

## Predicting thematic trends in sustainable marketing literature using BERTopic and deep learning

### Sürdürülebilir pazarlama literatüründe tematik eğilimlerin BERTopic ve derin öğrenme ile tahmini

İbrahim Budak<sup>1</sup> 

#### Abstract

This study aims to map academic trends in sustainable marketing both retrospectively and to develop future projections. To this end, the study presents thematic trend forecasts extending to 2030 by creating a quantitative map of the sustainable marketing literature around five main themes. Only English articles, conference papers, book chapters, and related document types were selected using the query "sustainable" AND "marketing" in the Elsevier Scopus database; abstracts of 17,747 records covering the period 1980–2025 were analysed. After preprocessing, BERTopic was applied using paraphrase-multilingual-MiniLM-L12-v2 embeddings. In line with prior BERTopic practice, the outlier/noise cluster (Topic = -1), accounting for 4.4% of documents, was excluded, and the remaining 209 topics were consolidated into five higher-level themes based on content similarity. The resulting themes cluster around axes such as green marketing, sustainable branding, marine and food-focused sustainability, green supply chain marketing, and AI-supported sustainability applications, quantitatively revealing the current backbone structure of the sustainable marketing literature. Annual publication counts were calculated for each theme and converted into time series, which were then forecasted for the 2026–2030 period using LSTM and GRU models. During the test period, LSTM demonstrated clearly superior performance compared to GRU, with lower errors (MAE and RMSE) and higher R<sup>2</sup> values. The findings clearly show that themes such as sustainable branding, green supply chain marketing, and AI-supported sustainability applications will experience a strong upward trend in the coming years. In contrast, areas such as sustainable packaging marketing will follow a relatively balanced, mature course. By combining topic modelling with deep learning-based time series forecasting, the study offers a unique methodological contribution to sustainable marketing research. It produces concrete, foresight-based insights for both the academic agenda and brand strategies.

**Keywords:** Sustainable Marketing, BERTopic, LSTM, GRU, Research Trend Prediction

**Jel Codes:** C45, M31, Q56

#### Öz

Bu çalışma, sürdürülebilir pazarlamadaki akademik eğilimleri hem geriye dönük olarak haritalandırmayı hem de gelecekteki projeksiyonları geliştirmeyi amaçlamaktadır. Bu amaçla, çalışma beş ana tema etrafında sürdürülebilir pazarlama literatürünün nicel bir haritasını oluşturarak 2030 yılına kadar uzanan tematik eğilim tahminlerini sunmaktadır. Elsevier Scopus veritabanında "sürdürülebilir" ve "pazarlama" sorgusu kullanılarak yalnızca İngilizce makaleler, konferans bildirimleri, kitap bölümleri ve ilgili belge türleri seçilmiştir; 1980-2025 dönemini kapsayan 17.747 kayıtlı özetleri analiz edilmiştir. Ön işleme sonrasında, BERTopic, paraphrase-multilingual-MiniLM-L12-v2 gömülü öğeleri kullanarak uygulanmıştır. Önceki BERTopic uygulamasına uygun olarak, belgelerin %4,4'ünü oluşturan aykırı değer/gürültü kümesi (Konu = -1) hariç tutulmuş ve kalan 209 konu, içerik benzerliğine göre beş üst düzey tema altında birleştirilmiştir. Ortaya çıkan temalar, yeşil pazarlama, sürdürülebilir markalaşma, deniz ve gıda odaklı sürdürülebilirlik, yeşil tedarik zinciri pazarlaması ve AI destekli sürdürülebilirlik uygulamaları gibi eksenler etrafında kümelenecek, sürdürülebilir pazarlama literatürünün mevcut omurga yapısını nicel olarak ortaya koymaktadır. Her tema için yıllık yayın sayısı hesaplandı ve zaman serisine dönüştürüldü, ardından LSTM ve GRU modelleri kullanılarak 2026-2030 dönemi için tahmin edildi. Test döneminde, LSTM, daha düşük hata (MAE ve RMSE) ve daha yüksek R<sup>2</sup> değerleri ile GRU'ya kıyasla açıkça daha iyi/üstün performans gösterdi. Bulgular, sürdürülebilir markalaşma, yeşil tedarik zinciri pazarlaması ve yapay zekâ destekli sürdürülebilirlik uygulamaları gibi temaların önümüzdeki yıllarda güçlü bir yükseliş eğilimi göstereceğini, sürdürülebilir ambalaj pazarlaması gibi alanların ise nispeten dengeli ve olgun bir seyir izleyeceğini açıkça göstermektedir. Konu modellemesini derin öğrenme tabanlı zaman serisi tahminleriyle birleştiren bu çalışma, sürdürülebilir pazarlama araştırmalarına benzersiz bir metodolojik katkı sunmakta ve hem akademik gündem hem de marka stratejileri için somut, öngörüye dayalı içgörüler üretmektedir.

**Anahtar Kelimeler:** Sürdürülebilir Pazarlama, BERTopic, LSTM, GRU, Araştırma Trendi Tahmini

**JEL Kodları:** C45, M31, Q56

<sup>1</sup> Dr., Kastamonu University, Kastamonu, Türkiye, [ibudak@kastamonu.edu.tr](mailto:ibudak@kastamonu.edu.tr)

ORCID: 0000-0001-7762-6114

Submitted: 31/12/2025

Revised: 20/02/2026

Accepted: 1/03/2026

Online Published: 25/03/2026

**Citation:** Budak, İ., Predicting thematic trends in sustainable marketing literature using BERTopic and Deep Learning, *bmij* (2026) 14 (1):113-128, doi:

<https://doi.org/10.15295/bmij.v14i1.2688>

## Introduction

Sustainability-focused approaches are rapidly moving to the centre of the marketing discipline amid the climate crisis, depletion of natural resources, and increasing stakeholder pressure. Sustainable marketing is defined as the holistic integration of environmental, social, and economic goals into marketing strategies and the consideration of long-term social welfare in businesses' value-creation processes, thereby distinguishing it from the classical understanding of marketing. Recent reviews and systematic studies reveal that sustainable marketing has become a rapidly developing subfield in the business literature, intersecting with many subfields such as consumer behaviour, brand management, digital marketing, and corporate sustainability (Madhavaram & Nirjar, 2025; White et al., 2025; Widiastuti et al., 2024).

As a natural consequence of this interest, the number of academic studies published under the heading of sustainable marketing has increased significantly over the past decade. Studies based on bibliometric mapping and citation analysis show that closely related concepts such as sustainable marketing, green marketing, sustainable consumption, and "marketing sustainability" have produced increasingly diverse themes across different journals and disciplines (İlmalhaq et al., 2024; Sono, 2023). While these studies reveal the leading countries, journals, research networks, and dominant keyword clusters in the field's development, they also report that themes such as the circular economy, digital sustainable marketing practices, stakeholder-based value creation, and consumer welfare have gained strength over time (Tian et al., 2025). Therefore, the current literature shows that sustainable marketing has enriched both theoretically and practically, and that the research field has become increasingly fragmented yet dynamic.

However, despite these comprehensive mapping efforts of the field's past and present state, quantitative and analytical studies on which themes will come to the fore in the future are limited. Current bibliometric and topic modelling studies are mostly descriptive, revealing past publication performance and current research clusters; they do not produce systematic predictions about which research axes will gain momentum in the coming years and which will lose relative importance. In a rapidly changing field such as sustainable marketing, there is a need for studies that statistically model the evolution of themes in the literature over time and provide trend forecasts for the future, enabling decision-makers, researchers, and policymakers to engage in strategic planning.

At this point, text mining and topic modelling techniques offer powerful tools for extracting latent themes from large text collections and tracking their evolution over time. Text mining, through the stages of preprocessing, vectorisation, and clustering, enables the transformation of numerous abstracts and article texts into structured information, revealing the developmental dynamics of research fields (Kokoç et al., 2024). Topic modelling, particularly through probabilistic approaches such as LDA, is widely used to uncover research trends and thematic clusters in sustainability and marketing literature (Jung & Kim, 2023; Yip et al., 2025). Neural network-based topic modelling methods developed in recent years, such as BERTopic, produce more consistent and interpretable topic representations using transformer-based embeddings, density-based clustering, and class-based TF-IDF (c-TF-IDF) approaches, yielding successful results on text sets across different domains. BERTopic is widely used across various application areas, including customer complaints, healthcare applications, and idea management, thereby providing a suitable framework for analysing large-scale texts in management and marketing literature (Cheddak et al., 2024).

