European Journal of Forest Research (2026) 145:72
https://doi.org/10.1007/s10342-026-01907-4

RESEARCH

Determinants of chemical wood pulp production in Europe: evidence from panel econometrics and machine learning

Hakan Aydogan[1]

Received: 10 September 2025 / Revised: 9 January 2026 / Accepted: 10 April 2026 / Published online: 26 May 2026

Abstract

Chemical wood pulp production plays a strategic role in Europe's forest-based bioeconomy, given its strong linkages with energy consumption, industrial structure, and international trade. Understanding the key determinants of production is therefore essential for designing policies that support both industrial competitiveness and sustainability. This study aims to identify the macroeconomic, structural, trade, and energy-related drivers of chemical wood pulp production across Europe, using a balanced panel of eleven countries over the period 2009–2023. To capture both structural relationships and potential nonlinearities, the analysis integrates panel data econometric techniques with machine learning methods. Specifically, conventional panel regression models are complemented by an Extreme Gradient Boosting (XGBoost) framework, with Shapley Additive Explanations (SHAP) employed to enhance model interpretability. The econometric results indicate that electricity consumption and chemical wood pulp exports are significant, positive determinants of production. A notable structural asymmetry emerges: countries characterized by a higher presence of small and medium-sized enterprises exhibit higher production levels, whereas greater industrial concentration in the hands of large firms is negatively associated with output. In addition, GDP is found to be negatively related to chemical wood pulp production in European economies, suggesting a partial decoupling of aggregate economic growth from resource-intensive manufacturing activities. The machine learning results corroborate these findings, demonstrating strong predictive performance and revealing nonlinear interactions that are not adequately captured by linear models. From a policy perspective, the results highlight the importance of securing reliable and affordable energy supply, supporting the resilience and adaptability of small and medium-sized enterprises, and fostering export-oriented strategies to sustain the long-term performance of Europe's chemical wood pulp industry.

Keywords Chemical wood pulp · Pulp and paper industry · Industrial structure · Energy intensity · Trade dynamics · Europe

Introduction

The pulp and paper industry represents a cornerstone of Europe's forest-based bioeconomy, combining high capital intensity, long investment cycles, and strong links to energy systems, trade, and industrial competitiveness. According to FAOSTAT, global chemical wood pulp (CWP) output increased from about 119.5 million tonnes in 2009–150.5 million tonnes in 2023 (an increase of roughly 26%). Over the same period, Europe's production rose more modestly, from 28.5 to 32.0 million tonnes (about 12%), and its share of global CWP production declined from approximately 23.8%–21.3%. European output also fell by 8.6% relative to its 2021 peak, underscoring the region's exposure to structural constraints and external shocks. In 2023, Sweden (8.29 million tonnes) and Finland (7.00 million tonnes) together accounted for nearly 48% of Europe's CWP production, indicating substantial concentration and cross-country heterogeneity that a multi-country framework should account for (FAO 2025). At the technological core of this industry, CWP is predominantly produced via the Kraft process, which remains the dominant global route due to

Communicated by Michaela Eder.

✉ Hakan Aydogan
hakanaydogan@kastamonu.edu.tr

[1] Faculty of Forestry, Department of Forest Industrial Engineering, Kastamonu University, Kastamonu, Türkiye

fibre strength, durability, and recyclability (Del Rio et al. 2022; Ramos et al. 2022; Urdaneta et al. 2024). Although Kraft pulping is energy intensive, internal energy recovery, most notably through black liquor combustion, has progressively improved energy self-sufficiency and reduced carbon intensity (Balkissoon et al. 2023; Evelyn et al. 2025). Consequently, CWP production is closely aligned with Europe's low-carbon bioeconomy objectives and longer-term industrial transition pathways (Lipiäinen et al. 2022; Lintunen et al. 2024).

Against this backdrop, the European pulp and paper sector faces increasing pressure from energy prices, environmental regulation, and global trade dynamics. Recent EU bioeconomy strategies emphasize the strategic role of forest-based industries in achieving climate neutrality while maintaining industrial competitiveness and value creation within Europe (CEPF 2025). At the same time, energy-intensive production processes and export-oriented market structures expose CWP producers to external shocks and structural constraints, reinforcing the need for robust empirical evidence on the drivers of production performance across European economies.

Despite its strategic importance, the empirical literature on CWP production remains fragmented. Existing studies often examine individual determinants, such as macroeconomic activity, energy use, or trade exposure (Kando and Buongiorno 2009; Ince et al. 2011), or rely on single-country settings (Fracaro et al. 2012; Shabbir et al. 2022) and linear time-series frameworks (McCarthy and Lei 2010), limiting their ability to capture cross-country heterogeneity and complex interactions. Moreover, European macro-panel applications pose well-documented econometric challenges, including cross-sectional dependence and heterogeneous production structures, which require careful methodological treatment to ensure valid inference. Recent applied research in energy and sustainability has therefore called for analytical frameworks that combine econometric identification with complementary data-driven approaches to better reflect real-world complexity (Islam et al. 2022; Ali et al. 2025b).

Against this background, this study aims to identify the macroeconomic, structural, trade-related, and energy-related determinants of chemical wood pulp (CWP) production across Europe using a multi-country panel framework. We analyse annual data for 11 European countries for 2009–2023. CWP output is modelled as the dependent variable, and four theoretically motivated covariates are considered: gross domestic product (GDP), the number of pulp and paper enterprises (MOPE), CWP exports (CWPE), and electricity consumption (EC). A distinctive feature of the study is its explicit treatment of industrial structure by distinguishing production systems characterised by a larger SME presence from those exhibiting greater concentration among large firms, allowing us to test for structural asymmetries.

Methodologically, the study adopts a dual analytical strategy. Panel econometric models provide interpretable parameter estimates while addressing key macro-panel challenges, whereas Extreme Gradient Boosting (XGBoost) is used as a complementary machine learning approach to capture potential nonlinearities and interaction effects. This combined strategy reflects a growing strand of empirical research that integrates econometric inference with machine learning validation to enhance robustness and policy relevance in complex production settings (Ali et al. 2024).

The results indicate that electricity consumption and exports exert significant positive effects on CWP production across Europe. In contrast, higher aggregate income levels are associated with lower domestic CWP output, suggesting a degree of structural decoupling between economic growth and resource-intensive manufacturing in higher-income economies. Industrial structure also matters: production is higher in countries with a greater SME presence, whereas greater concentration among large firms is associated with lower output.

The novelty of this study lies in providing an integrated assessment of European CWP production that jointly considers macroeconomic conditions, energy use, trade exposure, and industrial structure within a unified econometric–machine-learning framework. By explicitly addressing cross-sectional dependence and structural heterogeneity, the analysis offers new insights into the production dynamics of a key bioeconomy sector.

The remainder of the paper is structured as follows. "Literature review" reviews the theoretical and empirical literature and identifies the research gaps addressed by the study. "Data and Methods" describes the data, variables, and methodological framework. "Results and Discussion" presents and discusses the empirical results. "Discussion: Interpretation and Comparison of Econometric and Machine Learning Results" concludes with policy implications, the study's limitations, and directions for future research.

Literature review

Theoretical foundations of chemical wood pulp production

Chemical wood pulp production can be viewed through the lens of the forest-based bioeconomy, in which industrial output is jointly shaped by resource endowments, energy inputs, trade exposure, and industrial structure (Arnič et al. 2024; Smith-Hall et al. 2024). From a production-theoretic standpoint, pulp manufacturing is characterised by high capital intensity, pronounced scale economies, and strong reliance on energy inputs—particularly electricity and process heat (Klugman et al. 2007; Luhas et al. 2019). These characteristics place CWP production at the intersection of industrial organisation and the energy–production nexus literature (Klugman et al. 2007; Wang et al. 2025).

From a trade-theoretic perspective, export-oriented manufacturing can benefit from scale economies and learning-by-exporting, while gaining access to larger markets, especially in resource-based industries where domestic demand may be limited (Chaldun et al. 2024; Akhtar et al. 2025). At the same time, export exposure may amplify vulnerability to external shocks, energy-price volatility, and cross-country regulatory differences. Recent work in sustainability economics further suggests that, in higher-income economies, resource-intensive production may partially decouple from aggregate GDP growth as economic activity shifts towards services and knowledge-intensive sectors (Antonelli et al. 2023; Chaldun et al. 2024; Reis Mourão and Popescu 2025). This tension between growth, trade, energy use, and structural transformation motivates an integrated empirical assessment of CWP production.

Macroeconomic and trade determinants of pulp and paper production

A substantial body of empirical literature has examined the relationship between macroeconomic activity and pulp and paper production (Szabó et al. 2009; Ochuodho et al. 2017; Kuzminov and Lobanova 2021; Han et al. 2024; Xian et al. 2025). Studies focusing on major producer countries consistently report strong associations between output, income levels, and trade performance, underscoring the role of export demand in sustaining production capacity (Zhang et al. 2012; Tang et al. 2015; Jaunky and Lundmark 2017; Shen and Lovrić 2022). Export-oriented strategies have been shown to generate production multipliers in forest-based industries, reinforcing the relevance of international trade for sectoral growth (Chao and Buongiorno 2002; Haile et al. 2021).

Nevertheless, evidence from advanced European economies (e.g., Sweden and other EU countries) suggests that higher income levels do not necessarily translate into increased domestic pulp production (Karlsson 2025; Nikou and Sardianou 2025). Several studies, such as those by Karlsson (2025), document a weakening relationship between GDP growth and resource-intensive manufacturing output, largely attributed to structural rationalisation as older, emissions-intensive plants are phased out in favour of fewer, more efficient units. Furthermore, stringent environmental regulation can impose compliance costs and input-reallocation trade-offs; reducing undesirable outputs (e.g., pollution) often requires reallocating inputs, potentially dampening conventional productivity growth (Li et al. 2025). Consequently, trade openness may further decouple domestic production from emissions via offshoring and outsourcing channels (Mitić et al. 2024). These mixed findings indicate that macroeconomic effects on the pulp sector are context-specific, varying between developed and developing economies depending on their level of development, industrial structure, and regulatory environment (Su et al. 2025; Xian et al. 2025).

Energy use and industrial output in resource-based industries

Energy consumption constitutes a critical input in chemical pulping processes, with electricity demand playing a central role in determining production scale and cost competitiveness. Although improvements in energy efficiency and internal energy recovery have mitigated the sector's exposure to energy-market volatility (Lipiäinen et al. 2022; Shabbir et al. 2022), the relationship between energy costs and industrial output remains multifaceted. The theoretical literature offers two contrasting perspectives. Standard cost-competitiveness arguments, often discussed alongside the pollution haven hypothesis, suggest that rising energy and compliance costs erode competitiveness, leading to lower output or relocation (Copeland and Taylor 2004). Conversely, the Porter hypothesis posits that such cost pressures can spur innovation, potentially enhancing operational efficiency and long-run productivity (Porter and van der Linde 1995; Andre et al. 2023; Cali et al. 2023).

Recent empirical evidence indicates that these effects are heterogeneous across energy sources and time horizons. For instance, Cali et al. (2022) distinguish between energy sources and find that electricity price increases often impair manufacturing performance, whereas fuel price increases may, in some contexts, raise productivity by accelerating the replacement of inefficient capital. Moreover, the effects of energy shocks involve dynamic adjustment: short-run cost burdens can give way to efficiency gains as firms adapt their production and energy-management strategies (Andre et al. 2023).

Accordingly, the energy–output nexus is frequently nonlinear and asymmetric, challenging simple linear generalisations. Recent studies employing advanced econometric and machine-learning approaches also underscore the importance of flexible modelling strategies when energy variables interact with macroeconomic and risk factors (Ali et al. 2025a, c). Although these contributions focus on environmental indicators, the methodological implication is clear: assessing the role of energy inputs in industrial performance similarly requires analytical frameworks capable of capturing nonlinearities and heterogeneous effects, rather than relying solely on traditional linear models.

Industrial structure, firm size, and production outcomes

Beyond macroeconomic and energy factors, industrial structure is increasingly recognised as an important yet relatively understudied determinant of output and productivity in the pulp and paper industry. Firm-level studies document that scale economies, technological intensity, and organisational design are associated with productivity and output dynamics (Ottosson and Magnusson 2013; Johansson et al. 2021; Wang and Liu 2023). However, the direction of scale effects remains ambiguous: large integrated producers may benefit from capital deepening and technological advantages, whereas fragmented, SME-dominated industry structures may foster flexibility, innovation, and regional embeddedness.

Empirical evidence on the relationship between firm concentration and pulp production remains limited, particularly in multi-country settings. Existing studies often rely on firm-level or single-country data, which constrains the generalisability of findings across heterogeneous European production systems. As a result, the role of industrial structure, specifically the tension between concentration and fragmentation in shaping aggregate CWP output, remains insufficiently understood, underscoring the need for analyses that explicitly account for firm-scale regimes (e.g., SME-intensive versus large-firm-concentrated systems).

Methodological approaches and empirical gaps

Most of the existing literature relies on linear econometric models applied to time-series data or pooled panel settings. While these approaches offer interpretability, they often do not adequately account for cross-sectional dependence, slope heterogeneity, and nonlinear interactions that are common in macro-panel datasets. Recent contributions in energy and environmental economics therefore emphasise second-generation panel techniques and robustness diagnostics when analysing cross-country data (Nosheen et al. 2021; Aydin et al. 2023). This methodological rigour is illustrated by Islam et al. (2022), who explicitly address cross-sectional dependence and heterogeneity in their assessment of energy-related drivers of sustainable development in ASEAN economies.

