018). Es por esto que el nuevo factor denominado «Soft» es una variable resultante de la diferencia del factor de confort y las fibras objetables.

Modelo de Regresión Lineal Múltiple con el Enfoque del Descenso de Gradientes Estocástico

Se encontró que el modelo RLM propuesto con el enfoque SGD presenta una gran precisión en la predicción del FS. Se tomó como valor $k = 17$ de 34 datos en total de cada muestra (M1, M2, M3, M4, M5 y M6) de los cuales a partir de ellos se encontró el error (MSE, MAE, RMSE y r^2) El modelo realiza 1000 iteraciones y del total de datos solo usa 17 por iteración lo que permite generar todo tipo de combinación y obtener el mejor resultado.

Análisis de robustez del modelo RLM con el enfoque SGD

El modelo de IA propuesto realiza 1000 iteraciones teniendo un tiempo de respuesta promedio de 0.80 s. Los datos con poblaciones pequeñas se ajustaron al modelo con una excelente precisión sin mostrar sobre entrenamiento o poco entrenamiento de datos. Debido a esto, el modelo puede estar preparado para hacer una predicción muy precisa con datos nuevos. El modelo RLM convencional sin el proceso SGD muestra una precisión aún mayor para los datos, pero no estaría preparado para datos extraños, ya que al enfrentarse con datos nuevos la distancia

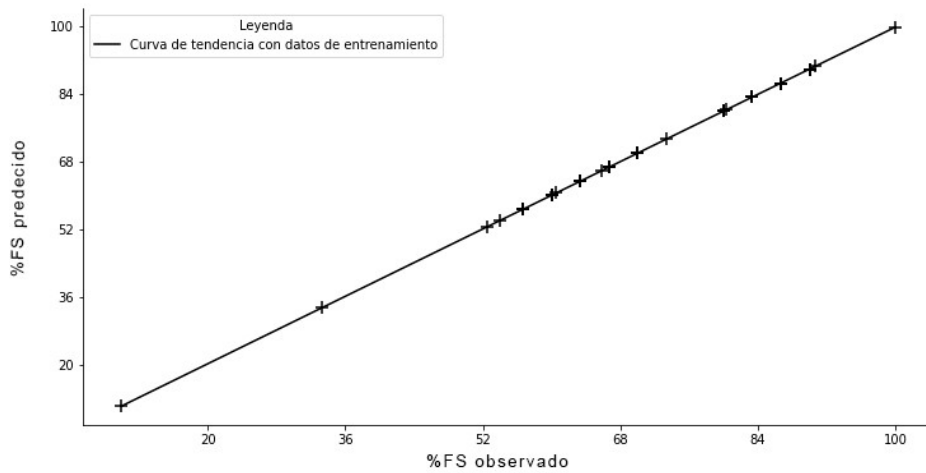


Figura 2. Curva de tendencia de los valores observados y los predichos de los datos de entrenamiento

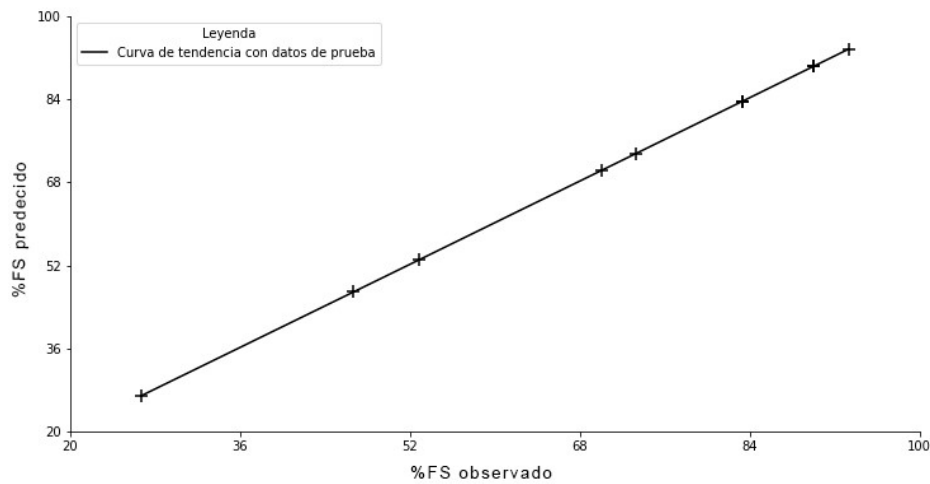


Figura 3. Curva de tendencia de los valores observados y los predichos de los datos de prueba

que habría entre la curva y el dato nuevo sería muy grande, obteniendo un error mayor, caso contrario cuando se utiliza SGD que permite determinar datos foráneos con más precisión.

