# POLITECNICO DI TORINO

Master's Degree in ENGINEERING AND MANAGEMENT (LM-31)

**Master’s Thesis**

*Patent Landscape on Hydrogen Technologies: Study of Hydrogen Technology Specialisation across European Regions*


Advisor
Prof. Giuseppe Scellato

Co-Advisor
Prof. Antonio De Marco

Candidate
Gian Nicolò Bernini

# INDEX

# INTRODUCTION

Climate change in recent years has emerged as the most critical threat to the stability of the earth and has forced governments of various countries to take urgent action to try to mitigate its negative effects so as to ensure prosperity and well-being for the future generations. CO2 emissions, mainly from the consumption of fossil fuels over the decade 2011-2020, caused a global temperature to rise of 1.09°C compared to pre-industrial levels (1850-1900), as highlighted in the 2020 report of the Intergovernmental Panel on Climate Change (IPCC, 2020). This temperature increase is accelerating the pace of climate change, with dramatic consequences such as extreme droughts, wildfires, more intense heat waves, melting glaciers and rising ocean levels and temperatures.

In 2015, 196 countries pledged to sign the Paris Agreements with the goal of limiting the increase in global average temperature well below 2°C thus limiting dramatic and irreversible effects on the planet. Countries set their sights on the ambitious goal of staying below a 1.5°C increase over pre-industrial levels (UNFCCC, 2015*: "Paris Agreement", United Nations*). In order to achieve these goals, it is strictly necessary to embark on a large-scale energy transition pathway, with an emphasis on the development and adoption of renewable energy sources on the one hand and the reduction and replacement of fossil fuel consumption on the other.

Within this perspective, hydrogen, in recent years, is emerging as a key element in facilitating the transition. Indeed, hydrogen represents a versatile substance with the potential to reduce CO2 emissions in many different industries. First, hydrogen as an energy carrier offers numerous advantages.

Hydrogen represents an inexhaustible resource and can be produced through the electrolysis of water, a process that breaks down the water molecule into hydrogen and oxygen molecules. If the energy used to promote the water electrolysis reaction comes from renewable sources, the hydrogen produced is a renewable resource (Ball and Wietschel, 2009). Specifically in this case, we speak of green hydrogen or clean hydrogen. Moreover, because of its versatility, hydrogen is an energy carrier that can be used in a variety of applications. First, hydrogen can contribute significantly to solving the problems related to energy storage and transportation, which, especially when considering renewable energy, pose major challenges. In fact, renewable sources such as solar and wind provide an intermittent supply that is not always in line with consumption needs. This means that the energy produced from these sources must be stored for later consumption to cover the time lapses in which energy production is low or nil. Batteries can store energy but have limitations in cost, useful life and storage capacity. Regarding transportation, renewable plants to date, are often located near the areas whose energy needs they cover. This is due to the difficulties of transporting energy for considerable distances without storing it. Hydrogen can be a solution to these challenges. Indeed, hydrogen has a high energy density, which implies that it can store a lot of energy in a small volume. It is also a relatively easy element to transport, either through pipelines or in specialized ships. Consequently, Hydrogen can facilitate the balance between energy production and demand, thus ensuring the integration and large-scale development of renewable energy resources.

In addition, hydrogen can be a sustainable solution for the transport industry. Hydrogen can be used in fuel cells to power light and heavy vehicles, offering an environmentally friendly alternative to fossil fuels. Hydrogen can also be used for domestic and commercial heating, providing a low-carbon alternative to fossil fuels.

Another sector that could benefit from hydrogen is the heavy industry, and in particular the iron and steel, which is a very high environmental impact industry. Hydrogen can replace natural gas as a heat source for blast furnaces and significantly reduce CO2 emissions.