In parallel, a growing strand of empirical research integrates econometric modelling with machine learning to enhance predictive accuracy and uncover complex relationships. Studies that combine traditional inference with machine-learning-based validation show that hybrid approaches can deliver more nuanced insights than either method alone (Ali et al. 2024). Despite their increasing prominence in the broader energy–industry literature, such integrated frameworks remain rare in analyses of pulp and paper production, particularly in European applications.

Research gap and contribution

In summary, the existing literature points to three interrelated gaps. First, comprehensive cross-country analyses of chemical wood pulp (CWP) production in European countries remain limited, particularly those that jointly consider macroeconomic conditions, energy inputs, trade exposure, and industrial structure. Second, the role of the firm-size distribution—specifically, SME prevalence versus large-firm concentration—in shaping aggregate CWP output is still insufficiently understood. Third, methodological integration remains limited, as few studies combine panel econometrics with machine-learning methods while explicitly addressing cross-sectional dependence and structural heterogeneity.

This study addresses these gaps by developing an integrated econometric–machine-learning framework to analyse CWP production across Europe. By combining interpretable structural inference with predictive modelling, the analysis provides a more granular account of how macroeconomic, trade, energy, and industrial-structure factors jointly shape production dynamics in a key forest-based bioeconomy sector.

Data and methods

Data description and variable construction

This study utilizes a balanced panel dataset covering 11 European countries from 2009 to 2023, compiled from Eurostat, FAOSTAT, and the World Bank. The analysis focuses on the determinants of CWP production.

The dependent variable is the natural logarithm of CWP output (in metric tons), allowing for elasticity interpretation and variance stabilization (Wooldridge 2016; Neykov et al. 2022; Dabbous et al. 2023; Farnsworth 2023; Montes-Pajuelo et al. 2024). GDP, expressed in constant 2015 USD, is also log-transformed to capture the non-linear relationship between economic size and industrial output (Baltagi 2008a). The use of logarithmic transformations further mitigates scale effects and potential heteroskedasticity, which are common in macroeconomic panel datasets. Table 1 reports the descriptive statistics of the variables used in the analysis.

Table 1 presents the original log-transformed variables in levels, reflecting the economic scale and dispersion of

Table 1 Descriptive statistics of log-transformed variables (levels)

Variable	Obs	Mean	Std. Dev	Min	25%	Median	75%	Max
ln(CWP)	165	5.596	1.01	3.684	4.695	6.074	6.21	6.945
ln(GDP)	165	11.786	0.511	10.608	11.521	11.718	12.266	12.568
MOPE	165	99.194	96.596	7	18	51	174	355
CWPE	165	53,708	120,486	0.051	101.224	6768	24,732	571,000
EC	165	195,376	147,804	54,148	82,425	125,004	310,251	510,774

Table 2 Definitions and descriptions of variables

Variable	Label	Unit	Description
CWP	ln_cwp	Tons	Natural logarithm of total annual domestic chemical wood pulp production
GDP	ln_gdp	Costant 2015 US$	Log-transformed real GDP at constant 2015 US$
Number of Pulp & Paper Enterprises	mope	Count	Manufacturing base size
Chemical Wood Pulp Exports	cwpe	Tons	Annual export volume of chemical wood pulp
Electricity Consumption	ec	GWh	Annual energy consumption

chemical wood pulp production and its key determinants across the sampled European countries. The statistics reveal substantial heterogeneity across countries, particularly in export volumes (CWPE) and electricity consumption (EC), underscoring the structural diversity of the European pulp and paper industry.

To summarize the dataset, Table 2 presents the definitions and measurement units of all variables used in this study.

Table 2 summarizes the dataset, while Fig. 1 illustrates the evolution of $\ln(CWP)$ from 2009 to 2023. The time series is notably stable, with average values decreasing slightly from 5.62 to 5.54, suggesting capital intensity and historical path-dependency typical of the pulp sector. Minor disruptions around 2009 and 2021 likely reflect the Eurozone debt crisis and the COVID-19 pandemic, respectively, further supporting the panel data approach, where cross-sectional variation dominates temporal change.

Variable selection reflects both theoretical and empirical considerations. Electricity consumption (EC) indicates operational scale in energy-intensive production systems (Virtanen et al. 2013; Fresner et al. 2017; Golmohamadi 2022). The logarithm of GDP serves as an indicator of a country's overall macroeconomic strength. Meanwhile, exports of chemical wood pulp (CWPE) reflect the sector's integration into global markets and its sensitivity to foreign demand. Firm count (MOPE) measures industrial

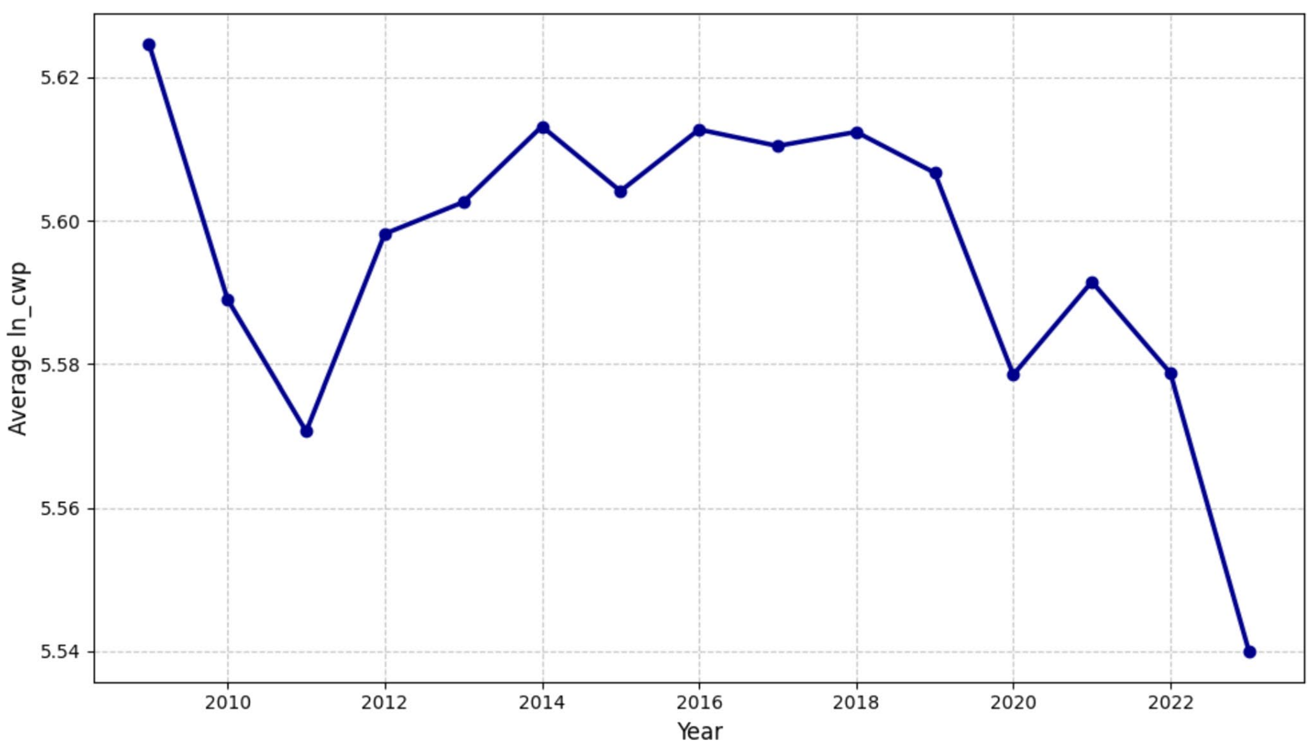

Fig. 1 Time trend of chemical wood pulp $(lnCWP)$ production

capacity but may obscure heterogeneity if treated as a continuous regressor. Therefore, we categorize countries as $MOPE^{low}$ ($\leq 100 firms$) and $MOPE^{high}$ (> 100) to capture scale-driven nonlinear effects.

Figure 2 shows that countries with $MOPE^{high}$ exhibit higher median $\ln(CWP)$, suggesting structural and technological distinctions, such as economies of scale, production chain integration, and energy efficiency intensity (Rapposelli et al. 2024). Stratification by firm count also enables analysis of asymmetric production responses, consistent with best practices in structural heterogeneity modelling (Mahapatra and Irfan 2024; Steinbrunner 2024; Wang et al. 2024a, b; Mishra and Acharya 2025).

Geographic scope and spatial variation of ln(CWP)

To place the empirical model in context, the map in Fig. 3 shows the geographical scope of the analysis, including eleven countries selected for consideration on the basis of available data and their respective output of CWP generation. The countries are shown in colours indicating the natural logarithm of their average annual CWP generation from 2009 to 2023, between around 8.5 and 16. This geographical division allows for comparisons of industrial intensity across the region.

There is a clear geographical pattern: High levels of $\ln(CWP)$ are seen for Finland and Sweden, expected from their internationally integrated pulp and paper sectors and large forest endowments. Italy, Switzerland, and Slovenia are found at the other end of the scale, and they represent a more restricted or diversified industrial production orientation. The Central European economies, especially Germany, Austria, and France, lie mid-range, indicating the characteristics of production to be moderate but consistent.

The bottom-right bar graph shows the nominal annual CWP output expressed in metric tonnes for the given time period. Compared to the log-transformed graph showing only orders of magnitudes, the bar graph shows absolute quantities, and the dominance of Finland and Sweden emerges, with their average annual outputs being more than 7 million tonnes. These two countries contribute significantly to the overall European production, an indication

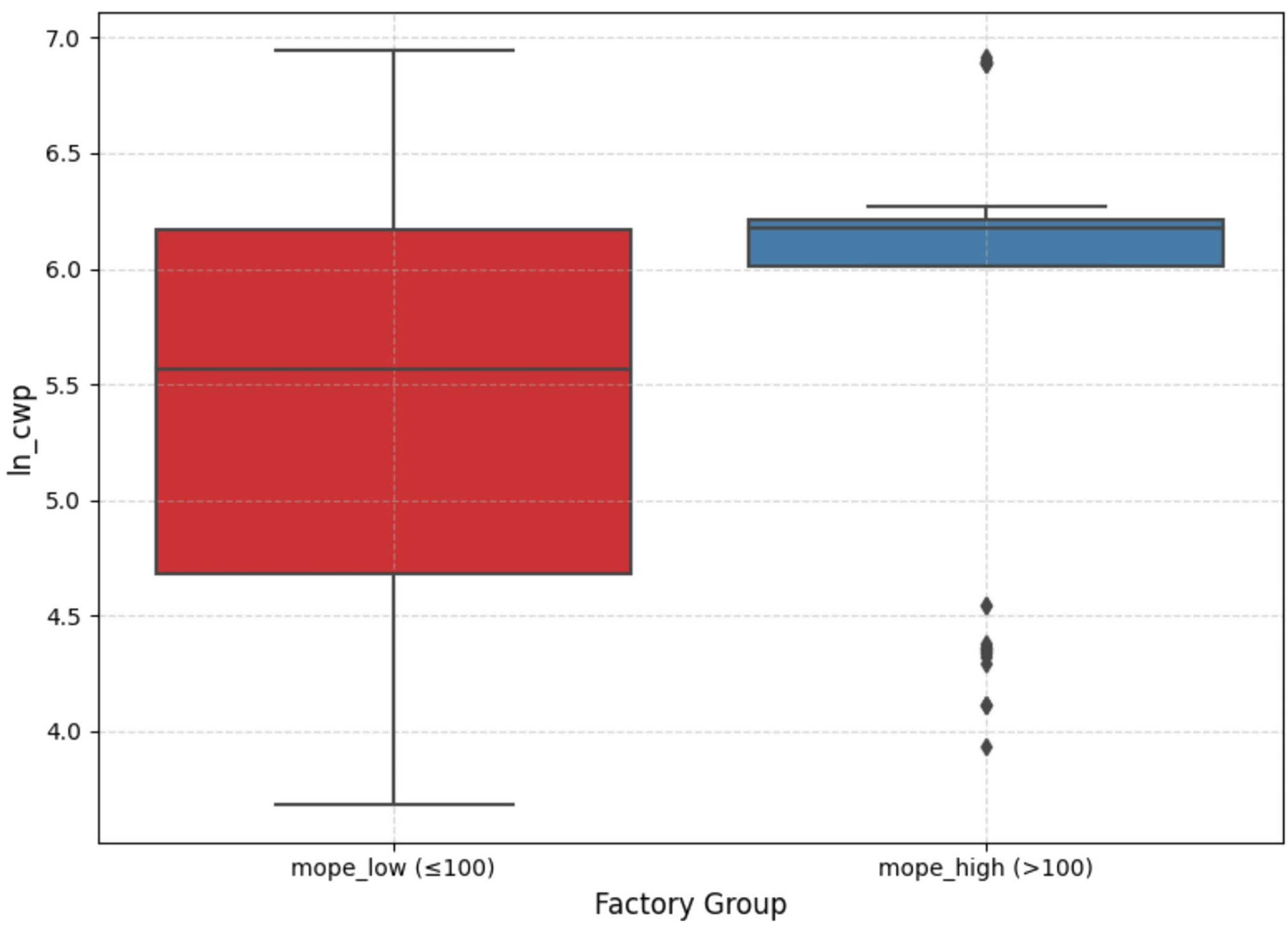

Fig. 2 Distribution of ln(CWP) across countries segmented by the number of pulp and paper enterprises (MOPE) The boxplot illustrates higher median and lower dispersion for the $MOPE^{high}$ group, suggesting a stronger and more concentrated production structure