Los resultados obtenidos por RLM con el algoritmo SGD son los siguientes: Para los datos de entrenamiento MSE = 0.1192, MAE = 0.2782, RMSE = 0.3452 y $r^2 = 0.9996$ y para los datos de prueba MSE = 0.1103, MAE =

0.2998, RMSE = 0.3321 y $R^2 = 0.9999$. Debido a la resultante de errores mínimos los datos observados son muy similares a los de predicción, de allí que en la curva de tendencia se observan los valores de predicción con distancias mínimas o muy cercanas a los datos observados. La curva de tendencia de los datos observados y de predicción en el entrenamiento y en la prueba se muestran en las figuras 2 y 3.

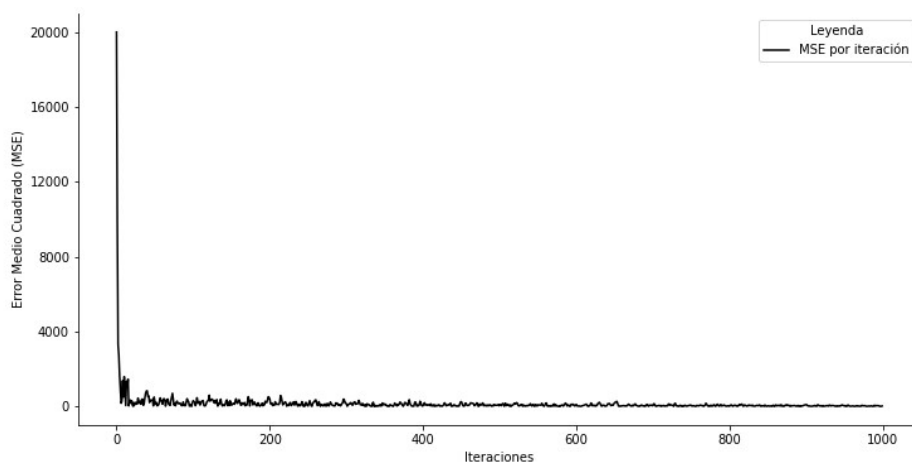


Figura 4. Características de convergencia MSE de rendimiento para el modelo de regresión lineal múltiple con el enfoque de gradiente descendiente estocástico

El MSE permite obtener la disminución de error de estimación por cada iteración y la correcta actualización de los pesos en la gradiente. Se calculan estos datos con el propósito de evaluar la convergencia del modelo. En otras palabras, el punto máximo en el cuál no se puede generalizar más la predicción de datos y pueda entregar las predicciones más acertadas y reales. Los errores se grafican en la Figura 4 por actualización de pesos en cada iteración.

CONCLUSIONES

- Existe una relación inversa y significativa entre el Factor Soft (FS) y las variables diámetro medio de fibra (DMF), diámetro medio de médula (DMM), porcentaje de modulación por volumen (MV) y porcentaje de fibras objetables (FO); y una correlación directa y fuerte con el factor de confort (FC), por lo que tienen el mismo efecto sobre la calidad de fibras.
- La técnica de regresión lineal múltiple (RLM) con el enfoque estocástico (SGD) satisface la predicción del FS, ya

que se ajusta a las variables seleccionadas como DMF, DMM, MV, FC y FO. El modelo RLM optimizado por SGD se adapta al tamaño muestral y a las variables mencionadas con gran precisión.

- Los datos de la predicción mostraron una gran robustez y con bajos valores en el error medio absoluto (MAE), el error medio cuadrado (MSE) y la raíz cuadrada del error medio cuadrado (RMSE) y un valor R^2 aceptable.
- Los datos mostrados por el equipo OFDA validaron la confiabilidad del modelo, de modo que este modelamiento de inteligencia artificial sea apto para la evaluación de fibras y ser un atributo computacional agregado en la industria textil.

