Econometric analysis of patent data: The case study of eternal pollutants and bioremediation (1988-2024)

Author: Dessislava Nedialkova

Abstract: The publication studies the correlation between the number of patent applications for bioremediation and for eternal pollutants (PFOS / PFAS), using econometric models as a tool for the calculations. The study is based on data from the European Patent Office, forming time series to which a vector autoregression model (VAR) was then applied. The results show that the number of applications for PFOS is related to the number of applications for bioremediation and the correlation is negative, i.e. when the number of applications for PFOS increases, the number of applications for bioremediation decreases.

Keywords: vector autoregression; patents; eternal pollutants; bioremediation

JEL: A11; C1

INTRODUCTION

The subject of this paper is eternal pollutants, also called “forever chemicals” (PFOS/PFAS). The problem of their impact on environment is still ongoing research and so far, it is still too distant from the focus of public interest. The data shows that, as of this moment, monitoring on the issue is vital. This is still far from anyone taking any decisive actions and from considerable involvement of the institutions that are required to exercise control, but the more reliable data and research is available, the more likely a positive outcome will be. We find it important not only to examine eternal pollutants as such, but also to place them in a broader context together with bioremediation as means for purification. The aim is to establish the interdependencies between patent activity in the field of eternal pollutants and in the field of bioremediation, using an econometric method for analysis.

The objective is achieved through the following research tasks:

1. To identify the object "eternal pollutants & bioremediation" as a polluting product and a means of purification.
2. To derive an econometric model and justify its selection.
3. To analyze the data using an autoregressive model and to identify the characteristics, benefits, advantages, and disadvantages of the results obtained.

This paper examines the influence of three variables:

- Patent applications for bioremediation
- Patent applications for PFAS
- Patent applications for PFOS

The relationships between these three variables have not been studied to date, which provides an opportunity to test whether there are dependencies between them and in what direction any correlations operate. This study uses the econometric model of vector autoregression (VAR) to demonstrate how different variables influence each other over time.

The research methodology is interdisciplinary, containing a systematic and comprehensive approach, based on data analysis theory and the possibility of creating forecasts. The research thesis defended in the article is that there is a connection between patent applications for eternal pollutants and patent applications related to bioremediation.

LIMITATIONS OF THE STUDY

The focus is on establishing a link between eternal pollutants (PFOA and PFOS) and bioremediation as a process for purifying the environment, through an analysis of patent applications. It is important to note that the article only examines the number of patent applications by year and does not conduct a patent study or research on the content of the applications themselves. PFOS and PFAS are examined as chemical substances, but not in their direct application as a result of the patent, and the same applies to bioremediation. It is considered as a purification process, but not as a direct application as a result of the registered patent.

METHOD

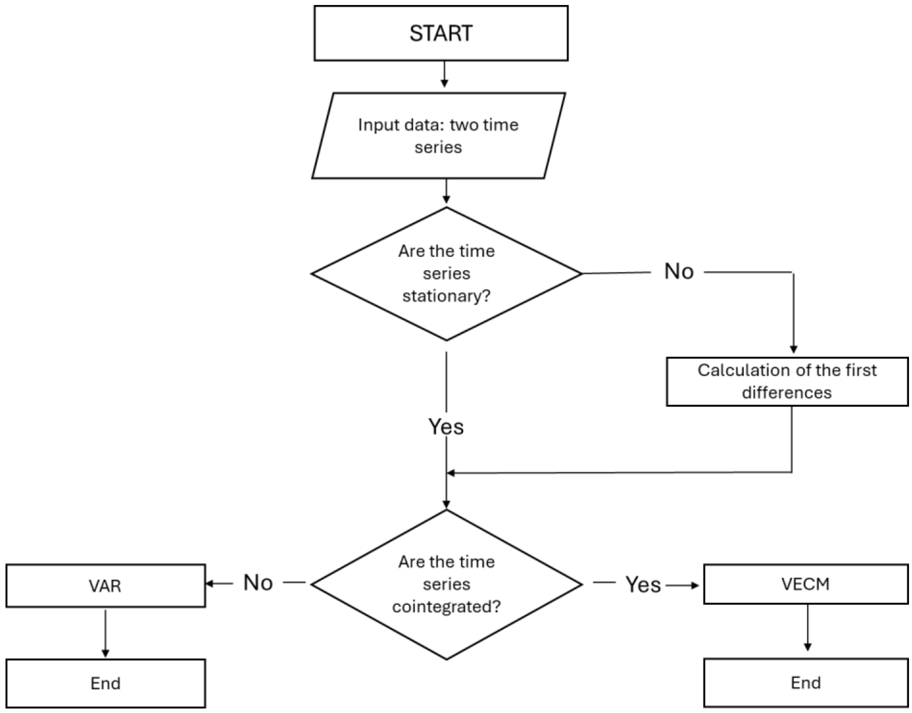
We will examine the correlations between the three variables by dividing them into pairs and tracking whether they influence each other and, if there is a correlation, which of the variables is a determinant and which is a result. Since the source data is in the form of time series, we can apply Granger's approach, which is used to check whether the dynamics of one series cause the changes present in another series, i.e., it shows how useful one variable is for predicting changes in another. For this purpose, Granger defines a simple causal two-variable model (Granger, 1969, p. 431). What Granger is actually saying through this model is that:

- If the current state of X depends on the past states of Y, but the current state of Y does not depend on the past states of X, then Y is causing X;
- If the current state of Y depends on the past states of X, but the current state of X does not depend on the past states of Y, then X is causing Y;
- If both the current state of X depends on the past state of Y, and the current state of Y depends on the past states of X, then there is mutual relationship between Y and X;
- If neither the current state of X depends on the past states of Y, nor the current state of Y depends on the past states of X, then there is no relationship between Y and X.(Haralampiev, 2025)

Main requirement of the model is that X and Y has to be stationary time series. The dynamic series derived from real data, in a large proportion, are non-stationary and that seriously impedes the correlation analyses based on the information contained therein. The scientific studies regularly recommend as a method for coming out of this conundrum to use the first differences instead of the level form. Series with a tendency to develop are called integrated series of a certain order, where the number of the first differences that must be applied to transform the series into stationary determines the order of integration. When a linear combination of integrated series is formed, it should also be integrated—i.e., the linear combination of non-stationary series should also be non-stationary.

However, there is a specific case when linear combination of non-stationary series becomes stationary. This specific case is described by Ivanov, L., Ovchinnikov, E. (2018) as cointegration between the dynamic series in the linear combination. When we have cointegration, we apply error correction models (e.g. Vector error correction model VECM). Thus, following the already described process we reach a point where we need to choose which model to employ. We will take the occasion to quote here assoc. prof Kaloyan Haralampiev (2025) who conveniently illustrates the process with a visual representation of the following algorithm:

Fig. 1 Visual representation of the Granger algorithm



Source: Assoc. prof K. Haralampiev “Who benefits from GDP growth in Bulgaria?”

Stationarity test in the study is completed through the Augmented Dickey-Fuller test (ADF), that one being the most popular and most employed for this particular purpose. To determine if the time series we have selected are cointegrated we apply the method of Johansen. As regards the selected model – vector autoregression gives us the ability to monitor the reaction of a variable, when another variable changes. It is considered as an Autoregressive model because, each variable (time Series) is modeled as a function of the past values, that is the predictors are nothing but the lags (time delayed value) of the series.

In other words, each variable is modeled as a linear combination of past values of itself and the past values of other variables in the system. That means that the past values of the series are used to forecast the present and the future. Since you have multiple time series that influence each other, it is modeled as a system of equations with one equation per variable (time series).

The primary difference of this model is that does not work uni-directionally, (the predictors influence the Y and not vice-versa), but gives us all the correlations between the variables simultaneously. Applying VAR the model can analyze and forecast and answer the following questions:

- What could happen to bioremediation applications if applications for PFOS and PFAS increase or decrease?
- Why is patent data important for measuring innovation?
- What is the significance of PFOS, PFAS, and bioremediation in an environmental and technological context?