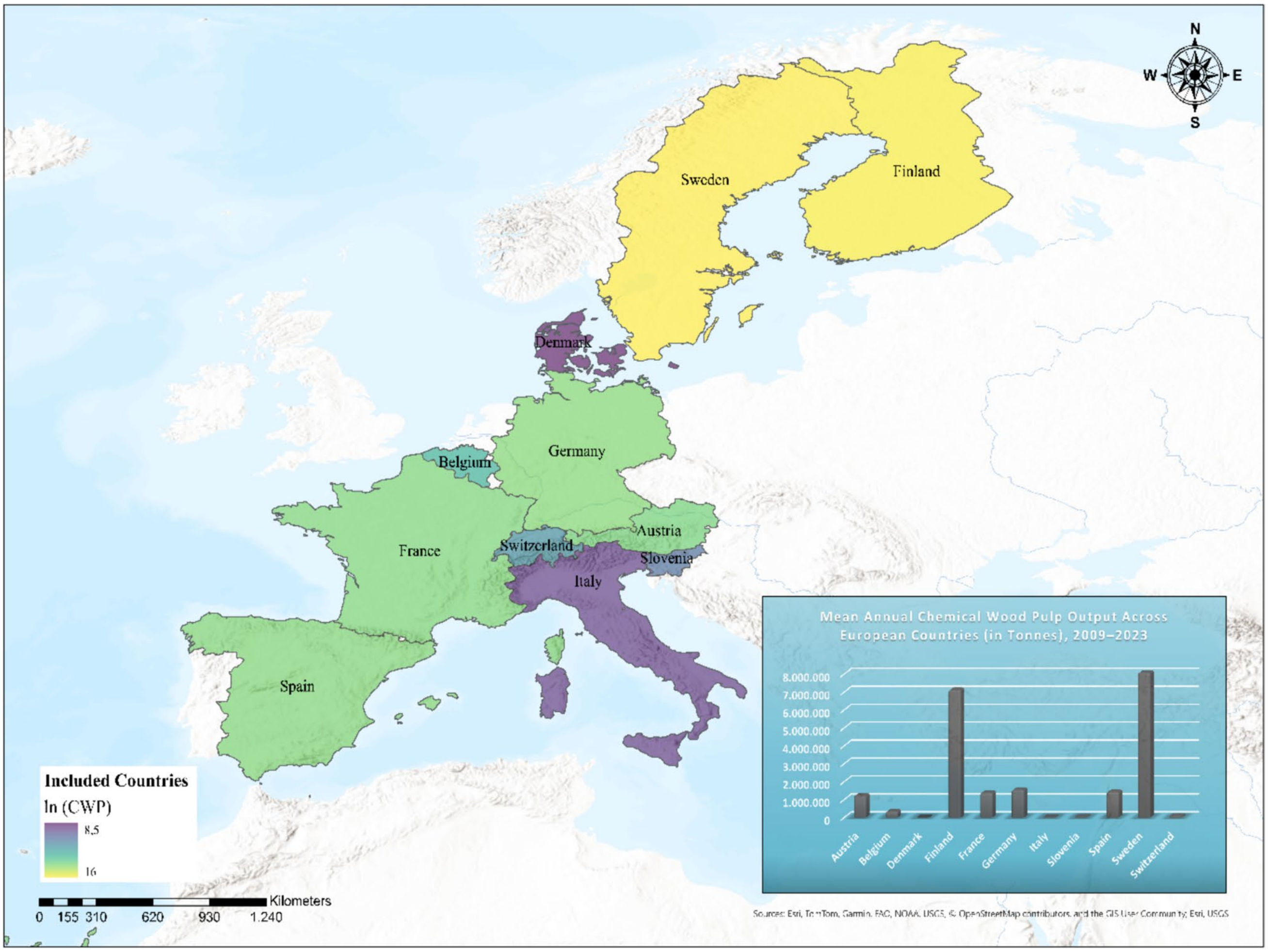

Fig. 3 Study area map depicting country-level classification of mean ln(CWP) and average annual output (tonnes), 2009–2023

easily seen visually, and this is important for the interpretation of the following econometric findings.

Econometric modelling and estimation techniques

The inherent structure of the dataset, encompassing observations from multiple countries over an extended period, renders it particularly well-suited for analysis using a panel data econometric model. This specific specification has the added advantage of allowing the model to handle unobserved heterogeneity between the countries and provide more accurate estimates using the cross-sectional and time dimensions of the dataset (Baltagi 2008a, 2021; Wooldridge 2010; Croissant and Millo 2019).

Panel data methods are preferred over pure cross-sectional or time-series approaches because they reduce omitted variable bias arising from unobserved, time-invariant country characteristics and improve estimation efficiency (Aydoğan and Yıldırım 2025).

We begin with the following log-linear panel model specification:

$$\begin{aligned}\ln\left(CWP_{it}\right) = & \delta + \varphi_1 \ln\left(GDP_{it}\right) + \varphi_2 MOPE_{it}^{low} + \varphi_3 MOPE_{it}^{high} \\ & + \varphi_4 CWPE_{it} + \varphi_5 EC_{it} + u_{it}\end{aligned} \tag{1}$$

where, $i = 1, \ldots, N$, indexes countries, $t = 1, \ldots, T$, indexes years. δ denotes the model's constant term. $\varphi_1, \ldots, \varphi_5$ represent the elasticities of the explanatory variables. The composite error term u_{it} is decomposed as follows:

$$u_{it} = \mu_i + \lambda_t + \varepsilon_{it} \tag{2}$$

Here, μ_i captures time-invariant country-specific effects, λ_t represents time-specific shocks common to all countries, and ε_{it} idiosyncratic disturbance term assumed to be that is independently and identically distributed with zero mean and constant variance, $\varepsilon_{it} \sim \text{IID } \left(0, \sigma^2\right)$. It is further assumed that the regressors are strictly exogenous for all countries and for all time periods with zero mean and constant variance. $E\left(\varepsilon_{it} X_{it}\right) = 0$ for all i and t.

This two-way error structure explicitly allows for both country-specific heterogeneity and common temporal shocks, which is particularly relevant in a European context characterised by shared economic cycles and policy regimes.

Preliminary diagnostics and trend adjustments

Prior to estimating the panel regression model, we perform a series of preliminary diagnostic tests to verify the validity of key econometric assumptions. These include tests for cross-sectional dependence, panel unit root and stationarity, heteroskedasticity, and serial correlation, each of which could compromise the consistency and efficiency of coefficient estimates if unaddressed. As emphasized by Baltagi (2008a) and Wooldridge (2010), failure to account for such properties may result in spurious relationships, biased standard errors, and potentially invalid policy implications. Conducting these diagnostics is essential because ignoring such properties may lead to biased coefficient estimates, inconsistent standard errors, and misleading statistical inference, especially in multi-country macro-panel settings.

Cross-sectional dependence

Given the shared institutional and economic setting of European countries, we test for cross-sectional dependence (CSD) using Pesaran's CD test (Pesaran 2021). The presence of cross-sectional dependence would indicate common shocks or spillover effects, thereby justifying the use of robust variance estimators or second-generation panel techniques.

Panel unit root tests and stationarity

To avoid spurious regressions, it is essential to assess the stationarity properties of the panel data, particularly when macroeconomic variables such as GDP or electricity consumption known to exhibit stochastic trends are involved (Wooldridge 2010). We employ two first-generation panel unit root tests to evaluate the presence of non-stationarity across variables: (i) the Levin et al. (2002) test, which assumes a common autoregressive root across units, and (ii) the Im et al. (2003) test, which allows for heterogeneous autoregressive dynamics across countries. Employing both tests allows us to assess stationarity under alternative assumptions regarding homogeneity and heterogeneity in autoregressive dynamics across countries.

The general form of the unit root test can be written as:

$$\Delta Z_{it} = \rho Z_{it-1} + \sum_{k=1}^{\rho} \emptyset_k \Delta Z_{it-k} + \varepsilon_{it} \tag{3}$$

where, Z_{it} represents the variable in levels, and ρ is the autoregressive coefficient under test, and ε_{it} is the error term. The null hypothesis of both test is $H_0 : \rho = 0$, i.e., implying the presence of a unit root (non-stationarity), while the alternative suggests stationarity.

Model estimation and selection procedure

Following the diagnostic stage, we estimate the panel regression model using pooled OLS, fixed effects (FE), and random effects (RE) estimators.

The pooled OLS model assumes full homogeneity across cross-sectional units and ignores unobserved heterogeneity, which may result in biased estimates when country-specific effects are present (Wooldridge 2010).

The FE estimator controls for all time-invariant unobserved heterogeneity via the within transformation, ensuring consistent estimation when regressors are correlated with country-specific effects.

The RE estimator is included for comparison purposes, as it offers efficiency gains under the strict exogeneity assumption. Model selection is guided by formal specification tests rather than ad hoc criteria.

To determine the appropriate estimator, we employ: (i) the Chow test, (ii) the Hausman test (Hausman 1978), and (iii) the Breusch–Pagan LM test.

Together, these tests ensure that the chosen econometric specification appropriately reflects the underlying data-generating process and provides both statistical validity and economic interpretability.

Panel regression models

We begin with traditional panel data regression models, including: (a) Pooled Ordinary Least Squares (Pooled OLS), (b) Fixed Effects (FE), (c) Random Effects (RE).

The Pooled OLS model assumes complete homogeneity across all country-year observations, ignoring unobserved heterogeneity and structural breaks. While computationally straightforward, the model is vulnerable to omitted variable bias in cases where country-specific effects influence the dependent variable but are not accounted for (Wooldridge 2010).

To solve this problem, we use the Fixed Effects (FE) model that controls for all the unobserved heterogeneity that persists over time using the application of the within transformation. The FE model is specified as:

$$\begin{aligned}\ln(CWP_{it}) - \overline{\ln(CWP_{it})} = \varphi_1 \left[\ln(GDP_{it}) - \overline{\ln(GDP_{it})}\right] \\ + \ldots + (u_{it} - \overline{u_i})\end{aligned} \tag{4}$$

This transformation removes δ and any unobserved effects μ_i, ensuring consistent estimation of φ parameters when unobserved effects are correlated with the regressors. Alternatively, the Random Effects model treats these unobserved effects μ_i as randomly distributed across countries and incorporates them into the composite error term. This model is more efficient than FE if the strict exogeneity assumption holds, i.e., if $Cov\left(x_{it}, \mu_i\right) = 0$.

To determine the appropriate model, we apply three standard specification tests:

(i) The Chow test for structural stability, assessing the validity of Pooled OLS relative to FE,
(ii) The Hausman test (Hausman 1978), evaluating whether the FE or RE estimator is consistent, and
(iii) The Breusch–Pagan Lagrange Multiplier (LM) test, which tests the null hypothesis of no significant panel effects.

Together, these tests determine whether the selected model correctly identifies the underlying data structure and offers both statistical precision and economic interpretability in describing disparities in the production of CWP in different nations.

Machine learning estimation: XGBoost

To account for non-linear interactions and saturation effects, we implement Extreme Gradient Boosting (XGBoost); a scalable, tree-based ensemble learning method (Chen and Guestrin 2016). Unlike linear panel estimators, XGBoost does not impose additive or linear functional forms and is therefore well suited to capturing complex interactions among macroeconomic, trade, and energy variables. The model specification is given by:

$$\begin{aligned}\ln\left(\widehat{CWP}_{it}\right) &= \sum_{k=1}^{K} f_k\left(x_{it}\right)\\ &= \sum_{k=1}^{K} f_k\left(\ln\left(GDP_{it}\right), MOPE_{it}^{low}, MOPE_{it}^{high}, CWPE_{it}, EC_{it}\right)\end{aligned} \tag{5}$$

Here; $\ln(\widehat{CWP_{it}})$ is the predicted value of $\ln(CWP_{it})$, f_k represents the k^{th} decision tree in the model, K is the total number of trees, and x_{it} is the vector of explanatory variables for country i and year t. Each function $f_k \in F$ belongs to the space of regression trees and is learned by minimizing a regularized objective function that penalizes both prediction error and model complexity.

The model is trained using 80% of the dataset and evaluated on the remaining 20% using the following metrics: RMSE- Root Mean Squared Error (6), MAE- Mean Absolute Error (7), and R^2.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\ln\left(CWP\right)_{it} - \ln\left(\widehat{CWP}\right)_{it}\right)^2} \tag{6}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\ln\left(CWP\right)_{it} - \ln\left(\widehat{\mathrm{CWP}}\right)_{\mathrm{it}}\right| \tag{7}$$

To address the common criticism that machine learning models lack interpretability, SHAP values are employed to decompose predictions into additive feature contributions, thereby linking predictive performance with economic interpretation and policy relevance.