Agradecimientos

Se agradece el apoyo a la Universidad Católica de Santa María de Arequipa (UCSM) y su Vicerrectorado de Investigación y al instituto de Investigación e Innovación en Energías Renovables y Medio Ambiente (INNOVERGY), para el desarrollo del proyecto de investigación

LITERATURA CITADA

1. **Allain D, Renieri C. 2010.** Genetics of fibre production and fleece characteristics in small ruminants, Angora rabbit and South American camelids. *Animal* 4: 1472-1481. doi: 10.1017/S175173-1110000029
2. **Amor N, Noman MT, Petru M, Sebastian N. 2022.** Comfort evaluation of ZnO coated fabrics by artificial neural network assisted with golden eagle optimizer model. *Sci Rep* 12: 6350. doi: 10.1038/s41598-022-10406-6
3. **Aylan-Parker J, McGregor BA. 2002.** Optimising sampling techniques and estimating sampling variance of fleece quality attributes in alpacas. *Small Ruminant Res* 44: 53-64. doi: 10.1016/S0921-4488(02)00038-X
4. **Berolatti G, Ruiz L, Cabrera F, Aliaga J, Quispe D, Quispe E. 2021.** Evaluación de la medulación de fibras de lanas y fibras especiales de algunas especies de animales. *Rev Inv Vet Perú* 32: e17639. doi: 10.15381/rirep.v32i5.17639
5. **Bonilla MQ, Serrano-Arriezu L, Trigo JD, Bonilla CQ, Gutiérrez AP, Peña EQ. 2022.** Application of artificial intelligence and digital images analysis to automatically determine the percentage of fiber medullation in alpaca fleece samples. *Small Ruminant Res* 213: 106724. doi: 10.1016/J.SMALLRUMRES.2022.106724
6. **Brimms MA, Peterson AD, Gherardi SG. 1999.** Introducing the OFDA2000 - For rapid measurement of diameter profile on greasy wool staples. IWTO, Raw wool group report RWG04, Florence, Italy.
7. **Bustanza V. 2001.** La alpaca, conocimiento del gran potencial andino. En: *La Alpaca*. Puno, Perú: Universidad Nacional del Altiplano. p 113-126..
8. **Carpio F. 2017.** La cadena de valor para optimizar la producción de fibra de alpaca en la empresa SAIS Sollocota Ltda. N° 5 - Perú. *Comuni@cción* 8: 125-136.
9. **Cordero A, Contreras J, Mayhua P, Jurado M, Castrejon M. 2011.** Correlación fenotípica entre características productivas en alpaca Huacaya. *Rev Inv Vet Perú* 22: 15-21.
10. **Cruz A, Morante R, Gutiérrez JP, Torres R, Burgos A, Cervantes I. 2019.** Genetic parameters for medullated fiber and its relationship with other productive traits in alpacas. *Animal* 13: 1358-1364. doi: 10.1017/S1751731118003282
11. **Fong SJ, Li G, Dey N, Gonzalez-Crespo R, Herrera-Viedma E. 2020.** Finding an accurate early forecasting model from small dataset: a case of 2019-nCoV novel coronavirus outbreak. *Int J Interact Multimed Artif Intell* 6: 132. doi: 10.9781/ijimai.2020.02.002
12. **Frank E. 2008.** Camélidos Sudamericanos. Producción de fibra, bases físicas e genéticas. *Rev Argen Prod Anim* 28: 119-122.
13. **Frank E, Hick M, Neyra L. 2017.** The Problem of prickling on fabrics of South American camelids fibers: possible approaches for mechanical solutions. *Asian Res J Agric* 5: 1-9. doi: 10.9734/arja/2017/34692.
14. **Frank EN, Hick MVH, Castillo M, Prieto A, Adot O. 2014.** Fiber-based components determining handle and skin comfort in fabrics made from dehaired and non dehaired llama fiber. *Int J Appl Sci Technol* 4: 51-66.
15. **Frank EN, Hick MVH, Gauna CD, Lamas HE, Renieri C, Antonini M. 2006.** Phenotypic and genetic description of fibre traits in South American domestic camelids (llamas and alpacas). *Small Ruminant Res* 61: 113-129. doi: 10.1016/J.SMALLRUMRES.2005.07.003
16. **González H, León C, Rosadio R, García W, Gavidia C. 2008.** Evaluación de un método numérico de medición del diámetro de la fibra de alpaca. *Rev Inv Vet Perú* 19: 1-8. doi: 10.15381/rirep.v19i1.641
17. **Gutiérrez L. 2011.** Las normas técnicas peruanas y su impacto en el desarrollo de los camélidos peruanos. Lima, Perú: Ministerio de Agricultura.