DATA

This study uses data for the period 1988 – 2024. We are studying two time series (number of patent applications and year) and three variables:

- Patent applications for bioremediation,
- Patent applications for PFOS,
- Patent applications for PFAS

The data was collected from the European Patent Office by searching the database using the relevant search criteria: PFAS, PFOS, bioremediation. The results obtained for exports were used to form time series for each of the three variables, which were used as input data in the analysis.

Without going into detail about organic chemistry, here is a working definition of the basic concepts used in relation to the topic of the article:

- Perfluorooctane sulfonic acid (PFOS) is used as a surfactant with the molecular formula C8F17SO3H. This chemical formula is scientifically proven to be an endocrine disruptor and pollutant, listed in Annex B of the Stockholm Convention on Persistent Organic Pollutants. It is derived mainly from perfluorooctane sulfonic acid C8F17SO2NH2 (PFOSA) and until the beginning of the century was the main ingredient in Scotchgard treatment, developed by 3M for impregnating fabrics, furniture, and carpets. It is also used to treat paper that comes into contact with food against water and grease stains, as well as for various other applications for the general public. Its traditional substitute is PFAS.
- PFAS (per- and polyfluoroalkyl substances), also known as "forever chemicals" due to their persistence in the environment, formerly also called perfluorinated compounds, are synthetic organic fluorine compounds containing one or more functional groups of perfluorinated or polyfluorinated alkyl groups (not all). They contain at least one perfluoroalkyl group. There are probably between six and seven million different PFAS. They have attracted the attention of researchers, regulatory authorities, and environmental NGOs due to their toxicity and ecotoxicity, their persistence as pollutants, and their now widespread presence in water, air, soil, and ecosystems (especially fauna), and in the blood of a large portion of the world's population. They are found in living organisms across the planet. Scientists and various administrations are calling for rapid "regulation, monitoring, and management" of PFAS. In 2025, the cost of cleaning them up in European waters and soils is estimated to be between $95 billion and $2 trillion over a 20-year period. PFAS are part of a "large family of over four thousand chemical compounds" with their anti-adhesive, water-repellent, and heat-resistant properties. Since the 1950s, "they have been widely used in various industrial fields and everyday products: textiles, food packaging, firefighting foams, non-stick coatings, cosmetics, plant protection products, etc." Some PFAS may also be found in food packaging such as burgers or agricultural raw materials, ski waxes, or non-stick pans. In this article, we will often refer to PFOS and PFAS collectively and use the term "forever chemicals."

- Bioremediation is the decontamination of polluted environments using techniques based on chemical degradation or the activity of living organisms. The technology is used to treat various types of waste (agricultural, food, etc.) and wastewater. In the context of this article, bioremediation is applied to the treatment of hazardous waste. The idea is to enable the purification of contaminated soil and wastewater. Bioremediation is a method that uses microorganisms in the soil to carry out a process of cleaning the environment, as close as possible to natural processes.

The patent study identified a large number of patent applications for eternal pollutants and for bioremediation processes, which gave rise to the idea of examining possible correlations and attempting to forecast their future development using econometric models and based on public opinion that the most important thing at the moment is to protect the environment and the atmosphere, air, water, and soil. We truly believe that the analysis is useful and could serve the interested institutions for future decision-making and taking concrete action.

RESULTS

The time series are tested for stationarity using the method of unit root, relying on the ADF. The results are presented in table 1.

Tab. 1 Unit root calculation

	level form			first differences		
	t	p	lag	t	p	lag
PFOS	0.30	0.923	0	**-8.23**	**0.000**	0
Bioremediation	-1.38	0.590	1	**-8.11**	**0.000**	0
PFAS	2.11	1.000	0	**-5.07**	**0.000**	0

Source: Author's own calculations

The results show that the original time series for the three variables have a unit root and are not stationary, but the time series from the first differences are stationary. These results mean that for all indicators, it is possible to work with the time series of the first differences.

To test for cointegration we use the method of Johansen. The results are presented in table 2. In the last column we mark the used model.