By integrating panel econometric models with machine learning techniques, the methodology combines inferential transparency with predictive flexibility. This hybrid approach allows the study to overcome the limitations of single-method strategies and provides a more comprehensive understanding of the drivers of chemical wood pulp production across Europe.

Results and discussion

Preliminary statistical relationships: correlation structure and visual diagnostics

Prior to the implementation of the panel regression framework, a Pearson correlation analysis was undertaken to explore the preliminary linear relationships between the dependent variable and the core explanatory factors. This step provides conceptual insights into the pairwise associations and serves as an initial diagnostic tool for detecting potential multicollinearity among the regressors. The correlation matrix is illustrated in Fig. 4 as a heatmap, visually representing the magnitude and direction of the bivariate coefficients.

The findings reveal a moderately strong positive association between $\ln(\mathrm{CWP})$ and EC $(r = 0.56)$, aligning with the well-documented energy-intensive characteristics of the pulp production industry (Posch et al. 2015; Lawrence et al. 2019; Andersson et al. 2021; Lipiäinen et al. 2022). In contrast, the association between $\ln(CWP)$ and $CWPE$ is weaker $(r = 0.29)$, suggesting that export volumes may play a secondary, though still relevant, role in explaining variation in output. Similarly, the link between $ln(CWP)$ and $ln(GDP)$ appears weak at $r = 0.22$, perhaps reflecting a decoupling of pulp production from aggregate economic growth in advanced industrial contexts.

A robust positive correlation is also observed between $\ln(\mathrm{GDP})$ and $MOPE^{high}$ $(r = 0.75)$, indicating that economically advanced countries typically host a greater number of large-scale pulp and paper firms. By contrast,

Fig. 4 Heatmap of pairwise Pearson correlations among model variables (non-detrended). The colour gradient illustrates the strength and direction of relationships, ranging from –0.65 to + 1.00

the strong negative correlation between $MOPE^{high}$ and $MOPE^{low}$ $(r = -0.65)$ underscores their functional divergence and supports their inclusion as distinct categorical variables in the regression model.

Although these bivariate findings offer preliminary insights, they do not account for unobserved heterogeneity, structural dynamics, or endogenous interactions among variables. Accordingly, the analysis proceeds to panel-specific multivariate estimations to rigorously evaluate these relationships within a controlled and temporally dynamic framework.

Table 3 Variance inflation factors (VIF)

Variable	VIF
ln(GDP)	2.85
$MOPE^{low}$	4.13
$MOPE^{high}$	3.68
CWPE	2.22
EC	1.58

Panel regression model results

Before proceeding to panel-specific diagnostic tests, potential multicollinearity among the explanatory variables is assessed using variance inflation factors (VIF) (Table 3).

All VIF values remain well below conventional threshold levels, indicating that multicollinearity does not pose a concern for coefficient stability or statistical inference.

Cross-sectional dependence and panel stationarity

Given the economic interconnectedness of European countries, CSD is likely. Pesaran's CD test confirms significant residual correlation across countries (Table 4), indicating shared shocks and spillovers that violate the independence assumption of first-generation unit root tests.

Indeed, both Levin-Lin-Chu (LLC) and Im-Pesaran-Shin (IPS) tests fail to reject the null of non-stationarity ($p>0.05$), suggesting the presence of unit roots under conventional assumptions. However, differencing is avoided to preserve long-term relationships and the interpretability of log-linear elasticities. Instead, a detrending strategy is employed to retain structural variation while removing deterministic trends.

Detrending strategy for stationarity and model integrity

To address CSD and non-stationarity without compromising model interpretability, we detrend the data (Phillips and Moon 2000; Breitung 2001; Westerlund 2015; Herwartz et al. 2023) by removing country-specific linear trends. This avoids the distortions associated with differencing and maintains elasticity interpretation. The final estimation model is expressed as:

$$\begin{aligned} n\left(CWP_{it}\right) =& \delta + \varphi_1 \ln\left(GDP_{it}\right) + \varphi_2 MOPE_{it}^{low} + \varphi_3 MOPE_{it}^{high} \\ &+ \varphi_4 CWPE_{it} + \varphi_5 EC_{it} + u_{it}\,(Detrended) \end{aligned} \quad (8)$$

Within this specification, u_{it} denotes the composite error term, capturing country-specific, time-specific, and idiosyncratic disturbances. This detrended formulation ensures that all included variables achieve stationarity while maintaining their logarithmic transformation and preserving the interpretability of coefficient estimates. This approach conforms to established methodological standards for macro-panel estimation (Baltagi and Kao 2001; Baltagi 2008b; Wooldridge 2010; Herwartz et al. 2023), and provides the empirical foundation for the estimation procedure.

Results from the detrended model

After implementing the detrending procedure to address issues of non-stationarity and cross-sectional dependence, the panel regression was re-estimated using the resulting stationary residual series. Among the three estimated specifications; Pooled Ordinary Least Squares (OLS), Fixed Effects (FE), and Random Effects (RE), the Pooled OLS model emerged as the preferred choice, based on formal statistical tests and superior model performance.

Specifically, the Chow test for structural stability ($F\text{-}stat = 1.07, p = 0.34$) did not reject the null hypothesis of parameter homogeneity across countries, thereby supporting the appropriateness of the pooled estimator. The Breusch–Pagan Lagrange Multiplier (LM) test ($\chi^2 = 2.45, p = 0.117$) also revealed no statistically significant panel-level random effects, diminishing the rationale for adopting the RE specification. Finally, the Hausman test comparing FE and RE estimators ($\chi^2 = 5.28, p = 0.071$) indicated that the differences in coefficients were not statistically significant, implying that while the RE model is not inconsistent, it also fails to offer efficiency gains under weak panel effects.

Taken together, the results strongly support the selection of the Pooled OLS specification as the most empirically robust model. It accounts for approximately 54.6% of the variation in the detrended CWP output ($adjusted R^2 = 0.546$), as shown in Table 5.

Model selection based on Chow, LM, and Hausman tests supports the pooled specification. All standard errors are heteroskedasticity-robust.

The estimated coefficient for $\ln(GDP_{it})$ is negative and statistically significant at the 1% level ($\varphi_1 = -0.483, p = 0.006$), This somewhat counterintuitive finding may suggest a structural transformation within high-income economies, where industrial activity is progressively shifting toward service-oriented sectors or less

Table 4 Results of cross-sectional dependence and unit root tests

Variable	Pesaran CD	*p* value (CD)	LLC Test (Z-stat)	*p* value (LLC)	IPS Test (W-stat)	*p* value (IPS)
ln(CWP)	85.66	0.0000	−1.12	0.131	−0.75	0.2262
ln(GDP)	88.74	0.0000	−0.87	0.194	−0.41	0.3410
MOPE_low	11.34	0.0068	−1.34	0.090	−0.98	0.1634
MOPE_high	12.68	0.0054	−1.01	0.157	−0.63	0.2672
CWPE	84.07	0.0000	−0.95	0.172	−0.59	0.2844
EC	82.45	0.0000	−1.2	0.115	−0.78	0.2143

All variables exhibit statistically significant cross-sectional dependence ($p<0.01$), while stationarity is not attained at levels based on the results of both Levin–Lin–Chu (LLC) and Im–Pesaran–Shin (IPS) tests

Table 5 Estimated coefficients from the detrended Pooled OLS regression model

Variable	Coef	Std.Err	t	$P>\|t\|$	[0.025	0.975]
const	−0.0000	0.0530	−0.0000	1.0000	−0.1047	0.1047
log_gdp_detrended	−0.4834	0.1758	−2.7494	0.0067	−0.8307	−0.1362
mope_low_detrended	0.0231	0.0029	7.9677	0.0000	0.0174	0.0288
mope_high_detrended	−0.0118	0.0015	−7.7638	0.0000	−0.0148	−0.0088
cwpe_detrended	4.94e-6	0.0000	7.5045	0.0000	0.0000	0.0000
ec_detrended	1.53e-6	0.0000	3.3853	0.0009	0.0000	0.0000

Table 6 Post-estimation diagnostic test results for the detrended pooled OLS model

Test	Test Statistic	*p* value
Breusch-Godfrey (Autocorr.)	96.3578	0.0000
Breusch-Pagan (Heterosk.)	9.8014	0.0810
Durbin-Watson	0.5333	–
Jarque–Bera	8.81	0.012

pulp-intensive manufacturing, thereby weakening the traditional link between GDP and CWP output.

A central contribution of this study lies in its decomposition of the manufacturing landscape into $MOPE^{low}$ and $MOPE^{high}$, categorizing countries by the scale of their pulp and paper sectors, those with 100 or fewer enterprises versus those exceeding that threshold. The coefficient on $MOPE^{low}$ is positive and highly significant $(\varphi_2 = 0.0231, p < 0.001)$, indicating that small and medium-sized enterprises play a constructive role in pulp production, possibly owing to their operational agility and greater responsiveness to regional market conditions. By contrast, the coefficient on $MOPE^{high}$ is negative and statistically significant $(\varphi_3 = -0.0118, p < 0.001)$, which may reflect diminishing marginal returns, coordination complexities, or regulatory rigidities typically associated with large-scale industrial structures.

Both $CWPE$ and EC yield statistically significant and positive coefficients ($p<0.001$), estimated at 4.94e-6 and 1.53e-6, respectively. Although the absolute magnitudes of these coefficients may seem small due to the units of measurement (tonnes and GWh, respectively), their economic implications are substantial. These findings underscore the critical importance of export orientation and energy availability as fundamental drivers of output within wood-based industrial systems.

Model assumption diagnostics

To evaluate the statistical soundness of the detrended Pooled OLS specification, a series of post-estimation diagnostic tests were conducted to examine adherence to key assumptions of panel regression namely, autocorrelation, heteroskedasticity, and error independence. The corresponding diagnostic test results are summarized in Table 6.

The Breusch–Godfrey LM test, employed to detect first-order serial correlation, yields a highly significant result (LM = 96.36, $p<0.001$thereby confirming the presence of autocorrelation in the residuals. This finding is further corroborated by the Durbin–Watson statistic, which registers a value of 0.533, substantially below the conventional benchmark of 2, indicating pronounced positive autocorrelation.

Regarding heteroskedasticity, the Breusch–Pagan test yields a test statistic of 9.80 with a p value of 0.081, hovering around the 10% significance threshold. While the result does not definitively confirm substantial heteroskedasticity, it does imply that residual variance may vary across the observed values.

The Jarque–Bera test applied to the residuals of the detrended pooled OLS model indicates a rejection of the null hypothesis of normality at the 5% significance level (JB = 8.81, $p = 0.012$). Deviations from normality are common in macro-panel data and do not invalidate the consistency of the OLS estimator. Importantly, the use of heteroskedasticity- and autocorrelation-robust standard errors ensures reliable statistical inference even under non-normal error distributions.

Taken together, the diagnostic results point to notable violations of the assumptions of homoskedasticity and residual independence, an outcome not entirely unexpected in macro-panel data characterized by persistence and inter-country spillovers. As such, employing robust standard error estimators is warranted to mitigate the effects of these assumption breaches. Subsequent estimations rely on Driscoll–Kraay standard errors, which are specifically designed to account for both heteroskedasticity and autocorrelation in panel datasets exhibiting cross-sectional dependence (Hoechle 2007).

Robust estimationtion

To address autocorrelation and potential heteroskedasticity identified in diagnostics, three alternative standard error estimators were applied: classical (unadjusted), Driscoll–Kraay, and country-clustered (Table 7). Results across all methods are highly consistent both in sign and statistical significance, indicating robust inference.

Table 7 Robust estimation results

Variable	Coefficient	Std. Err. (Unadjusted)	Std. Err. (Driscoll-Kraay)	Std. Err. (Clustered)
ln(GDP)	−0.483	0.176	0.189	0.196
MOPE_low	0.0231	0.0029	0.0031	0.0033
MOPE_high	−0.0118	0.0015	0.0016	0.0017
CWPE	4.94e-6	6.58e-7	6.72e-7	7.01e-7
EC	1.53e-6	4.53e-7	4.75e-7	4.89e-7

Table 8 Test performance of the XGBoost Model

Metric	Value
RMSE	0.130
MAE	0.065
R^2	0.979

The coefficient on $\ln(GDP)$ remains negative (−0.483) under all specifications, confirming the previously identified structural shift in high-income economies. While its significance is modestly reduced with Driscoll–Kraay and clustered errors (with standard errors of 0.189 and 0.196, respectively), the direction and economic interpretation remain stable.

$MOPE^{low}$ retains its strong and significant positive effect ($\varphi_2 = 0.0231$), reinforcing the role of small-to-medium enterprises in CWP output. Conversely, $MOPE^{high}$ remains negative and significant ($\varphi_3 = -0.0118$), suggesting persistent inefficiencies or diminishing scale effects among larger producers.

Although the coefficients for $CWPE$ and EC are numerically small due to their units (tonnes and GWh), they are consistently positive and statistically significant across all estimators, underlining the importance of export integration and energy access in pulp production.

Overall, the robustness of these results across multiple error specifications validates the empirical integrity of the linear model and establishes a solid benchmark for the machine learning analysis that follows.

Machine learning estimation: XGBoost prediction model results

To enrich the panel regression analysis and capture potential nonlinearities in predictor interactions, we employed the XGBoost algorithm to model $\ln(CWP)$ output. The model retained the same five explanatory variables as in the econometric framework: ($\ln(GDP)$, $MOPE^{low}$, $MOPE^{high}$, $CWPE$, EC).