However, although hydrogen holds all the necessary characteristics to become a key component of the energy system of the future, there are still major challenges to be overcome in the coming years. Hydrogen-related technologies still require improvements in cost, efficiency and infrastructure. For example, the production of clean hydrogen is currently more expensive than hydrogen produced from fossil fuels and the infrastructure for hydrogen transport and storage is still under development. In addition, there are significant constraints at the geopolitical level that could hinder the global-scale development of the hydrogen value chain. Despite these challenges, forecasts, suggest an increase in hydrogen production and consumption in the coming decades; a trend driven by continuous scientific research, technological improvements, major investments and favourable policies. At the European level, for example, the European Commission has recognized hydrogen's key role in the energy transition. The European Green Deal aims to increase ''green'' hydrogen production to 1 million tons per year by 2024 (European Commission, 2020*: "A hydrogen strategy for a climate-neutral Europe"*). This represents a significant step toward decarbonizing the European energy system.

To better understand the innovation landscape in the field of hydrogen, this master thesis aims to analyse in detail the patent data related to hydrogen and related technologies. Indeed, patents are an essential tool in the study of innovation output as they protect new inventions and processes in various fields, including energy (*OECD, 2020: "Patents and innovation: Trends and policy challenges"*). A patent provides the owner with the exclusive right to commercially exploit an invention for a specified period of time, normally 20 years. This exclusive right stimulates innovation because it gives inventors the security that their creative efforts will be protected and can be exploited economically. Furthermore, patents are an invaluable source of technical information, as each patent must describe the invention clearly and completely so that an expert in the field can reproduce it (European Patent Office, 2020*: "What is a patent?"*).

Analysis of patent data, known as patent landscaping, can reveal technology trends, identify market leaders, potential competitors and possible collaboration opportunities. It can also provide insight into emerging research areas and key technologies in a specific field.

This work aims to use patent data as a tool to study innovation in the field of hydrogen. By integrating this information with statistical studies, a more comprehensive analysis of the domain will be possible and help to outline past and present trends, but also to anticipate future directions of innovation in the hydrogen field.

In conclusion, the goal of this thesis is to contribute to the understanding of the role of hydrogen in the energy transition through an analysis of the patent activity that has developed over the years. Consequently, this work will fit into all the research that aims to provide valuable insights to guide future research and development strategies, energy policies, and investments in the hydrogen sector.

This master thesis is structured into multiple chapters. Initially, the Literature Review will focus on why it is important to study innovation and innovative activities, which are at the basis for the development of economic development, growth and prosperity; secondly it will explain, through the analysis of previous studies how and why it is appropriate to study innovation through the output, and more specifically using patents, considering both advantages and disadvantages of using this method; last it will report studies that have been conducted, especially in recent years on green patented technologies and more specifically hydrogen technologies, illustrating their findings. Following the review of past literature, a chapter will be dedicated to the representation of the hydrogen value chain: from the production to the storage, distribution, transformation, and end-uses of this extremely versatile and promising resource. As the analysis that have been conducted, an individual paragraph will also be dedicated to explaining what patents are and how the patenting procedure of an innovation is conducted. Subsequently, the following chapter, "Data & Methodologies" will explain how the research perimeter of hydrogen technologies has been defined, how the dataset has been identified and extracted from the adopted database. After the cleaning and refinement of the dataset, a series of statistics and analysis have been run to better understand the nature, characteristics, and peculiarities of hydrogen technologies. More specifically, the studies that will be presented in the "Descriptive statistics of the Dataset" chapter reveal interesting insights regarding the evolution in time, value-chain development, geographical distribution and evolution, complexity, collaborativity among innovators, top innovators, technological domain, and greenness/sustainability on hydrogen patented technologies. In the last part of the thesis, a study on hydrogen technology specialisation in European regions will be presented. Technological specialisation in the field of hydrogen will be assessed through patents granted at the European Patent Office (EPO). After describing the patent concentration in Europe on a regional level, two research studies will be developed. In the first, the presence of spatial autocorrelation of hydrogen technology specialisation will be investigated and possible clusters of specialisation will be analysed. In the second, statistical inference will be conducted through an econometric study to analyse which characteristics influence hydrogen technology specialisation in Europe. The thesis will close with conclusions, limitations and possible new developments of the study.

