

RESOURCE ARTICLE **OPEN ACCESS**

Harmony, Conflict, and Evolution of the Common Agricultural Policy in Europe: A Text Mining Survey

Shuang Liu¹ | Jasmin Wehner²  | Jan-Henning Feil³  | Xiaohua Yu¹ 

¹Department of Agricultural Economics and Rural Development, University of Göttingen, Göttingen, Germany | ²Dyson School of Applied Economics and Management, Cornell University, Ithaca, New York, USA | ³Martin-Luther-University Halle-Wittenberg, Institute of Agricultural and Nutritional Sciences, Halle (Saale), Germany

Correspondence: Xiaohua Yu (xyu@uni-goettingen.de)

Received: 27 November 2024 | **Revised:** 15 June 2025 | **Accepted:** 29 June 2025

Keywords: agricultural development stages | clustering | common agricultural policy | policy dimensions | text mining | trend analysis

ABSTRACT

The Common Agricultural Policy (CAP) represents one of the main initiatives of the European Union (EU) to develop and enhance agricultural sectors. This paper applies text mining and clustering to analyse changes in the CAP's main policy objectives based on 495 legislative texts adopted by the European Parliament. We develop an SSCIR-3F analysis framework that identifies five core CAP policy concerns—food security (S), sustainability (S), competitiveness (C), farmer income (I), and rural development (R), abbreviated as SSCIR—illustrates their internal conflicts using the conflict triangle, and explains their changing importance over time using agricultural development stage theory (food, farm, and future, 3F). The text mining results show that the importance of these five dimensions has gradually shifted. Sustainability has become the most prominent dimension in recent years, followed by farmer income, food security, and rural development. Clusters of documents further support this trend, with two clusters emphasising biodiversity, ecosystems, forests, and climate change. The changing trend of budget allocation also supports our results that the major concerns of CAP have been shifting over time. Overall, the evolving SSCIR priorities indicate that the CAP is moving towards a fairer, greener, and more flexible direction.

JEL Classification: Q01, Q18

1 | Introduction

Agriculture and agri-food sectors hold significant importance for European citizens, largely influenced by the Common Agricultural Policy (CAP), a central policy of the European Union (EU) established in 1962. Initially, CAP consumed 66% of the EU budget in the early 1980s, declining to 37.8% during 2014–2020, and further to 33.1% (EUR 55.71 billion) for the 2021–2027 period. CAP currently dedicates 76.8% (EUR 40.4 billion) to direct payments and market measures (pillar 1) and 23.2% (EUR 15.3 billion) to rural development (pillar 2).¹ The evolution of CAP has attracted ample research, though many questions have not been well answered. This study addresses

these gaps by identifying key CAP concerns and tracing their evolution.

Over its 50-year history, CAP's objectives have notably shifted. Initially focused on agricultural productivity, farmers' incomes, and market stabilisation, CAP employed guaranteed minimum prices, import levies, and export subsidies [1]. The 1992 Mac-Sharry reform significantly reduced price supports by introducing area payments coupled to production [2]. Agenda 2000 reforms divided CAP into two pillars: market/income support and rural development, including 'cross compliance' measures [3, 4]. The transformative Fischler reforms in 2003 decoupled area payments from production, based on historical farm

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Modern Agriculture* published by Wiley-VCH GmbH.

Summary

- This study introduces SSCIR-3F framework to examine the evolving priorities of the Common Agricultural Policy (CAP) based on 495 legislative texts analysed using text mining and machine learning techniques.
- We find a clear shift towards sustainability, with increasing attention to biodiversity and climate change, supported by both policy rhetoric and budget allocations.
- This study bridges qualitative and quantitative policy analysis by offering a novel, data-driven approach to assess the alignment of the CAP with agricultural development stages using unstructured governmental documents.

activities [5, 6]. The 2013 reforms further recalibrated direct income support, emphasised payment convergence, and introduced ‘Greening,’ obliging member states to allocate 30% of direct payments to environmental benefits [3, 7]. With the new CAP effective from 2023, the policy emphasises ambitious climate, environmental, and animal welfare goals.²

Some research examined CAP’s evolution. Lowe et al. [8] identified redirection from commodity support towards rural and environmental dimensions post-Agenda 2000. Greer [9] observed CAP’s resistance to substantial budgetary and instrumental changes since 2010. Alons [10] observed environmental concerns influencing CAP debates, and Fusco [11] highlighted increased environmental sustainability and rural development objectives through keyword analysis. Generally, CAP evolved from primarily supporting farmer livelihoods and productivity towards balanced territorial development [12]. However, the literature has inadequately defined CAP dimensions or explained their theoretical underpinnings over time.

Methodologically, machine learning (ML) techniques have emerged as valuable tools in economic and policy analysis [13, 14]. For example, text mining, which combines information extraction, ML, and data mining techniques, builds a bridge between quality analysis and quantity analysis [15, 16]. Several studies have shown the potential of text mining in effectively addressing human limitations in time and cognition in various disciplines [17, 18]. Cej et al. [19] applied quantitative text analysis to more than 11,000 abstracts related to agricultural economics literature and found 14 themes such as consumer behaviour and food quality. Cooper et al. [20] analysed food security literature, observing thematic expansions from economic policies to livelihoods, health, and environmental concerns.

Although ML approaches are employed in CAP trend analysis [11, 21] and Fusco [11], existing studies primarily use library analysis of academic abstracts, neglecting official policy documents. Notably, K. Erjavec and Erjavec [22] conducted frame analysis on only six European Commission (EC) strategic communications (1991–2017), identifying frames like economic, trade, and societal concerns. Juventia et al. [17] used text mining to analyse national commitments on agrobiodiversity but did not consider broader CAP dimensions.

This paper makes the following marginal contributions. First, we develop an SSCIR-3F framework that identifies five core CAP policy concerns—food security (S), sustainability (S), competitiveness (C), farmer income (I), and rural development (R), abbreviated as SSCIR—illustrates their internal conflicts using the conflict triangle, and explains their changing importance over time using agricultural development stage theory (food, farm, future, 3F). Second, in terms of data and methodology, we analyse 495 CAP-related legislative texts from the European Parliament (EP), providing a document-based perspective. Using machine learning and text mining, we bridge qualitative and quantitative analysis to offer a data-driven understanding of CAP reforms. Third, we cross-validate the identified policy shifts with actual CAP budget expenditures from 1980 to 2021, thereby enhancing the robustness and methodological rigour of our analysis.

Our findings indicate dynamic shifts among CAP’s primary concerns, notably elevating sustainability, followed by farmer income, food security, and rural development. K-means clustering of texts identified four distinct clusters—namely, the ‘farmer,’ ‘forest,’ ‘biodiversity,’ and ‘general concern’ clusters—with two of these highlighting the increasing emphasis on biodiversity, ecosystems, forests, and climate change. Our expenditure cross-check further confirms a consistent EU emphasis on farmer income and growing prioritisation of sustainability, signalling a transition towards high-quality sustainable agricultural development.

The structure of this paper is as follows: Section 2 presents our analytical framework, Section 3 details data and methodologies, Section 4 discusses results, and Section 5 concludes with policy implications.

2 | Analysis Framework

Figure 1 presents the SSCIR-3F framework for analysing the CAP, highlighting its multiple policy dimensions, inherent budgetary constraints, and evolutionary stages of agricultural development. CAP has multiple policy objectives and comprises a wide range of agricultural and rural measures, from agricultural market interventions to agri-environmental payments and rural development measures [23].

The left part of Figure 1 identifies five interrelated policy dimensions of the CAP, summarised as SSCIR: ensuring food security (S), promoting sustainability (S), enhancing agricultural competitiveness (C), supporting farmer incomes (I), and fostering rural development (R). The categorisation of these dimensions is grounded in the financial support structure of the CAP, reflecting budgetary allocations and structural reforms over time [24]. Initially financed through a single fund—the European Agricultural Guidance and Guarantee Fund (EAGGF)—the CAP transitioned to a dual-fund system in 2007 consisting of the European Agricultural Guarantee Fund (EAGF, the ‘first pillar’)³ and the European Agricultural Fund for Rural Development (EAFRD, the ‘second pillar’).⁴ The EAGF supports basic payment schemes, green direct payments, and assistance for young farmers, while the EAFRD funds Rural

