AEM JOURNAL.ORG

Role of Semantic Feature Mapping of Clothing Descriptions in Ming and Qing Novels in Improving the Air Permeability of Textile Materials

Jingzi Han
Department of Education, A ba Vocational College, Aba Tibetan and Qiang Autonomous Prefecture 623200, Sichuan, China
Corresponding author: Jingzi Han (e-mail: aki6110208@163.com)

ABSTRACT Bridging cultural semantics and engineering-oriented textile design remains a significant challenge in the development of intelligent functional materials. This study proposes a three-level mapping framework that quantitatively translates clothing descriptions in Ming and Qing novels into textile engineering parameters through semantic deconstruction, perceptual evaluation, and material characterization. Historical semantic clusters associated with lightweight, ventilation, and moisture management are extracted and linked to user perception and fabric properties using Word2Vec, Delphi analysis, partial least squares regression, and structural equation modeling. The resulting model achieves a prediction coefficient (R^2) of 0.81 for air permeability and establishes a quantitative interface between literary semantics and engineering indicators. Guided by the semantic concept of Summer Plain Yarn, a regenerated cellulose fabric reaches an air permeability of 215 mm/s, while a multifunctional single-layer fabric inspired by the concept of lightweight comfort achieves 192 mm/s together with a UPF above 50 and an antibacterial rate exceeding 98%. Beyond providing a culturally informed strategy for breathable textile development, the proposed semantic-to-parameter mapping framework offers a novel methodology for the design of functional materials in wearable electronic and electromagnetic textile systems, where optimized transport characteristics and human-centered performance are critical.

INDEX TERMS semantic feature mapping, functional textiles, air permeability, wearable electromagnetic textiles, material parameter optimization

I. INTRODUCTION

As functional textiles develop towards sensory design and cultural empowerment, how to transform users' cultural perceptions of wearing comfort into engineered material performance indicators has become a cutting-edge topic in interdisciplinary research between textile science and the humanities [1-3]. Current research focuses on physical testing or physiological indicator-driven breathability optimization, but generally ignores the rich clothing perception semantics contained in historical texts, resulting in a disconnect between material design and users' deep cultural expectations [4-6]. In particular, in the development of highly breathable fabrics, the lack of systematic decoding of traditional aesthetic and functional expressions such as lightness, transparency, and coolness has restricted the synergistic improvement of products in terms of cultural identity and comfort experience [7,8].

Existing research can be roughly divided into three categories:

The first category is the study of clothing oriented towards literature and history. For example, Wang L constructed a textual research method for clothing materials by cross-checking the text of Dream of the Red Chamber with the weaving archives of the Qing Dynasty [9]; Lu C systematically sorted out the design feature system of the shape, pattern and fabric of Jiangnan clothing in the late Ming Dynasty [10]. Although such research effectively reconstructs the cultural context and aesthetic logic of historical clothing, it remains largely qualitative and does not establish a quantitative mapping between perceptual semantics (e.g., lightness and transparency) and engineering parameters such as porosity and air permeability, thereby limiting its integration into closed-loop design frameworks for modern functional textiles.

The second category focuses on mapping modern semantics with engineering performance, but ignores the particularity of historical context and cultural semantics. Nawaz Z et al. used Long Short-Term Memory/Convolutional Neural Network (LSTM/CNN) to identify sentiment polarity from e-commerce platform reviews [11]; Kodžoman D et al. established an association between visual texture and semantic adjectives using the Kansei scale [12]; Li X et al. constructed a physical prediction model for fabric coolness based on multivariate regression [13]. Although these methods have demonstrated effectiveness in modern contexts, they either rely on structured short review data, are limited to predefined adjectives, or fail to capture the culturally embedded metaphorical expressions found in classical texts, such as "sweating does not stick to the body" and "coolness through the body". In addition, although the Outfit Generative Adversarial Network (OutfitGAN) clothing generation framework proposed by Zhou D et al. applied semantic masks to improve the rationality of matching [14], its semantics are derived from modern fashion labels, and do not involve functional descriptions in historical texts, nor do they establish any mapping with physical properties such as breathability; Xing B's dual-scene feel evaluation study revealed cognitive bias [15], but it relied on static semantic scales and expert subjective judgments, and did not establish a quantitative mechanism from perceptual descriptions to fabric physical parameters, and could not support material innovation driven by cultural semantics.

The third category attempts to digitize traditional culture, such as Yuan H et al.'s construction of a knowledge graph of ancient clothing to support intelligent question answering [16]. Although such work enhances the retrievability of cultural knowledge, it relies on expert-annotation static graphs and does not provide a dynamically quantitative mapping between literary semantics and physical properties such as fabric breathability. Although a few linguistic studies focus on the word formation rules or phrase semantics of clothing terms (such as [17,18]), their analysis is limited to the language structure itself and does not establish any engineering connection with material functions or user perception.

In summary, existing research is either limited to historical research without parameter interfaces, or rooted in modern contexts but lacking cultural depth, or remaining static knowledge without dynamic translation. Consequently, a comprehensive mapping path from classical literary semantics to user perception and further to textile engineering parameters remains underexplored. To address this gap, this study proposes a three-level mapping model that connects literary semantics with user perception and material parameters. Through a closed-loop integration of semantic deconstruction, sensory verification, and physical testing, this study establishes a quantitative mapping between culturally embedded semantics (e.g., light and thin, transparent, sweat-resistant) and engineering metrics such as air permeability, thereby contributing a methodological framework for culturally informed functional textile design.

II. AN INTERDISCIPLINARY METHOD FRAMEWORK BASED ON SEMANTIC DECONSTRUCTION AND MULTIDIMENSIONAL MAPPING

A. CONSTRUCTION AND PREPROCESSING OF THE COSTUME CORPUS OF MING AND QING NOVELS

Ten Ming and Qing dynasty novels featuring extensive social descriptions and detailed costumes are selected as the original text sources: The Dream of the Red Chamber, The Golden Lotus, The Story of Awakening the World, The Scholars, The Heroes of the Children, Strange Stories from a Chinese Studio, Stories of the World, Strange Stories from an Erke Paian (Erke Paian Jingqi), The Casual Commentaries of Taowu, and The Flowers of Shanghai. All texts are converted to plain text using UTF-8 (Unicode Transformation Format – 8-bit) encoding using authoritative collated editions, with page numbers, annotations, and non-text content removed.