On the other hand, for predicting time-dependent trends, deep learning-based time series models offer significant advantages. Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) are widely used in many fields such as financial markets, energy demand, macroeconomic indicators, and epidemiological processes because they offer higher prediction accuracy compared to classical statistical methods in series with non-linear and complex dependency structures (Naufal & Wibowo, 2023; Torres et al., 2021). However, studies integrating subject modelling outputs with LSTM/GRU-based time-series models to model the thematic evolution of academic production in sustainable marketing and forecast future trends remain quite limited.

This study aims to fill this gap and develop an empirical forecasting framework for the future directions of the sustainable marketing literature. In this context, first, abstracts of articles on sustainable marketing from the Scopus database are collected to create a textual database spanning a broad period of the field. Second, meaningful and interpretable themes are extracted from these abstracts using the BERTopic algorithm, revealing the thematic structure of the sustainable marketing literature. Third, the publication frequency of each theme is calculated over the years, analysing the rise and fall of these themes. Finally, the annual frequency series of the selected prominent themes is fed into LSTM and

GRU models to generate quantitative predictions of academic trends in sustainable marketing for the coming years. Thus, the study aims to shed light on the current state of the sustainable marketing literature and to predict future research themes by offering a holistic approach that integrates text mining, modern topic modelling, and deep learning-based time-series forecasting.

## **Literature review**

Sustainable marketing is a holistic approach that requires businesses to design their marketing activities not solely around profit maximisation, but within the framework of the triple bottom line, encompassing people, planet, and profit. This perspective addresses reducing environmental impacts, considering social justice and stakeholder welfare, and creating long-term economic value across the processes of product and service development, pricing, distribution, and communication. Recent textbooks and conceptual studies emphasise that sustainable marketing has been shaped around components such as sustainable product design, responsible communication, circular business models, and stakeholder-based value creation, reinterpreting the classic marketing mix (Trojanowski, 2022).

The growing strength of this conceptual framework is also reflected in the number of bibliometric studies conducted in the field. Analyses conducted on closely related topics such as green marketing, sustainable marketing mix, and sustainable business performance mostly use citation analyses, co-word analyses, and network visualisations to identify leading journals, influential authors, countries, and collaboration networks (Wang et al., 2023). However, a significant portion of these studies has limitations, including the use of keywords in different contexts, overly general word-based clusters, and the presentation of theme evolution over time solely through static maps. These limitations of co-word and classical science mapping approaches indicate that they may be insufficient for capturing deeper semantic structures and dynamic trends (Hafizoglu & Tuzlukaya, 2023).

In this context, text mining and topic modelling methods have increasingly been used in academic trend analysis. Early studies primarily utilised probabilistic/linear models, such as Latent Dirichlet Allocation (LDA), Non-negative Matrix Factorisation (NMF), and related approaches, to uncover hidden themes in large text collections. These models statistically decompose the term-document matrix, enabling each document to be represented as a mixture of specific topics. They are widely preferred in bibliometric applications aimed at mapping the thematic structure of research fields (Hankar et al., 2025). However, it has been emphasised that they are limited in capturing conceptual nuances and emerging sub-themes over time, as they consider word order and context to a limited extent and often rely on a static period.

The BERTopic approach, developed in recent years, is considered an important step towards overcoming these limitations in topic modelling. BERTopic converts texts into dense vectors using transformer-based sentence embeddings, identifies topics by grouping these embeddings with density-based clustering algorithms, and produces the distinctive keywords of each cluster using class-based TF-IDF (c-TF-IDF) (Grootendorst, 2022; Li & Hu, 2025). This architecture enables the extraction of topic clusters that are clearly distinct from one another and exhibit higher semantic consistency than those produced by classical LDA/NMF models. Indeed, recent applications across different disciplines have demonstrated that BERTopic can reveal topic structures consistent with predefined themes in qualitative datasets, enable detailed visualisation of the thematic structure and main research axes of research areas, and serve as an effective tool for examining changing research trends over time, particularly in technological fields (Samsir et al., 2023).

The need for research trends to not only describe the current situation but also to generate predictions for the future has brought time-series-based approaches to the forefront of bibliometric studies. However, a review of the literature reveals that most science mapping and big data-based bibliometric analyses have been limited to static analyses of publication counts, citation relationships, and keyword networks; studies that systematically model publication performance series or topic frequencies and make forward-looking predictions remain relatively scarce (Monroy & Diaz, 2018). In parallel, some studies that approach research trends indirectly have developed predictive models for various socio-economic variables using search volume indicators such as Google Trends; it has been reported that Google Trends data can improve the performance of traditional time series models in predictions related to tourism demand, unemployment rates, or the course of epidemics (Jun et al., 2018). However, most of these studies focus not on the sustainable marketing literature itself, but rather on economic or social indicators; examples of systematically predicting academic topic trends remain limited.

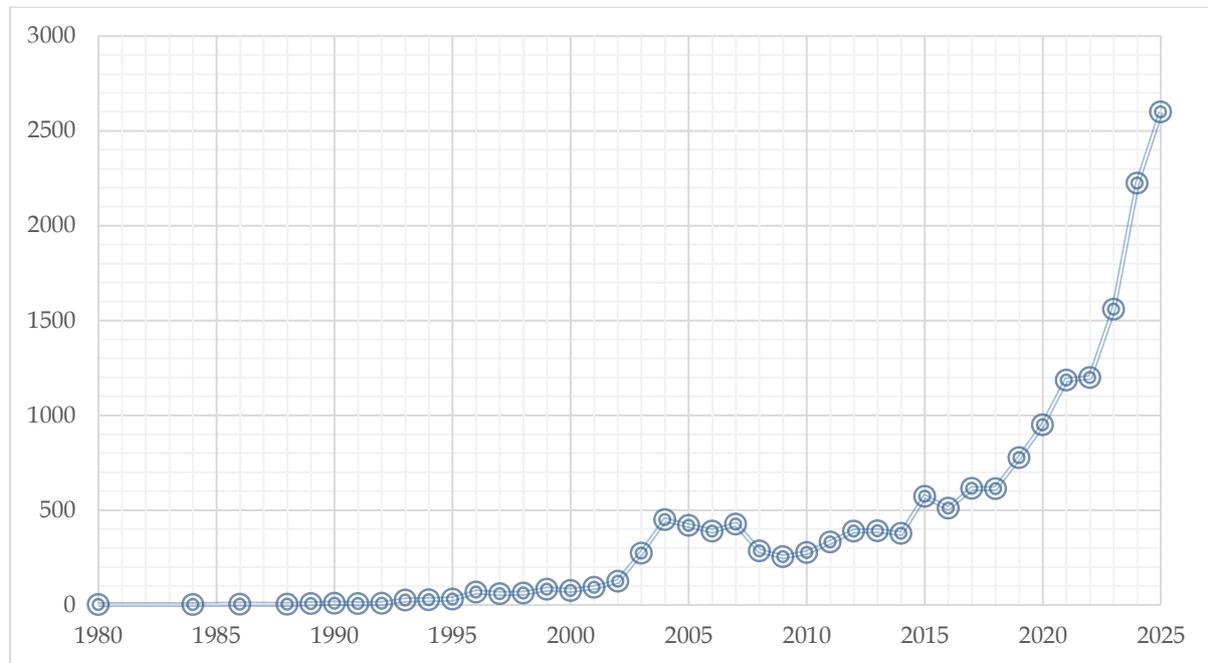
Deep learning methods in time series forecasting stand out for their flexibility in handling data with complex, non-linear dependencies. Long Short-Term Memory (LSTM) networks overcome the chronic "forgetting" problem of recurrent neural networks through gate mechanisms, enabling them to learn long-term dependencies; while gated recurrent units (GRU) implement a similar gate structure with

fewer parameters, offering lighter models that achieve similar accuracy levels in most cases. Recent comparative studies show that LSTM and GRU models achieve higher predictive performance than traditional statistical methods across financial markets, energy demand, climate indicators, and related fields (Arokodare et al., 2024; Yunita et al., 2025). Furthermore, it has been reported that LSTM-based models can successfully predict, for example, user evaluation dynamics of mobile banking applications or migration movements using Google Trends data in text-based or user behaviour-based numerical series (Golenvaux et al., 2020; Kiliç, 2025). These findings indicate that annual series derived from topic frequencies in the academic field can similarly be predicted using deep learning models, adding a predictive dimension to bibliometric trend analysis.