18. **Guillen RIC, Garate ACC, Anco M, Ramos LMM, Pacheco HGJ. 2023.** Optimización de la curva de teñido de la de fibra de alpaca (*Vicugna pacos*) empleando alcohol graso etoxilado. *Veritas* 23: 65-70. doi: 10.35286/veritas.-v23i1.338
19. **Hunter L, Smuts S, Botha AF. 2013.** Characterizing visually objectionable and nonobjectionable medullated fibers in mohair. *J Nat Fibers* 10: 112-135. doi: 10.1080/15440478.2013.763483
20. **Infoalpaca. 2013.** Perfil de mercado y competitividad exportadora de prendas de alpaca. [Internet]. Disponible en: http://infoalpacas.com.pe/wp-content/uploads/2013/04/Tejido_Prendas_de_Alpaca.pdf
21. **IWTO. 1998.** IWTO-57 Determination of medullated fibre content of wool and mohair samples by opacity measurements using an OFDA. International Wool Textile Organization, Brussels. Brussels.
22. **IWTO. 2017a. IWTO-57-200.** Determination of medullated fibre content of wool and mohair samples by opacity measurements using an OFDA. International Wool Textile Organization, Brussels. Brussels.
23. **IWTO. 2017b. IWTO-8-2011.** Method of determining fibre diameter distribution parameters and percentage of medullated fibres in wool and other animal fibres by the projection microscope. International Wool Textile Organization, Brussels.
24. **Krizhevsky A, Sutskever I, Hinton G. 2012.** ImageNet classification with deep convolutional neural networks. En: Proc. XXV Annual Conference on Neural Information Processing Systems. Spain.
25. **Kulkarni P, Chellam S. 2010.** Disinfection by-product formation following chlorination of drinking water: artificial neural network models and changes in speciation with treatment. *Sci Total Environ* 408: 4202-4210. doi: 10.1016/J.SCITOTENV.2010.05.040
26. **Lupton CJ, McColl A, Stobart RH. 2006.** Fiber characteristics of the Huacaya Alpaca. *Small Ruminant Res* 64: 211-224. doi: 10.1016/J.SMALLRUMRES.2005.04.023.
27. **Machaca V, Bustinza AV, Arizapana FAC, Paucara V, Quispe EE, Machaca R. 2017.** Fiber characteristics of Huacaya alpaca at Cotaruse, Apurímac, Peru. *Rev Inv Vet Perú* 28: 843-851. doi: 10.15381/rivep.v28i4.13889
28. **Mamani RH, Frank EN, Prieto A, Castillo MF, Condori-Rojas N, Hick MVH. 2022.** Effect of fibre diameter, prickly factor and coarse fibre bias on yarn surface hairiness in South American camelids (SAC) fibre. *Fibers* 10: 18. doi: 10.3390/fib10020018
29. **McGregor B. 1997.** The quality of fibre grown by Australian alpacas. In: Seminar of shaping the future. Proc International Alpaca Industry 1997. Melbourne, Australia.
30. **McGregor B. 2006.** Production attributes and relative value of alpaca fleeces in southern Australia and implications for industry development. *Small Ruminant Res* 61: 93-111. doi: 10.1016/j.smallrumres.2005.07.001
31. **McGregor BA, Butler KL. 2004.** Sources of variation in fibre diameter attributes of Australian alpacas and implications for fleece evaluation and animal selection. *Aust J Agric Res* 55: 433-442. doi: 10.1071/AR03073
32. **Merrick NC. 1998.** Measurement of density and medullation in wool. Thesis Master. New Zealand: Lincoln University. 87 p.
33. **[MIDAGRI] Ministerio de Desarrollo Agrario y Riego. 2021.** Análisis de mercado 2016-2020. Tops de alpaca, hilados y prendas de vestir. Lima, Perú: MIDAGRI. 100 p.
34. **Moore KE, Blache D, Maloney SK. 2011.** Fibre diameter and insulation in alpacas: the biophysical implications. *Small Ruminant Res* 96: 165-172. doi: 10.1016/J.SMALLRUMRES.2011.-01.012