# LITERATURE REVIEW

This chapter aims at investigating the previous literature with respect to the existing relationship between the development of economic growth and prosperity in the long term and innovation. While there may be more intuitive and immediate metrics to account for a country's (e.g., GDP pro capita, life expectancy) wealth and progress, it may be more challenging and complex to measure innovation. Previous literature suggests that an effective way to account for innovation is by measuring the research outputs, and more specifically, patents. Therefore, the second goal of this chapter is to present previous studies that explain why patents represent an accurate proxy of innovation activity. Since the main topic of this thesis is the analysis of sustainable hydrogen technologies, the third objective of this review will be to present some significant examples of studies of green innovations through the use of data provided by patents. Finally, previous examples of patent analysis and patent landscape with specific reference to hydrogen and related innovations will be shown. As much as possible, scientific papers and articles will be presented in chronological order so as to highlight evolution and trends in research.

## Section 1: Innovation and Economic Development

The first section of this literature review concentrates on studies that demonstrate the connection between innovation activities, studies, and investments with long-term growth and prosperity of nations. It is essential to investigate the various aspects of innovation and their effects on countries to gain a deeper understanding of the important economic growth and prosperity factors. Several prominent scientific papers in this field will be discussed, to cast light on their findings and establish a thorough understanding of the topic. In the beginning, it is significant to delve into the link between the role of innovation and its impact on the economic development of an industry or a country.

In this regard, a first meaningful example is provided by the Schumpeter's Theory of Economic Growth (1911). Schumpeter (1911), one of the greatest Austrian economists of the 20$^{th}$ century, widely regarded as the founding father and precursor of modern economic growth theories, introduced the concept of creative destruction, highlighting the role of innovation in propelling economic development (Schumpeter, 1911). According to Schumpeter, new innovations decimate old industries and spawn new ones, thereby fostering economic growth and prosperity. The work of Schumpeter (1911) established the groundwork for comprehending the importance of innovation in shaping the economic landscape.

Afterwards, the topic of innovation, was addressed by Solow (1956) who created the *Neoclassical Growth Model*. The model emphasized the significance of technological advancement in nurturing economic development (Solow, 1956). His model demonstrated that investments in physical and human capital could only result in transient development, whereas technological innovation was the true source of long-term growth. This seminal work had a significant impact on how economists comprehend and research economic growth. Among these economists stands Romer (1990), who extended the neoclassical growth model, proposed by Solow (1956), by integrating endogenous technological change, contending that investments in research and development and human capital also account as important growth drivers (Romer, 1990). Romer's model stresses the significance of knowledge spill overs and rising returns to scale, showing that public policy can play an important role in nurturing innovation and economic growth.

Solow's Neoclassical Growth Model (1956) also served as a baseline for Mankiw et al. (1992), which focused deeply on the influence of a specific factor. Indeed, Mankiw et al. (1992) added human capital as an additional factor of production to the neoclassical development model (Solow, 1956). Human capital, as proxied by education, has a significant impact on economic expansion, according to their empirical research. This study highlights the significance of investing in human capital and education as a mean to foster innovation and long-term economic prosperity.

Numerous studies have also been done regarding the determinants of the development of innovative activities and subsequent growth of the economy. Grossman and Helpman (1991) proposed a conceptual framework for comprehending the role of innovation in propelling global economic development (Grossman and Helpman, 1991). Their work highlights the significance of international trade and R&D investments, demonstrating how these factors contribute to technological advancement and economic expansion. Additionally, their findings underline the importance of international cooperation and policy coordination for fostering global innovation.