## Data and methodology

### Data collection

The dataset used in this study was obtained from Elsevier's Scopus database to capture the sustainable marketing literature as comprehensively as possible. The search query was structured to cover the title, abstract, and keyword fields. It used the phrase "sustainable" AND "marketing": ( TITLE-ABS-KEY ('sustainable') AND TITLE-ABS-KEY ("marketing") ) AND ( EXCLUDE ( PUBYEAR , 2026 ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) OR LIMIT-TO ( DOCTYPE , "cp" ) OR LIMIT-TO ( DOCTYPE , "ch" ) OR LIMIT-TO ( DOCTYPE , "re" ) OR LIMIT-TO ( DOCTYPE , "bk" ) OR LIMIT-TO ( DOCTYPE , "cr" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) ). Thus, only articles, conference papers, book chapters, and complementary review articles, books, and conference reviews published in English were included in the study. This query yielded 17,747 records: 10,319 articles, 2,991 conference papers, 2,275 book chapters, 1,492 reviews, 482 books, and 188 conference evaluations.



**Figure 1:** Publication Trends on Sustainability and Marketing (1980–2025)

**Source:** Produced by the author.

Looking at the distribution by year in Figure 1, it can be seen that the first record in the data set dates back to 1980, publication numbers remained in the single digits during the 1980–1990 period, and a clear upward trend emerged in the 2000s. For example, the number of records, which was 76 in 2000, rose to 276 in 2010, 572 in 2015, 949 in 2020, and 2,224 in 2024; in 2025, it has already risen to 2,599. In this study, the main analysis period was 2000–2025 to track the evolution of themes in the sustainable marketing literature. Still, early period records dating back to the 1980s were also retained in the descriptive statistics to show the historical beginning of the trend.

### Preprocessing

In the study, all abstracts obtained from Scopus were first processed through a standard text-mining pipeline to provide input to the BERTopic model. In this context, the abstracts were converted to lowercase, punctuation marks and numbers were removed, and noisy expressions such as URLs were removed, leaving only meaningful words. Next, English stopwords—high-frequency words such as "the," "and," and "of" that do not contain distinctive information—were compiled and removed from

the texts. At this stage, domain-specific words that occur very frequently but weaken thematic differentiation, such as "sustainable," "marketing," "study," and "paper," were added to an additional stopword list to ensure the model focused on contextual differences rather than these concepts.

Following stopword cleaning, the texts were subjected to lemmatisation; verbs, nouns, and adjectives were reduced to their basic dictionary forms (e.g., "practices," "practicing" → "practice"). This prevented different inflexions of the same concept from being represented as separate terms, thereby increasing conceptual density. Using frequency-based filtering, low-frequency words that fell below a certain threshold of the total number of documents (e.g., appearing in at least five documents or in a certain percentage of documents) were removed from the corpus; this reduced rare noise terms and prevented the topic model from operating in an overly sparse word space.

The cleaned summaries obtained after preprocessing were converted into numerical vectors using a sentence-transformers model to generate sentence-level embeddings. In this study, despite the sustainability marketing literature being largely in English, the "paraphrase-multilingual-MiniLM-L12-v2" model, which also allows for multilingual use, was preferred, and a fixed-size (384-dimensional) embedding was produced for each summary. The resulting embedding matrix served as the primary input for density-based clustering and class-based TF-IDF calculations within the BERTopic framework; thus, the preprocessing steps enabled both the acquisition of semantically consistent topic clusters and the execution of subsequent annual frequency and trend analyses on a more reliable basis.

Although the corpus analysed in this study is English-only, we selected paraphrase-multilingual-MiniLM-L12-v2 primarily for computational efficiency and scalability (384-dimensional embeddings) and to keep the pipeline extensible for potential multilingual expansions in future bibliometric updates. We acknowledge that English-specialised encoders such as all-mpnet-base-v2 (768 dimensions) may yield higher semantic resolution in strictly monolingual corpora; therefore, comparing embedding backbones constitutes a meaningful robustness extension for future work. Alternative embedding backbones were not systematically benchmarked in this submission; we therefore frame encoder comparison as a robustness extension for future work.

### **Topic extraction with BERTopic**

To perform topic extraction on Scopus abstracts using BERTopic, a model was first trained on the embeddings generated in the 'Preprocessing' section. For this purpose, the paraphrase-multilingual-MiniLM-L12-v2 embedding model was integrated into the BERTopic framework, the language parameter was set to 'multilingual' to align with the selected encoder and keep the pipeline extensible for potential multilingual updates, and the model was run on all abstracts with default clustering and c-TF-IDF settings. It's a standard pipeline that applies UMAP for dimensionality reduction and HDBSCAN for density-based clustering; we used these components without overriding the library's default configurations. The only explicitly set BERTopic arguments were the embedding backbone (paraphrase-multilingual-MiniLM-L12-v2) and (plus during model fitting). Thus, instead of fixing the number of topics in advance, it was determined by the data using BERTopic's density-based clustering; documents that are not confidently assigned to any coherent cluster are labelled as outliers/noise (Topic = -1) and excluded from downstream analyses, resulting in 209 distinct topics. No automated topic reduction (e.g., topic merging/reduction) was applied at this stage; the 209 topics reflect the model's raw clustering output after excluding outliers, and thematic consolidation was performed only at the higher-level (five-theme) aggregation step. In our case, the outlier share was 4.4%, well below 20%, suggesting that the default clustering configuration produced an acceptable topic structure for downstream trend analysis. We did not conduct an exhaustive hyperparameter sensitivity grid for BERTopic in this submission; however, the low outlier ratio (4.4%) suggests that the default clustering configuration yields a stable topic structure for downstream trend analysis. A systematic sensitivity analysis (e.g., varying minimum topic size and UMAP neighbourhood parameters) is left for future robustness work. In the analyses, as is common in the literature, the -1 cluster was excluded, and only meaningful topic clusters were evaluated (Koruyan, 2022).

The model outputs are organised in a summary table (topic info) containing the words and sample summaries that best represent each topic. In this table, each topic is represented by a list of approximately 10–15 words ranked according to c-TF-IDF scores; for example, words such as "green, marketing, environmental, products, consumers, purchasing" point to the theme of "green marketing and environmental products," while words such as "brand, authenticity, reputation, corporate, image, loyalty" highlight the axis of sustainable brand management and corporate reputation. The research team examined these word lists and representative summaries together and assigned meaningful and literature-consistent topic tags to each cluster. In this process, clusters with high conceptual proximity

were grouped into higher-level headings, yielding five broader themes – Green Marketing, Sustainable Branding, Marine and Food-Focused Sustainability, Green Supply Chain Marketing, and AI-Supported Sustainability Applications – from over 200 detailed topic clusters for discussion and trend analysis. Analyses throughout the study were deepened along these five themes. This hierarchical structure preserved the fine-grained thematic diversity of the sustainable marketing literature while enabling the results to be presented in a more readable format for management and policy contexts.

Each of the non-noise BERTopic clusters was assigned to exactly one of the five higher-level themes to preserve mutually exclusive annual time series (i.e., no multi-theme overlaps). The aggregation was based on (i) c-TF-IDF top keywords, (ii) representative documents, and (iii) conceptual proximity to established sustainable marketing sub-streams. Excluding the outlier/noise cluster (Topic = -1), a total of 208 topics were consolidated into five themes with the following topic-level distribution: Green Marketing (n = 37; 17.79%), Sustainable Branding (n = 35; 16.83%), Marine and Food-Focused Sustainability (n = 38; 18.27%), Green Supply Chain Marketing (n = 50; 24.04%), and AI-Supported Sustainability Applications (n = 48; 23.08%).

To connect topic extraction with forecasting, we convert document-level topic assignments into yearly topic-prevalence series. Specifically, for each year, we aggregate the number (or proportion) of documents assigned to each topic, yielding a multivariate time series that captures the temporal evolution of thematic attention. These yearly topic-prevalence trajectories are then modelled using LSTM/GRU to forecast future topic dynamics.