The corpus was pre-filtered to focus exclusively on passages containing garment descriptions and seasonal settings. Only sentences referencing clothing items, wearing scenarios, or seasonal contexts were retained for semantic analysis. This pre-filtering ensured that all subsequent semantic clustering was strictly relevant to breathability-related apparel contexts, avoiding noise from unrelated narrative sections.

Based on research findings from the history of clothing, three categories of keywords are identified: material-based keywords (e.g., gauze, silk, rabdosia, and hemp), perceptual keywords (e.g., light, thin, transparent, refreshing, cool, and sweat-resistant), and seasonal/scene-based keywords (e.g., summer months, summer clothes, and wearing a single layer in hot weather). This resulted in an initial keyword list of 87 items. Regular expressions are developed using the Python re module to match complete sentences containing at least one keyword, expanding the context window to 50 characters before and after to preserve semantic integrity. The initial extracted results are manually verified item by item, removing false matches and duplicate fragments, ultimately retaining over 2,000 valid clothing descriptions. Each record is structurally annotated using predefined fields, including: novel title, chapter number, character identity (scholar, lady, servant, monk, etc.), season (spring, summer, autumn, winter, unknown), clothing type (top, skirt, pants, underwear, outerwear, etc.), material description, color, pattern, and perceptual semantic keywords. The annotation work is independently completed by two researchers with backgrounds in ancient Chinese clothing history. Third-party experts are applied to adjudicate on items with disagreements. The final Kappa consistency coefficient is 0.86, indicating that the annotation results had high reliability [19,20].

The specific statistical information of each novel in terms of the number of entries, the proportion of summer clothing and high-frequency material words is shown in Table 1.

B. EXTRACTION AND QUANTIZATION CODING OF SEMANTIC FEATURES RELATED TO AIR PERMEABILITY

Based on documents such as General History of Chinese Textiles and Research on Ming and Qing Clothing, a semantic label system for thermal and moisture comfort is con-

TABLE 1. Structured Corpus Statistics

Novel Title	Valid Entries	Summer Costume (%)	Top 3 Material Terms
Dream of the Red Chamber	312	68.3	sha, luo, chou
The Golden Lotus	287	72.1	ge, sha, ma
Marriages to Awaken the World	245	63.7	sha, ge, si
The Scholars	198	60.1	luo, sha, bu
The Tale of Heroic Sons and Daughters	186	66.7	sha, luo, ge
Strange Tales from a Chinese Studio	172	58.1	sha, ma, chou
Models for Life	165	64.2	ge, sha, luo
Amazing Tales: Second Series	158	61.4	sha, si, ma
The Idle Criticism of Taowu	142	59.2	ge, bu, sha
The Sing-song girls of Shanghai	183	70.5	sha, luo, si
Total	2048	64.4	—

Note: sha = gauze, luo = leno/ro gauze, ge = ramie fabric, ma = hemp, chou = silk satin, bu = cloth, si = silk.

structed, and breathability-related descriptions are divided into three functional labels: light and thin type (emphasizing the lightness of the fabric and the sparse structure, such as "thin as a cicada's wing" and "light as a weight"), ventilated type (emphasizing the sense of air circulation, such as "transparent body" and "wind penetrates the texture"), and moisture-absorbing type (emphasizing sweat management, such as "sweat does not stick to the body" and "dry and non-sticky").

All clothing descriptions in the corpus are segmented into Chinese words (using the Jieba word segmentation tool), and stop words and function words are filtered out, retaining semantically significant content words and phrases. Subsequently, the Term Frequency–Inverse Document Frequency (TF-IDF) algorithm is used to calculate the weight of each term in the corpus [21-23]. Its calculation formula is as follows:

$$\text{TF-IDF}(t,d) = tf(t,d) \cdot \log\left(\frac{N}{df(t)}\right) \tag{1}$$

t is a term, d is a document (i.e., a sentence describing clothing), N is the total number of documents in the corpus, $tf(t,d)$ is the frequency of term t in document d, and $df(t)$ is the number of documents containing term t. Based on this formula, terms with a TF-IDF value greater than 0.35 are selected as candidate feature words.

In addition, word embeddings are trained on the segmented corpus using the Word2Vec model [24,25] (Skip-gram architecture). Cosine similarity clustering (threshold $\geq$ 0.7) is used to identify semantically similar word clusters, forming semantic clusters such as {tou, touti, shengliang}, {qing, bao, chanyi, wuwu}, and {huxibuzhan, shuang, ganshuang}.

To address polysemy, we implemented context-aware disambiguation by analyzing the 50-character context window surrounding each keyword. The functional labeling process considered the immediate lexical environment to assign the most probable meaning. For synonymy handling, the Word2Vec cosine similarity clustering grouped semantically equivalent terms into unified clusters. Experts then verified these clusters during the Delphi process, ensuring that synonymous expressions received consistent quantitative weights regardless of their surface form.

To quantify semantic strength, five experts (two clothing historians, two textile comfort researchers, and one expert in Kansei engineering) conducted three rounds of Delphi consultation. They are asked to rate the perceived strength of breathability for each semantic item on a scale of 0–1 (1 being the strongest). Scores were iteratively adjusted after each round until consensus was achieved (coefficient of variation < 0.15). The mean score for each semantic item is used as its quantitative weight, calculated as follows:

$$w_i = \frac{1}{K}\sum_{k=1}^{K} s_{ik} \tag{2}$$

where w_i is the quantized weight of the i-th semantic term, s_{ik} is the score assigned by the k-th expert, and $K = 5$ is the total number of experts. The resulting structured semantic feature vector set contains the semantic word, its functional type, Word2Vec cluster, and quantized perceptual strength value. Table 2 shows an excerpt of representative results.

C. PERCEPTUAL EVALUATION EXPERIMENT: VERIFICATION OF THE BRIDGE BETWEEN SEMANTIC VECTORS AND USER PERCEPTION

The experiment uses the 12 high-frequency breathability semantic terms (e.g., "sweat-free" and "light as a cicada's wing") generated in Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability" as sensory stimuli. Each term carries a corresponding functional label (lightweight/ventilated/moisture-absorbing), a Word2Vec semantic cluster, and a Delphi quantitative strength value (ranging from 0 to 1). Twenty fabric samples with significantly different physical properties (covering fibers such as cotton and linen, and structures such as plain weave, gauze, and mesh) are simultaneously selected as perceptual objects.