Overall predictive accuracy

To evaluate the out-of-sample performance of the XGBoost model, 80% of the dataset was allocated for training, while the remaining 20% served as a holdout sample. Model performance was assessed using standard metrics, including $RMSE$, MAE, and the coefficient of determination R^2, as summarised in Table 8.

The model yielded an $RMSE$ of 0.130, an MAE of 0.065, and an R^2 of 0.979, indicating that approximately 98% of the variation in $\ln(CWP)$ was accurately explained by the predictive algorithm. These metrics suggest a remarkably good fit, with minimal average deviation between the predicted and actual values.

To further evaluate model calibration, Fig. 5 presents a scatter plot comparing predicted and actual $\ln(CWP)$ values within the test sample.

The resulting distribution exhibits a dense clustering around the 45-degree identity line, confirming the close alignment between predicted and actual values. Minor discrepancies appear predominantly at the lower end of the production spectrum, where the model tends to slightly underpredict output for observations with exceptionally low production levels. Nevertheless, the overall symmetry and low dispersion of predictions underscore the model's capacity to effectively extract latent patterns embedded within the dataset.

Overfitting control

To determine the extent to which the XGBoost model is prone to overfitting, we examined the trajectory of $RMSE$ across successive boosting iterations for both the training and test subsets. Figure 6 illustrates the $RMSE$ trends across 100 boosting rounds, separately for the training and test sets.

The results reveal a sharp decline in $RMSE$ for both datasets during the initial 20 boosting rounds, followed by a plateau phase in which the curves level off and maintain a nearly parallel trajectory. While, the training $RMSE$ continues to decline gradually, the test $RMSE$ stabilizes at a relatively low and consistent level. This behaviour is characteristic of a well-calibrated model that effectively captures underlying structure in the data without succumbing to overfitting or spurious noise.

The absence of any substantial divergence between the training and test error curves supports the conclusion that the model achieves an optimal bias variance trade-off, thereby retaining strong generalisation performance throughout the boosting process.

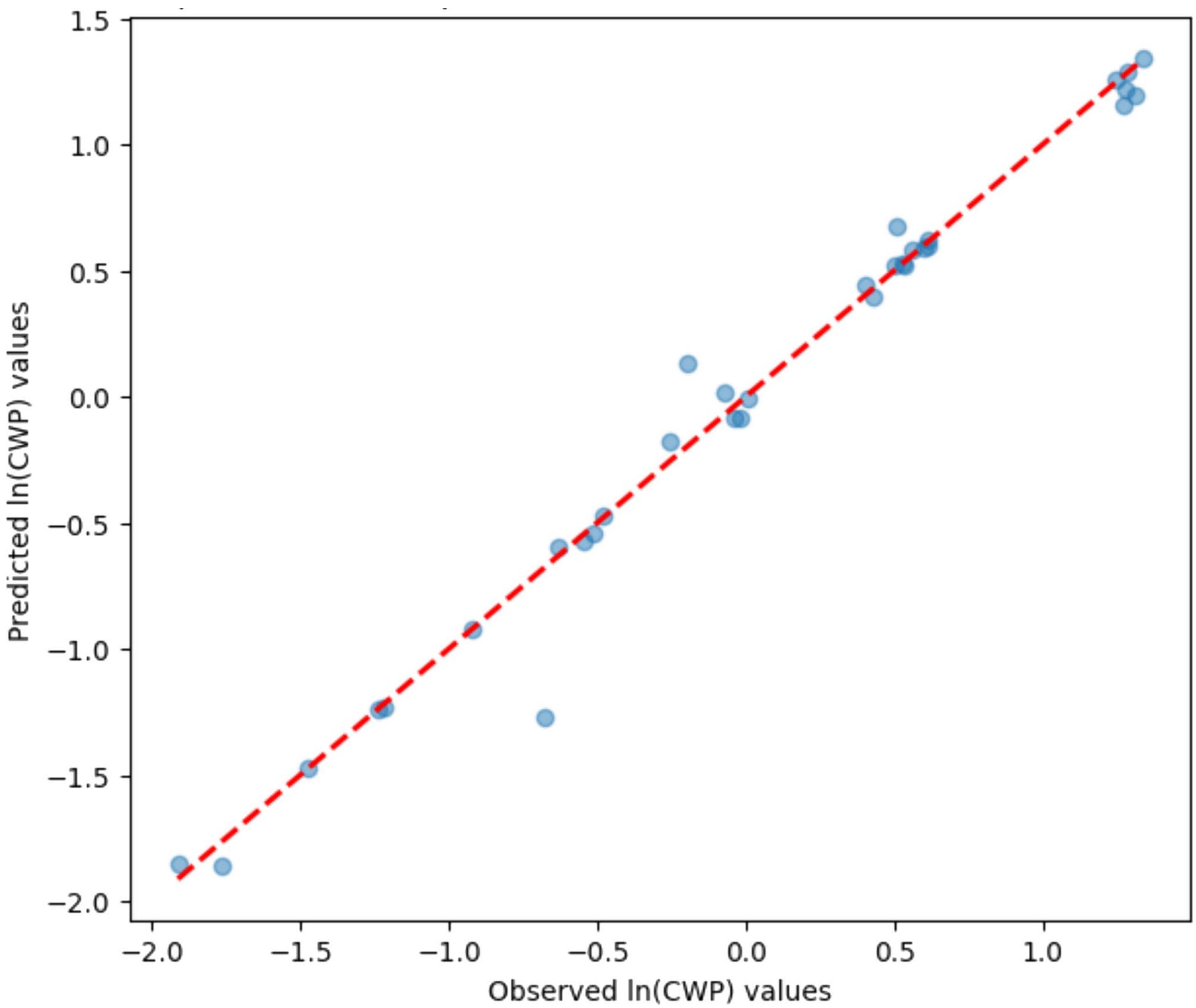

Fig. 5 Scatter plot of XGBoost predictions vs observed $\ln(CWP)$ in the test sample

Cross-validation stability

To assess the model's ability to generalise across varying data segments, a five-fold cross-validation procedure was implemented, allowing performance assessment across distinct training test splits. Table 9 provides a summary the $RMSE$, MAE, and R^2 values obtained for each cross-validation fold.

Although Folds 2 and 5 demonstrate slightly reduced R^2 values ranging between $0.87 - 0.92$, the model exhibits a consistent level of predictive performance across all partitions. The mean $RMSE$ of 0.225 and the average R^2 of 0.942 suggest that the model sustains a high level of predictive accuracy and exhibits limited error variability, even within the more challenging data folds.

Such robustness across validation folds highlights the underlying stability of the model's learned structure and reinforces its suitability as a dependable forecasting tool for chemical pulp production modelling.

Explaining model predictions with SHAP values

To enhance interpretability and address the limitations of black-box modelling, SHAP is used to quantify each predictor's contribution to the XGBoost outputs. Grounded in cooperative game theory, SHAP decomposes model predictions into additive feature contributions, offering both global and observation-specific insight (Lundberg and Lee 2017).

Figure 7 illustrates the SHAP summary plot, where each point represents a country–year observation. The horizontal axis shows the SHAP value, i.e., the marginal effect of each predictor, while the colour gradient indicates feature magnitude (blue: low, red: high). Variables are ranked by their mean absolute SHAP values.

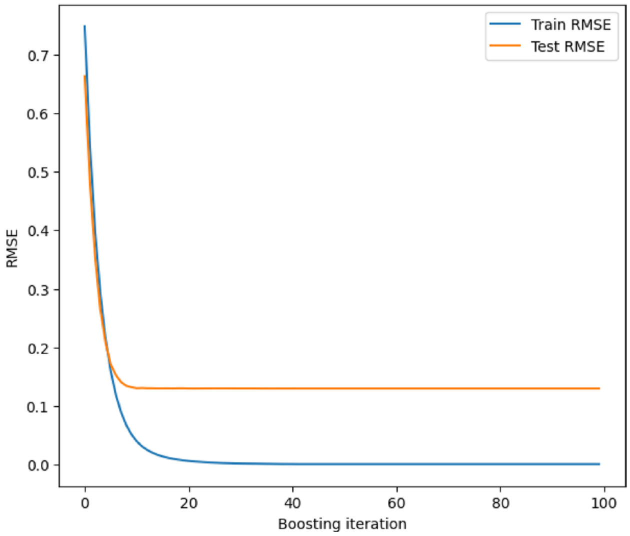

Fig. 6 Comparison of train and test RMSE over the number of boosting iterations in the XGBoost model

Table 9 Five-fold cross-validation results for the XGBoost model

Fold	*RMSE*	*MAE*	R^2
1	0.1302	0.0649	0.9755
2	0.3065	0.1299	0.9170
3	0.1527	0.0720	0.9763
4	0.1830	0.0792	0.9692
5	0.3535	0.1339	0.8703
Mean	0.2252	0.0960	0.9425

EC and $CWPE$ are the most influential predictors, with higher values driving up predicted output. $\ln(GDP)$ and $MOPE^{low}$ also exert meaningful, though less dominant, effects. Conversely, $MOPE^{high}$ and year contribute minimally, aligning with earlier panel results that suggested diminished marginal influence.

SHAP thus not only ranks feature relevance but also reveals directionality and nonlinear dynamics, offering a powerful interpretative complement to the model's predictive accuracy, critical for informed policy formulation in industrial strategy.

Robustness of model performance to hyperparameter settings

To test the generalisability of the XGBoost model, we assessed its performance across a wide range of hyperparameter combinations, focusing on the number of boosting rounds $(n\_estimators)$ and the maximum tree depth $(\max\_depth)$ (Fig. 8)A randomized grid search was conducted, averaging across multiple learning rates to account for interaction effects.. .

The heatmap demonstrates that $RMSE$ values remain consistently low and stable, ranging from 0.178 to 0.202, across a substantial portion of the hyperparameter space. The optimal configuration yielding the lowest $RMSE$ (0.178) corresponds n_estimators = 171, max_depth = 4, and a learning rate of 0.067. Notably, incremental performance improvements taper off beyond roughly 150 boosting iterations, indicating a saturation threshold in marginal predictive returns.

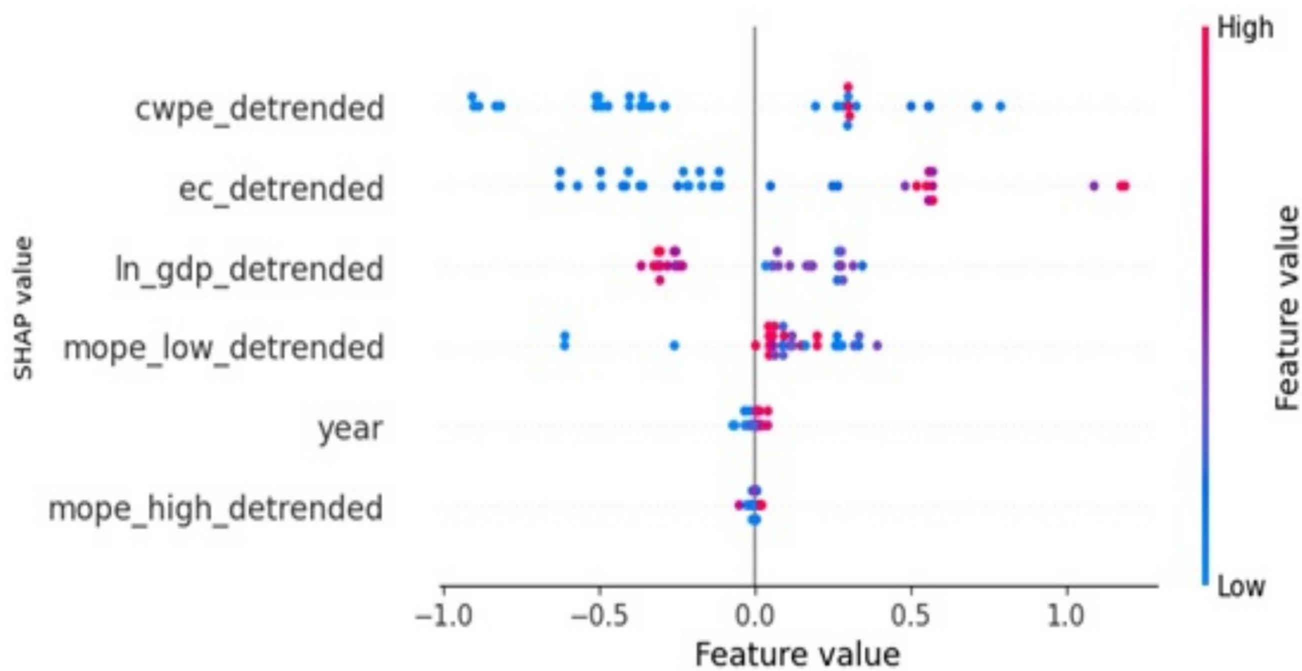

Fig. 7 SHAP summary plot showing predictor contributions to $\ln(CWP)$ across country-year observations