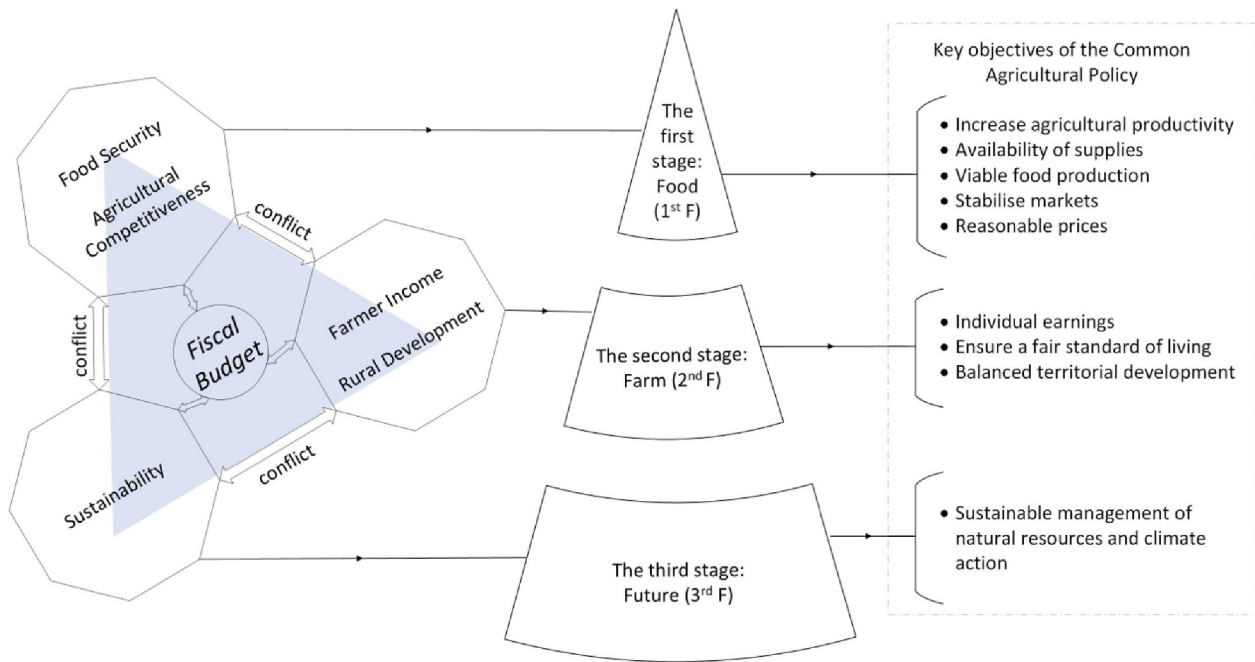


FIGURE 1 | The SSCIR-3F analysis framework. ‘SSCIR’ represents five cardinal CAP policy domains: food security (S), sustainability (S), agricultural competitiveness (C), farmer income (I), and rural development (R). ‘3F’ demarcates the triphasic progression of agricultural development: Food, Farm, and Future. The key objectives of the CAP in SSCIR-3F analytical framework come from Pe’er et al. [26]. Created by the author.

Development Programmes (RDs), targeting agricultural competitiveness, sustainable resource management, climate action, and rural community development. Thus, the five dimensions of CAP emerge directly from this structured funding framework, ensuring coherence and validity in their identification.

Given the fiscal constraints as stipulated by the Multiannual Financial Framework (MFF),⁵ the total budget can be seen as the summary or the budget for each dimension, which is:

$$B_{\text{total}} = B_{\text{export subsidies}} + B_{\text{other market support}} + B_{\text{coupled direct payments}} + B_{\text{decoupled direct payments}} + B_{\text{greening}} + B_{\text{RD environment and climate}} + B_{\text{total rural development}} \quad (1)$$

Restricted to $B_{\text{total}} = B_{\text{EAGF}} + B_{\text{EAFRD}} = \alpha \cdot Y_{\text{MMF}}$. Where B_{total} denotes the budget allocation to CAP, $B_{\text{(export subsidies)}}$ and $B_{\text{(other market support)}}$ represent allocations for food security and agricultural competitiveness, $B_{\text{(coupled direct payments)}}$ and $B_{\text{(decoupled direct payments)}}$ for farmer income support, B_{greening} for sustainability, $B_{\text{(RD environment and climate)}}$ and $B_{\text{(total rural development)}}$ for rural development and sustainability. B_{EAGF} and B_{EAFRD} denote the budget allocation to two funds, and Y_{MMF} denotes the total amount of EU long-term budget. α denotes the share of CAP in total EU budget, it equals 73% in 1985, and then gradually decreases to 32% in the recent EU MMF 2021-27. With the fixed total budget, the increase in budget allocation in one dimension means decreases in other dimensions. Beyond budgetary

constraints, the CAP faces inherent tensions among its multiple policy objectives, which often involve trade-offs and are not always defined in a SMART (specific, measurable, achievable, relevant, and time-bound) manner. As a result, CAP priorities have shifted over time [25, 26].

The formulation of agricultural policy frameworks, constrained by limited budgets and inherent tensions among objectives, requires policymakers to prioritise non-conflicting goals aligned with specific economic development stages, resource endowments, macroeconomic conditions, and international market contexts [27]. Thus, Figure 1 also integrates the ‘3F’ theory—a triphasic progression of agricultural development comprising the stages of Food, Farm, and Future. Schultz [28] initially characterised agricultural development by two stages: the ‘food problem,’ typical in developing countries, emphasising food security through increased production, and the ‘farm problem,’ prevalent in developed and middle-income countries, which prioritises farmer income stabilisation and agricultural competitiveness. Extending this, Yu and Zhao [29] introduced a third stage—the ‘future problem’—emphasising sustainable development.

Figure 1 reflects our central hypothesis: that CAP policy adaptations evolve in tandem with the challenges of each developmental stage, as outlined by the ‘3F’ theory. Specifically, we hypothesise that early CAP policies focused on food security, mid-period reforms emphasised farmer support and competitiveness, and recent reforms increasingly prioritise sustainability and multifunctionality.

3 | Data and Method

3.1 | Data From EP

This study utilises legislative texts related to CAP obtained from the EP website. As one of the EU's principal legislative bodies, the EP enacts legislation—primarily based on EC proposals—in collaboration with the Council of the EU. Using ‘common agricultural policy’ as a search term in the EP's adopted texts portal, we identified 495 relevant texts in PDF spanning six parliamentary terms (see Table 1). The accessibility and systematic coherence of these texts provide a robust foundation for our analysis. It is important to note that under the Lisbon Treaty, agriculture became subject to the co-decision procedure on December 1, 2009. However, the EP exercised full co-decision authority over the CAP only after 2013 [30]. Limitations of this data source are outlined in the conclusion.

3.2 | Method

3.2.1 | Keyword Identification for CAP SSCIR Dimensions

We use the EC's common monitoring and evaluation framework (CMEF) performance indicators⁶ (e.g., farming income support, environment) as a benchmark for keyword identification. Recognising that the CMEF has evolved over time, we have further enriched our keyword set by incorporating terms from CAP-related academic research—most notably Fusco [11], which provided a comprehensive bibliometric list—and from studies by Bureau and Swinnen [31], Koemle, Lakner et al. [32], Koemle, Zinngrebe et al. [33], Radić et al. [34], Graskemper, Yu, and Feil [35] and Graskemper, Yu, and Feil [36]. Our keyword identification approach integrates policy frameworks and academic perspectives, ensuring a comprehensive and historically grounded analysis of CAP priorities.⁷ Table 2 presents the CAP dimensions and associated keywords, which form a predefined dictionary guiding text mining.⁸ Notably, ‘Fiscal Budget’ is excluded from the table, as it is analysed separately by cross-

checking actual CAP expenditures with policy shifts detected through text mining.

3.2.2 | Text Mining Techniques

Building on previous keyword identification, we apply text mining to compute keyword frequencies in 495 EP legislative PDFs, enabling us to identify prevalent terms and track temporal trends. Figure 2 outlines our workflow: employing R's ‘tm’ package [37] to tokenise and preprocess the texts—by converting to lowercase, removing numbers, punctuation, stop words,⁹ extra whitespace, and applying stemming¹⁰—which produces a Document-Term Matrix (DTM) capturing term frequencies [38, 39]. While big data and ML efficiently process large corpora and reveal policy trends, they are limited by a potential loss of nuance and sensitivity to preprocessing choices, such as the selection of predefined keywords. Nonetheless, it enriches our study by confirming anticipated patterns, uncovering novel insights into evolving legislative priorities, and providing a quantitative analysis of unstructured text data to complement traditional qualitative methods.

3.2.3 | Word Combination Detection Method

While the ‘tm’ package efficiently processes individual words, it falls short of capturing significant word combinations crucial to our analysis. To address this, we employ the ‘pdftools’ package [40], which leverages a robust C++ library to extract text from PDFs and compute term frequencies for word combinations in R. Briefly, the word detection method is similar to text mining but incorporates word combinations, and the visualised procedure for detecting these combinations is presented in Supporting Information S1: Figure A1 in the Appendix. However, due to its longer computational time and lower efficiency, we use it solely as a supplement for analysing word combinations.

3.2.4 | Word Frequency Correlation Analysis

We computed Pearson's correlation coefficients between high-frequency keywords based on their occurrences across the 495 texts. A strong positive coefficient indicates that two keywords tend to co-occur, suggesting an underlying connection or potential harmony. Conversely, weak or negative coefficients imply little association or potential conflicts. This approach quantifies the strength and direction of word relationships, revealing co-occurrence patterns and contextual connections within the dataset. By analysing these correlations, we identify significant word pairs and possible thematic clusters, providing insights into dominant themes and their interconnections in a straightforward and interpretable manner.

3.2.5 | Clustering Analysis

To analyse similarities and differences in policy priorities across legislative texts, we employ *K-means* clustering to

TABLE 1 | Number of legislative texts related to Common Agricultural Policy (CAP) from 1994 to 2021 obtained from the European Parliament (EP).

Year	Number	Year	Number	Year	Number
1994	7	2004	18	2013	16
1995	13	2005	16	2014	14
1996	12	2006	30	2015	17
1997	12	2007	23	2016	13
1998	17	2008	25	2017	18
1999	6	2009	24	2018	23
2000	1	2010	21	2019	23
2002	29	2011	18	2020	19
2003	28	2012	22	2021	30
Total			495		

TABLE 2 | Dimensions and the words related to CAP.