Instead, Aghion and Howitt (1990), taking up the concepts developed by Schumpeter (1956), constructed a new development theory (Aghion and Howitt, 1990) that illustrates how creative destruction stimulates economic development by encouraging firms to invest in R&D and innovate. The model believes that entrepreneurs can invest in research and development to come up with new technologies that are more efficient than the current ones. However, these new technologies are not quickly adopted by all companies. Instead, new firms with innovative technologies join the market and replace the current ones, leading to an increase in output and economic growth. In addition, the authors emphasize the role of competition in stimulating innovation, offering a determined view on the intricate relationship existing between market competition, innovation, and economic development.

Furthermore, Fagerberg (1987) carried out a country-level comparative analysis between growth rates. Fagerberg (1987) investigated the relationship between innovation, catching up, and economic development, emphasizing the significance of innovation in fostering growth in both developed and developing nations (Fagerberg, 1987). The author demonstrated that countries that invest in research and development, human capital, and technology adoption have a greater chance of catching up to dominant economies. The work of Fagerberg (1987) emphasizes the significance of innovation in reducing income disparities and fostering a path towards international inclusion.

Another countrywide comparative research was conducted by Hall and Jones (1999). The two American economists examined the role of social infrastructure in promoting economic development and argued that robust institutions and a supportive social infrastructure are essential for fostering innovation and technological advancement (Hall and Jones, 1999). Their empirical findings demonstrate the significance of factors such as property rights, the rule of law, and government efficacy in fostering innovation and long-term economic prosperity.

Acemoglu et al. (2001) also delved into the connection between institutions, innovation, and economic growth (Acemoglu et. al, 2001). The authors argue that inclusive institutions that level the playing field and encourage innovation led to long-term economic growth. Their research emphasizes the importance of institutions in nurturing an environment that is conducive to innovation, and thus, to economic growth.

Furthermore, previous literature also includes relevant studies regarding policies about innovation and the role of governments in fostering the establishment of innovative activities. Lundvall (1992) investigated the idea of a National Innovation System (NIS), focusing on the interactive learning process among system players (Lundvall, 1992). According to Lundvall (1992), a country's economic performance is dependent on its capacity for innovation, which is the product of collaborative learning amongst a variety of stakeholders, including businesses, governments, and academic institutions. The findings of Lundvall's study (1992) emphasize the value of group learning and information exchange in fostering innovation at the national level. Another meaningful analysis within this field is provided by Nelson (1993) who explored innovation and economic development (Nelson, 1993), by investigating how countries stimulate innovation and technical development using the National Innovation Systems (NIS) paradigm, previously illustrated by Lundvall (1992). By comparing NIS in Japan, Germany, and the United States, Nelson (1993) demonstrated the importance of coordinated innovation promotion by businesses, universities, and government institutions. Infrastructure, education, and political institutions determine the viability of a NIS, which can sustain economic development.

Considering the role of government institutions in promoting the development of an innovation ecosystem, a relevant contribution is given by the work of Arrow (1962) who emphasized the importance of government intervention in fostering innovation and generating economic welfare (*Arrow, 1962*). Arrow (1962) argued that market failures impede innovation because firms tend to underinvest in R&D due to the prevalence of externalities and the character of knowledge as a public benefit. In this study, the author also focused on the role of patents (as mentioned before, patents will be deepened in the next sections of the literature review) in incentivizing R&D and emphasized the importance of patents in providing firms with incentives to invest in R&D and engage in innovative activities. The research also highlights the limitations of the patent system, focusing on the trade-offs between disclosure and monopoly power. The influence of Arrow's work (1962) on innovation-related public policy has been significant, leading to the development of government policies and programs such as research grants, fiscal incentives, and funding for fundamental research. These initiatives seek to encourage innovation and provide firms with the resources and incentives necessary to invest in R&D, thereby contributing to economic growth and development.

The same topic was also addressed more recently by Mazzucato (2013). The author argues that the state can play a crucial role in advancing technological development by operating as an entrepreneurial actor (Mazzucato, 2013). The author suggests that public investments in research and development, infrastructure, and human capital, can foster innovation and thus long-term economic growth. The author also emphasizes the significance of public-private partnerships, which can help leverage public investments and facilitate collaboration between government, industry, and academia, to advance innovation.