### Converting topic frequencies to time series

BERTopic outputs include both the topic tag associated with each abstract and the probabilities of related topics. In this study, the dominant topic (the topic tag with the highest probability) was first determined for each record, and this tag was used together with the publication year to create a year-topic frequency matrix. In other words, the number of publications published in each year and on each topic that carried that topic as the dominant theme was calculated; thus, a "raw frequency" matrix was obtained, with years in the rows and topic clusters defined by BERTopic in the columns. Then, to balance for the increase in total publication numbers over the years, the raw frequency values for each year-topic combination were normalised by dividing them by the total number of publications for that year, creating a second "ratio" matrix. These two matrices were used as complementary indicators reflecting the absolute production volume and relative thematic weight, respectively.

The obtained frequency and ratio series were converted into annual time series for each topic, covering the 2000–2025 period. These series for each topic were first visually examined, then evaluated for long-term trends and possible structural breaks using classical time series decomposition (trend + random component). Since the data are annual, no significant seasonality is expected; however, a strong, non-linear increase has been observed in many topics, especially after 2010. The time series obtained at this stage served as input for LSTM- and GRU-based prediction models discussed in the next section to produce future projections of thematic trends in sustainable marketing.

Before scaling, we assessed stationarity of the raw annual topic-frequency series using the Augmented Dickey-Fuller (ADF) test with automatic lag selection (AIC) (Dickey & Fuller, 1981). For the primary modeled series (Topic 0, annual publication counts, 2000–2025, n = 26), the ADF test indicated strong evidence of non-stationarity: under the constant specification, ADF = 2.615, p = 0.999 (lags = 9, nobs = 16), and under the constant-plus-trend specification, ADF = 1.542, p = 1.000 (lags = 9, nobs = 16). Consistent with the pronounced growth patterns observed across themes, similar non-stationarity was observed for other dominant topic-count series (Topics 1–4; p-values  $\geq 0.84$  under both specifications). Therefore, model evaluation relied on a strictly forward (time-preserving) train/test split to avoid information leakage, and performance metrics are interpreted as out-of-sample predictive accuracy on non-stationary annual counts rather than as variance explained by a stationary process.

### Forecast modeling

In this study, deep learning-based recurrent neural networks were used to predict thematic trends in sustainable marketing. First, the annual publication counts for each topic were scaled to the 0–1 range using Min–Max scaling, and model inputs were generated using a sliding window approach. Each sample consists of an input vector containing the scaled values from the past three years (window width = 3) and a target value representing the frequency for the following year. This captured a short but sequential context of the topic's annual development, establishing a structure consistent with studies emphasising that LSTM can model long-term dependencies with gate mechanisms and outperform classical methods, especially in non-linear time series (Mojtahedi et al., 2025; Omoware et al., 2023). In

practice, a single-layer LSTM layer with 32 hidden units and a "tanh" activation function was used, followed by a single-neuron output layer. The model was trained for up to 200 epochs using the "adam" optimisation algorithm and the mean squared error (MSE) loss function; an early stopping mechanism monitoring validation loss was implemented to prevent overfitting.

Both LSTM and GRU were trained under identical settings using Adam (learning rate = 0.001), batch size = 4, and up to 200 epochs with a time-preserving validation split (validation\_split = 0.2) and early stopping (monitor = val\_loss, patience = 20, restore\_best\_weights = True); the GRU followed the standard library implementation with update/reset gates and no custom gate modifications.

Window length was selected by balancing temporal context with effective sample size in a short annual series. With 26 annual observations (2000–2025), increasing the look-back horizon substantially reduces the number of trainable sequences (e.g., from 19 sequences at  $w=3$  to 15 sequences at  $w=7$  when reserving the last three years for testing), which can increase estimation variance and overfitting risk in deep recurrent models. Therefore,  $w=3$  was used as the primary specification to preserve a workable number of training sequences while still providing a sequential context for year-to-year dynamics. To enhance transparency, we report the effective sample-size implications for alternative window sizes ( $w = 3, 5, 7$ ) in Table 1.

**Table 1:** Effective Number of Sequences under Alternative Window Sizes (Annual Series; Test Horizon = Last 3 Years)

Window size ( $w$ )	Total sequences ( $N-w$ )	Train sequences ( $N-w-3$ )	Test sequences
3	23	20	3
5	21	18	3
7	19	16	3

In parallel with LSTM, the Gated Recurrent Unit (GRU) architecture, which has fewer parameters but can similarly capture temporal dependencies, was also implemented. Various comparative studies report that the gate structure and update mechanism of the GRU provide comparable, and in some cases even better, prediction performance than the LSTM in many applications; it is trained faster, especially for short or medium-length series (Yunita et al., 2025). Therefore, in this study, the same input window (3 years), the same number of hidden units (32), and the same training settings were used for both LSTM and GRU, allowing for a fair comparison of the two models' performance. For both models, the training and testing split was performed by separating the last three years of the time series as the test set and using the previous years as the training set; this approach is based on the "forward-looking" evaluation principle, which ensures that future observations do not leak into the training process in time series forecasting (Svetunkov, 2026). Additionally, 20% of the training data was reserved as a validation set (using a time-preserving split), and the model's hyperparameters were indirectly adjusted via early stopping. Given the computational load and data loss of fully rounded-start (rolling origin) cross-validation schemes for short annual series, this structure offers a practical and literature-consistent compromise.

All model and training hyperparameters (including batch size, learning rate, early stopping, and validation protocol) are summarised in Table 2 for reproducibility.

**Table 2:** Model and Training Hyperparameters for LSTM and GRU

Item	Setting
Input scaling	Min-Max scaling to [0, 1]
Window size (primary)	3 years
Test split	Last 3 years held out as test set
Validation	20% of training sequences (time-preserving; last portion of training set)
LSTM architecture	1× LSTM layer, 32 hidden units, tanh; Dense(1) output
GRU architecture	1× GRU layer, 32 hidden units; Dense(1) output
Loss	Mean Squared Error (MSE)
Optimizer	Adam
Learning rate	0.001 (Keras Adam default)
Batch size	4
Max epochs	200
Early stopping	monitor=val_loss; patience=20; restore_best_weights=True

Model performance was evaluated using the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination ( $R^2$ ) metrics on observations from the test period; these metrics are widely used indicators in both deep learning-based time series studies and LSTM applications for forecasting research trends (Gür, 2024). In the final stage,

five-year forward forecasts were generated for each topic using the final LSTM and GRU models trained with the most recent three years of data; thus, the potential future publication dynamics of prominent themes in the sustainable marketing literature were quantitatively revealed.

To communicate forecast uncertainty, we constructed approximate 95% prediction intervals for the 2026–2030 point forecasts using the empirical error scale from the held-out test period ( $PI \approx \hat{y} \pm 1.96 \times RMSE_{test}$ ).