Tab. 2. Cointegration test

		Trace Statistics	p	
PFOS	Bioremediation	8.96	0.376	VAR
PFAS	Bioremediation	14.52	0.069	VAR

Source: Author's own calculations

The null hypothesis is that there is no cointegration. The results show that there is no cointegration for either of the dependencies we are examining (for PFOS p=0.376, for PFAS p=0.069). In this sense, the null hypothesis is proven and we should proceed to apply the vector autoregression model.

The results from applying the VAR model are presented in table 3.

Tab. 3 VAR results

	ΔPFOS					ΔBioremediation		
	coefficien	t	p			coefficien	t	p
Const	1.42	1.90	0.067		Const	8.18	**2.06**	**0.047**
ΔPFOS(-1)	-0.34	-2.01	0.053		ΔPFOS(-1)	-2.51	**-2.81**	**0.009**
ΔBioremediation(-1	-0.01	-0.19	0.854		ΔBioremediation(-1	-0.45	**-2.53**	**0.017**
Adjusted R-squarec	0.066				Adjusted R-squarec	0.316		
F(2, 32)	2.20				F(2, 32)	**8.87**		
P-value(F)	0.127				P-value(F)	**0.001**		
	ΔPFAS					ΔBioremediation		
	coefficien	t	p			coefficien	t	p
Const	5.06	1.38	0.179		Const	7.71	1.75	0.090
ΔPFAS(-1)	0.12	0.69	0.496		ΔPFAS(-1)	-0.28	-1.32	0.197
ΔBioremediation(-1	0.10	0.62	0.539		ΔBioremediation(-1	-0.53	**-2.82**	**0.008**
Adjusted R-squarec	-0.033				Adjusted R-squarec	0.192		
F(2, 32)	0.45				F(2, 32)	**5.04**		
P-value(F)	0.641				P-value(F)	**0.013**		

Source: Author's own calculations

The results show that the number of PFOS applications is related to the number of bioremediation applications in a negative correlation i.e. when the number of applications for PFOS increases, the number of applications for bioremediation decreases.

The interpretation of the results is based on the following principles:

- If all results are zero, they do not influence each other;
- Values of p above 0.05 are not statistically significant;

As can be seen from Table 3, the change in bioremediation, measured by the number of patent applications, relative to PFOS, depends on itself in the previous year (p=0.017) and on PFOS in the previous year (p=0.009). For PFAS, bioremediation depends only on itself (p=0.008), but not on PFAS for previous periods (p=0.197).

We also examine the following statistics:

- Adjusted R-squared is a measure that shows us how well the model explains the changes in a given variable, taking into account how many variables we have included in it.
- We use P-value to help us decide whether the relationship we see in the data is "real" or random. A small P-value (e.g., below 0.05) means: "Yes, this relationship is likely to be significant." A large P-value (e.g., above 0.10) means: "This effect is probably not significant." In vector autoregression analyses, the p-value is used to see whether one variable actually influences another—or appears to influence it, but this is statistically insignificant.
- F-statistic – used to test whether one group of variables has a significant effect on another in the model. The F-statistic helps us understand whether the model "makes sense" or not.

The results in Table 3 show that, when examining the above indicators, the right-hand side models are adequate to explain the changes in the respective variables, while the left-hand side models are not.

The negative correlation between PFOS applications and bioremediation suggests a technological and strategic divergence. Chemical remediation methods remain cheaper and more readily accessible, leading to a substitution effect where industries prioritize these over more complex biotechnological solutions. At the same time, bioremediation often demands significant upfront investment and longer time horizons before results can be demonstrated, which weakens its competitive standing in markets driven by faster returns. On a practical level, bioremediation is naturally more effective against organic pollutants such as hydrocarbons and solvents—compounds that microorganisms have adapted to over millennia—whereas “forever chemicals” like PFOS only entered widespread use in the last half-century, leaving microbes ill-equipped to degrade them without deliberate engineering. As a result, current remediation strategies for PFOS are largely misaligned with bioremediation approaches, raising the critical policy question of whether greater research efforts should be directed toward developing microbial strains capable of addressing these persistent pollutants.