Dimensions	CAP indicators	Words
Food security	Productivity	Agricultural productivity, agricultural production, agricultural product, agri-food, crop, food production, input, livestock farming, output, yield, arable land
	Food and health quality protection	Affordable food, food quantity, food quality, food safety, food security, organic production
	Adding value	Geographical indication, protected designation of origin, short supply chain, local food
Agricultural competitiveness		Agricultural competitiveness, agricultural exports, agricultural trade, agri-food exports, agri-food trade, food exports, food trade, agricultural imports, agri-food imports, food imports
Farmer income	Farming income support	Basic payment, decoupling, direct payment, farm, farm income, farmer, farming, farmland, first pillar, income support, small farm, medium farm, farm size, coupling, farmer income, indirect payment, redistributive income, redistributive payment, single area payment, subsidy system, support farmer
	Market orientation	Market measure, market monitoring, market orientation, agricultural market
Rural development	Jobs and growth in rural areas	Young farmer, European innovation partnership, living condition, rural area, rural development, rural development programme, rural economy, second pillar, smart village, genuine farmer, living standard, reasonable living, territorial balance, territorial development
Sustainability	Environment and climate action	Afforestation, agri-environment, agroecology, agroforestry, cross compliance, ecology, eco scheme, ecosystem, ecosystem service, environmental protection, forest, greenhouse gas, greening, landscape, natura 2000, natural resource, organic farming, sustainable development, sustainable management, soil, animal health, animal welfare, wildlife, habitat fragmentation, habitat loss, climate action, climate change, biodiversity, biological diversity, bird, species richness, ecological focus area

group 495 texts based on the frequency of high-frequency keywords, identifying meaningful subsets that capture variations in policy focus as reflected by word frequency [41]. In this approach, each legislative text x_i is assigned to the nearest centroid μ_k by minimising the squared Euclidean distance ($z_i \leftarrow \operatorname{argmin} \|\mu_k - x_i\|_2^2$). The centroids μ_k are iteratively updated by minimising the sum of distances between μ_k and the text x_i assigned to the cluster k , across all clusters: $\mu_k \leftarrow \operatorname{argmin} \sum_{i: z_i = k} \|\mu_k - x_i\|_2^2$. A proper number of clusters are crucial for *K-means* performance, as too few clusters oversimplify document patterns, while too many lead to mixed content [41]. The optimal number of clusters is determined using the elbow method, which selects the point where additional clusters yield minimal reduction in within-cluster variance. Although the *K-means* method efficiently processes continuous data, it is sensitive to initialisation and may produce centroids that are not present in the original dataset. To address these limitations, we also performed a robustness check using the Partitioning Around Medoids (PAM) method.

4 | Results

4.1 | Most Frequent Words of the CAP

Figure 3 presents the top 10 most frequent words with the highest frequency in the adopted legislative texts related to CAP. The results show that the top 10 most frequent words in the EP adopted legislative texts between 1994 and 2021 are ‘farmer’, ‘biodiversity’, ‘rural area’, ‘rural development’, ‘forest’, ‘climate change’, ‘farm’, ‘farming’, ‘agricultural product’, ‘ecosystem’ (Figure 3, top left). The above words cover the four dimensions, that is, food security, safety, and quality; supporting farmers; EU agriculture and environmental sustainability; and securing the future of rural areas. These findings are important because they are representative of the major concerns of CAP between 1994 and 2021. Taking the CAP reforms into account, Figure 3 also displays the most frequent words during the different periods, namely, 1994–2002 (top right), 2002–2012 (bottom left), and 2013–2021 (bottom right). Overall, in addition to the economic aspect of supporting the farmers, objectives linked to

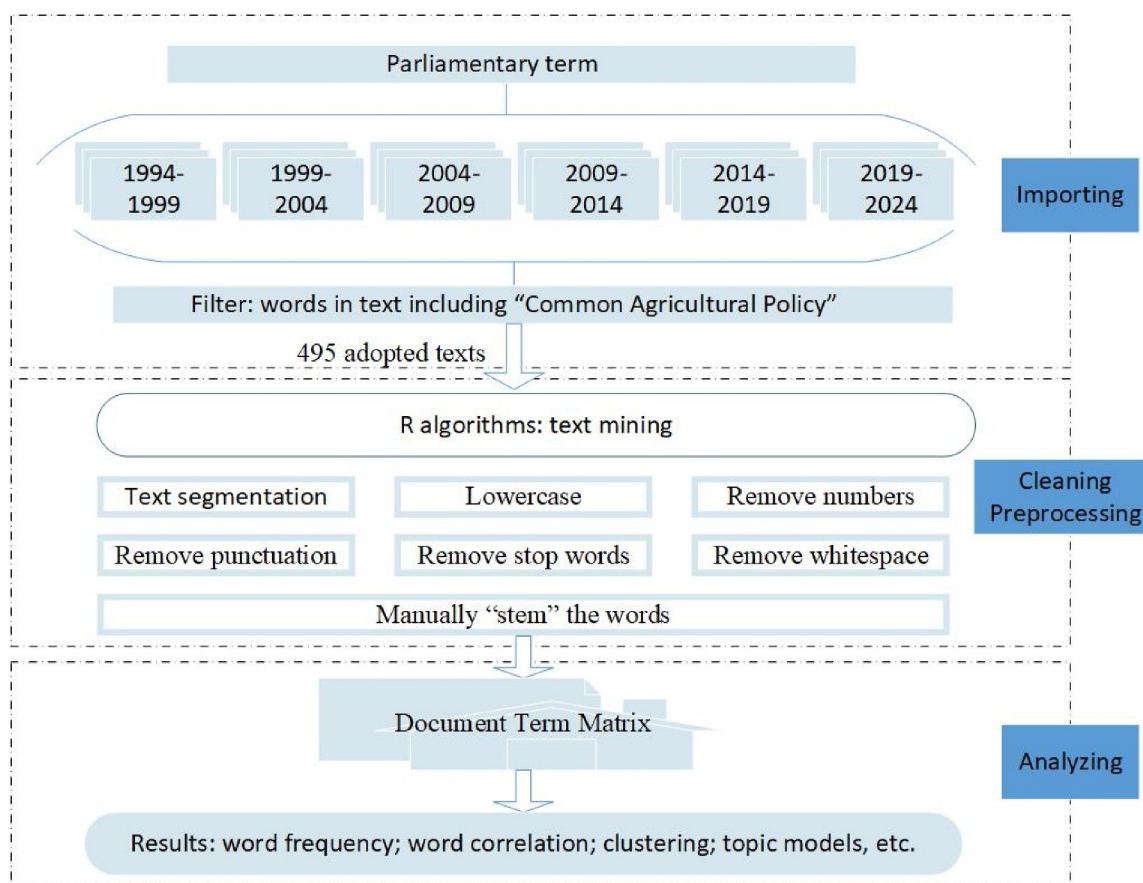


FIGURE 2 | Order of text analysis operations for data preparation and analysis. Created by the author.

environmental sustainability and rural development have been introduced in all the periods.

The term ‘farmer’ figured prominently in all periods, which demonstrates the impression that support for farm incomes has always been the main concern of the CAP. According to the latest Eurobarometer, ‘supporting farmers’ has been widely supported by EU citizens.¹¹ Recognising many ways in which farmers contribute to society, a growing number of citizens believe that income support for farmers should increase over the next 10 years – 56% held this opinion in 2020, a rise of 12% compared to 2017 and 27% compared to 2007 [42]. This result is highly consistent with other studies. K. Erjavec and Erjavec [22] found five frames in EU communication, while the farmers’ economic frame retained its primacy and continuity (from 1991 to 2017), demonstrating the power of the farmers’ lobbies and conservative member states.

In addition, the concepts of ‘active farmers,’ ‘genuine farmers,’ ‘small farmers’ and ‘young farmers’ have been brought under the spotlight. The ‘active farmer’ criterion, introduced in the 2014–2020 CAP reform, aims to exclude beneficiaries who engage in minimal or no agricultural activity while receiving subsidies. Similarly, the EC Draft Regulation on CAP 2021–2027 proposes that the concept of ‘genuine farmers’ shall be defined to ensure that support is granted to those whose primary business activity is agriculture [34]. Small farmers are the main components of the EU agricultural sector. More than

three-quarters of farm holdings in the EU are below 10 ha.¹² Rivera et al. [43] found that small farms play an important role in food production and regional food availability. However, the importance of small farms is still not established and recognised in Europe, as agriculture is largely industrialised and thus characterised by large farms. Glowinkel et al. [44] showed that the current way of granting subsidies promotes a structural change that constantly pushes small- and medium-sized companies out of the agricultural market. Guth et al. [45] showed that the distribution of CAP income support favoured the largest farms and increased disparities within the sector. Actually, the decrease in the number of small agricultural farms has been witnessed under the dynamic processes of structural changes [46]. To ensure a fairer distribution of financial support for farmers and workers across the EU, the EU countries apply the Small Farmers Scheme (SFS), a simplified income support scheme granting a one-off payment to farmers who choose to participate. This scheme will be continued in the New CAP 2023–27. The influence of young farmers’ support system has been examined in the present studies, including both direct payments for young farmers and rural development measures. Based on the data collected from Lithuania, Balezentis et al. [47] found that the young farmers’ support system has the most substantial perceived effect on income support.