To conclude, the first chapter of the literature review has provided an overview of the most significant scientific articles discussing the role of innovation, studies, and investments in innovation in promoting economic growth and prosperity among nations. These studies highlight the significance of technological advancement, human capital, institutions, and supportive policies in promoting long-term economic expansion. Moreover, governments play a crucial role in correcting market failures and fostering an environment conducive to innovation. As the global economy continues to evolve, it will be essential for policymakers and researchers equally to comprehend these relationships to shape a prosperous and equitable future.

## Section 2: The use of patents in the analysis of innovative activities

As illustrated in the previous section, the study of innovation has long been of interest to academics, policymakers, and business leaders. Patent data is one of the most important sources of information on innovation, providing valuable insights into the character and patterns of technological change. Patent data and patent analysis have arisen as a popular and useful method for studying innovation, and a vast body of literature has accumulated on the subject. Indeed, since patents contain a lot of semi-structured data, researchers have been able to create and use a variety of quantitative indices and metrics to examine and compare patent data. This section of the literature review seeks to summarize the major contributions and debates in the field, as well as discuss the advantages and disadvantages of using patent data to research innovation.

To begin this second part of the literature review, the reasons why it is relevant to study innovation through patent data are presented by using a selection of notable publications. A primary argument for using patent data to study innovation is that patents are often adopted as indicators of innovation and innovative activity. This assumption is well supported by Griliches (1990), who examined the potential advantages and disadvantages of utilizing patent data in economic research (Griliches, 1990). The author emphasized the vast amount of valuable information that can be derived from patent data, and the data’s capacity to offer valuable insights into the process of innovation, the economic ramifications of technological progress, and the competitive environment of diverse industries. The study also highlights the prospective uses of patent data in figuring out the connections between R&D expenditures, technological advancement, and productivity development, underlining the versatile applications of patent data in studying economic phenomena. The Griliches (1990) recognized that patent data are susceptible to biases and incongruities, including, but not limited to, variations in patent systems across nations, differences in the inclination to patent among different industries, and the impact of legal and institutional factors on patenting practices.

The value of patents in studying innovation was later addressed even by Hall et al. (2001). The authors published a scientific paper focused on the National Bureau of Economic Research (NBER) patent citations data (Hall et al. 2001). The NBER is a private, non-profit organization in the United States that conducts and spreads economic research for academics, public policymakers, and companies. Hall and her co-authors, stressed the indispensable resource for patent analysis research of the NBER patent citations data file. The authors emphasized the role of patents in providing legal protection to inventors and encouraging information disclosure, which facilitates the systematic study of innovation. Hall et al. (2001) presented the various methodological tools and techniques used to construct the data file, including patent citation matching and cleaning procedures, and discussed the data's limitations and potential biases. Hall et al. (2001) also highlights several empirical applications of the NBER patent citations data file, demonstrating its utility in addressing a variety of research questions pertaining to innovation, technology diffusion, and the influence of public research.

Additionally, the interest in patents as an attractive source for investigating innovation is due to the easy availability, quality and potential insights provided by patent data. In this sense there are a variety of reports and studies available that analyze these aspects.

The World Intellectual Property Organization's annual report *"World Intellectual Property Indicators 2022"* (WIPO, 2022) provides, for instance, an exhaustive overview of global patent trends, highlighting the vast amount of information available on inventions, inventors, and the intellectual property landscape. The report underscores the increasing interest in studying innovation through patent data. According to this publication, the growing curiosity in analyzing patent data can be attributed to different factors:

1. *The ability to track technological progress*
   Patent data is an abundant source of information that can assist researchers, policymakers, and businesses in understanding the evolution of technology and innovation across multiple domains.
2. *The potential for cross-disciplinary analysis*
   Patent data can be used to examine the relationships between innovation, economic growth, and social development, making it a valuable resource for economists, managers, and sociologists.
3. *The opportunity for international comparisons*

The standardization of patent data permits cross-country comparisons and reveals the global distribution of innovation activities and technology transfer.