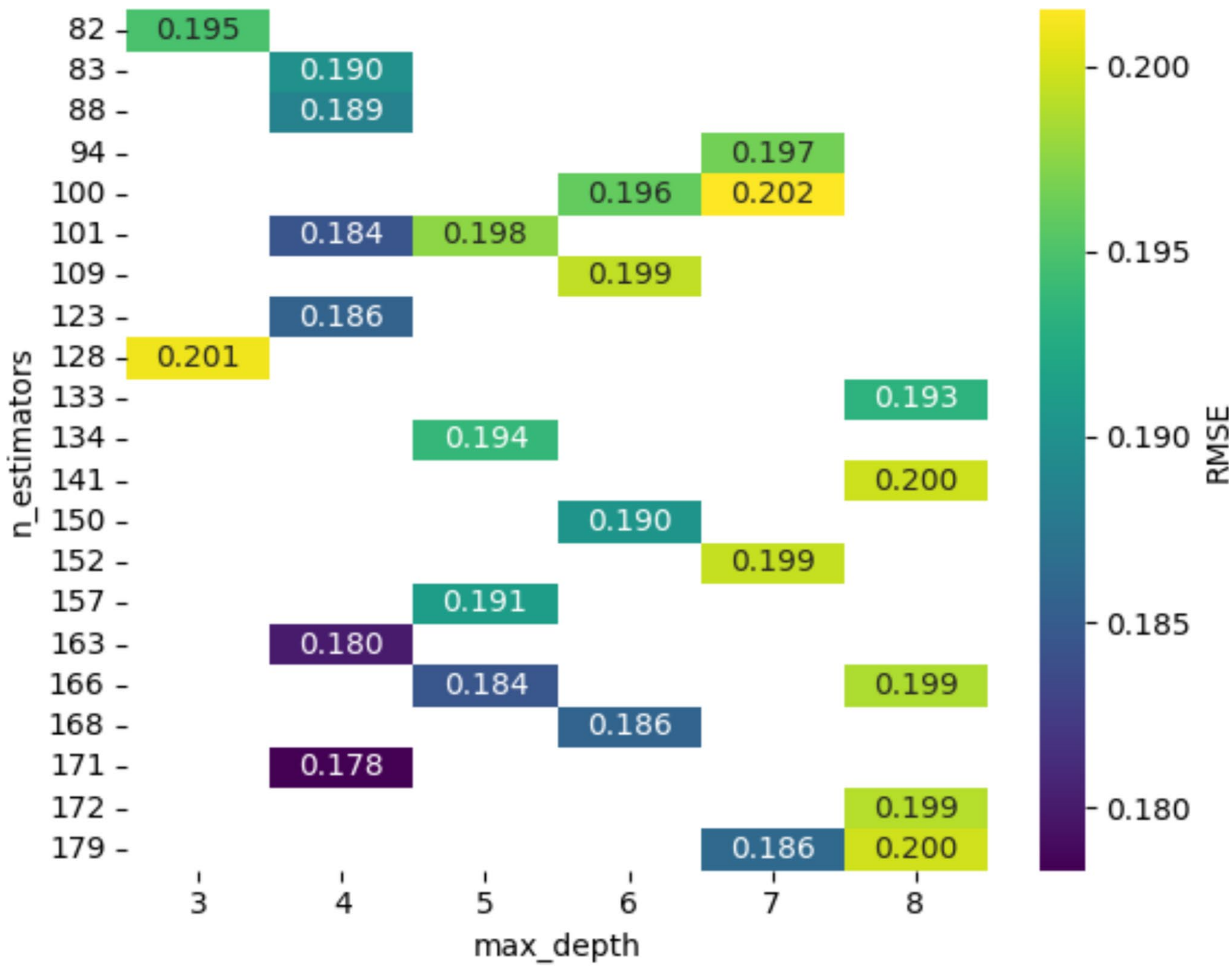

Fig. 8 Cross-validated RMSE heatmap for XGBoost performance across varying values of $n\_estimators$ and max_depth averaged over different learning rates

Furthermore, no signs of overfitting are detected, even with increased depth or complexity. Performance peaks with moderate complexity (max_depth between 4 and 6), while shallow or excessively deep trees offer diminishing returns.

These results confirm that XGBoost's performance is robust to parameter tuning, supporting its use in forecasting and bolstering the reliability of the study's insights.

Discussion: interpretation and comparison of econometric and machine learning results

This section interprets the empirical findings by explicitly relating each estimated determinant of chemical wood pulp (CWP) production to existing empirical and theoretical research, while also contrasting insights derived from econometric and machine learning approaches. In doing so, the discussion highlights both areas of convergence with the literature and instances where the present results diverge, offering context-specific explanations grounded in Europe's structural, institutional, and policy environment.

The econometric results reveal a negative and statistically significant relationship between $\ln(GDP)$ and $\ln(CWP)$ production, a finding that contrasts with traditional growth–industry narratives which often posit a positive association between aggregate economic expansion and industrial output (McCarthy and Lei 2010; Hujala et al. 2013; Charles et al. 2024; Liang et al. 2025). However, this result is strongly aligned with recent literature emphasizing structural transformation, deindustrialization, and economic decoupling in advanced economies (Van Neuss 2018; Montagna et al. 2025). Studies focusing on European and high-income contexts document that GDP growth increasingly reflects expansion in services, digital activities, and high-value manufacturing rather than scale-intensive, resource-based industries (Lagakos and Shu 2023; Aituar et al. 2025). In this sense, the negative GDP coefficient observed in this study supports the decoupling hypothesis identified in multiple European-focused analyses, where material-intensive production stagnates or declines despite continued economic growth (Huttmanová et al. 2024; Ziemblinska et al. 2025). While some earlier studies report positive GDP–industrial output linkages (McCarthy and Lei 2010; Hujala 2011), such discrepancies can be attributed to differences in development stages, sectoral composition, and policy regimes. In Europe, stringent environmental regulation, high labour costs, and the strategic shift toward a circular bioeconomy appear to weaken the traditional GDP–output nexus for pulp production.

Firm structure emerges as a critical determinant of CWP output, with the econometric estimates indicating a positive and significant effect of $MOPE^{low}$ and a negative effect of $MOPE^{high}$. This pattern is broadly consistent with firm-level and sectoral literature emphasizing the role of small and medium-sized enterprises (SMEs) in driving adaptability, incremental innovation, and regional embeddedness (Sikandar et al. 2024; Albuquerque et al. 2025). Studies examining scale effects similarly note that smaller firms often outperform large incumbents in environments characterized by regulatory complexity and rapidly evolving market conditions (Liao et al. 2025; Mac Clay et al. 2025). Conversely, the negative association between large firm concentration ($MOPE^{high}$) and output aligns with research documenting diminishing returns to scale, coordination inefficiencies, and reduced innovation responsiveness in highly concentrated industrial structures (Aghion et al. 2024; Fan et al. 2025; Sarin et al. 2025). While some studies as Lah and Kotnik (2025) find positive scale effects in capital-intensive industries, the divergence observed here can be explained by Europe's institutional context, where environmental compliance costs, energy price volatility, and labour rigidities disproportionately burden large, centralized producers (Kalantzis et al. 2025).

Export orientation, captured by $CWPE$, exhibits a positive and statistically robust relationship with CWP production, corroborating a substantial body of trade-oriented literature. Research on global pulp markets consistently emphasizes that access to international demand is a key driver of production sustainability, particularly in regions facing stagnant or declining domestic consumption (Hayafune and Tachibana 2024; Heydenreich 2024). The present findings by Allevi et al. (2024) align closely with these studies, suggesting that European producers rely on export channels to maintain output levels amid structural changes in internal markets. Where discrepancies arise in the literature such as weaker export effects in emerging economies, they can largely be attributed to differences in trade openness, logistics infrastructure, and market power, all of which favour European producers in global pulp value chains (Wang et al. 2024a, b).

Electricity consumption (EC) also displays a consistently positive and significant effect on CWP production, reinforcing the characterization of chemical pulping as an energy-intensive industrial process. This result is fully consistent with prior empirical evidence linking industrial output to energy availability and cost structures (Anser et al. 2024). Notably, the persistence of this relationship after detrending suggests that energy dependence is structural rather than cyclical. Although recent empirical inquiries suggest that the 'technique effect' and efficiency mandates can decouple energy use from industrial growth (Caglar et al. 2024; Gajdzik et al. 2024), and specifically reduce energy intensity in the pulp and paper sector through technological upgrades (Chauhan and Mohanasundari 2025), the European pulp sector's reliance on continuous-process technologies and electrified production systems likely sustains this dependence, particularly under increasingly volatile energy markets shaped by geopolitical tensions and decarbonization policies.

Comparing econometric and machine learning results further enriches the interpretation of these findings. The panel regression framework provides clear elasticity-based interpretations and facilitates hypothesis testing, making it well suited for identifying average marginal effects and policy-relevant relationships. In contrast, the XGBoost model achieves substantially higher predictive accuracy and captures nonlinear interactions and threshold effects that linear models cannot easily detect. This distinction mirrors findings in the broader macroeconomic and industrial forecasting literature, where machine learning models consistently outperform traditional econometric approaches in prediction tasks (Furukawa et al. 2024; Sofianos et al. 2025). Importantly, the SHAP-based interpretation confirms that EC and CWPE are the most influential predictors of CWP production, thereby reinforcing the econometric conclusions rather than contradicting them. At the same time, SHAP results reveal heterogeneity in the influence of GDP and firm structure across observations, offering a plausible explanation for why average econometric effects may mask context-specific dynamics.

Overall, the convergence between econometric inference and machine learning interpretability strengthens confidence in the study's findings. Where discrepancies with existing research arise, they can largely be explained by Europe's unique socioeconomic and political context, characterized by advanced stages of structural transformation, stringent environmental regulation, integrated energy markets, and strong export orientation. By explicitly comparing results across methods and against the relevant literature, the analysis provides a nuanced and context-aware understanding of the drivers of chemical wood pulp production in Europe.

Conclusion and policy implications

This study examined the macroeconomic, structural, and trade-related determinants of chemical wood pulp (CWP) production across eleven European economies between 2009 and 2023 by integrating econometric panel estimation with machine learning–based predictive modelling. The joint analysis of GDP, electricity consumption, export intensity, and enterprise-scale heterogeneity provides a comprehensive and contextually grounded understanding of the production dynamics of the European pulp sector.

The econometric results show that electricity consumption and export orientation exert persistently positive and statistically significant effects on CWP output, confirming the structurally energy-intensive and trade-dependent nature of the industry. The decomposition of firm structure into $MOPE^{low}$ and $MOPE^{high}$ reveals an asymmetric scale pattern whereby smaller and more distributed production structures contribute positively to output, while highly concentrated industrial configurations are associated with declining marginal performance. The negative relationship observed between GDP and pulp production suggests a decoupling process characteristic of mature European economies, where economic growth increasingly derives from service-based and knowledge-intensive sectors rather than resource-dependent manufacturing activities.

These findings are reinforced by the machine learning analysis. The XGBoost model demonstrates strong predictive accuracy, while SHAP-based interpretation confirms the dominant role of energy consumption and export intensity and reveals heterogeneous marginal effects of GDP and firm structure across countries. The convergence between inferential and predictive evidence strengthens the robustness of the empirical results and enhances their policy relevance.

Taken together, the results indicate that the sustainability and competitiveness of the European pulp sector are closely linked to energy security, export-market integration, and the structural configuration of the industrial base. Policies that stabilise electricity costs, support energy-efficient production systems, and mitigate exposure to volatility in European energy markets appear essential for maintaining industrial output capacity. Likewise, the evidence suggests that reinforcing the operational resilience of small and medium-sized enterprises, rather than promoting excessive concentration among large incumbents, may foster productivity, innovation responsiveness, and regional industrial stability. At the same time, the strong export dependence of the sector underscores the importance of trade logistics, market diversification strategies, and continued integration into global value chains within the framework of the EU's bioeconomy and decarbonisation agenda.

Despite its contributions, this study is subject to several methodological and empirical constraints. The reliance on macro-panel indicators limits the ability to capture plant-level technological heterogeneity and organisational dynamics, while potential bidirectional interactions between industrial activity and energy use cannot be fully ruled out. Moreover, the findings reflect the institutional and policy environment of Europe and should be interpreted cautiously when extrapolated to emerging or resource-abundant economies.

Future research could therefore extend the present framework by incorporating firm- or plant-level microdata to explore productivity mechanisms in greater depth, by employing causal identification strategies to address endogeneity more explicitly, and by conducting comparative analyses across world regions to determine whether the observed decoupling, energy dependence, and export-orientation patterns represent Europe-specific behaviour or a broader global industrial transition.

Acknowledgements The author would like to express their sincere gratitude to Dr. Emre Akturk for his valuable insights and unwavering encouragement throughout the research process. Appreciation is also extended to Kastamonu University for providing a supportive academic environment that fostered the development of this study. Lastly, the author acknowledges the Council of Higher Education of Türkiye (CoHE) for their institutional support and commitment to advancing higher education and research in the country.