We want to emphasise that the term ‘young farmers’ occurs with a stable and increasing frequency during the last 3 decades. Support for young farmers is sometimes combined with rural

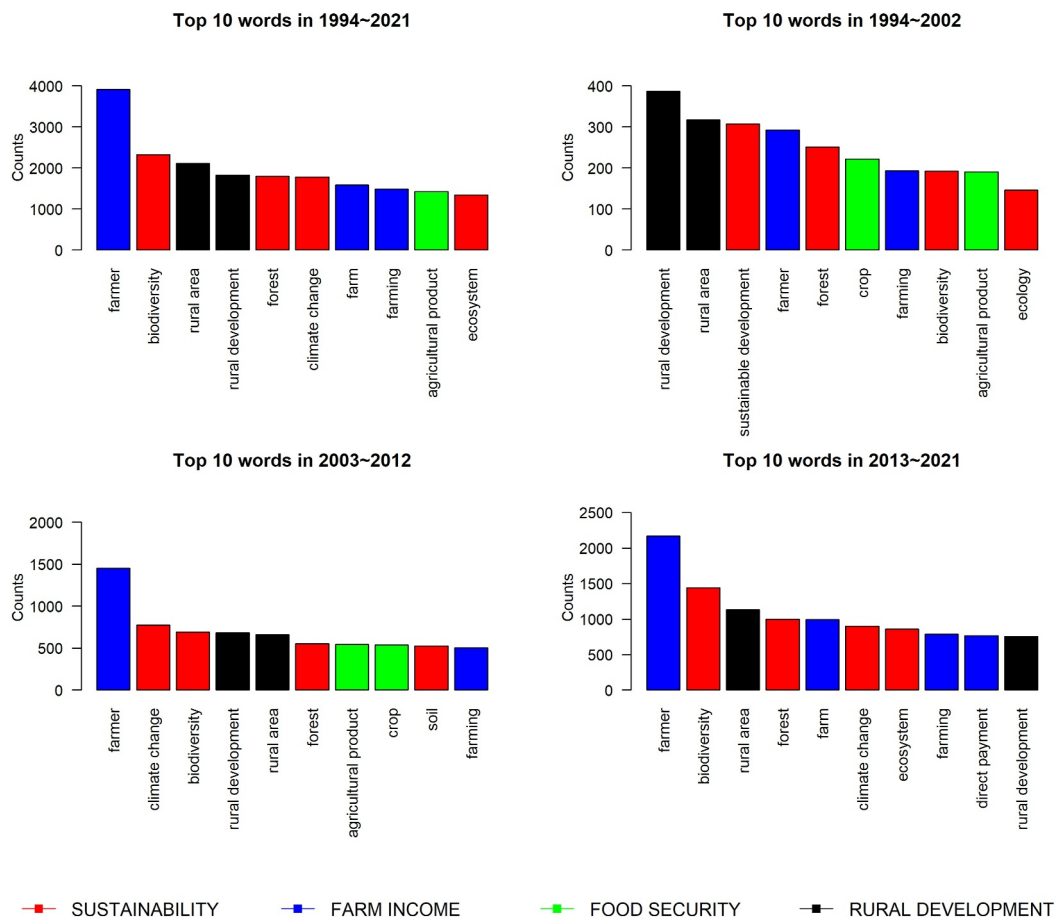


FIGURE 3 | Top 10 most frequent words in EP adopted texts related to CAP. Created by the author.

development, and it is crucial to achieving the sustainability of rural regions under the CAP [47]. With the goal of restructuring and redistribution, the rural sustainability concept should consider the role of young farmers. The new fairer CAP 2023–27 directs support to young farmers, calling EU countries to distribute at least 3% of their direct payments budget towards young farmers, including income support, investment support, and start-up aid.

In the period 1994–2002, the term ‘biodiversity’ was barely present and had a relatively lower term frequency. However, it has gained importance in the recently adopted legislative texts with a higher frequency during 2013–2021. Pe’er, Dicks et al. [48] found that the CAP before the 2013 reform had not been successful in preserving farmland biodiversity. The environmental prescriptions in CAP were so diluted that they are unlikely to benefit biodiversity [49].

After the 2013 reform of the CAP, biodiversity concerns gained more attention. One specific measure to protect biodiversity in the EU is the ecological focus area (EFA) scheme; these areas should provide sufficient habitat for wild plant and animal species and facilitate the dispersal of species across the landscape [50]. In the CAP, it assigned a ‘green payment’ measure under Pillar 1, in which 3%–5% of EU farmland should be managed as EFA to halt biodiversity loss, to be increased to 7%

until 2017 [51]. Hristov et al. [52] showed that the 2013 CAP ‘greening’ reform had the potential to improve biodiversity and ecosystem services on farms through the EFA measure. Cormont et al. [53] empirically assessed the implications of this policy and found that even landscapes with 3%–7% of natural elements harboured generally 37%–75% of maximum species richness, indicating the excellent potential of implementing the CAP target to preserve farmland biodiversity.

‘Rural area’ and ‘rural development’ nearly kept their positions during the 3 decades. As the ‘second pillar’ of CAP, ‘rural development’ reinforces the ‘first pillar’ of income support and market measures by strengthening rural areas’ social, environmental, and economic sustainability. A shift from the ‘first’ to the ‘second’ pillar after the 2000s implies a reduction of the funds allocated for market support and an increase of the funds available for the second pillar, which is often presented as a strengthening of rural policy [54]. Diversified modern rural areas have become attractive and viable, providing advantages for professional farmers and their families and reducing or constraining the need for indefinite agricultural support [54]. Promoting more inclusive rural development is one of the main aims of CAP for 2014–2020. The innovations introduced in the rural development policy, the LEADER approach, create local partnerships for bottom-up governance of rural areas [55]. The EAFRD supports the CAP’s contribution to the EU’s rural

development objectives, whose budget amounts to €95.5 billion for CAP 2021-27. Under the new CAP, rural development actions respond more to current and future climate change and generational renewal challenges. It will also make a solid contribution to the European Green Deal and the long-term vision for rural areas.¹³

4.2 | Correlations of the Most Frequent Words

We analyse word correlations among the top 10 most frequent words in legislative texts from 1994 to 2021, visualised as a heatmap in Figure 4. As described in the methodology section, each cell in the heatmap is colour-coded based on the Pearson correlation coefficient calculated from keyword occurrences across 495 adopted texts. The colour gradient ranges from blue (indicating a strong positive correlation) to red (indicating a negative correlation). Notably, the absence of red cells suggests no negative correlations among the key terms, reflecting a balanced representation of CAP priorities over the past 2 decades. This finding underscores that the CAP is not a singular, product-focused policy but rather a comprehensive framework encompassing diverse measures. Designed to promote regional economic growth and support the EU agricultural sector, the CAP has evolved to address a broad spectrum of concerns while

allowing flexibility in implementation across member states and regions [56].

Specifically, ‘farmer’ is mostly associated with ‘farm’ and ‘farming,’ followed by ‘rural area’ and ‘rural development’ and is slightly associated with ‘agricultural product.’ ‘biodiversity’ is highly associated with ‘ecosystem,’ followed by ‘climate change’ and ‘forest.’ These correlations are reasonable and meaningful since agriculture has multifunctional roles besides food production, and farms are the essential components in rural areas. Rural areas are a place of living and work for farmers, and agricultural production is essential for the national economy [46]. Thus, farmers are highly associated with rural development. Keeping rural areas alive and restructuring the agricultural sector requires the CAP to provide more targeted support to small farms, young farmers, and genuine farmers [34]. Environment protection in the European Green Deal and Farm to Fork strategy is directly connected to farming practices [57, 58]. Saraceno [54] pointed out that the sustainability of rural areas does not depend on the farming sector alone but on the diversity and attractiveness of the rural resources. Farmers, rural, and the environment are all connected together; thus, the CAP covers all those important issues. As Guth et al. [45] pointed out, CAP is a comprehensive policy that comprises smart growth, greener resources, and a competitive economy accompanied by social and territorial cohesion.