## Results

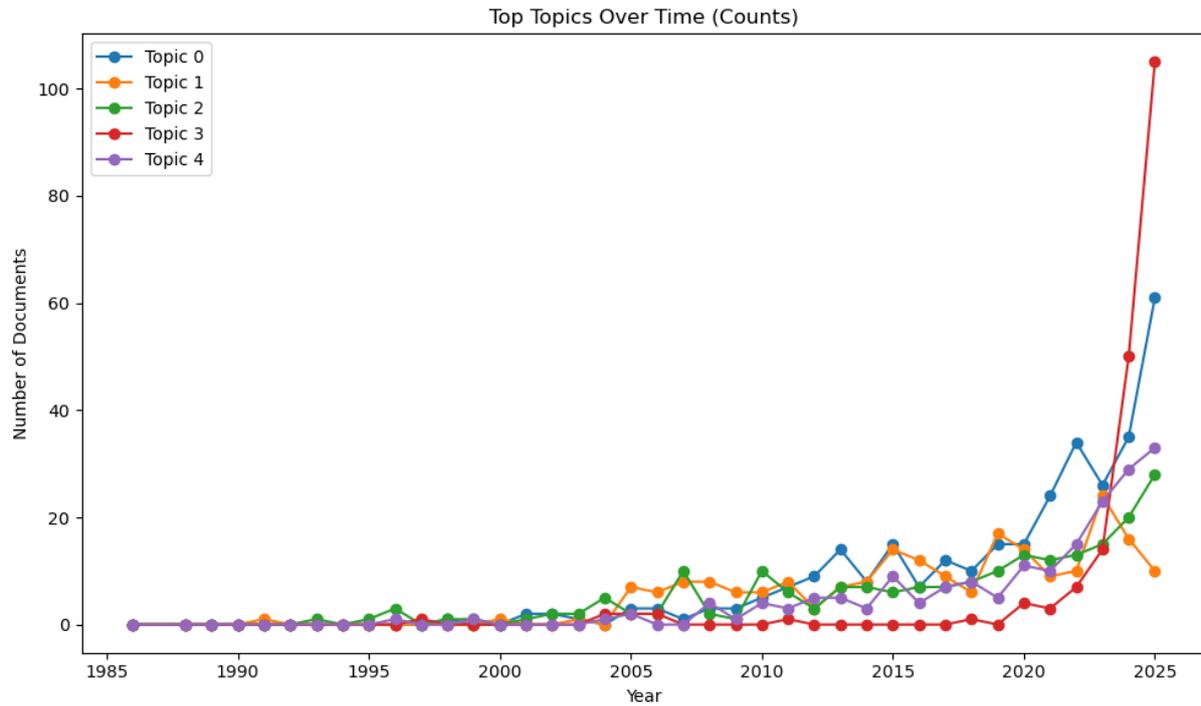
### BERTopic findings

Based on the BERTopic analysis, the top five topics emerging in the sustainable marketing literature were examined in detail, focusing on both content and their development over time. Considering word lists and representativeness, Topic 0 aligns with the classic green marketing and environmental product design axis, with terms such as "green," "eco-labelling," "eco-friendly," "carbon-neutrality," and "green-branding." Topics 1 and 2 are shaped around marine biodiversity, sustainable fishing, and ocean-based food supply chains, with overlapping word clusters such as "fish," "aquaculture," "marine species," "ocean-sustainability," "sustainable seafood," and "coastal management." Topic 3 concerns artificial intelligence and data-driven sustainability applications, including terms such as "AI," "predictive-modelling," "data-driven," "green-technologies," "eco-optimisation," and "impact-forecasting." Topic 4 highlights corporate sustainability, responsibility reporting, and governance themes, including "sustainability," "corporate-responsibility," "environmental-governance," "social-value," "sustainable-practices," "impact-assessment," and "climate-alignment."

**Table 3:** The Most Distinctive Keywords and c-TF-IDF Scores for the Top Five Topics Obtained with BERTopic

Topic 0		Topic 1		Topic 2		Topic 3		Topic 4	
Word	Score	Word	Score	Word	Score	Word	Score	Word	Score
green	0,032	fish	0,025	fish	0,025	ai	0,067	sustainability	0,016
eco-labeling	0,015	aquaculture	0,018	aquaculture	0,018	predictive-modeling	0,019	corporate-responsibility	0,014
environmental	0,014	marine-species	0,014	marine-species	0,014	data-driven	0,018	environmental-governance	0,013
sustainable-products	0,013	biodiversity	0,013	biodiversity	0,013	ethical-tech	0,017	social-value	0,012
ecological-impact	0,012	habitat-protection	0,013	habitat-protection	0,013	automation	0,016	sustainable-practices	0,012
eco-friendly	0,011	traceability	0,012	traceability	0,012	analytics	0,015	long-term-strategy	0,011
resource-efficiency	0,010	ocean-sustainability	0,011	ocean-sustainability	0,011	green-technologies	0,014	reporting	0,011
carbon-neutrality	0,010	sustainable-seafood	0,011	sustainable-seafood	0,011	eco-optimisation	0,014	circularity	0,010
lifecycle-design	0,009	coastal-management	0,010	coastal-management	0,010	digital-efficiency	0,013	impact-assessment	0,010
green-branding	0,009	renewable-stocks	0,009	renewable-stocks	0,009	impact-forecasting	0,012	climate-alignment	0,009

Table 3 shows the top ten keywords that best represent each topic and their corresponding c-TF-IDF scores. The scores indicate how distinctive the relevant word is within that topic cluster; higher scores indicate that the word is part of the topic's conceptual core. The concentration of Topic 0 around "green" and "eco-labelling" and Topic 3 around "AI (artificial intelligence)" and "predictive-modelling" reveals that traditional green marketing and the new wave based on data/technology coexist in the sustainable marketing literature.



**Figure 2:** Publication Counts by Year for the Five Most Dominant Bertopic Subjects in Sustainable Marketing Literature

Source: Produced by the author.

Figure 2 shows the annual document counts for the top five topics with the highest total frequency during the 1985–2025 period. The graph reveals that Topics 0 and 4 have exhibited a steady upward trend since the mid-2000s, becoming the "backbone themes" of sustainable marketing literature. In contrast, Topics 1 and 2, which focus on marine and food sustainability, became particularly prominent in the 2010s and have since followed a relatively flat trajectory.

**Deep learning modelling results**

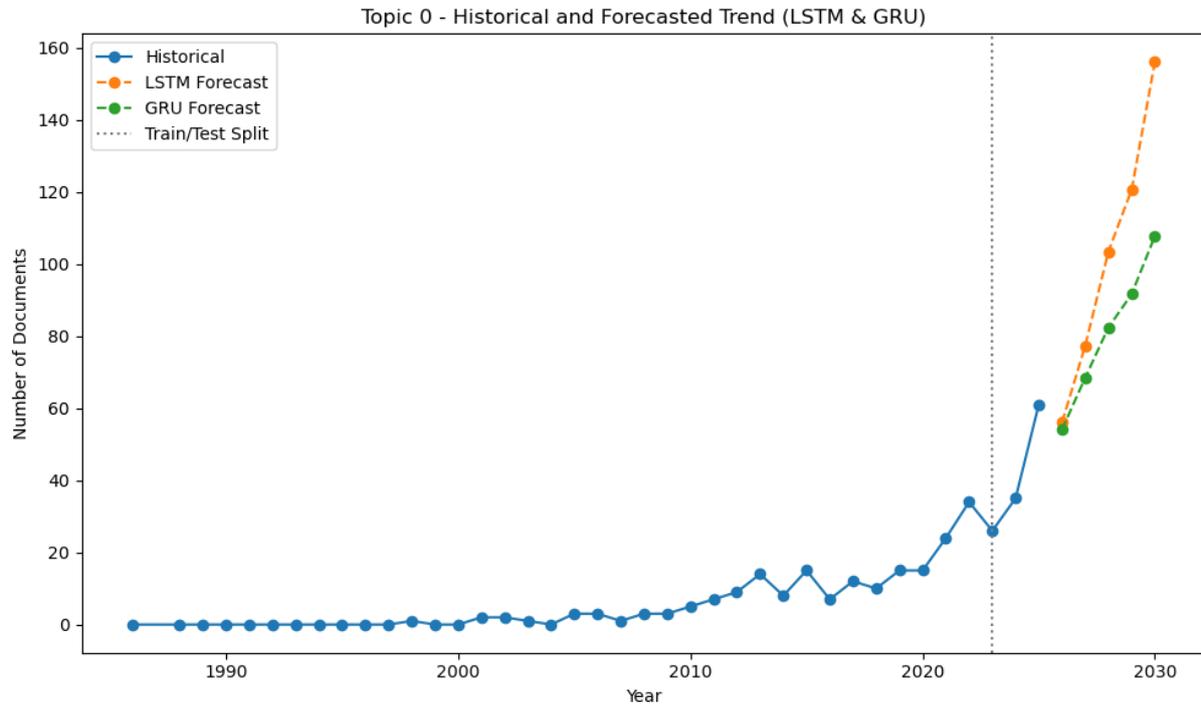
Deep learning modelling results indicate that LSTM and GRU models exhibit similar but not identical performance in predicting thematic trends in the sustainable marketing literature. The error metrics obtained for Topic 0 during the test period (2023–2025) are summarised in Table 4. Accordingly, the LSTM model yields better results in terms of absolute error, with lower MAE (8.64 vs 9.41) and RMSE (9.10 vs 11.47) values than the GRU. In terms of MAPE, the two models are quite close (LSTM 21.57%; GRU 21.51%); however, the coefficient of determination ( $R^2$ ) was 0.624 for LSTM and 0.402 for GRU, indicating that LSTM explained a larger portion of the variance in the test period.

**Table 4:** Comparison of LSTM and GRU Test Performance Metrics for Topic 0 (2023–2025)

Model	MAE	RMSE	MAPE (%)	$R^2$
LSTM	8,637	9,097	21,57	0,624
GRU	9,412	11,472	21,51	0,402

Table 4 shows the error metrics obtained during the test period on a selected topic (Topic 0). Lower MAE and RMSE values, and a higher  $R^2$  value, indicate that the LSTM model captures changes in annual publication numbers better than the GRU.