4. *The capacity to inform policy*

   Patent data may assist in the formulation and evaluation of innovation policies and strategies at the national, regional, and international levels, thereby assisting governments and organizations in fostering innovation-driven growth.

Furthermore, the *"OECD Patent Statistics Manual"* (OECD, 2009), from the Organisation for Economic Co-Operation and Development provides guidelines for the gathering, analysis, and interpretation of patent data. The manual underlines the standardized format of patent data, making comparisons and analyses across countries and time periods more straightforward. It also provides an overview of patent-based indicators and methodologies employed in the study of innovation, such as patent counts, citation analysis, and co-invention networks. In addition to delineating best practices for working with patent data, the manual discusses potential obstacles and hazards that researchers may encounter, including biases in patenting behaviour, data quality issues, and the limitations of certain patent-based indicators. The manual also emphasizes the significance of taking into account the legal, economic, and institutional contexts in which patents are granted and enforced, as these variables can impact the interpretation of patent data. The OECD Patent Statistics Manual (OECD, 2009) is an indispensable resource for researchers, policymakers, and practitioners who seek to better comprehend the dynamics of technological change and its impact on economic and social development.

Over the years, diverse methods have been established for analysing patent data and deriving innovation-related insights. Among the most prevalent methods there are: Patent Citation Analysis, Patent Co-classification Analysis and Network Analysis.

<u>Patent Citation Analysis</u>

Patent citation analysis is a widely employed technique that entails examining the references made by patents to previous patents. This method can assist scholars in comprehending the transfer of knowledge between inventions and identifying influential patents and emerging technologies. (OECD, 2009).

Narin and Noma (1985) introduced the concept of measuring the flow of knowledge and the relationship between science and technology using patent citations (Narin and Noma, 1985). The authors demonstrated the utility of patent citation analysis for comprehending the connections between scientific research and technological innovation, laying the groundwork for future studies employing this methodology. Narin and Noma (1985) analysed the patterns of scientific citations in patents and demonstrated how the extent of such citations reflects the integration of scientific knowledge into technological development.

An exhaustive overview of patent citation analysis, discussing numerous methods, applications, and empirical findings is given by Jaffe and Trajtenberg (2022). The authors highlighted the importance of patent citation analysis in the study of knowledge flows, technology diffusion, and the influence of public research on innovation. Jaffe and Trajtenberg (2022) covers a vast array of topics, including the theoretical foundations of patent citation analysis, the construction and validation of patent citation measures, and the econometric techniques used to analyze patent data. In addition, Jaffe and Trajtenberg (2022) examine the limitations and potential biases associated with patent citations, emphasizing the need for cautious interpretation of results.

Patent Co-classification Analysis

Patent co-classification analysis involves examining the distribution of patents across different technology classes. This method can be used to identify technological diversification and specialization patterns, as well as to monitor the evolution of technological fields. Researchers can gain insight into the development of new technological combinations and the convergence of formerly distinct disciplines by analysing the co-occurrence of technology classes within individual patents. (OECD, 2009)

Breschi et al. (2003) utilized patent co-classification analysis to investigate the technological diversification patterns of companies. The authors discovered that firms tend to diversify within related technological fields and emphasized the significance of knowledge-relatedness in determining firms' innovation strategies. In particular, Breschi et al. (2003) analysed a large dataset of patents filed by European and American companies, focusing on the distribution of patents across various technology classifications. The authors built a measure of knowledge-relatedness based on the co-occurrence of technology classes within patents and used it to examine the extent to which firms diversified their technological activities within or across related disciplines. The study revealed that firms were more likely to expand their technological portfolio into areas with a higher degree of knowledge-relatedness, indicating that existing technological capabilities played a substantial role in determining firms' diversification decisions.