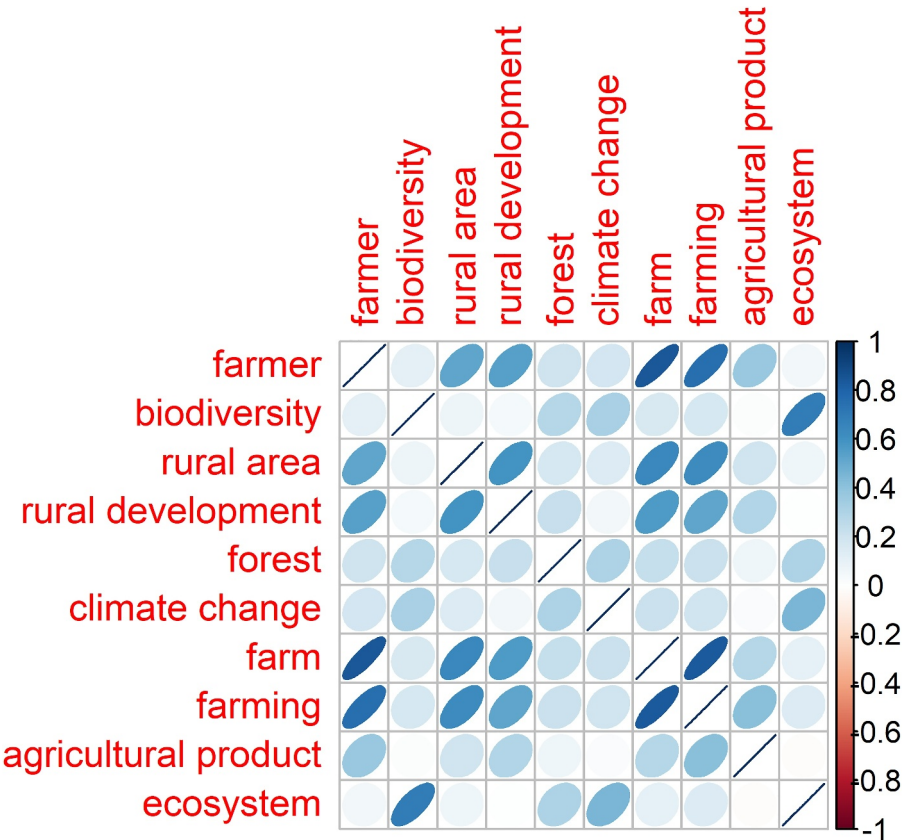


FIGURE 4 | Correlations of the top 10 most frequent words. Pearson correlation coefficients among the top 10 most frequent words were calculated across 495 documents. Strong positive coefficients suggest frequent co-occurrence and potential thematic connections, while weak or negative coefficients indicate minimal association or possible conflicts. Each cell in the heatmap is colour-coded based on these coefficients using a gradient from blue (strong positive) to red (strong negative). Created by the author.

4.3 | Changing Trend of SSCIR

Figure 5 displays the five dimensions' annual trend in the term frequency. The term frequency for each dimension is calculated by summarising the term frequency of the corresponding words in Table 2. It is possible to affirm that the term frequency of each dimension shares the same vibration rise trend over the period considered (1994–2021), except for 'agricultural competitiveness'. It is important to underline some obvious peaks, such as the year 1998, the year 2003, the year 2013, the year 2018, and the year 2021. Those years have a relatively high term frequency for all dimensions, which implies that important texts covering food production, farmer income, sustainability, and rural development were adopted. In addition, these pick points also validate the accuracy of our text mining method since they exactly match the critical reform windows: agenda 2000, 2003 reform, 2013 reform, 2018 proposal for new CAP, and the 2021 agreement on the reform of the CAP.

We want to highlight the changing trend of the dimension 'Sustainability,' which has registered exponential growth in recent years. It is the dimension that has attracted the most attention in CAP policy during the last few years, showing the ambitions of the new greener CAP, which aligns with the findings from Feindt [59]. The environmental and climate legislation requires each EU country to display better environmental and climate action and update the plan when climate and environmental legislation is modified. In addition, at least 25% of the budget for direct payments will be allocated to eco-schemes, 35% of funds will be allocated to measures to support climate, biodiversity, environment, and animal welfare, and 40% of the CAP budget will have to be climate-relevant.¹⁴ However, the high term frequency of sustainability does not necessarily translate into good sustainability practices in EU countries. Šumrada et al. [60] examined the

degree of biodiversity policy integration into the CAP and found that the integration is limited to instruments that remunerate above-standard farming practices. Pe'er, Dicks et al. [48] suggested that the CAP should encourage EU countries to start moving towards more sustainable agriculture, such as maintaining or enhancing the agri-environment climate schemes budget.

4.4 | Cluster Identification Results

In this section, we applied *K-means* clustering to Group 495 legislative texts based on the frequencies of the top 10 keywords, thereby elucidating similarities and differences in policy priorities. The optimal number of clusters was determined to be four using the elbow method, based on the within-cluster sum of squares and second-order difference [41] (Supporting Information S1: Figure A2 in the Appendix). Figure 6 visualises the clustering results for four clusters. The left panel illustrates the distribution of texts across clusters, while the right panel presents a radar chart of mean term frequency. In the radar diagram, the inner circle represents a low term frequency (count = 0), while the outer circle indicates a high term frequency (count = 250). The detailed mean term frequency for the radar diagram can be found in Supporting Information S1: Table A1 in the Appendix, and the example contents of the adopted texts in each cluster can be found in Supporting Information S1: Table A2 in the Appendix.

The clustering analysis reveals three key thematic concerns: 'farmer' in the first cluster, 'forest' in the second, and 'biodiversity' in the third. The fourth cluster, characterised by relatively low mean term frequencies for the top 10 words, likely represents a general concern for the EU agricultural sector across all sectoral documents.

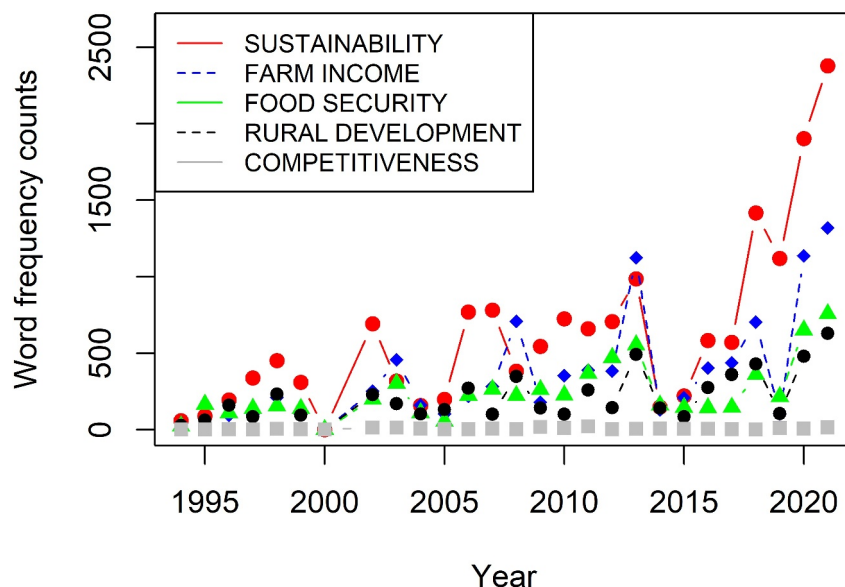


FIGURE 5 | Annual trend in the term frequency of the five dimensions, 1994–2021. The words have very low term frequencies at the beginning years because we select the words mostly based on the academic studies in the recent year. This approach has led us to choose more words that have emerged in recent years. We are not concerned with the term frequency for each dimension but with the comparison of the term frequencies among the five dimensions. Therefore, this systematic bias will not have a significant impact on the results of this paper. Created by the author.

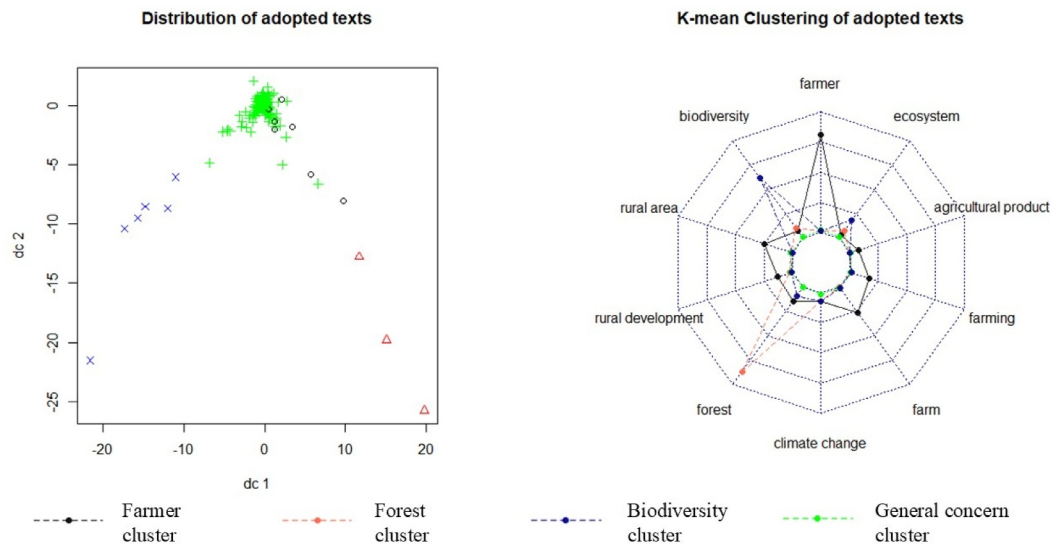


FIGURE 6 | Distribution of data points and radar diagram of mean term frequency for four clusters. The inner circle of the radar diagram indicates the low expression of the term frequency (0), while the outer circle indicates the high expression of the term frequency (250). Created by the author.

The first cluster consists of seven adopted texts, where ‘farmer’ exhibits the highest mean term frequency, followed by ‘farm,’ ‘rural area’ and ‘farming.’ These texts primarily focus on supporting farmers and rural development, underscoring the central role of farmers in the CAP and its alignment with rural development objectives. This finding aligns with previous analyses, including the top 10 term frequency rankings and the positive correlation between ‘farmer’ and ‘rural area,’ demonstrating consistency between supervised (top 10 words frequencies) and unsupervised (clustering) ML approaches. The second cluster contains three adopted texts, where ‘forest’ emerges as the most prominent term. These texts primarily address forest protection, climate change, and natural resource management, highlighting the significant emphasis on forests in the European Green Deal. This result underscores the critical role of forests as a key natural resource in EU environmental policy. The third cluster comprises six adopted texts, characterised by a high frequency of ‘biodiversity’ and a relatively high frequency of ‘ecosystem.’ The main topics in this cluster include biodiversity loss, biodiversity conservation, and the EU Biodiversity Strategy. These findings reaffirm the prominence of biodiversity concerns within CAP, consistent with previous analyses.