The experiment employs a semantically guided perceptual rating paradigm: Each round, participants are presented with a semantic term (e.g., "wind-through silk clothing") and all 20 fabric samples. Participants are asked to rate the degree to which each fabric fits the semantic term based on the overall

TABLE 2. Semantic feature quantification results

Cultural Semantic Term (Original Chinese + Pinyin)	Functional Type	Word2Vec Cluster (Top Neighbors)	Perceived Breathability Intensity (Mean ± SD)
Hàn bù zhān shēn	Moisture-wicking	{shuǎng, gān shuǎng, bù nián}	0.92 ± 0.03
Qīng bó rú chán yì	Lightweight	{báo, wú wù, qīng yíng}	0.89 ± 0.05
Tòu tǐ shēng liáng	Ventilation	{tòu, fēng, liáng}	0.87 ± 0.04
Shǔ yuè zhe gé	Moisture-wicking	{gé, liáng, xià yī}	0.87 ± 0.04
Fēng tòu luó yī	Ventilation	{fēng, tòu, luó}	0.83 ± 0.06
Qīng ruò wú wù	Lightweight	{qīng, báo, wú gǎn}	0.88 ± 0.05
Dān yī dù xià	Lightweight	{dān, xià, báo}	0.85 ± 0.05
Luó bó tòu fū	Ventilation	{luó, tòu, báo}	0.86 ± 0.04
Gě bù shēng liáng	Moisture-wicking	{gé, liáng, shuǎng}	0.86 ± 0.04
Yán tiān zhe shā	Lightweight	{shā, xià, báo}	0.84 ± 0.05
Tòu fēng bù rè	Ventilation	{tòu, fēng, liáng}	0.85 ± 0.05
Gān shuǎng bù nián	Moisture-wicking	{shuǎng, gān, bù zhān}	0.90 ± 0.03

Note: For clarity, semantic terms are represented in pinyin. Clustering reflects the top co-occurrence term (cosine similarity $\geq$ 0.7). The intensity score comes from the Delphi consensus (0-1 scale) of 5 experts.

wearing image evoked by the semantic term on a 7-point Likert scale (1 = not at all fitting, 7 = very fitting) [26,27]. The experiment is conducted in a standard temperature and humidity environment (25±1 °C, 60±5% RH). Thirty healthy subjects (half male and half female, aged 20–35 years, without color vision or tactile impairments) are recruited and provided written informed consent. A double-blind design with Latin square randomization is used to control for experimenter bias and sequence effects.

After the original scores were normalized by Z-score, the average perception score was calculated by aggregation of semantic items, and a consistency test is performed with the quantitative intensity value of the corresponding semantics (Pearson $r = 0.81$, $p < 0.01$). Furthermore, the average perception score is subjected to Spearman rank correlation analysis with the physical parameters of the fabric (porosity, surface density, thickness, and air permeability) to construct a semantic-perception-parameter correlation matrix. This matrix serves as the core input of the user perception layer in the three-level mapping model, directly bridging the abstract expression of literary semantics and the engineering indicators of material performance, providing an empirical basis for subsequent Partial Least Squares Regression (PLSR) and Structural Equation Modeling (SEM) modeling [28]. The Spearman rank correlation coefficient is defined as follows:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{3}$$

where n is the number of samples (here $n = 20$), and d_i is the difference in rank between the i-th sample in two variables (e.g., "cooling through the body" score and porosity).

D. CONSTRUCTION OF THE THREE-LEVEL MAPPING MODEL OF LITERARY SEMANTICS—USER PERCEPTION—MATERIAL PARAMETERS

The breathability perception intensity value in the semantic feature vector is used as the literary semantic layer variable, the average Likert rating of each fabric under the corresponding semantics in Section "Perceptual Evaluation Experiment: Verification of the Bridge Between Semantic Vectors and User Perception" is used as the user perception layer variable, and the fabric air permeability, porosity, surface density, and thickness measured in Section "Fabric Sample Preparation and Physical Properties" Testing are used as the material parameter layer variables. Min-Max normalization [29,30] is performed on these three variables to eliminate dimensional differences. The mathematical expression of Min-Max normalization is as follows:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

where x is the original variable value, x_{min} and x_{max} are the minimum and maximum values of the variable in all samples, respectively, and x' is the standardized value, ranging from [0, 1].

On this basis, PLSR is used to establish a prediction model for the user perception layer to the material parameter layer [31,32]. The number of latent variables was determined via 10-fold cross-validation. PLSR extracts latent variables while maximizing the covariance between the independent and dependent variables. It is suitable for modeling small sample, high-dimensional, and collinear data. Its basic form is:

$$Y = XB + E \tag{5}$$

$X \in R^{n\times p}$ represents the user perception rating matrix ($n = 20$ samples, $p = 5$ perceptual dimensions), $Y \in R^{n\times q}$ represents the material parameter matrix ($q = 4$, including air permeability, porosity, surface density, and thickness), B represents the regression coefficient matrix, and E represents the residual matrix. The model uses air permeability as the primary response variable, and the remaining physical parameters as auxiliary response variables. The number of latent variables is determined through 10-fold cross-validation, with a retention criterion of $Q = 0.9$.

In addition, a structural equation model (SEM) [33-35] is constructed, in which literary semantics is specified as the exogenous latent variable, user perception as the mediating latent variable, and air permeability as the endogenous observed variable. Path relationships are established based on theoretical hypotheses. Maximum likelihood estimation is used for model fitting. Evaluation metrics included Chi-square to degrees of freedom ratio (χ^2/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA).

Figure 1 illustrates the data flow and method modules of the entire modeling process, presenting a three-level mapping path from historical text to engineering parameters: the literary semantic layer on the left, which achieves semantic quantification based on corpora and Delphi scoring; the user perception layer in the middle, which establishes semantic-fabric matching through sensory experiments; and the material parameter layer on the right, which provides engineering validation based on physical testing. PLSR and SEM together form the mapping core, achieving a closed-loop transformation from cultural expression to performance indicators.

Finally, the PLSR regression coefficients and SEM path coefficients are jointly encoded to form a callable mapping function: given any literary semantic word and its perceptual intensity, the model automatically outputs a recommended air permeability target range and the corresponding porosity and surface density constraints. This mapping model is implemented using Python's scikit-learn and lavaan (R interface) and encapsulated as a parametric design interface to support semantic-driven selection in subsequent fabric development.