Author contributions H.A. conceived the study, collected and analyzed the data, and wrote the manuscript. The author reviewed and approved the final version of the manuscript.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Data availability The datasets analyzed during the current study are publicly available and can be accessed from the following sources: Chemical Wood Pulp (CWP) Production: FAOSTAT available at: https://www.fao.org/faostat/en/#data/FO Economic structure (GDP-costant 2015 US$): World bank available at: https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.KD&country=# Energy consumption (EC): Eurostat available at: https://ec.europa.eu/eurostat/databrowser/view/ten00121/default/table?lang=en&category=t_nrg.t_nrg_indic Manufacture of Pulp and Paper Enterprises (MOPE): Eurostat https://ec.europa.eu/eurostat/databrowser/view/sbs_sc_sca_r2__custom_16326449/default/table?lang=en Chemical Wood Pulp Export (CWPE): FAOSTAT available at: https://www.fao.org/faostat/en/#data/FO

Declarations

Conflict of interest The authors declare no competing interests.

References

Aghion P, Bergeaud A, Lequien M, Melitz MJ (2024) The heterogeneous impact of market size on innovation: evidence from French firm-level exports. Rev Econ Stat 106:608–626. https://doi.org/10.1162/rest_a_01199

Aituar A, Bolatbay O, Abdukadyrov N, Abdukhamet S (2025) Forks in the road: globalization, deindustrialization, and economic growth pathways. Economies 13:348. https://doi.org/10.3390/economies13120348

Akhtar R, Hussain J, Li L, Cheng B, Cuong T, Haider S et al (2025) Unlocking emerging economies: exploring the influence of international trade on strategic options through logistic regression approach. J Knowl Econ 16:7845–7873. https://doi.org/10.1007/s13132-024-02072-y

Albuquerque F, Ferrão J, dos Santos PG (2025) Labour productivity in European non-financial corporations: the roles of country, sector, and size. J Risk Financ Manag 18:647. https://doi.org/10.3390/jrfm18110647

Ali MI, Islam MM, Ceh B (2024) Growth-environment nexus in Canada: revisiting EKC via demand and supply dynamics. Energy Environ. https://doi.org/10.1177/0958305X2412638

Ali MI, Islam MM, Ceh B (2025a) Assessing the impact of three emission (3E) parameters on environmental quality in Canada: a provincial data analysis using the quantiles via moments approach. Int J Green Energy 22:551–569. https://doi.org/10.1080/15435075.2024.2421329

Ali MI, Islam MM, Ceh B (2025b) Interaction between decomposed energy utilization and environmental health in Canada: a cointegration and counterfactual analysis approach. Int J Energy Res. https://doi.org/10.1155/er/1173970

Ali MI, Rahaman MA, Ali MJ, Rahman MF (2025c) The growth–environment nexus amid geopolitical risks: cointegration and machine learning algorithm approaches. Discov Sustain 6:78. https://doi.org/10.1007/s43621-025-00872-z

Allevi E, Gnudi A, Konnov IV, Oggioni G, Riccardi R (2024) A spatial equilibrium problem for the European pulp and paper industry under the Emission Trading System. J Optim Theory Appl 203(2):1767–1793. https://doi.org/10.1007/s10957-023-02291-4

Andersson E, Dernegård H, Wallén M, Thollander P (2021) Decarbonization of industry: implementation of energy performance indicators for successful energy management practices in kraft pulp mills. Energy Rep 7:1808–1817. https://doi.org/10.1016/j.egyr.2021.03.009

Andre C, Costa H, Demmou L, Franco G (2023) Rising energy prices and productivity: short run pain, long term gain? OECD Economics Department Working Papers, No. 1755. https://doi.org/10.1787/2ce493f0-en

Anser MK, Ali S, Umair M, Javid R, Mirzaliev S (2024) Energy consumption, technological innovation, and economic growth in BRICS: a GMM panel VAR framework analysis. Energy Strateg Rev 56:101587. https://doi.org/10.1016/j.esr.2024.101587

Antonelli C, Orsatti G, Pialli G (2023) The knowledge-intensive direction of technological change. Eurasian Bus Rev 13:1–27. https://doi.org/10.1007/s40821-022-00234-z

Arnič D, Loizou E, Ščap Š, Prislan P, Juvančič L (2024) Evaluating alternative transformation pathways of wood-based bioeconomy: application of an input–output model. Forests 15:2084. https://doi.org/10.3390/f15122084

Aydin M, Koc P, Tumay M (2023) Investigating the environmental Kuznets curve hypothesis with recovered paper consumption, human development index, urbanization, and forest footprint. Int J Environ Sci Technol 20:9053–9064. https://doi.org/10.1007/s13762-023-05049-7

Aydoğan H, Yıldırım H (2025) Effects of productivity, R&D and income inequality on economic growth: an empirical estimation using panel data. Verimlilik Dergisi 59:47–60. https://doi.org/10.51551/verimlilik.1510434

Balkissoon S, Andrew J, Sithole B (2023) Dissolving wood pulp production: a review. Biomass Conv Bioref 13:16607–16642. https://doi.org/10.1007/s13399-022-02442-z

Baltagi BH (2008a) Forecasting with panel data. J Forecast 27:153–173. https://doi.org/10.1002/for.1047

Baltagi BH (2008b) Time-series analysis. In: Econometrics. 355–377. https://doi.org/10.1007/978-3-540-76516-5_14

Baltagi BH (2021) Introduction: econometric analysis of panel data, Springer texts in business and economics, 6th edn. Springer, Cham

Baltagi BH, Kao C (2001) Nonstationary panels, cointegration in panels and dynamic panels: A survey. In: Nonstationary panels, panel cointegration, and dynamic panels. Emerald Group Publishing Limited, pp 7–51. https://doi.org/10.1016/S0731-9053(00)15002-9

Breitung J (2001) The local power of some unit root tests for panel data. Nonstationary panels, panel cointegration, and dynamic panels. Emerald Group Publishing Limited, UK, pp 161–177. https://doi.org/10.1016/S0731-9053(00)15006-6

Caglar AE, Daştan M, Bulut E, Marangoz C (2024) Evaluating a pathway for environmental sustainability: the role of competitive industrial performance and renewable energy consumption in European countries. Sustain Dev 32(3):1811–1824. https://doi.org/10.1002/sd.2755

Calì M, Cantore N, Iacovone L, Pereira-López M, Presidente G (2022) Too much energy: the perverse effect of low fuel prices on firms. J Environ Econ Manag 111:102587. https://doi.org/10.1016/j.jeem.2021.102587

Calì M, Cantore N, Marin G, Mazzanti M, Nicolli F, Presidente G (2023) Energy prices and the economic performance of firms in emerging countries. Struct Change Econ Dyn 66:357–366. https://doi.org/10.1016/j.strueco.2023.04.019

CEPF (2025) CEPF position on the new EU bioeconomy strategy. Confederation of European Forest Owners, Brussels. https://www.cepf-eu.org/sites/default/files/2025-09/CEPF_position_EU_Bioeconomy_Strategy.pdf. Accessed 20 December 2025

Chaldun ER, Yudoko G, Prasetio EA (2024) Developing a theoretical framework of export-oriented small enterprises: a multiple case study in an emerging country. Sustainability 16:11132. https://doi.org/10.3390/su162411132

Chao WS, Buongiorno J (2002) Exports and growth: a causality analysis for the pulp and paper industries based on international panel data. Appl Econ 34(1):1–13. https://doi.org/10.1080/00036840010027531

Charles M, Narayan KB, Edmonds J, Yu S (2024) The role of the pulp and paper industry in achieving net zero US CO2 emissions in 2050. Energy Clim Change 5:100160. https://doi.org/10.1016/j.egycc.2024.100160

Chauhan A, Mohanasundari T (2025) The impact of the perform, achieve, and trade (PAT) scheme on energy intensity: evidence from the pulp and paper industry. Energy Sour B Econ Plann Policy 20(1):2588253. https://doi.org/10.1080/15567249.2025.2588253
Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 785–794. https://doi.org/10.1145/2939672.2939785
Copeland BR, Taylor MS (2004) Trade, growth, and the environment. J Econ Lit 42:7–71. https://doi.org/10.1257/002205104773558047
Croissant Y, Millo G (2019) Panel data econometrics with R. Wiley, UK. https://doi.org/10.1002/9781119504641
Dabbous A, Barakat KA, Kraus S (2023) The impact of digitalization on entrepreneurial activity and sustainable competitiveness: a panel data analysis. Technol Soc 73:102224. https://doi.org/10.1016/j.techsoc.2023.102224
Del Rio DF, Sovacool BK, Griffiths S, Bazilian M, Kim J, Foley AM, Rooney D (2022) Decarbonizing the pulp and paper industry: a critical and systematic review of sociotechnical developments and policy options. Renew Sustain Energy Rev 167:112706. https://doi.org/10.1016/j.rser.2022.112706
Evelyn E, Okewale AO, Owabor CN (2025) Optimization, kinetics and thermodynamic modeling of pulp production from plantain stem using the Kraft process. Clean Chem Eng 11:100129. https://doi.org/10.1016/j.clce.2024.100129
Fan Z, Zhang X, Zhang Y, Kang M, Cui Q (2025) The impact of digital–real economy integration on ESG performance: the moderating role of organizational inertia. Sustainability 17(23):10670. https://doi.org/10.3390/su172310670
FAO (2025) FAOSTAT statistical database: forestry production and trade. https://www.fao.org/faostat/en/#data/FO
Farnsworth DL (2023) Transforming data in a first course in statistics. Int J Math Educ Sci Technol 54(6):1146–1152. https://doi.org/10.1080/0020739X.2022.2086499
Fracaro G, Vakkilainen E, Hamaguchi M, Melegari de Souza SN (2012) Energy efficiency in the Brazilian pulp and paper industry. Energies 5:3550–3572. https://doi.org/10.3390/en5093550
Fresner J, Morea F, Krenn C, Uson JA, Tomasi F (2017) Energy efficiency in small and medium enterprises: lessons learned from 280 energy audits across Europe. J Clean Prod 142:1650–1660. https://doi.org/10.1016/j.jclepro.2016.11.126
Furukawa K, Hisano R, Minoura Y, Yagi T (2024) A nowcasting model of industrial production using alternative data and machine learning approaches. Jpn World Econ 71:101271. https://doi.org/10.1016/j.japwor.2024.101271
Gajdzik B, Nagaj R, Wolniak R, Bałaga D, Žuromskaitė B, Grebski WW (2024) Renewable energy share in European industry: analysis and extrapolation of trends in EU countries. Energies 17(11):2476. https://doi.org/10.3390/en17112476
Golmohamadi H (2022) Demand-side management in industrial sector: a review of heavy industries. Renew Sustain Energy Rev 156:111963. https://doi.org/10.1016/j.rser.2021.111963
Haile A, Gebino G, Tesfaye T, Mengie W, Ayele M, Abuhay A, Yilie D (2021) Utilization of non-wood biomass for pulp manufacturing in paper industry: case of Ethiopia. Biomass Conv Biorefinery 13:7441–7459. https://doi.org/10.1007/s13399-021-01424-x
Han QL, Zhao HL, Wei GX, Zhu YW, Li T, Xu M, Liu HQ (2024) Sustainable papermaking in China: assessing provincial economic and environmental performance of pulping technologies. ACS Sustain Chem Eng 12(11):4517–4529. https://doi.org/10.1021/acssuschemeng.3c07611
Hausman JA (1978) Specification tests in econometrics. Econometrica 46:1251–1271. https://doi.org/10.2307/1913827
Hayafune M, Tachibana S (2024) Construction and transfiguration of the raw material procurement system of woodchips in the Japanese pulp and paper industry. J for Res 29(4):260–269. https://doi.org/10.1080/13416979.2024.2315539
Herwartz H, Maxand S, Walle YM (2023) Forward detrending for heteroskedasticity-robust panel unit root testing. Econom Rev 42(1):28–53. https://doi.org/10.1080/07474938.2022.2135495
Heydenreich T (2024) Trade and decoupling of fossil fuel use embedded in EU consumption. J Clean Prod 464:142702. https://doi.org/10.1016/j.jclepro.2024.142702
Hoechle D (2007) Robust standard errors for panel regressions with cross-sectional dependence. Stata J 7(3):281–312. https://doi.org/10.1177/1536867X0700700301
Hujala M (2011) The role of information and communication technologies in paper consumption. Int J Bus Inf Syst 7(2):121–135. https://doi.org/10.1504/IJBIS.2011.038507
Hujala M, Arminen H, Hill RC, Puumalainen K (2013) Explaining the shifts of international trade in pulp and paper industry. For Sci 59(2):211–222. https://doi.org/10.5849/forsci.11-078
Huttmanová E, Chovancová J, Petruška I, Sallaku K (2024) Circular economy in action: examining the decoupling of economic growth and material use across EU countries. Eur J Sustain Dev 13(4):13–22. https://doi.org/10.14207/ejsd.2024.v13n4p13
Im KS, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. J Econometrics 115:53–74. https://doi.org/10.1016/S0304-4076(03)00092-7
Ince PJ, Kramp AD, Skog KE, Yoo DI, Sample VA (2011) Modeling future US forest sector market and trade impacts of expansion in wood energy consumption. J for Econ 17:142–156. https://doi.org/10.1016/j.jfe.2011.02.007
Islam MM, Ali MI, Ceh B, Singh S, Khan MK, Dagar V (2022) Renewable and non-renewable energy consumption driven sustainable development in ASEAN countries: do financial development and institutional quality matter? Environ Sci Pollut Res 29:34231–34247. https://doi.org/10.1007/s11356-021-18488-x
Jaunky VC, Lundmark R (2017) Dynamics of wood pulp production: evidence from OECD countries. Forests 8(4):107. https://doi.org/10.3390/f8040107
Johansson MT, Broberg S, Ottosson M (2021) Energy strategies in the pulp and paper industry in Sweden: interactions between efficient resource utilisation and increased product diversification. J Clean Prod 311:127681. https://doi.org/10.1016/j.jclepro.2021.127681
Kalantzis F, Khalid S, Solovyeva A, Wolski M (2025) Firms' response to climate regulations: empirical investigations based on the European Emissions Trading System. Energy Policy. https://doi.org/10.1016/j.enpol.2025.114612
Kando H, Buongiorno J (2009) Efficiency in wood and fiber utilization in OECD countries. J for Res 14:321–327. https://doi.org/10.1007/s10310-009-0142-8
Karlsson L (2025) Structural change as the primary driver of water pollution reductions: re-evaluating the role of command and control in the Swedish pulp and paper industry (1969–1989). Bus Hist:1–25. https://doi.org/10.1080/00076791.2025.2532788
Klugman S, Karlsson M, Moshfegh B (2007) A Scandinavian chemical wood pulp mill. Part 1. energy audit aiming at efficiency measures. Appl Energy 84:326–339. https://doi.org/10.1016/j.apenergy.2006.07.003
Kuzminov IF, Lobanova PA (2021) Economic and geographical analysis of the current state of the pulp and paper industry of European Russia: how is text mining helpful? Reg Res Russ 11:477–489. https://doi.org/10.1134/S2079970521040092
Lagakos D, Shu M (2023) The role of micro data in understanding structural transformation. Oxf Dev Stud 51:436–454. https://doi.org/10.1080/13600818.2023.2278601