In summary, these results define three pronounced clusters: the ‘farmer cluster,’ the ‘forest cluster,’ and the ‘biodiversity cluster.’ The farmer cluster confirms a harmonious relationship between the farmer income and rural development dimensions—consistent with our SSCIR-3F framework (which situates these issues within the farm problem stage)—while the forest and biodiversity clusters underscore the evolving emphasis on sustainable development. The clustering results remain robust when using the alternative clustering method, PAM (see Supporting Information S1: Table A3). Key concerns such as ‘farmer,’ ‘rural forest’ and ‘biodiversity’ are consistently emphasised, further validating the robustness of our findings.

4.5 | Cross-Check Between Actual Expenditure With Text Mining Results

We are aware that the frequency of certain terms of the policy texts is not necessarily related to the ultimate policy effects. Sometimes, words may be used to convey an impression that is not in line with what the policy actually aims for. For example, some critics contend that ‘cross compliance’ is used as a pretext to continue income support for farmers rather than to genuinely promote sustainability [61]. Although the objectives and scope of cross compliance are well defined in the documents, its implementation is complicated by the fact that it encompasses a multitude of regulations and allows member states considerable flexibility in setting standards [62]. This complexity partly motivated the 2013 reform, which introduced a more easily verifiable greener criterion while retaining cross compliance and good farming conditions. Similarly, despite claims that the new CAP is ‘greener,’ the environmental prescriptions often remain so diluted that their benefits for biodiversity are limited [48].

Although there are still many challenges in analysing policy intentions through word frequency statistical methods in text analysis, we still believe that a systematic analysis of policy documents is meaningful. Words with high word frequency can reflect the focus of current policies to a certain extent, and even if these policies cannot be realised at present, they may be achieved in the future.

To verify that the results of the textual analysis reflect the main concerns of CAP policy, we compare the results of the textual analysis with the detailed allocation of CAP funding from 1980 to 2021. Figure 7 shows the funds allocated to different dimensions of the CAP since 1980 under the MFF, and it also shows the share of funding for each dimension. Before the MacSharry reform in 1992, the export subsidies and other market support accounted for the vast majority of CAP’s budget.

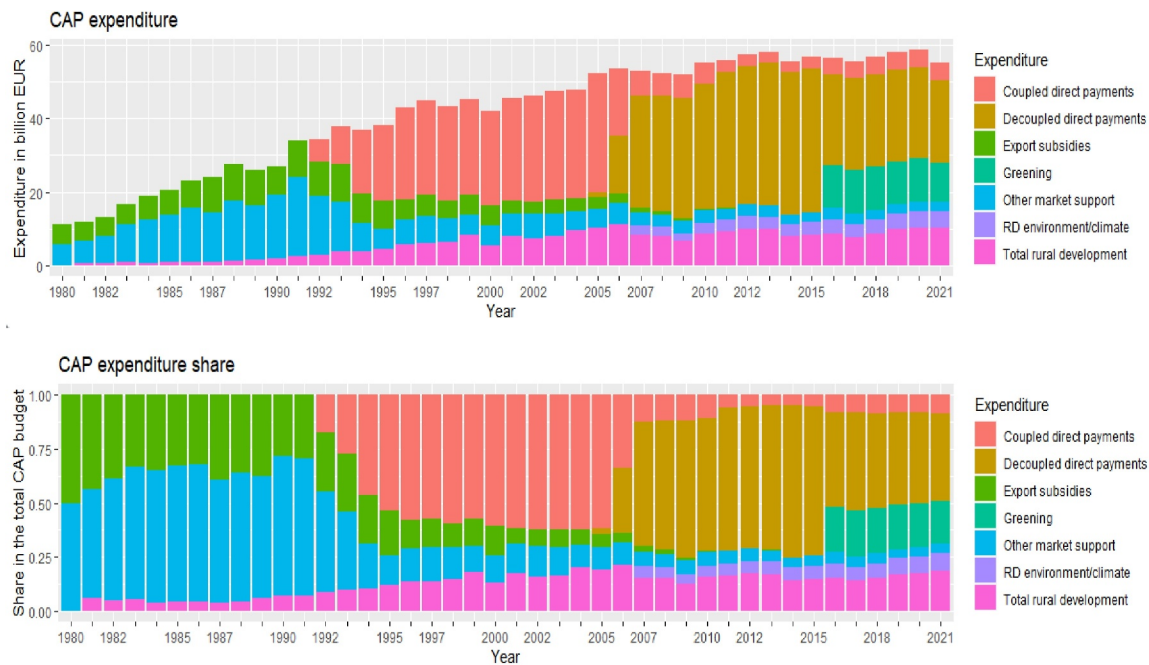


FIGURE 7 | CAP expenditure and expenditure share, 1980-2021 (current prices). 2021 budget amounts, coupled direct payments including POSEI and SAI direct payment components, and Annex I Regulation 1305/2013. The authors thank Dirk De Smet affiliated with EC, DG Agriculture and Rural Development, for sharing a very useful CAP expenditure table. CAP expenditure for past years is sourced from EC, DG Agriculture and Rural Development (Financial Report). The figure is created by the author.

However, they have been gradually eliminated since 1980. The budget for rural development (RD) is increasing gradually, which confirms the text mining result of changing trends for rural development. Direct payments for farmers occupied a large proportion of the total CAP budget from 1992 onwards. The coupled direct payments increased gradually after the Fischler reforms in 2003, and the decoupled direct payments outweigh the coupled direct payments after the 2013 reforms. The budget allocated to sustainability dominions included the 'RD environment/climate' and 'Greening'. The focus on environmental sustainability is shown by the fact that the new greening payment accounts for 30% of direct aid to farmers and that at least 30% of the Rural Development envelope must be reserved for environmental/climate related actions.

Overall, the changing trend of direct payments and environmental sustainability confirms the EU's constant concerns about the farmers' income and increasing concerns about sustainability. The changing trends of the budget allocation match the text mining results.

5 | Conclusions and Policy Implications

The CAP, a foundational pillar of the European project since 1962, has undergone substantial transformations in response to evolving agricultural, economic, and environmental challenges. In this study, we analyse 495 legislative texts adopted by the EP from 1994 to 2021 using text mining and ML methods (clustering) to examine the harmonies, conflicts, and evolutions within CAP. We develop the SSCIR-3F analytical framework,

which identifies five key policy concerns, including food security, sustainability, agricultural competitiveness, farmer income, and rural development (SSCIR), which are interpreted within the budget constraints and agricultural development stages theory (food, farm, and future, 3F). Our findings reveal that while farmer income has remained a consistent priority, sustainability has emerged as the most critical concern in recent years, indicating a move towards a fairer, greener, and more flexible policy framework.

The textual analysis demonstrates that the term 'farmer' dominates CAP discourse, yet the rising prominence of terms such as 'biodiversity' and 'ecosystem' signals an increasing focus on environmental sustainability. Clustering analysis further distinguishes thematic groups, with one cluster emphasising farmer and rural issues and others highlighting environmental concerns aligned with the third stage of agricultural development. Additionally, a cross-check between the evolution of policy rhetoric and CAP funding allocations from 1980 to 2021 confirms that the EU's budgetary decisions are increasingly consistent with its stated priorities, underscoring a coherent shift in policy orientation.

This study offers several valuable policy recommendations for future agricultural policy design. Policymakers should tailor CAP interventions to the specific stage of agricultural development, balancing the need to support farmer income with a growing commitment to sustainability. Given the inherent conflicts among multiple policy objectives under fiscal constraints, the EU should adopt an integrated policy framework that aligns financial commitments with strategic goals. As CAP

continues to evolve, evidence-based policy design and data-driven monitoring mechanisms will be essential to ensure its effectiveness. Moving forward, greater integration of digital technologies and data-driven policy evaluation can further enhance CAP's responsiveness to emerging challenges. As the EU transitions into the Future stage of agricultural development, CAP must remain flexible, forward-looking, and environmentally sustainable, ensuring that it continues to support resilient agricultural systems and rural communities in an era of climate change and economic uncertainty.

This study has several limitations. First, it relies solely on EP texts, while the EC and the Council of the EU also shape the CAP. Although the EP only gained full co-decision powers over CAP in 2013, it was actively involved in CAP discussions and legislative processes well before that, as evidenced by adopted texts from the 1990s onwards. Nevertheless, a more comprehensive analysis incorporating documents from all three institutions could provide a more complete picture of CAP's legislative evolution. Second, differences in document types—such as EP amendments to Commission proposals versus EP critiques of Council decisions—may influence keyword prominence and thematic focus. While we avoided arbitrary exclusions to maintain completeness, refining the text selection process could offer deeper insights into CAP's policymaking dynamics. Future research should explore how different document types shape CAP discourse and affect text-mining conclusions. Third, the EP Agriculture Committee, with many members of the EP linked to the farming lobby, may introduce institutional biases. Further studies could assess how these affiliations influence CAP discussions and whether lobbying affects policy rhetoric. Finally, with regard to our methodology, our reliance on a predefined dictionary for keyword identification may overlook emerging terms and nuanced policy shifts not fully captured in our selected sources. Although cross-checking with CAP budget allocations partially mitigates this bias, future studies could benefit from dynamic dictionary updates and the integration of additional data sources.