III. EXPERIMENT

A. FABRIC SAMPLE PREPARATION AND PHYSICAL PROPERTIES TESTING

To cover the types of light, thin, and breathable fabrics frequently mentioned in Ming and Qing novels, 20 fabric samples are prepared, covering four types of fiber composition (cotton, linen, regenerated cellulose, and polyester) and five weave structures (plain weave, twill, gauze, openwork mesh, and honeycomb jacquard). The warp and weft densities range from 40–120 threads/inch, and the surface density is controlled at 40–150 g/m^2 to match the physical range corresponding to the semantics of light and transparent. All samples are conditioned under standard atmospheric conditions (20$\pm$2 °C, 65$\pm$2% RH) for 48 hours before testing. Air permeability is measured in accordance with the national standard GB/T 5453–2023, "Determination of Air Permeability of Textile Fabrics". Air flow per unit area is measured using an FX3300 air permeability meter at a pressure differential of 200 Pa (unit: mm/s). Five locations are tested on each sample and the average value is taken. Porosity is determined using optical microscopic image analysis: fabric surface images are captured using a digital microscope ($\times$50) and binarized using ImageJ software. The percentage of pore area to total area is calculated. Sample thickness (using a YG(B)141D digital fabric thickness gauge, pressure 2 kPa) and areal density (10 cm $\times$ 10 cm specimens are weighed using an electronic balance with a precision of 0.001 g) are simultaneously recorded. All physical parameter data are verified and stored in the fabric performance database as input variables for the material parameter layer described in Section "Construction of the Three-Level Mapping Model of Literary Semantics—User Perception—Material Parameters".

B. IMPLEMENTATION OF SENSORY EVALUATION EXPERIMENT

The experiment is conducted in a standard environmental chamber (25 $\pm$ 1 °C, 60 $\pm$ 5% RH). The experimental subjects are the 20 fabric samples prepared in Section "Fabric Sample Preparation and Physical Properties Testing". Each sample is cut into 15 cm $\times$ 15 cm pieces and the edges are overlocked to prevent tactile interference.

30 healthy subjects (15 men and 15 women, aged 20–35 years, with no history of skin allergies or sensory dysfunction) are recruited. Before the experiment, they signed an informed consent form and received brief training to familiarize themselves with the 7-point Likert scale. The evaluation dimensions included cool touch, breathability, lightness, dryness, and overall comfort, each corresponding to a semantic functional label defined in Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability" (e.g., lightness corresponds to thin and lightweight).

The experiment uses a double-blind design: subjects are unaware of the experimental objective and sample number, and the experimenter is not involved in guiding the scoring. The 20 fabrics are presented one by one in a random order. Before evaluating each dimension, the system randomly prompts one of the 12 core semantic terms from Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability" (e.g., "thin as a cicada's wing" and "cool through the body") and asks participants to rate the fabric based on this semantic association. All ratings are collected in real time via an electronic questionnaire system. After removing outliers using the Tukey method, the raw data were aggregated by semantic category to calculate the average perception score for each fabric across each semantic dimension, which served as the user perception layer variable.

C. MODEL INPUT DATA INTEGRATION

Outlier detection is performed on the physical parameters in Section "Fabric Sample Preparation and Physical Properties Testing" and the perceptual scores in Section "Implementation of Sensory Evaluation Experiment" (Tukey's rule, IQR (Interquartile Range) $\pm$1.5). Min-Max normalization is then performed on all variables, mapping the values to the [0, 1] interval to eliminate dimensional differences.

Based on this, three types of data are associated using fabric samples as indexes: the literary semantic layer: the quantized perceptual strength values of the 12 semantic terms output in Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability" (e.g.,

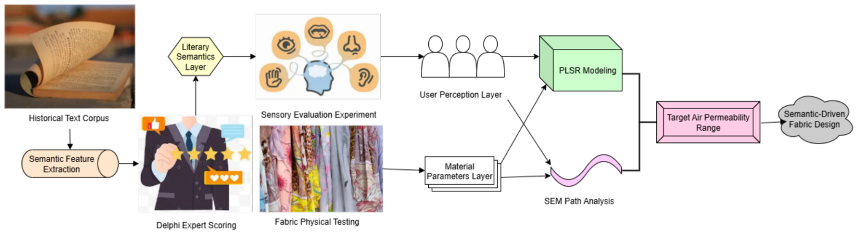

FIGURE 1. Framework of the three-level mapping model

transparent = 0.85); the user perception layer: the average Likert ratings of the corresponding semantic cues in Section "Implementation of Sensory Evaluation Experiment"; and the material parameter layer: the air permeability, porosity, surface density, and thickness measured in Section "Fabric Sample Preparation and Physical Properties Testing". The resulting structured dataset contained N = 20 samples and M = 18 feature dimensions (12 semantic strength items + 5 perceptual score items + 1 air permeability as the target variable). The data are divided into training set and validation set in a ratio of 7:3 for fitting and validating PLSR and SEM models.

IV. RESULT ANALYSIS

A. VERIFICATION OF THE CORRELATION BETWEEN SEMANTIC FEATURES AND BREATHABILITY OF MING AND QING DYNASTY CLOTHING

Based on the structured corpus constructed in Section "Construction and Preprocessing of the Costume Corpus of Ming and Qing Novels", a sliding window method (window width ± 3 sentences) is used to count the co-occurrence frequencies of semantic words with material words (such as si, luo, ge), seasonal words (such as xia, re), and structural descriptive words (such as dan, bao). Pointwise mutual information (PMI) is calculated to identify significant co-occurrence combinations (PMI > 1.5).

The word cloud on the left in Figure 2 highlights high-frequency breathability semantic terms such as light and thin, transparent, and sweat-free. The co-occurrence network on the right demonstrates strong correlations between semantic terms and material/structural terms (PMI > 1.5). For example, the nodes transparent and gauze, without lining, and light and thin and yarn, single clothing are densely connected, confirming the quantitative results of text co-occurrence analysis.

Based on the sensory experiment data from Section "Perceptual Evaluation Experiment: Verification of the Bridge Between Semantic Vectors and User Perception", the Spearman rank correlation coefficients (n = 20) are calculated between three representative semantic terms and four fabric physical properties. The probability of co-occurrence between semantic terms and material is also calculated. The results are shown in Table 3.

As can be seen, transparent and cool is strongly positively correlated with air permeability (r = 0.78**) and porosity (r = 0.72**), and significantly negatively correlated with thickness (r = -0.58**). "Thin as a cicada's wing" also shows an association with low surface density and high porosity. Furthermore, the results show that thin co-occurs with yarn, single clothing, and summer at 82.3%, 76.5%, and 89.1%, respectively. The transparent co-occurs with silk, without lining, and wind, at 78.6%, 71.2%, and 65.4%, respectively.