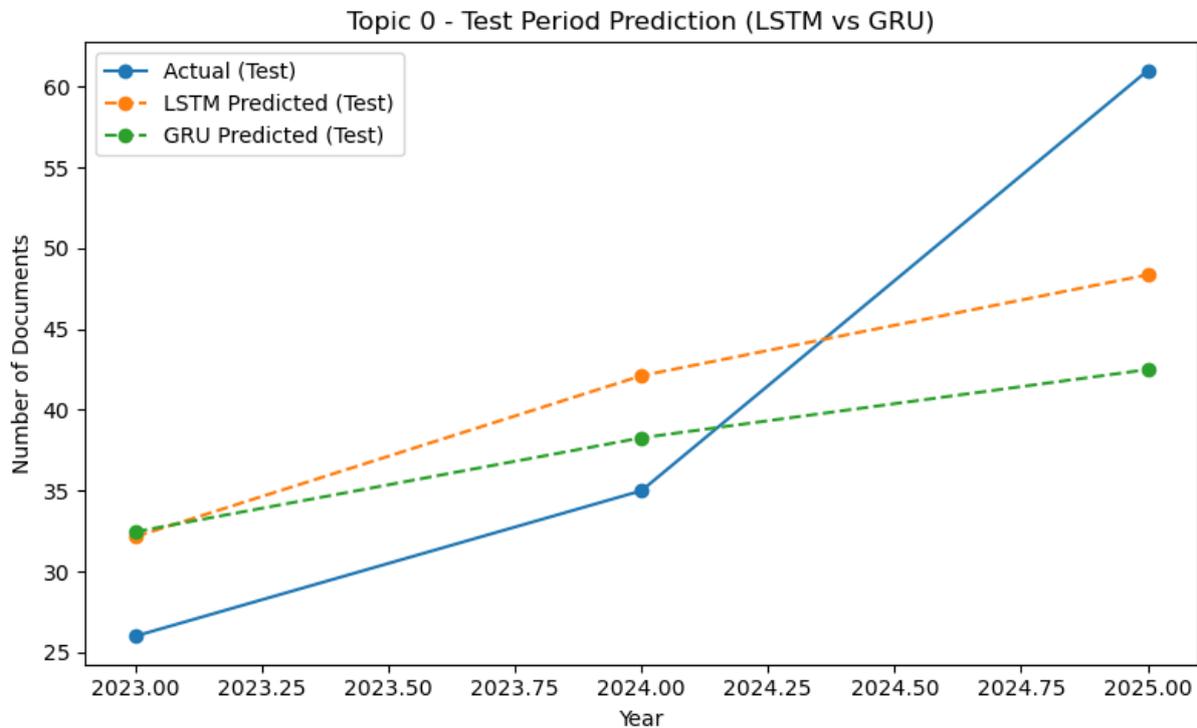
Using the held-out RMSE as an empirical error scale, the approximate 95% prediction-interval half-width for Topic 0 corresponds to about  $\pm(1.96 \times \text{RMSE})$ , i.e., approximately  $\pm 18$  publications for LSTM and  $\pm 22$  publications for GRU.



**Figure 3:** Historical and Projected Publication Trend for Topic 0 (LSTM and GRU Comparison)

Source: Produced by the author.

Figure 3 shows the historical publication counts for Topic 0 alongside predictions generated using LSTM and GRU models for the 2026–2030 period. The vertical dashed line indicates the training/test split, while the blue line up to this point represents historical data, and the subsequent orange (LSTM) and green (GRU) dashed lines represent forward projections. Both models project a continued increase in Topic 0 for 2026–2030; however, the LSTM projection shows a steeper slope than the GRU projection, particularly after 2028.



**Figure 4:** Comparison of Actual and Predicted Values for Topic 0 during the Test Period (2023–2025)

Source: Produced by the author.

Figure 4 compares the actual publication counts (blue line) with the LSTM (orange dashed) and GRU (green dashed) predictions throughout the test period. While both models correctly capture the overall upward trend, the LSTM produces forecasts that are closer to the actual values, particularly in 2024 and

2025. The GRU reflects the trend relatively more moderately and smooths out sudden increases somewhat; this is consistent with the higher RMSE and lower  $R^2$  values in Table 4.

The finding that the LSTM model is more successful is also reflected in the projections generated for different themes. For example, in the series labelled "Sustainable Branding," the model predicts that the upward trend observed since the 2020s will continue until 2030, with the number of annual publications rising gradually but steadily. In contrast, the "Environmentally Friendly Packaging Marketing" theme, despite showing a significant increase in previous years, is seen to fluctuate within a more limited range in the forward-looking period, i.e., it follows a relatively stable line. For studies grouped under the heading "Green Supply Chain Marketing," LSTM predictions indicate that the curve will gain upward momentum, especially after 2025, signalling that the supply chain and logistics dimension will increasingly feature on the sustainable marketing agenda. These results provide quantitative evidence that sustainable brand management and green supply chain marketing will rank high among research priorities in the near future. At the same time, relatively mature sub-areas, such as eco-friendly packaging, will achieve a more stable rate of growth.

## Discussion

The findings of this study largely confirm existing bibliometric and conceptual studies in the sustainable marketing literature. Still, unlike them, it also quantitatively reveals the future-oriented dynamics of the themes. Recent comprehensive reviews show that sustainable marketing research has focused heavily on integrating environmental and social impacts into marketing strategies, consumer behaviour, brand management, and supply chains, with the volume of publications increasing rapidly, especially after 2010 (Buvaneswari & Aishwaryaa, 2024). The themes obtained with BERTopic in this study—Green Marketing, Sustainable Branding, Marine and Food-Focused Sustainability, Green Supply Chain Marketing, and AI-Supported Sustainability Applications—present a thematic map consistent with prior reviews. However, LSTM/GRU predictions not only reveal the current importance of these themes but also identify which areas will gain momentum by 2030, adding a dynamic "trend prediction" dimension to bibliometric analyses, which are often static in the literature. Overall, when error and fit metrics are combined, the LSTM model performs better than the GRU at predicting the number of thematic publications in this study.

To aid interpretation of the forecasting accuracy reported in Table 4, we translate the MAPE (~21.6%) into approximate absolute deviations for the 2026–2030 horizon. In practical terms, an MAPE of ~21.6% implies that, on average, the forecasted annual publication counts may deviate from the realised values by roughly one-fifth. For example, a point forecast of 100 publications corresponds to an average absolute deviation of approximately  $\pm 22$  publications, whereas forecasts of 200 and 300 publications correspond to deviations of approximately  $\pm 43$  and  $\pm 65$  publications, respectively. To make this uncertainty explicit for the 2026–2030 horizon, we report approximate absolute deviations alongside the point forecasts (Hyndman & Koehler, 2006).

The results indicate that themes surrounding "sustainable branding" and responsible brand narratives, in particular, will continue to strengthen in the coming years. Increased research on sustainable brand management in recent years emphasises that a brand identity consistent with environmental and social sensitivity has positive effects on long-term brand value, reputation, and loyalty; conversely, exaggerated or unfounded environmental claims (greenwashing) erode trust (Badhusha et al., 2025). In our study, both the past and future growth trends of this theme indicate that sustainable branding will remain a permanent research agenda in the marketing literature. Similarly, although environmentally friendly and sustainable packaging issues follow a relatively more "stable" course in the samples tested, current studies indicate that this area still has a significant impact on purchase intention, quality perception, and brand loyalty (Imran et al., 2024). This suggests that even in areas that appear mature as themes, research depth can be increased through micro-level consumer segmentation, cross-cultural comparisons, and experimental designs.

LSTM predictions indicate that themes related to green supply chain and circular economy marketing will also gain momentum in the coming period. Bibliometric analyses in recent years emphasise that the topics of "green supply chain management" and "circular economy" are rapidly rising in both management and marketing literature and offer important opportunities for future research (Wahyuni et al., 2024). The increasing trend of these themes in our study indicates that the pressure on businesses to create competitive advantage through carbon footprint reduction, resource efficiency, and circular business models is also strongly reflected in the academic agenda. Furthermore, the rapid rise of topics related to "AI-driven sustainability marketing" and data-driven green technologies from virtually zero in recent years demonstrates that sustainable marketing is intertwined with the fourth industrial

revolution. Indeed, recent studies show that AI-supported campaigns can increase green product sales and sustainable consumption intentions, and that personalised messages and predictive segmentation enable brands to align their sustainability goals with customer expectations better (Garg et al., 2025; Sohaib et al., 2025).