Lah LM, Kotnik Ž (2025) Drivers of environmental regulation costs and their perception by firms: evidence from Slovenia. J Clean Prod 501:145274. https://doi.org/10.1016/j.jclepro.2025.145274

Lawrence A, Nehler T, Andersson E, Karlsson M, Thollander P (2019) Drivers, barriers and success factors for energy management in the Swedish pulp and paper industry. J Clean Prod 223:67–82. https://doi.org/10.1016/j.jclepro.2019.03.143

Levin A, Lin CF, Chu CSJ (2002) Unit root tests in panel data: asymptotic and finite-sample properties. J Econometrics 108:1–24. https://doi.org/10.1016/S0304-4076(01)00098-7

Li Y, Chan HK, Zhang T (2025) Environmental production and productivity growth: evidence from European paper and pulp manufacturing. Ann Oper Res 349:477–494. https://doi.org/10.1007/s10479-018-3126-2

Liang X, Hong M, Man Y, He Z (2025) Dynamics and scenarios of embodied carbon emissions in China's pulp and paper industry. Resour Conserv Recycl 220:108364. https://doi.org/10.1016/j.resconrec.2025.108364

Liao C, Parkhurst G, Parker CJ (2025) Future flight's dynamic capabilities and circular economy opportunities. Sustain Technol Entrep. https://doi.org/10.1016/j.stae.2025.100114

Lintunen J, Kohl J, Buchert J, Asikainen A, Jyske T, Maunuksela J, Lehto J (2024) 2035 Vision: Doubling the Value Added of Finland´s Forest sector. Natural resources and bioeconomy studies 15/2024. Natural Resources Institute Finland, Helsinki. p. 21

Lipiäinen S, Kuparinen K, Sermyagina E, Vakkilainen E (2022) Pulp and paper industry in energy transition: towards energy-efficient and low carbon operation in Finland and Sweden. Sustain Prod Consum 29:421–431. https://doi.org/10.1016/j.spc.2021.10.029

Luhas J, Mikkilä M, Uusitalo V, Linnanen L (2019) Product diversification in sustainability transition: the forest-based bioeconomy in Finland. Sustainability 11:3293. https://doi.org/10.3390/su11123293

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 30.

Mac Clay P, Bort G, Feeney R, Sellare J (2025) Industrial dynamics and business strategies in the emergence of agricultural biological inputs. Bus Strateg Environ 34(6):6472–6487. https://doi.org/10.1002/bse.4314

Mahapatra B, Irfan M (2024) Asymmetric adjustments between energy and labour efficiencies in India: new evidence using sectoral panel data analysis. Qual Quant 58(4):3921–3948. https://doi.org/10.1007/s11135-024-01855-4

McCarthy P, Lei L (2010) Regional demands for pulp and paper products. J for Econ 16(2):127–144. https://doi.org/10.1016/j.jfe.2009.11.001

Mishra L, Acharya RH (2025) The asymmetric interaction between oil price change and stock returns of the renewable energy companies in India: a panel NARDL approach. Asia-Pac Financ Mark 32:103–116. https://doi.org/10.1007/s10690-024-09447-w

Mitić P, Kojić M, Minović J, Stevanović S, Radulescu M (2024) An EKC-based modelling of CO2 emissions, economic growth, electricity consumption and trade openness in Serbia. Environ Sci Pollut Res 31:5807–5825. https://doi.org/10.1007/s11356-023-31617-y

Montagna S, Huang L, Long Y, Yoshida Y (2025) Sectoral economic complexity and environmental degradation: a sectoral perspective on the EKC hypothesis. Humanit Soc Sci Commun 12(1):1–24. https://doi.org/10.1057/s41599-025-04820-0

Montes-Pajuelo R, Rodríguez-Pérez ÁM, López R, Rodríguez CA (2024) Analysis of probability distributions for modelling extreme rainfall events and detecting climate change: insights from mathematical and statistical methods. Mathematics 12(7):1093. https://doi.org/10.3390/math12071093

Neykov N, Krišťáková S, Antov P, Halalisan AF, Hajduchova I, Sedliačiková M, Šišák L (2022) Capital structure determinants of forest enterprises: empirical study based on panel data analysis from the Czech Republic, Slovakia, and Bulgaria. Forests 13(5):749. https://doi.org/10.3390/f13050749

Nikou V, Sardianou E (2025) Determinants of paper packaging waste: evidence across 28 EU countries. J Mater Cycles Waste Manag 27:1541–1556. https://doi.org/10.1007/s10163-025-02195-7

Nosheen M, Iqbal J, Abbasi MA (2021) Do technological innovations promote green growth in the European Union? Environ Sci Pollut Res 28:21717–21729. https://doi.org/10.1007/s11356-020-11926-2

Ochuodho TO, Johnston CMT, Withey P (2017) Assessing economic impacts of internet adoption through reduced pulp and paper demand. Can J for Res 47:1381–1391. https://doi.org/10.1139/cjfr-2017-0014

Ottosson M, Magnusson T (2013) Socio-technical regimes and heterogeneous capabilities: the Swedish pulp and paper industry's response to energy policies. Technol Anal Strat Manag 25(4):355–368. https://doi.org/10.1080/09537325.2013.774349

Pesaran MH (2021) General diagnostic tests for cross-sectional dependence in panels. Empir Econ 60:13–50. https://doi.org/10.1007/s00181-020-01875-7

Phillips PC, Moon HR (2000) Nonstationary panel data analysis: an overview of some recent developments. Econom Rev 19(3):263–286. https://doi.org/10.1080/07474930008800473

Porter ME, van der Linde C (1995) Toward a new conception of the environment-competitiveness relationship. J Econ Perspect 9:97–118. https://doi.org/10.1257/jep.9.4.97

Posch A, Brudermann T, Braschel N, Gabriel M (2015) Strategic energy management in energy-intensive enterprises: a quantitative analysis of relevant factors in the Austrian paper and pulp industry. J Clean Prod 90:291–299. https://doi.org/10.1016/j.jclepro.2014.11.044

Ramos DNM, Rangel AS, Azevedo KS, Melo MGB, Oliveira MC, Watanabe CMU, Aguiar A (2022) Characteristics and treatment of Brazilian pulp and paper mill effluents: a review. Environ Monit Assess 194(9):651. https://doi.org/10.1007/s10661-022-10331-1

Rapposelli A, Birindelli G, Modina M (2024) The relationship between firm size and efficiency: why does default on bank loans matter? Qual Quant 58(4):3379–3401. https://doi.org/10.1007/s11135-023-01810-9

Reis Mourão PJ, Popescu IA (2025) Regional analysis of business agglomeration patterns in knowledge-intensive service sectors in Romania. Ann Reg Sci 74:1–36. https://doi.org/10.1007/s00168-024-01331-9

Sarin S, Im S, Di Benedetto A, Sego T, Chitturi R (2025) Contingent influences of innovation characteristics and firm capabilities on early-mover advantage: A conceptual framework for high-tech environments. AMS Rev. https://doi.org/10.1007/s13162-025-00327-8

Shabbir I, Mirzaeian M, Sher F (2022) Energy efficiency improvement potentials through energy benchmarking in pulp and paper industry. Clean Chem Eng 3:100058. https://doi.org/10.1016/j.clce.2022.100058

Shen X, Lovrić M (2022) Structural determinants of global trade in graphic paper and pulp products. Forest Policy Econ 134:102629. https://doi.org/10.1016/j.forpol.2021.102629

Sikandar H, Kohar UHA, Sanda G et al (2024) Eco-innovation in small and medium enterprises (SMEs): a systematic literature review. J Knowl Econ 15(2):5841–5863. https://doi.org/10.1007/s13132-023-01367-w

Smith-Hall C, Piplani M, Pyakurel D (2024) Theorising and analysing the forest-based bioeconomy through a global production network lens. Forest Policy Econ 159:103128. https://doi.org/10.1016/j.forpol.2023.103128

Sofianos E, Alexakis C, Gogas P, Papadimitriou T (2025) Machine learning forecasting in the macroeconomic environment: the case of the US output gap. Econ Change Restruct 58(1):9. https://doi.org/10.1007/s10644-024-09849-w

Steinbrunner P (2024) Is larger really better? Productivity and firm size in European electricity generation sectors. J Clean Prod 446:141382. https://doi.org/10.1016/j.jclepro.2024.141382

Su X, Raza K, Waseem LA, Waheed R (2025) Fostering a green tomorrow: exploring the impact of economic fitness on CO2 reduction along the environmental Kuznets curve with capital and renewable energy. Pol J Environ Stud. https://doi.org/10.15244/pjoes/188187

Szabó L, Soria A, Forsström J, Keränen JT, Hytönen E (2009) A world model of the pulp and paper industry: demand, energy consumption and emission scenarios to 2030. Environ Sci Policy 12:257–269. https://doi.org/10.1016/j.envsci.2009.01.011

Tang S, Song W, Perez-Garcia J, Eastin IL (2015) An empirical analysis of China's wood pulp and recovered paper imports using an augmented gravity model approach. For Prod J 65(7–8):381–386

Urdaneta F, Kumar R, Marquez R, Vera RE, Franco J, Urdaneta I, Gonzalez RW (2024) Evaluating chemi-mechanical pulping processes of agricultural residues: high-yield pulps from wheat straw for fiber-based bioproducts. Ind Crops Prod 221:119379. https://doi.org/10.1016/j.indcrop.2024.119379

Van Neuss L (2018) Globalization and deindustrialization in advanced countries. Struct Chang Econ Dyn 45:49–63. https://doi.org/10.1016/j.strueco.2018.02.002

Virtanen T, Tuomaala M, Pentti E (2013) Energy efficiency complexities: a technical and managerial investigation. Manag Account Res 24(4):401–416. https://doi.org/10.1016/j.mar.2013.06.002

Wang Y, Liu K (2023) The effect of environmental regulation on firm productivity: evidence from pulp and paper industry in China. Econ Res-Ekon Istraz 36(1):264–295. https://doi.org/10.1080/1331677X.2022.2076142

Wang F, Meng X, Liu F, Cheng B, Tian M (2024a) Market power of China's chemical wood pulp imports: based on PTM model. SSRN. https://doi.org/10.2139/ssrn.4962349

Wang X, Chang BH, Uche E, Zhao Q (2024b) The asymmetric effect of income and price changes on the consumption expenditures: evidence from G7 countries using nonlinear bounds testing approach. Port Econ J 23(1):35–53. https://doi.org/10.1007/s10258-022-00229-8

Wang F, Meng X, Cheng B, Tian M (2025) Market power of China's chemical wood pulp imports based on the pricing-to-market model. For Prod J 75:404–415. https://doi.org/10.13073/FPJ-D-25-00008

Westerlund J (2015) The effect of recursive detrending on panel unit root tests. J Econom 185(2):453–467. https://doi.org/10.1016/j.jeconom.2014.06.015

Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT Press

Wooldridge JM (2016) Introductory econometrics: a modern approach 6rd ed. Cengage learning.

Xian Y, Ma P, Wang K (2025) Emission patterns and marginal abatement cost of CO2 of world's largest pulp and paper production sector: an enterprise based estimation of China. Appl Energy 397:126344. https://doi.org/10.1016/j.apenergy.2025.126344

Zhang JY, LaMarche R, Lin W, Tung W, Jiang C (2012) An investigation into wheat straw paper and wood pulp paper use at UBC. University of British Columbia. https://doi.org/10.14288/1.0108600

Ziemblińska K, Urbaniak M, Jinlong H, Olejnik J, Kundzewicz ZW (2025) Decoupling of economic growth and CO2 emissions in 11 European Union Member States in Central and Eastern Europe. Sustain Energy Technol Assess 82:104482. https://doi.org/10.1016/j.seta.2025.104482