B. FITTING EFFECT AND PREDICTION ABILITY OF THE THREE-LEVEL MAPPING MODEL

Based on the training dataset compiled in Section "Model Input Data Integration", a prediction model for the relationship between user perception score and breathability is constructed using partial least squares regression. Five sensory ratings (coolness, breathability, lightness, dryness, and overall comfort) are used as independent variables, and the measured air permeability of the fabric (mm/s) is used as the dependent variable. A 10-fold cross-validation analysis determines the optimal number of latent variables to be three. The final model achieved an R^2 of 0.81, with a mean absolute error of 9.706 mm/s and a root mean square error of 11.88 mm/s, indicating good predictive performance for air permeability.

Figure 3 demonstrates the model's prediction accuracy. All data points are closely distributed around the ideal prediction line ($y = x$), with an R^2 of 0.81, further validating the user perception ratings' ability to effectively represent breathability performance.

To further verify the plausibility of the theoretical pathway linking literary semantics, user perception, and breathability, exploratory SEM was conducted. It should be noted that due to the limited sample size (N = 20), this analysis serves only as a hypothesis-generating exploration, not a rigorous model validation. Conventional SEM typically requires a sample size of at least 100 to ensure parameter estimation stability and statistical power.

The SEM used the Delphi-scaled intensity of the 12 semantic terms as the exogenous variable, the average Likert perception score of the corresponding semantic terms as the mediating variable, and the measured air permeability as the endogenous observed variable. The model fit indices and path coefficients are summarized in Table 4.

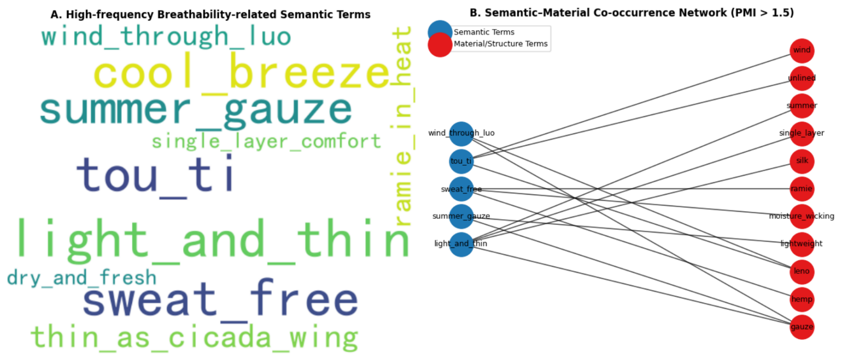

FIGURE 2. Semantic word cloud and co-occurrence network of Ming and Qing clothing

TABLE 3. Joint analysis of semantic-material co-occurrence rate and semantic-physical performance correlation (Spearman r)

Semantic Term	Co-occurring Term	Co-occurrence Rate (%)	Physical Parameter	Spearman r (n=20)
Qīng bó rú chán yì	gauze	82.3	Air Permeability	0.71**
	single-layer	76.5	Porosity	0.68**
	summer	89.1	Areal Density	−0.70**
Tòu tǐ shēng liáng	leno	78.6	Air Permeability	0.78**
	unlined	71.2	Porosity	0.72**
	wind	65.4	Thickness	−0.58**
Hàn bù zhān shēn	ramie	—	Air Permeability	0.63**
	hemp	—	Areal Density	−0.49*

Note: Co-occurrence data are from a corpus of 2048 Ming and Qing Dynasty novels (sliding window ±3 sentences, PMI > 1.5); Spearman correlation is based on sensory ratings and physical tests of 20 fabric samples; "—" indicates that the co-occurrence combination is not reported in Section "Verification of the Correlation Between Semantic Features and Breathability of Ming and Qing Dynasty Clothing"; ** $p < 0.01$, * $p < 0.05$.

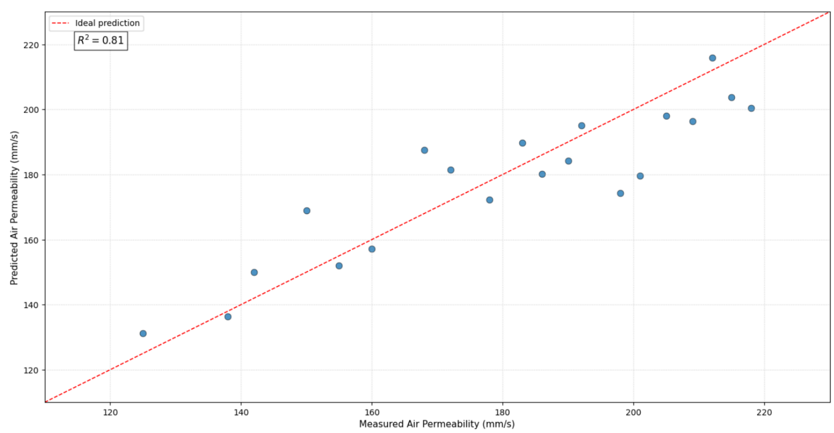

FIGURE 3. Scatter plot of PLSR model prediction vs. measured air permeability

TABLE 4. Exploratory SEM Results (N=20)

Component	Parameter	Estimate	p-value	Interpretation
Model Fit Indices	χ^2/df	1.83	—	< 3 = acceptable fit
	CFI	0.96	—	> 0.95 = good fit
	TLI	0.94	—	> 0.90 = acceptable
	RMSEA	0.062	—	< 0.08 = reasonable fit
Path Coefficients (Standardized)	Literary Semantics → User Perception	0.79	< 0.001	Strong positive effect
	User Perception → Air Permeability	0.68	< 0.01	Moderate positive effect
Indirect Effect (Point Estimate)	Literary Semantics → User Perception → Air Permeability	0.537	—	Product of path coefficients

Note: This SEM is conducted as an exploratory analysis due to the small sample size (N = 20), which is far below conventional requirements for confirmatory SEM (typically N $\geq$ 100). Fit indices and path coefficients are reported for hypothesis-generating purposes only. The indirect effect is a point estimate (0.79 × 0.68); no confidence interval is reported because Bootstrap resampling is not performed.

The results show that the model achieved an acceptable approximate fit (χ^2/df = 1.83, CFI = 0.96, TLI = 0.94, RMSEA = 0.062). Path analysis reveals that the standardized path coefficients for literary semantics → user perception and user perception → permeability are 0.79 and 0.68, respectively, with a point estimate of the indirect effect of $0.79 \times 0.68 = 0.537$.