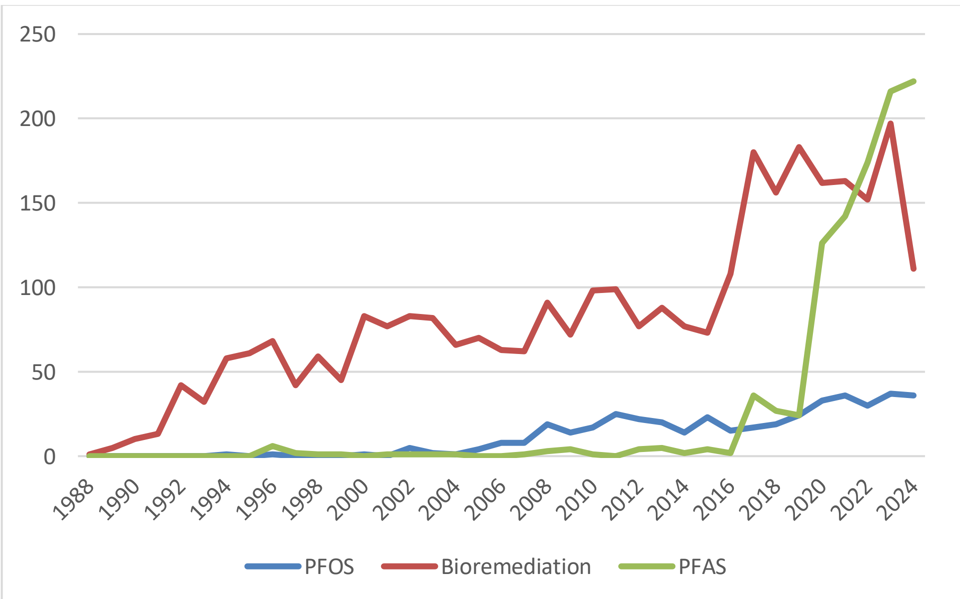
DISCUSSION

In this paper we have set out to establish the interdependencies between patent activity in the field of eternal pollutants and in the field of bioremediation, using an econometric method for analysis. As the process above demonstrates we have used the first differences for our calculations, because of the unit root (Tab. 1). But so far, we have not discussed the raw data. Here we will turn our attention to that and establish a trend for each of the three variables, trying to outline their development over the years, how they compare to each other and what does this signify.

We are using the same time series as above to establish a trend for each of the three variables and then compare them. Again, we look upon the “forever chemicals” in terms of patent applications, mainly for industrial needs and on the bioremediation in terms of patent applications mainly for processes for purification of water, soil and air.

On the graph (Fig. 2) we present the number of patent applications for the period 1988–2024 for the three variables (PFAS, PFOS and bioremediation).

Fig. 2 Establishing trends



Source: Author’s own calculations based on data from Espacenet-Patent Search

We can see in the graph displayed the dynamics of the scientific and technological development in the context of strengthened international and European regulations on eternal pollutants. In the beginning of the nineties there is no activity in PFOS и PFAS and the second part of the decade changes that only slightly. In fact, a slight rise in applications for PFOS is not observed until 2008 and for PFAS all the way until 2017, when we witness a true surge in applications. Bioremediation on the other hand is climbing throughout the whole period with a big upswing in 2016 and a sharp drop at the end of the studied period. This trend seems easy to explain as bioremediation is specialized in cleaning petrochemicals and is yet to prove effective against eternal pollutants.

All this coincides with a period of intense regulatory changes aimed at restricting or completely banning these PFOS and PFAS substances. In fact, it is the busiest period for the International and European Regulatory Framework and the Stockholm Convention on Persistent Organic Pollutants (2009, 2019, 2023). This international convention gradually adds PFOS (2009), PFOA (2019), and a number of other PFAS (2023) to the list of substances subject to restriction or elimination. EU Member States are required to develop substitution and disposal technologies, which stimulates patent activity in the field of environmentally friendly solutions. Major milestone here is Regulation (EC) 1907/2006 (REACH). The REACH system regulates the production and use of chemicals in the EU. PFOS is banned by Directive 2006/122/EC, and PFOA and its salts are banned by Regulation (EU) 2019/1021 on persistent organic pollutants. From 2023, the European Chemicals Agency (ECHA) has proposed a comprehensive ban on over 10,000 PFAS substances. All these actions have led to rapid growth in innovation aimed at alternative materials and processes that meet the new requirements. As part of its Zero Pollution for Europe policy, the European Commission is focusing on sustainable chemicals management, clean technologies, and soil and water remediation, included in the European Green Deal.