Author Contributions

Shuang Liu: conceptualization, data curation, formal analysis, investigation, methodology, software, writing – original draft, writing – review and editing. **Jasmin Wehner:** conceptualization, data curation, formal analysis, writing – original draft. **Jan-Henning Feil:** conceptualization, data curation, formal analysis. **Xiaohua Yu:** conceptualization, formal analysis, investigation, methodology, supervision, writing – original draft, writing – review and editing, project administration, supervision.

Acknowledgements

The authors gratefully acknowledge the efficient review coordination of the Managing Editor, Dr. Taolan Zhao, and the useful comments of two anonymous reviewers. The authors also thank Prof. Stefan Tangermann and Prof. Achim Spiller of the University of Göttingen for their constructive comments on this paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Supporting information data and code for this article are provided with the publication of the paper.

Endnotes

- ¹ Refers to 'Fact Sheets on the European Union: Financing of the CAP' at <https://www.europarl.europa.eu/factsheets/en/sheet/106/financing-of-the-cap>, Last access on February 25th, 2022.
- ² Refer to 'The new common agricultural policy: 2023-27' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/new-cap-2023-27_en#keyareasofreform. Last access on March 22, 2022.
- ³ Refers to 'European agricultural guarantee fund (EAGF)' at https://ec.europa.eu/info/funding-tenders/find-funding/eu-funding-programmes/european-agricultural-guarantee-fund-eagf_en. Last access on June 28th, 2022.
- ⁴ Refers to 'European Agricultural Fund for Rural Development (EAFRD)' at https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/european-agricultural-fund-rural-development-eafrd_en. Last access on June 28th, 2022.
- ⁵ Refers to 'Common agricultural policy funds' at https://agriculture.ec.europa.eu/common-agricultural-policy/financing-cap/cap-funds_en. Last access on June 28th, 2022.
- ⁶ Refers to 'Common monitoring and evaluation framework' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/cmef_en, last access on March 8, 2022.
- ⁷ While our approach integrates both policy frameworks and academic research to enhance keyword coverage, we acknowledge that no predefined dictionary can fully capture the complexity and evolving nature of CAP priorities. The reliance on predefined keywords may exclude emerging terms or policy nuances not explicitly reflected in the selected sources. However, we argue that potential biases and the robustness of our results based on keyword identification can be partially addressed by cross-checking with CAP budget allocations.
- ⁸ Notably, text mining can compute word frequencies in policy documents without a predefined dictionary. However, this approach demands substantial computational resources and produces results that are difficult to interpret. Typically, dictionary-free text mining is applied to keyword sections in large academic corpora [11]. In our study, policy documents lack the structured keywords found in academic articles; therefore, it is both meaningful and methodologically sound to first define a set of targeted keywords using a predefined dictionary. This strategy is also common in previous research [63], where keywords were selected for each topic to maximise the precision of topic identification.
- ⁹ Stop words refers to words designated in advance to be of no interest, and which are therefore discarded prior to analysis, such as 'the'.
- ¹⁰ Stemming of the words deletes the word suffixes, creating the base word or what is known as the radical.
- ¹¹ Refer to 'Public opinion on the common agricultural policy' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/cap-glance/eurobarometer_en. Last access on June 29, 2022.
- ¹² Refer to 'EU support for small farms' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/income-support/additional-optional-schemes/small-farmers-scheme_en. Last access on March 19, 2022.
- ¹³ Refer to 'CAP support for rural development' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/rural-development_en#eip. Last access on March 19, 2022.
- ¹⁴ Refer to 'The new common agricultural policy: 2023-27' at https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/new-cap-2023-27_en#keyareasofreform. Last access on March 22, 2022.

References

1. C. Folmer, M. Keyzer, M. Merbis, H. Stolwijk, and P. Veenendaal, *The Common Agricultural Policy Beyond the Mac-Sharry Reform, Volume 230 of Contributions to Economic Analysis* (Elsevier Science, 1995).
2. A. Swinbank, "CAP Reform, 1992," *JCMS: Journal of Common Market Studies* 31, no. 3 (1993): 359–372, <https://doi.org/10.1111/j.1468-5965.1993.tb00469.x>.
3. G. Alons and P. Zwaan, "22 the Common Agricultural Policy: common Dynamics of Policy Change in an Uncommon Policy Domain," in *The Routledge Handbook of European Public Policy* N. Zahariadis and L. Buonanno, eds. (Taylor & Francis Ltd, 2017), 245–301.
4. A. Swinbank, "CAP Reform and the WTO: Compatibility and Developments," *European Review of Agricultural Economics* 26, no. 3 (1999): 389–407, <https://doi.org/10.1093/erae/26.3.389>.
5. N. Beard and A. Swinbank, "Decoupled Payments to Facilitate CAP Reform," *Food Policy* 26, no. 2 (2001): 121–145, [https://doi.org/10.1016/S0306-9192\(00\)00041-5](https://doi.org/10.1016/S0306-9192(00)00041-5).
6. J. Swinnen, "1. The Political Economy of the 2014–2020 Common Agricultural Policy: Introduction and Key Conclusions," in *The Political Economy of the 2014–2020 Common Agricultural Policy* J. Swinnen, ed. (Rowman & Littlefield International Ltd., 2015), 1–30.
7. G. Anania and M. R. P. D'Andrea, "2. The 2013 Reform of the Common Agricultural Policy," in *The Political Economy of the 2014–2020 Common Agricultural Policy* J. Swinnen, ed. (Rowman & Littlefield International, 2015), 33–87.
8. P. Lowe, H. Buller, and N. Ward, "Setting the Next Agenda? British and French Approaches to the Second Pillar of the Common Agricultural Policy," *Journal of Rural Studies* 18 (2002): 1–17, [https://doi.org/10.1016/S0743-0167\(01\)00025-0](https://doi.org/10.1016/S0743-0167(01)00025-0).
9. A. Greer, "The Common Agricultural Policy and the EU Budget: Stasis or Change?," *European Journal of Government and Economics* 2 (2013): 119–136, <https://doi.org/10.17979/ejge.2013.2.2.4291>.
10. G. Alons, "Environmental Policy Integration in the EU's Common Agricultural Policy: Greening or Greenwashing?," *Journal of European Public Policy* 24, no. 11 (2017): 1604–1622, <https://doi.org/10.1080/13501763.2017.1334085>.
11. G. Fusco, "Twenty Years of Common Agricultural Policy in Europe: A Bibliometric Analysis," *Sustainability* 131, no. 19 (2021): 10650, <https://doi.org/10.3390/su131910650>.
12. M. Dumangane, M. Freo, S. Granato, A. Lapatinas, and G. Mazarella, *The Regional Dimension of the CAP: 2007–2018*, Vol. EUR 30878 EN (Publications Office of the European Union, 2021): ISBN 978-92-76-43170-1, <https://doi.org/10.2760/60203,JRC125450>.
13. K. Isoaho, D. Gritsenko, and E. Mäkelä, "Topic Modeling and Text Analysis for Qualitative Policy Research," *Policy Studies Journal* 49, no. 1 (2021): 300–324, <https://doi.org/10.1111/psj.12343>.
14. H. Storm, K. Baylis, and T. Heckeley, "Machine Learning in Agricultural and Applied Economics," *European Review of Agricultural Economics* 47, no. 3 (2020): 849–892, <https://doi.org/10.1093/erae/jbz033>.
15. H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganeh, "Text Mining in Big Data Analytics," *Big Data and Cognitive Computing* 4 (2020): 1, <https://doi.org/10.3390/bdcc4010001>.
16. F. Shi, L. Chen, J. Han, and P. Childs, "A Data-Driven Text Mining and Semantic Network Analysis for Design Information Retrieval," *Journal of Mechanical Design* 139, no. 11 (2017), <https://doi.org/10.1115/1.4037649>.
17. S. Juventia, S. Jones, M.-A. Laporte, R. Remans, C. Villani, and N. Estrada-Carmona, "Text Mining National Commitments Towards Agrobiodiversity Conservation and Use," *Sustainability* 12, no. 2 (2020): 715, <https://doi.org/10.3390/su12020715>.
18. F. Sebastiani, "Machine Learning in Automated Text Categorization," *ACM Computing Surveys* 34 (2002): 1–47, <https://doi.org/10.1145/505282.505283>.
19. L. Cei, E. Defrancesco, and G. Stefani, "What Topic Modelling Can Show About the Development of Agricultural Economics: Evidence From the Journal Citation Report Category Top Journals," *European Review of Agricultural Economics* 49, no. 2 (2022): 289–330, <https://doi.org/10.1093/erae/jbab055>.
20. M. W. Cooper, M. E. Brown, M. T. Niles, and M. M. ElQadi, "Text Mining the Food Security Literature Reveals Substantial Spatial Bias and Thematic Broadening Over Time," *Global Food Security* 26 (2020): 100392, <https://doi.org/10.1016/j.gfs.2020.100392>.
21. E. Erjavec and M. Lovec, "Research of European Union's Common Agricultural Policy: Disciplinary Boundaries and beyond," *European Review of Agricultural Economics* 44, no. 4 (2017): 732–754, <https://doi.org/10.1093/erae/jbx008>.
22. K. Erjavec and E. Erjavec, "Framing Agricultural Policy Through the EC's Strategies on CAP Reforms (1992–2017)," *Agricultural and Food Economics* 9 (2021): 1–18, <https://doi.org/10.1186/s40100-021-00178-4>.
23. B. Camaioni, R. Esposti, F. Pagliacci, F. Sotte, One Policy, Many Policies: The Spatial Allocation of First and Second Pillar CAP Expenditure. (2014).
24. EP, *European Parliament: Fact Sheets on the European Union* (Financing of the CAP, 2023), May 16, 2023.
25. G. Pe'er, A. Bonn, H. Bruelheide, et al., "Action Needed for the EU Common Agricultural Policy to Address Sustainability Challenges," *People and Nature* 2 (2020): 305–316, <https://doi.org/10.1002/pan3.10080>.
26. G. Pe'er, S. Lakner, R. Müller, et al., *Is the CAP Fit for Purpose. An Evidence-Based, Rapid Fitness-Check Assessment* (German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 2017).
27. S. Quan and X. Yu, "China's Agricultural Policy System and its International Competitiveness," *REFORM* (2016): 130–138: (in Chinese).
28. T. W. Schultz, "The Economic Organization of Agriculture," *Geographical Review* 43, no. 4 (1953): 577, <https://doi.org/10.2307/212041>.
29. X. Yu and G. Zhao, "Chinese Agricultural Development in 30 Years: A Literature Review," *Frontiers of Economics in China* 4 (2009): 633–648, <https://doi.org/10.1007/s11459-009-0034-y>.
30. Tangermann, S., von Cramon-Taubadel, S., Agricultural Policy in the European Union: An Overview. (2013)
31. J.-C. Bureau and J. Swinnen, "EU Policies and Global Food Security," *Global Food Security* 16 (2018): 106–115, <https://doi.org/10.1016/j.gfs.2017.12.001>.
32. D. Koemle, S. Lakner, and X. Yu, "The Impact of Natura 2000 Designation on Agricultural Land Rents in Germany," *Land Use Policy* 87 (2019): 104013, <https://doi.org/10.1016/j.landusepol.2019.05.032>.
33. D. Koemle, Y. Zinngrebe, and X. Yu, "Highway Construction and Wildlife Populations: Evidence From Austria," *Land Use Policy* 73 (2018): 447–457, <https://doi.org/10.1016/j.landusepol.2018.02.021>.
34. T. Radić, O. Mikuš, A. Matin, and R. Franić, "Defining the Concept of 'Genuine Farmer' Under the Common Agricultural Policy after 2020," *Agriculture Conspectus Scientificus* 86 (2021): 265–270.
35. V. Graskemper, X. Yu, and J.-H. Feil, "Farmer Typology and Implications for Policy Design—An Unsupervised Machine Learning Approach," *Land Use Policy* 103 (2021): 105328, <https://doi.org/10.1016/j.landusepol.2021.105328>.