C. SEMANTIC-DRIVEN BREATHABLE FABRIC DESIGN: TAKING SUMMER PLAIN YARN AS AN EXAMPLE

To validate the driving power of the constructed three-level mapping model in actual material development, this study selected Summer Plain Yarn as the semantic driver. This semantic is categorized as a composite label of lightweight and ventilated in Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability". Its Delphi method-quantified breathability perceived strength reached 0.89, constituting a clear perceptual design directive.

After inputting this semantic and its strength into the mapping model established in Section "Construction of the Three-Level Mapping Model of Literary Semantics—User Perception—Material Parameters", the system automatically outputs semantically driven target engineering parameter ranges: air permeability $\geq$ 200 mm/s, porosity > 16%, surface density < 75 g/m^2, and thickness < 0.25 mm.

Based on this semantically driven parameter constraint, a regenerated cellulose fabric imitating the solid yarn structure of the Ming and Qing dynasties is designed: 75D regenerated cellulose filaments are used, the warp and weft density is set to 68×52 threads/inch, and the weave structure used is a leno weave to form a stable through-hole structure.

As shown in Figure 4, after trial weaving on a sample loom, the fabric achieves an air permeability of 215 mm/s, a porosity of 17.8%, an areal density of 71.3 g/m^2, and a thickness of 0.23 mm. All indicators fall precisely within the model's recommended ranges. For the three key indicators of air permeability, areal density, and porosity, the measured values closely approached or exceeded the target lower limits.

D. SEMANTIC-DRIVEN COLLABORATIVE OPTIMIZATION OF BREATHABILITY OF MULTIFUNCTIONAL FABRICS: BASED ON THE CONCEPT OF SINGLE-LAYER MEANS COMFORT

This study incorporates the core concept of Ming and Qing dynasty clothing, single layer equals comfort, as a semantic driver. This semantic is categorized as lightweight in Section "Extraction and Quantization Coding of Semantic Features Related to Air Permeability". A Word2Vec cluster analysis revealed its semantic neighbor, "wind-through robe", as the literal strength. Therefore, its Delphi score of 0.83 is used as its quantitative strength. Its underlying meaning points to a philosophy of comfort characterized by simplified structure and unburdened wear, reflecting a contrast with the modern multifunctional fabric layering paradigm.

After inputting this semantic into the mapping model, the system outputs semantically driven collaborative design constraints: while maintaining necessary functionality, the surface density should be < 80 g/m^2 and the air permeability should be > 180 mm/s. The qualitative single-layer comfort concept was mapped to three quantifiable Level 3 parameters: areal density below 80 g/m^2, air permeability exceeding 180 mm/s, and thickness constrained by the structural requirement of a single-layer weave. These parameter thresholds were derived from the PLSR model's regression coefficients linking the semantic intensity of 0.83 to the material performance space. Based on this, a semantically driven reconstruction of a commercially available double-layer coated UV-resistant/antibacterial polyester fabric (initial air permeability 125 mm/s) is conducted. By eliminating the independent functional layers and adopting a co-spinning process, nano-ZnO (2.0 wt%) and a quaternary ammonium antimicrobial agent (1.5 wt%) are directly embedded into the regenerated cellulose spinning solution to produce a single-layer multifunctional fiber. This fiber is then woven into a plain weave structure (70 × 60 strands/inch). Test results show that the new fabric achieved an air permeability of 192 mm/s, while its surface density is reduced to 76 g/m^2. Furthermore, its UPF value reached 52.3 and its antibacterial rate reached 98.7%, significantly improving breathability without sacrificing functionality.

Figure 5 compares the performance of a traditional double-layer fabric with the single-layer fabric developed in this paper across five key metrics: air permeability, UPF, antimicrobial rate, surface density, and thickness. As can be seen, while maintaining a UPF > 50 and an antimicrobial rate > 98%, the new fabric significantly improves air permeability (125 → 192 mm/s), while simultaneously reducing surface density and thickness. This overall performance profile approaches the ideal of high functionality, high breathability, and lightweight.

V. CONCLUSION

This paper constructs a three-level mapping model from classical literary semantics–user sensory perception–textile engineering parameters, achieving a quantitative translation of historical clothing semantics into engineering indicators. Through combined PLSR and SEM modeling, it verifies a strong correlation between semantic matching and air permeability, and establishes a mapping interface that can be used in engineering applications. Case studies demonstrate that a gauze-like fabric developed based on Summer Plain Yarn achieves an air permeability of 215 mm/s; a multifunctional fabric reconstructed based on the single-layer comfort concept achieves an air permeability of 192 mm/s with a UPF > 50 and an antibacterial rate > 98%, indicating the potential of semantic-driven approaches in functional textile design. The reliance on literary texts introduces potential limitations regarding corpus bias. The novels predominantly reflect the attire and perceptions of specific social strata, potentially underrepresenting commoner clothing. The sample size of 2,048 descriptions, while substantial, may not capture the full semantic spectrum of historical clothing. Future