Another significant work involving the usage of patent co-classification analysis to examine the development and structure of technological fields was carried out by Fleming and Sorenson (2001). The study demonstrated that technological fields exhibit the characteristics of a complex adaptive system in which the emergence of new fields is driven by the recombination and integration of prior knowledge. Fleming and Sorenson (2001) examined the distribution of patents across various technology classes and the relationships between them using a large dataset of patent data. The study showed that co-classification patterns were consistent with the concept of technologies as building elements that can be recombined in novel ways to produce new innovations. Additionally, Fleming and Sorenson (2001) discovered that technological innovation is characterized by both incremental advances within existing fields and the creation of completely new fields through the integration of previously distinct areas of knowledge. This finding emphasizes the significance of understanding the interplay between existing technologies and the potential for inter-disciplinary collaboration in fostering technological progress.

Network Analysis

Network analysis is a collection of methodologies designed to examine the interactions between entities, which are typically depicted as nodes, and the relationships or associations between them, which are depicted as edges or arcs. By applying network analysis to patent data, researchers have examined the connections between inventors, organizations, and technologies. In patent network analysis, nodes can represent several entities such as inventors, organizations, technology classes and so on, whereas edges can represent a variety of relationships, such as co-invention, co-assignment, citation connections, and shared technology classes. These techniques are useful for obtaining insights into the structure and organization of innovation networks, as well as for understanding the roles of various actors within these networks. By analysing the structure of these networks, researchers are able to identify patterns of collaboration, knowledge transmission, technological and innovative activity distribution among various actors and domains (OECD, 2009).

The effectiveness and the potential of network analysis techniques applied to patent data is showcased by Verspagen (2007). Indeed, by analyzing patent citation networks, Verspagen (2007) investigated the history of fuel cell research using network analysis techniques and identified key insights into the evolution of fuel cell technology and the roles performed by various organizations, inventors, and nations in determining the trajectory of this research field. Specifically, the author analysed patent data from various sources and constructed an exhaustive citation network to depict the intricate relationships between all the different entities involved in the technological development of fuel cells.

Moreover, an extensive overview to network analysis including its application to patent data and the study of innovation networks is available from Newman (2010). The author presented an in-depth examination of various network analysis methods, their theoretical foundations, and their practical applications, demonstrating their potential for examining innovation networks and revealing the structure and dynamics of the innovation process. Newman, in particular, discussed a wide variety of topics, such as graph theory, centrality measures, community detection, and network evolution, as well as the application of these techniques to real-world problems, including the analysis of patent data. Newman (2010) emphasized the significance of understanding network properties and patterns in order to reveal the fundamental mechanisms that drive innovation and the factors that contribute to the success or failure of various technologies.

As widely demonstrated so far, the study of innovation and innovative activities through the use and analysis of patent data has many pros but at the same time as mentioned in a few articles, it also involves some limitations (Griliches, 1990). The following studies highlight the flaws arising from the use of patents in the analysis of technological development.

Levin et al. (1987) analysed the effectiveness of different mechanisms for appropriating the returns from industrial research and development. Levin et al. (1987) argued that patent data may not capture the complete scope of innovative activity due to the fact that not all innovations are patentable due to factors such as the character of the invention or the strategic decisions of the inventing firm. In addition to patent data, the authors emphasized the importance of considering other innovation indicators, such as R&D expenditures, product announcements, and alternative forms of intellectual property protection. Thus, the findings of Levin et al. (1987) highlighted the need for a more comprehensive approach to measuring innovation than patent data alone.