36. V. Graskemper, X. Yu, and J.-H. Feil, "Values of Farmers—Evidence From Germany," *Journal of Rural Studies* 89 (2022): 13–24, <https://doi.org/10.1016/j.jrurstud.2021.11.005>.
37. K. Welbers, W. Van Atteveldt, and K. Benoit, "Text Analysis in R," *Communication Methods and Measures* 11, no. 4 (2017): 245–265, <https://doi.org/10.1080/19312458.2017.1387238>.
38. F. Ingo and H. Kurt, "Tm: Text Mining Package," [Computer software manual], R package version 0.7-8, (2020), <https://CRAN.R-project.org/package=tm>.
39. C. Lesmeister, *Mastering Machine Learning with R* (Packt Publishing Ltd, 2015).
40. J. Ooms, "Pdfutils: Text Extraction, Rendering and Converting of Pdf Documents," [Computer software manual], (2022), <https://CRAN.R-project.org/package=pdfutils>.
41. A. Ghatak, *Machine Learning with R* (Springer, 2017).
42. EC, "Special Eurobarometer 504—October 2020 Report," *Europeans, Agriculture and the CAP* (2020), <https://europa.eu/eurobarometer/surveys/detail/2229>.
43. M. Rivera, A. Guarín, T. Pinto-Correia, et al., "Assessing the Role of Small Farms in Regional Food Systems in Europe: Evidence From a Comparative Study," *Global Food Security* 26 (2020): 100417, <https://doi.org/10.1016/j.gfs.2020.100417>.
44. M. Glowinkel, M. Mocan, and M. Külkens, "Survey of German Farmers Towards the Impact of the Common Agricultural Policy on Their Businesses," *Bulgarian Journal of Agricultural Science* 27 (2021): 646–655.
45. M. Guth, K. Smędzik-Ambroży, B. Czyżewski, and S. Stepień, "The Economic Sustainability of Farms Under Common Agricultural Policy in the European Union Countries," *Agriculture* 10, no. 2 (2020): 34, <https://doi.org/10.3390/agriculture10020034>.
46. S. Zawisza, P. Prus, and S. Beben, "Development of Agricultural Farms in Terms of Common Agricultural Policy Support in the Opinion of Farmers," *Economic science for rural development* 50 (2019): 264–271, <https://doi.org/10.22616/esrd.2019.033>.
47. T. Balezentis, E. Ribasauskiene, M. Morkunas, A. Volkov, D. Streimikiene, and P. Toma, "Young Farmers' Support Under the Common Agricultural Policy and Sustainability of Rural Regions: Evidence From Lithuania," *Land Use Policy* 94 (2020): 104542, <https://doi.org/10.1016/j.landusepol.2020.104542>.
48. G. Pe'er, L. V. Dicks, P. Visconti, et al., "EU Agricultural Reform Fails on Biodiversity," *Science* 344, no. 6188 (2014): 1090–1092, <https://doi.org/10.1126/science.1253425>.
49. G. Pe'er, Y. Zinngrebe, F. Moreira, et al., "A Greener Path for the EU Common Agricultural Policy," *Science* 365, no. 6452 (2019): 449–451, <https://doi.org/10.1126/science.aax3146>.
50. EC. The CAP towards 2020: Meeting the Food, Natural Resources and Territorial Challenges of the Future. (2010), https://www.europarl.europa.eu/doceo/document/A-7-2011-0202_EN.html.
51. EC. CAP Reform—An Explanation of the Main Elements. (2013), https://ec.europa.eu/commission/presscorner/detail/en/MEMO_13_621.
52. J. Hristov, Y. Clough, U. Sahlin, et al., "Impacts of the EU's Common Agricultural Policy 'Greening' Reform on Agricultural Development, Biodiversity, and Ecosystem Services," *Applied Economic Perspectives and Policy* 42, no. 4 (2020): 716–738, <https://doi.org/10.1002/aep.13037>.
53. A. Cormont, H. Siepel, J. Clement, et al., "Landscape Complexity and Farmland Biodiversity: Evaluating the CAP Target on Natural Elements," *Journal for Nature Conservation* 30 (2016): 19–26, <https://doi.org/10.1016/j.jnc.2015.12.006>.
54. E. Saraceno, "Rural Development Policies and the Second Pillar of the Common Agricultural Policy," in *Policy Vision for Sustainable Rural Economies in an Enlarged Europe* (Verlag der ARL—Akademie für Raumforschung und Landesplanung, 2003), 197–222: ISBN 3-88838-230-0.
55. C. Nazzaro and G. Marotta, "The Common Agricultural Policy 2014–2020: Scenarios for the European Agricultural and Rural Systems," *Agricultural and Food Economics* 4 (2016): 1–5, <https://doi.org/10.1186/s40100-016-0060-y>.
56. Dumangane, M., Freo, M., Granato, S., Lapatinas, A., Mazzarella, G., . *An Evaluation of the CAPs Impact: A Discrete Policy Mix Analysis*, EUR 30880 EN, (Publications Office of the European Union, 2021) ISBN 978-92-76-43291-3, <https://doi.org/10.2760/72177>, JRC125451.
57. EC, *Communication from the Commission* (European Green Deal, 2019), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN>.
58. EC. Working With Parliament and Council to Make the CAP Reform Fit for the European Green Deal. (2020), https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/key_policies/documents/factsheet-cap-reform-to-fit-european-green-deal_en.pdf.
59. P. H. Feindt, "Policy-Learning and Environmental Policy Integration in the Common Agricultural Policy, 1973–2003," *Public Administration* 88, no. 2 (2010): 296–314, <https://doi.org/10.1111/j.1467-9299.2010.01833.x>.
60. T. Šumrada, M. Lovec, L. Juvančič, I. Rac, and E. Erjavec, "Fit for the Task? Integration of Biodiversity Policy Into the Post-2020 Common Agricultural Policy: Illustration on the Case of Slovenia," *Journal for Nature Conservation* 54 (2020): 125804, <https://doi.org/10.1016/j.jnc.2020.125804>.
61. D. Blandford and T. Josling, "Should the Green Box Be Modified," *International Food & Agricultural Trade Policy Council (IPC) Discussion Paper. March* (2007).
62. European Court of Auditors. Is Cross Compliancean Effective Policy, 12, Rue Alcide De Gasperi, 1615 Luxembourg. (2008).
63. K. Krawczyk, T. Chelkowski, D. J. Laydon, et al., "Quantifying Online News Media Coverage of the COVID-19 Pandemic: Text Mining Study and Resource," *Journal of Medical Internet Research* 23, no. 6 (2021): e28253, <https://doi.org/10.2196/28253>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.