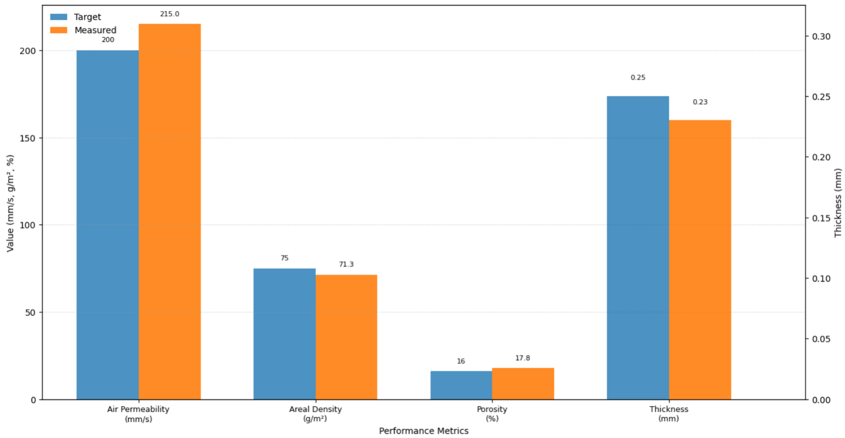

FIGURE 4. Performance comparison of imitation plain yarn fabrics

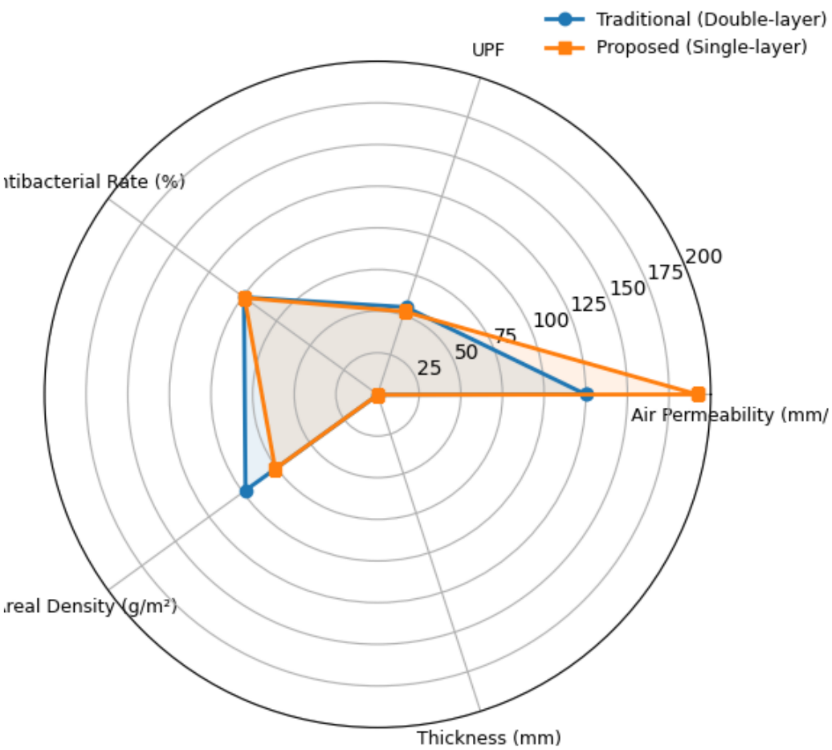

FIGURE 5. Performance radar chart of the multifunctional fabric under the single-layer comfort concept

work should incorporate archaeological textile data and expand the corpus to include more diverse socioeconomic sources to mitigate these biases.

AUTHOR CONTRIBUTIONS

Jingzi Han designed, collected and analyzed the data, and drafted the manuscript. Jingzi Han conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Jingzi Han participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] S. Li, S. Jiang, M. Tian, Y. Su, and J. Li, "Mapping the research status and dynamic frontiers of functional clothing: a review via bibliometric and knowledge visualization," International Journal of Clothing Science and Technology, vol. 34, no. 5, pp. 697-715, 2022, doi: 10.1108/IJCST-10-2021-0151.

[2] Y. Zhang, X. Xia, K. Ma, G. Xia, M. Wu, Y. H. Cheung, et al., "Functional textiles with smart properties: their fabrications and sustainable applications," Advanced functional materials, vol. 33, no. 33, Art. no. 2301607, 2023, doi: 10.1002/adfm.202301607.

[3] M. G. Tadesse, C. Loghin, I. Dulgheriu, and E. Loghin, "Comfort evaluation of wearable functional textiles," Materials, vol. 14, no. 21, pp. 6466, 2021, doi: 10.3390/ma14216466.

[4] X. LIU and Y. MA, "Metaphors on "Sleeve": A Cognitive Perspective on Chinese Characters and Traditional Clothing," Journal of Literature and Art Studies, vol. 15, no. 1, pp. 11-20, 2025, doi: 10.17265/2159-5836/2025.01.003.

[5] Q. K. Ding and H. E. Liang, "Digital restoration and reconstruction of heritage clothing: a review," Heritage Science, vol. 12, no. 1, pp. 225, 2024, doi: 10.1186/s40494-024-01349-4.

[6] Y. Gao, X. Feng, T. Zhang, E. Rigall, H. Zhou, L. Qi, et al., "Wallpaper texture generation and style transfer based on multi-label semantics," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1552-1563, 2021, doi: 10.1109/TCSVT.2021.3078560.

[7] X. Huang, S. Kettley, S. Lycouris, and Y. Yao, "Autobiographical design for emotional durability through digital transformable fashion and textiles," Sustainability, vol. 15, no. 5, pp. 4451, 2023, doi: 10.3390/su15054451.

[8] N. Chuprina, T. Remenieva, I. Frolov, and O. H. Tereshchenko, "Design of the contemporary garments on the basis of the transformation of stylistic and artistic-compositional characteristics of traditional decorative art," Art and design, vol. 3, no. 4, pp. 30-44, 2021, doi: 10.30857/2617-0272.2021.3.3.

[9] L. Wang, "Chinese silk fine art in early Qing dynasty observations from historical literature," New Techno Humanities, vol. 3, no. 1, pp. 36-40, 2023, doi: 10.1016/j.techum.2023.01.001.

[10] C. Lu and N. H. Ismail, "THE DESIGN CHARACTERISTIC OF LATE 15TH CENTURY MING DYNASTY COSTUMES IN THE JIANGNAN REGION," Jurnal Tuah, vol. 4, no. 1, pp. 179-188, 2023, [Online]. Available: https://www.unimel.edu.my/journal/index.php/JTuah/article/view/1581.

[11] Z. Nawaz, C. Zhao, F. Nawaz, A. A. Safeer, and W. Irshad, "Role of artificial neural networks techniques in development of market intelligence: a study of sentiment analysis of ewom of a women's clothing company," Journal of Theoretical and Applied Electronic Commerce Research, vol. 16, no. 5, pp. 1862-1876, 2021, doi: 10.3390/jtaer16050104.

[12] D. Kodžoman, A. Hladnik, A. P. Čuden, and V. Čok, "Assessment and semantic categorization of fabric visual texture preferences," AUTEX research journal, vol. 23, no. 2, pp. 279-291, 2023, doi: 10.2478/aut-2022-0006.

[13] X. Li, S. Wang, Y. Li, X. Jin, L. Ma, W. Tian, et al., "Evaluation model of fabric transient cooling sensation based on multiple stepwise regression analysis," Journal of Engineered Fibers and Fabrics, vol. 18, no. 4, Art. no. 15589250221144014, 2023, doi: 10.1177/15589250221144014.

[14] D. Zhou, H. Zhang, K. Yang, L. Liu, H. Yan, X. Xu, et al., "Learning to synthesize compatible fashion items using semantic alignment and collocation classification: An outfit generation framework," IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 4, pp. 5226-5240, 2022, doi: 10.1109/tnnls.2022.3202842.