This creates a favorable environment for scientific research and patent activity in the field of bioremediation, which is considered a promising technology for the long-term elimination of PFAS contamination. Bioremediation has shown steady growth since the 1990s, but its importance has increased particularly since 2015, when EU policy documents began to promote environmentally friendly methods of depollution. Among the most important initiatives are:

- The Water Framework Directive (2000/60/EC)
- The upcoming Soil Directive, aimed at improving soil resilience and ensuring sustainable management of contaminated sites (adoption procedure in the Council of the EU is concluded, the European Parliament is expected to hold its final vote. Member states will have three years after entry into force to transpose the new rules into national law).
- The EU Strategy for Sustainable Chemicals (2020)

All of these encourage the development of integrated biotechnological solutions to reduce PFAS pollution, which is clearly reflected in the increased number of patents since 2018.

From the above, we can conclude that there is a direct link between legislative restrictions (regulations) and innovation activity. Following the introduction of bans and restrictions on PFOS/PFAS, the industry responded by intensively developing new technologies.

The temporal evolution of patent activity highlights a clear shift in environmental innovation priorities over the past three decades. Early interest in bioremediation technologies (1990s–2010s) reflects a foundational period of developing biological solutions for pollutant removal. However, since approximately 2018, there has been a pronounced surge in patents related to PFAS, signaling an emergent technological frontier driven by regulatory pressure and global concern over persistent “forever chemicals.”

The relatively stable yet modest growth of PFOS-specific patents suggests a transition from compound-specific remediation approaches toward holistic PFAS treatment strategies. The sharp rise in PFAS-related patents post-2018 likely corresponds to increased detection of PFAS in water systems, policy interventions in North America and Europe, and industrial investment in novel degradation methods (chemical, photolytic, and biological).

CONCLUSION

In this publication we studied the correlation between the number of patent applications for bioremediation and for eternal pollutants (PFOS / PFAS). As the time series was not stationary, we used the first differences to apply vector autoregression model. The results show that the number of applications for PFOS is related to the number of applications for bioremediation and the correlation is negative, i.e. when the number of applications for PFOS increases, the number of applications for bioremediation decreases. We interpret this result as a technological and strategic divergence – technology in bioremediation is blooming and is considered a path to a more sustainable future, but is not yet ready to substitute traditional remediation methods as they remain cheaper and more readily accessible.

At the same time, our analyses of the trends of the raw data, indicate that the regulation boom in these years shape innovation activity and suggest a paradigm shift from general bioremediation toward targeted remediation of recalcitrant fluorinated compounds, aligning technological development with evolving environmental priorities.

REFERENCES:

1. Haralampiev, K. (2025), "Who benefits from GDP growth in Bulgaria?", 50 years Department Economic Sociology : Collection of reports from a scientific conference, UNWE;
2. Ivanov, L., Ovchinnikov, E. (2018), "Applied aspects of econometric modeling of dynamic dependencies, Almanac of Scientific Research, anniversary edition, vol. 26;
3. Granger, C. (1981). Some properties of time series data and their use in econometric model specification. Journal of Econometrics, 23: 121-130;
4. Granger, C., & Weiss, A. (1983). Time series analysis of error-correction models. Studies in econometrics, time series, and multivariate statistics. Academic Press;
5. Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. Econometrica, 37(3), 424-438;
6. Granger, C. (2003). Time Series Analysis, Cointegration, and Applications. Nobel Lecture, December 8, 360-366.
7. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC
8. Regulation (EU) 2019/1021 of the European Parliament and of the Council of 20 June 2019 on persistent organic pollutants (recast)
9. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy
10. European Chemicals Strategy for Sustainability (CSS) as adopted by the European Commission on 14 October 2020