35. **Naebe M, McGregor B, Dowling M, Tester D. 2018.** Prickle discomfort assessment of commercial knitted wool garments. *Int J Cloth Sci Tech* 30: 73-81. doi: 10.1108/IJCST-03-2017-0023
36. **Naylor GRS, Hansford K. 1999.** Fibre end diameter properties in processed top relative to the staple for wool grown in a Mediterranean climate and shorn different seasons. *Wool Tech Sheep Bree* 47: 107-117.
37. **Onal L, Korkmaz M, Tutak M. 2007.** Relations between the characteristics of Angora rabbit fiber. *Fiber Polym* 8: 198-204. doi: 10.1007/BF0287
38. **[ONUDI] Organización de las Naciones Unidas para el Desarrollo Industrial. 2006.** Producción textil de fibras de camélidos sudamericanos en el área altoandina de Bolivia, Ecuador y Perú. UNIDO. 48 p.
39. **Peleato NM, Legge RL, Andrews RC. 2018.** Neural networks for dimensionality reduction of fluorescence spectra and prediction of drinking water disinfection by-products. *Water Res* 136: 84-94. doi: 10.1016/j.watres.2018.02.052
40. **Pillaud-Vivien L, Rudi A, Bach F. 2018.** Statistical optimality of stochastic gradient descent on hard learning problems through multiple passes. *Adv Neural Inf Process Syst* 2018: 8114-8124. doi: 10.48550/arXiv.1805.10074
41. **Pinares R, Gutiérrez GA, Cruz A, Morante R, Cervantes I, Burgos A, Gutiérrez JP. 2018.** Heritability of individual fiber medullation in Peruvian alpacas. *Small Ruminant Res* 165: 93-100. doi: 10.1016/j.smallrumres.-2018.04.007
42. **Professor EF. 2021.** The prickle effect comes from fabrics made of South American camelid (alpaca and lama) fibers. Mechanical and/or genetic solutions. *Eur J Appl Sci* 9: 46-57. doi: 10.14738/aivp.93.10130
43. **Shishoo RL. 1995.** Importance of mechanical and physical properties of fabrics in the clothing manufacturing process. *Int J Clothing Sci Technol* 7: 35-42. doi: 10.1108/09556229510087137
44. **Tadesse MG, Loghin C, Dulgheriu I, Loghin E. 2021.** Comfort evaluation of wearable functional textiles. *Materials (Basel)* 14: 6466. doi: 10.3390/ma1-4216466
45. **Tao X. 2001.** Smart fibres, fabrics and clothing. Elsevier. 336 p.
46. **Uyak V, Toroz I, Meriç S. 2005.** Monitoring and modeling of trihalomethanes (THMs) for a water treatment plant in Istanbul. *Desalination* 176: 91-101. doi: 10.1016/J.DESAL.2004.10.023
47. **Wang H, Liu X, Wang X. 2005.** Internal structure and pigment granules in coloured Alpaca fibers. *Fibers Polym* 6: 263-268.
48. **Xing W, Liu Y, Deng N, Xin B, Wang W, Chen Y. 2020.** Automatic identification of cashmere and wool fibers based on the morphological features analysis. *Micron* 28: 102768. doi: 10.1016/j.micron.2019.102768
49. **Zang L, Xin B, Deng N. 2021.** Identification of overlapped wool/cashmere fibers based on multi-focus image fusion and convolutional neural network. *J Nat Fibers* 19: 6715-6726. doi: 10.1080/15440478.2021.1932669
50. **Zhang T. 2004.** Solving large scale linear prediction problems using stochastic gradient descent algorithms. In: *Proc Twenty-First Int Conf Mach Learn ICML 2004*. doi: 10.1145/1015330.-1015332
51. **Zhong Y, Lu K, Tian J, Zhu H. 2016.** Wool/cashmere identification based on projection curves. *Text Res J* 87: 1730-1741. doi: 10.1177/0040517516658516