[15] B. Xing, Q. Shao, X. Zeng, L. Koehl, and J. Wang, "Investigating the effects of cognitive scenarios between real fabric and fabric visual representations," Textile Research Journal, vol. 95, no. 13-14, pp. 1576-1588, 2025, doi: 10.1177/00405175241296522.

[16] H. Yuan, Y. Li, B. Wang, K. Liu, and J. Zhang, "Knowledge graph-based intelligent question answering system for ancient Chinese costume heritage," npj Heritage Science, vol. 13, no. 1, pp. 198, 2025, doi: 10.1038/s40494-025-01776-x.

[17] M. N. Shodmurodova, "DECODING THE STRUCTURAL AND SEMANTIC FEATURES OF PHRASEOLOGICAL UNITS IN THE DOMAIN OF CLOTHING IN ENGLISH," Western European Journal of Linguistics and Education, vol. 2, no. 5, pp. 80-84, 2024, [Online]. Available: https://westerneuropeanstudies.com/index.php/2.

[18] S. Jumanova, "EXPLORING THE STRUCTURAL FEATURES OF COMPOUND NOUNS REFERRING TO THE NAMES OF CLOTHES," Mental Enlightenment Scientific-Methodological Journal, vol. 5, no. 02, pp. 131-137, 2024, [Online]. Available: https://mentaljournal-jspu.uz/index.php/mesmj/article/view/417.

[19] M. Li, Q. Gao, and T. Yu, "Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters," BMC cancer, vol. 23, no. 1, pp. 799, 2023, doi: 10.1186/s12885-023-11325-z.

[20] B. Więckowska, K. B. Kubiak, P. Jóźwiak, W. Moryson, and B. Stawińska-Witoszyńska, "Cohen's kappa coefficient as a measure to assess classification improvement following the addition of a new marker to a regression model," International journal of environmental research and public health, vol. 19, no. 16, Art. no. 10213, 2022, doi: 10.3390/ijerph191610213.

[21] W. N. I. Al-Obaydy, H. A. Hashim, Y. A. Najm, and A. A. Jalal, "Document classification using term frequency-inverse document frequency and K-means clustering," Indonesian Journal of Electrical Engineering and Computer Science, vol. 27, no. 3, pp. 1517-1524, 2022, doi: 10.11591/ijeecs.v27.i3.pp1517-1524.

[22] M. T. Mohammed and O. F. Rashid, "Document retrieval using term term frequency inverse sentence frequency weighting scheme," Indones. J. Electr. Eng. Comput. Sci, vol. 31, no. 3, pp. 1478, 2023, doi: 10.11591/ijeecs.v31.i3.pp1478-1485.

[23] N. Kulkarni, R. Vaidya, and M. Bhate, "A comparative study of word embedding techniques to extract features from text," Turkish Journal of Computer and Mathematics Education, vol. 12, no. 12, pp. 3550-3557, 2021, doi: 10.17762/turcomat.v12i12.8101.

[24] S. J. Johnson, M. R. Murty, and I. Navakanth, "A detailed review on word embedding techniques with emphasis on word2vec," Multimedia Tools and Applications, vol. 83, no. 13, pp. 37979-38007, 2024, doi: 10.1007/s11042-023-17007-z.

[25] G. Di Gennaro, A. Buonanno, and F. A. N. Palmieri, "Considerations about learning Word2Vec," The Journal of Supercomputing, vol. 77, no. 11, pp. 12320-12335, 2021, doi: 10.1007/s11227-021-03743-2.

[26] B. Tanujaya, I. Prahmana R, and J. Mumu, "Likert scale in social sciences research: Problems and difficulties," FWU Journal of Social Sciences, vol. 16, no. 4, pp. 89-101, 2022, doi: 10.51709/19951272/Winter2022/7.

[27] I. Kusmaryono, D. Wijayanti, and H. R. Maharani, "Number of response options, reliability, validity, and potential bias in the use of the likert scale education and social science research: A literature review," International Journal of Educational Methodology, vol. 8, no. 4, pp. 625-637, 2022, doi: 10.12973/ijem.8.4.625.

[28] P. Schober, E. J. Mascha, and T. R. Vetter, "Statistics from A (agreement) to Z (z score): a guide to interpreting common measures of association, agreement, diagnostic accuracy, effect size, heterogeneity, and reliability in medical research," Anesthesia & Analgesia, vol. 133, no. 6, pp. 1633-1641, 2021, doi: 10.1213/ANE.0000000000005773.

[29] M. Mazziotta and A. Pareto, "Normalization methods for spatio-temporal analysis of environmental performance: Revisiting the Min–Max method," Environmetrics, vol. 33, no. 5, pp. e2730, 2022, doi: 10.1002/env.2730.

[30] M. Ali P J, "Investigating the Impact of min-max data normalization on the regression performance of K-nearest neighbor with different similarity measurements," ARO-The Scientific Journal of Koya University, vol. 10, no. 1, pp. 85-91, 2022, doi: 10.14500/aro.10955.

[31] W. Putra, "Problems, common beliefs and procedures on the use of partial least squares structural equation modeling in business research," South Asian Journal of Social Studies and Economics, vol. 14, no. 1, pp. 1-20, 2022, doi: 10.9734/SAJSSE/2022/v14i130367.

[32] M. Sarstedt, F. Hair J, M. Pick, and B. D. Liengaard, "Progress in partial least squares structural equation modeling use in marketing research in the last decade," Psychology & Marketing, vol. 39, no. 5, pp. 1035-1064, 2022, doi: 10.1002/mar.21640.

[33] W. Cheung G, D. Cooper-Thomas H, S. Lau R, and L. C. Wang, "Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations," Asia pacific journal of management, vol. 41, no. 2, pp. 745-783, 2024, doi: 10.1007/s10490-023-09871-y.

[34] L. J. Jobst, M. Bader, and M. Moshagen, "A tutorial on assessing statistical power and determining sample size for structural equation models," Psychological Methods, vol. 28, no. 1, pp. 207, 2023, doi: 10.1037/met0000423.

[35] F. Magno, F. Cassia, and C. M. Ringle, "A brief review of partial least squares structural equation modeling (PLS-SEM) use in quality management studies," The TQM Journal, vol. 36, no. 5, pp. 1242-1251, 2024, doi: 10.1108/TQM-06-2022-0197.