The sharp post-2020 increase observed for the AI/data-driven theme is consistent with the broader diffusion of AI-enabled analytics, personalisation, and optimisation in sustainability-oriented marketing and management research. Strategic marketing frameworks emphasise that AI increasingly supports data-driven decision-making and personalised communication, which can strengthen sustainability communication and stakeholder engagement and may explain the accelerated scholarly attention to this cluster (Garg et al., 2025; Huang & Rust, 2021).

From an industrial applications perspective, the findings offer a series of strategic implications for brands and managers. First, while sustainability-focused brand strategies appear to have strong long-term value creation potential, increasing regulatory pressures and public sensitivity have made transparency and verifiability fundamental requirements; recent cases demonstrate that misleading green claims generate reputational and legal risks rather than short-term gains (Kohli, 2024). Second, results in the supply chain and packaging areas reveal that sustainability investments are not only a cost but also a competitive factor in terms of market positioning and customer loyalty; therefore, managers need to strengthen their R&D and collaboration strategies in areas such as green logistics, circular supply models, and eco-friendly packaging (Siuda & Grębosz-Krawczyk, 2025).

For mature themes such as eco-friendly packaging, the relatively stable forecast suggests rebalancing R&D from broad exploration toward execution and validation. Concretely, brand managers can prioritise life-cycle assessment (LCA), material performance testing, supplier qualification, and claims substantiation/certification to reduce regulatory and reputational risk, while maintaining a smaller "optionality" budget for breakthrough materials and pilot trials. In parallel, controlled consumer experiments can be used to optimise packaging cues and messaging for willingness-to-pay and trust outcomes.

Finally, the trend of AI-powered sustainable marketing also brings data governance, algorithmic transparency, and ethical issues to the forefront. When using AI-based personalisation and impact optimisation, businesses should consider both energy consumption and data privacy; they should ensure that their sustainability narrative is consistent not only at the content level but also in the environmental footprint of the digital infrastructure they use.

Some methodological limitations should also be considered when interpreting this study's findings. First, the dataset was compiled solely from Elsevier's Scopus database; therefore, studies included in Web of Science, PubMed, and other field indexes but not indexed in Scopus could not be included in the analysis. This may have led to the exclusion of sustainability marketing research from certain disciplines or regional journals.

Second, because of the data collection strategy, only English-language publications were considered; studies published in other languages were excluded. Although English-language publications largely shape the sustainable marketing literature, it should not be overlooked that studies published in other languages that make important contributions in local contexts may not be reflected in the analysis. In future research, the simultaneous use of multiple databases and the creation of multilingual data sets will both increase thematic coverage and allow for a more comprehensive presentation of the regional and cultural diversity of the sustainable marketing literature.

The moving time window used in this study was set at 3 years, and the findings were interpreted in this short- to medium-term context. While this choice allows for more accurate capture of seasonal trends and turning points, it also carries the risk of failing to adequately reflect longer-term trends. Testing longer windows, such as 5 or 10 years, in future studies and comparing results across different window lengths would help assess the temporal robustness of the findings.

Another limitation concerns interpretability: although LSTM/GRU achieved higher predictive accuracy, these recurrent networks operate as black-box models. Future work can integrate Explainable AI (XAI) techniques (e.g., SHAP and LIME) to quantify which topic descriptors or term-level signals contribute most to forecasted increases, thereby improving transparency and managerial interpretability.

As an additional extension, latent thematic structures could be explored by computing semantic similarity between topic descriptors (e.g., cosine similarity over c-TF-IDF vectors) and then performing secondary clustering (e.g., k-means) to test alternative thematic taxonomies.

## **Conclusion**

This study proposes an integrated framework that both maps the sustainable marketing literature thematically and quantitatively predicts the future trajectory of these themes. By using BERTopic-based topic modelling together with deep learning-based time-series models such as LSTM and GRU, scenarios have been generated for how research themes derived from Scopus abstracts may evolve until 2030. In the literature, there are very few studies that combine topic modelling with deep learning-based publication count prediction on this scale in the field of sustainable marketing; in this respect, the study offers an original methodological contribution by adding a "predictive" dimension to bibliometric analyses.

Unlike traditional static bibliometric maps that are inherently retrospective, our predictive framework enables strategic foresight by estimating which sustainability themes are likely to accelerate, plateau, or decline over the next five years. This shifts the use of bibliometric evidence from "what has happened" to "what is likely to happen," supporting proactive agenda-setting for researchers and forward-looking resource allocation for decision makers.

From an academic perspective, the study provides three key contributions. First, the main themes of the sustainable marketing literature were detailed using BERTopic; axes such as green marketing, sustainable branding, marine and food-focused sustainability, green supply chain marketing, and AI-supported sustainability applications were structured in line with current research clusters in the field. Second, time series at annual frequencies for these themes were predicted using LSTM and GRU models, yielding quantitative findings on which themes are emerging, which are mature, and which are relatively stagnant. Thirdly, the comparison of LSTM and GRU performance provides an empirical reference for which model has an advantage under which conditions, in short, non-linear academic series.

From a practical perspective, the findings are instructive for both academic managers and brand and marketing professionals. The prediction that areas such as sustainable branding, green supply chain marketing, and AI-supported sustainability applications will gain momentum in the coming period provides concrete support for prioritising research funding and corporate R&D investments in these areas. For businesses, strong signals indicate that developing data-driven and transparent solutions that integrate sustainability discourse with corporate governance, supply chain, and digital marketing strategies will create a competitive advantage.

A two-pronged agenda can be proposed for future research. First, the LSTM/GRU structures used in this study can be compared with more advanced machine learning-based "topic evolution" models (e.g., dynamic topic models, attention-based sequence models, and Bayesian nonparametric approaches), thereby enabling the simultaneous modelling of both the evolution of themes and publication volume. Second, it is important to increase studies examining sector-based variations in sustainable marketing themes: for example, by creating separate corpora for specific sectors such as energy, food, fashion, tourism, or finance, it is possible to analyse which sustainability themes stand out in each field and how they will evolve in the future.

For high-impact and high-volatility sectors (e.g., energy and fashion), future studies can operationalise a sector-specific construct by (i) constructing sector-filtered corpora, (ii) estimating sector-level theme prevalence and volatility, (iii) producing rolling-origin forecasts with uncertainty bounds, and (iv) translating outputs into early-warning indicators for emerging sustainability themes and regulatory/technology-sensitive shifts. Such a roadmap would enable sector-tailored strategic foresight rather than generic cross-sector averages. Such studies will contribute to both a more granular view of the field and the establishment of more robust empirical foundations for sector-focused policies and strategies.

### **Peer-review:**

Externally peer-reviewed

### **Conflict of interests:**

The author has no conflict of interest to declare.

### Grant Support:

The author declared that this study has received no financial support.

### Ethics Committee Approval:

This study used publicly available data obtained from the Scopus database. Since the study did not involve human participants, animals, or confidential personal data, ethics committee approval was not required.

### References

- Arokodare, O., Egun, O., & Kim, A. (2024). Leveraging Machine Learning for Enhanced Predictive Accuracy in Time Series Forecasting: A Comparative Analysis of LSTM and GRU Models. <https://doi.org/10.13140/RG.2.2.30014.80965>
- Badhusha, M. H. N., Pandey, G., Yadav, L. N., Chadalavada Lakshmi Nath, D. A., & Chandra, S. (2025). Sustainability narratives in branding: The role of ethical consumerism in shaping purchase decisions. *Advances in Consumer Research*, 2, 4308–4317.
- Buvaneswari, P. S., & Aishwaryaa, V. (2024). Sustainable marketing: Insights on current and future research directions—A bibliometric analysis. *Journal of Business & Tourism*, 10(2), 1–21. <https://doi.org/10.34260/jbt.v10i02.298>
- Cheddak, A., Ait Baha, T., Es-Saady, Y., El Hajji, M., & Baslam, M. (2024). BERTopic for enhanced idea management and topic generation in brainstorming sessions. *Information*, 15(6), 365. <https://doi.org/10.3390/info15060365>
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4), 1057–1072.
- Imran, N. F. N. B. A., Effendee, N. F. B. N., Faris, N. H. H. B. M., Puad, N. I. S. B. M., Nitin, N., Verma, V., Mayur, K., Lim, H. L., Ng, W. C., Baliyan, M., Kee, D. M. H. (2024). The effect of eco-friendly packaging on consumer purchase intention: A study of beverage sustainable packaging practices. *Journal of Community Development Asia*, 7(1), 91–104. <https://doi.org/10.32535/jcda.v7i1.2539>
- Garg, V., Bohara, S., & Srivastav, A. (2025). AI-driven sustainability marketing transforming consumers' perception toward eco-friendly brands. *Discover Sustainability*, 6(1), 984. <https://doi.org/10.1007/s43621-025-01934-y>
- Golenvaux, N., Alvarez, P. G., Kiossou, H. S., & Schaus, P. (2020). An LSTM approach to forecast migration using Google Trends. *arXiv*. <https://doi.org/10.48550/arXiv.2005.09902>
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*. <https://doi.org/10.48550/arXiv.2203.05794>
- Gür, Y. E. (2024). Comparative analysis of deep learning models for silver price prediction: CNN, LSTM, GRU and hybrid approach. *Akdeniz İİBF Dergisi*, 24(1), 1–13. <https://doi.org/10.25294/auibfd.1404173>
- Hafizoglu, M., & Tuzlukaya, Ş. (2023). Social network theory in project management: A bibliometric analysis. *Başkent Üniversitesi Ticari Bilimler Fakültesi Dergisi*, 7(2), 63–82.
- Hankar, M., Kasri, M., & Beni-Hssane, A. (2025). A comprehensive overview of topic modeling: Techniques, applications and challenges. *Neurocomputing*, 628, 129638. <https://doi.org/10.1016/j.neucom.2025.129638>
- Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49(1), 30–50. <https://doi.org/10.1007/s11747-020-00749-9>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>

- Ilmalhaq, A., Pradana, M., & Rubiyanti, N. (2024). Sustainable consumption on marketing concept: A bibliometric analysis (2003–2023). *International Journal of Sustainable Development & Planning*, 19(9), 3595–3601. <https://doi.org/10.18280/ijstdp.190927>
- Jun, S. P., Yoo, H. S., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change*, 130, 69–87. <https://doi.org/10.1016/j.techfore.2017.11.009>
- Jung, Y. J., & Kim, Y. (2023). Research trends of sustainability and marketing research, 2010–2020: Topic modeling analysis. *Heliyon*, 9(3), e14208. <https://doi.org/10.1016/j.heliyon.2023.e14208>
- Kilinç, M. (2025). LSTM-based time series forecasting of user-derived quality signals in mobile banking systems. *Systems*, 13(11), 949. <https://doi.org/10.3390/systems13110949>
- Kohli, V. (2024). Greenwashing in the fashion industry: A manipulative facade in the name of corporate environment responsibility. *Jus Corpus Law Journal*, 4(3), 28–41.
- Kokoç, M., Kokoç, M., & Tuncer, Ö. (2024). Using text mining to identify research trends in management information systems theses: A topic modeling approach. *Ordu Üniversitesi Bilim ve Teknoloji Dergisi*, 14(2), 293–306. <https://doi.org/10.54370/ordubtd.1598387>
- Koruyan, K. (2022). Classification of customer complaints using BERTopic topic modelling technique. *Izmir Journal of Social Sciences*, 4(2), 66–79. <https://doi.org/10.47899/ijss.1167719>
- Li, C., & Hu, X. (2025). Medical artificial intelligence in scholarly and public perspective: BERTopic-based analysis of topic-sentiment collaborative mining. *Data Science and Informetrics*, 5(1), 33–42. <https://doi.org/10.1016/j.dsim.2025.05.001>
- Madhavaram, S., & Nirjar, A. (2025). Capability development for sustainable marketing: A theoretical framework. *AMS Review*, 15, 157–190. <https://doi.org/10.1007/s13162-025-00299-9>
- Mojtahedi, F. F., Yousefpour, N., Chow, S. H., & Cassidy, M. (2025). Deep learning for time series forecasting: Review and applications in geotechnics and geosciences. *Archives of Computational Methods in Engineering*, 32, 3415–3445. <https://doi.org/10.1007/s11831-025-10244-5>
- Monroy, S. E., & Diaz, H. (2018). Time series-based bibliometric analysis of the dynamics of scientific production. *Scientometrics*, 115(3), 1139–1159. <https://doi.org/10.1007/s11192-018-2728-4>
- Naufal, G. R., & Wibowo, A. (2023). Time series forecasting based on deep learning CNN-LSTM-GRU model on stock prices. *International Journal of Engineering Trends and Technology*, 71(6), 126–133. <https://doi.org/10.14445/22315381/IJETT-V71I6P215>
- Omoware, J. M., Abiodun, O. J., & Wreford, A. I. (2023). Predicting stock series of Amazon and Google using long short-term memory (LSTM). *Asian Research Journal of Current Science*, 5(1), 205–217.
- Samsir, S., Saragih, R. S., Subagio, S., Aditiya, R., & Watrianthos, R. (2023). BERTopic modeling of natural language processing abstracts: Thematic structure and trajectory. *Jurnal Media Informatika Budidarma*, 7(3), 1514–1520. <https://doi.org/10.30865/mib.v7i3.6426>
- Siuda, D., & Grębosz-Krawczyk, M. (2025). The role of pro-ecological packaging in shaping purchase intentions and brand image in the food sector: An experimental study. *Sustainability*, 17(4), 1744. <https://doi.org/10.3390/su17041744>
- Sohaib, O., Alshemeili, A., & Bhatti, T. (2025). Exploring AI-enabled green marketing and green intention: An integrated PLS-SEM and NCA approach. *Cleaner and Responsible Consumption*, 100269. <https://doi.org/10.1016/j.clrc.2025.100269>
- Sono, M. G. (2023). Bibliometric analysis of the term "marketing sustainability". *West Science Interdisciplinary Studies*, 1(6), 314–325. <https://doi.org/10.58812/wsis.v1i6.105>
- Svetunkov, I. (2026, February 22). Rolling origin. In *greybox* [Package vignette]. Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/web/packages/greybox/vignettes/ro.html>
- Tian, Y., Kamran, Q., & Henseler, J. (2025). Sustainability in marketing: A review using multiple correspondence analysis. *Cogent Business & Management*, 12(1), 2493389. <https://doi.org/10.1080/23311975.2025.2493389>
- Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: A survey. *Big Data*, 9(1), 3–21. <https://doi.org/10.1089/big.2020.0159>

- Trojanowski, T. (2022). The triple bottom line concept in sustainable marketing mix activities of food industry enterprises. *WSEAS Transactions on Business and Economics*, 19, 1296–1302. <https://doi.org/10.37394/23207.2022.19.116>
- Wahyuni, H., Haryati, T., & Asnawi, Y. H. (2024). Trends in green supply chain management: Insights from bibliometric analysis (2014–2023). *Airlangga Journal of Innovation Management*, 5(3), 543–554. <https://doi.org/10.20473/ajim.v5i3.59014>
- Wang, S., Liu, M. T., & Pérez, A. (2023). A bibliometric analysis of green marketing in marketing and related fields: From 1991 to 2021. *Asia Pacific Journal of Marketing and Logistics*, 35(8), 1857–1882. <https://doi.org/10.1108/APJML-07-2022-0651>
- White, K., Cakanlar, A., Sethi, S., & Trudel, R. (2025). The past, present, and future of sustainability marketing: How did we get here and where might we go? *Journal of Business Research*, 187, 115056. <https://doi.org/10.1016/j.jbusres.2024.115056>
- Widiastuti, E., Sukesi, S., & Sarsiti, S. (2024). Sustainable marketing in the digital age: A systematic review of the latest strategies and tactics. *International Journal of Economics, Business and Accounting Research (IJEBAR)*, 8(1). <https://doi.org/10.29040/ijebar.v8i1.12158>
- Yip, W. S., To, S., Zhou, H., & Ren, J. (2025). Text mining in sustainable manufacturing for topic modeling. In W. S. Yip, S. To, H. Zhou, & J. Ren, *Sustainable machining and micro-machining* (pp. 63–78). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-82986-4\\_5](https://doi.org/10.1007/978-3-031-82986-4_5)
- Yunita, A., Pratama, M. I., Almuzakki, M. Z., Ramadhan, H., Akhir, E. A. P., Mansur, A. B. F., & Basori, A. H. (2025). Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, 15, 103462. <https://doi.org/10.1016/j.mex.2025.103462>