The limitations of employing patents to study innovative activity have also been proved Cohen et al., (2000). The authors investigated the differences in patenting propensity across industries, countries, and companies, identifying potential biases and comparability issues in patent data, and thus, demonstrating the need for careful interpretation of patent-based innovation indicators. In particular, by analyzing survey data from U.S. manufacturing firms, Cohen et al., (2000) identified several factors that influence patenting behaviour, including industry-specific appropriability conditions, the efficacy of alternative mechanisms for protecting intellectual assets, and the strategic goals of individual firms. The study revealed that patenting propensity can vary considerably based on these factors, resulting in potential biases when comparing patent data across contexts.

In conclusion, patent data and patent analysis have emerged as important instruments for analyzing innovation, providing valuable insights into technological change's patterns, dynamics, and determinants. The literature on this topic is vast and varied, and the studies presented in this Master’s thesis represent only a fraction of the significant contributions and debates in the field. Even though its many advantages, researchers must be aware of the limitations of patent data utilization in investigating innovation, such as the incompleteness, the disparities in patent propensity and in the strategic patenting behavior. Despite these obstacles, patent data and patent analysis continue to be an indispensable resource for scholars, policymakers, and industry leaders attempting to comprehend and foster innovation in today's complex and quickly changing world.

## Section 3: Patent Analysis of Green Technologies and Hydrogen-related Technologies

This third section of the literature review, seeks to investigate the increasing volume of research that examines green technology using patent data, with an emphasis on hydrogen-related technologies. With the intensifying urgency to address climate change and achieve sustainable development, there is a growing interest in understanding the dynamics of green innovation, and hydrogen-based technologies have emerged as a promising pathway to ensure and facilitate the energy transition to carbon neutrality and a more sustainable future. Through an analysis of relevant scientific papers and articles, this section intends to provide a comprehensive overview of the main findings and trends in the study and analysis of patent data on green technologies, as well as hydrogen-related technologies such as green hydrogen production, storage, distribution, and end-use applications. By examining the methodologies, results, and implications of these studies, this third part attempts to shed light on the factors driving innovation in these fields and their potential contribution to a more sustainable future.

The first part of this third section will present studies on the progress of sustainable technologies through the use of patent data while the second part will present significant studies on the status and development of hydrogen-related technologies again through patent analysis.

### Green Technologies and Patent Analysis

To begin with, Barbieri et al. (2020) investigated the distinctions between green and non-green technologies in terms of their knowledge sources and their influence on future innovations. The authors used a citation-based methodology to trace the roots of both green and non-green patents, revealing significant patterns. In particular, green technologies exhibited a distinct level of technological specialization, which likely reflects their multidisciplinary complexity. In addition, these technologies drew from a vast array of knowledge sources, indicating the diverse influences that shape green innovation. In contrast, non-green technologies were found to rely on a more uniform set of knowledge sources, indicating that their innovation process is more limited. Moreover, Barbieri et al. (2020) discovered that green patents produce a greater influence on subsequent inventions, highlighting their profound impact and the growing importance of sustainability in the domain of invention and innovation. The outcomes of Barbieri et al. (2020) emphasize the unique role of green technologies in propelling technological progress and the critical importance of supporting green innovation as a catalyst for sustainable and broad-based development.

Aldieri et al. (2019) conducted another interesting study concerning the analysis of patent data in the field of sustainability and green technologies (Aldieri et al., 2019). Specifically, the authors, using patent data, delved into the complex relationship between environmental innovation and sustainable development, concentrating on the crucial role of technological proximity between different industries. Aldieri et al. (2019) found out that industries with greater technological proximity demonstrated a greater propensity for environmental innovation. This finding suggests that technological origins that are similar or shared can facilitate collaborative innovation, leading to more sustainable solutions. Consequently, technological proximity appears to be a driving force behind the development of green technologies. In addition, the Aldieri et al. (2019) highlighted the significance of technological relatedness and knowledge spillovers in promoting the development and widespread adoption of environmentally friendly technologies. This suggests that the flow of knowledge between industries can substantially influence the rate and direction of sustainable innovation.

Green technologies through the use of patent data have also been studied by Nomaler and Verspagen (2019). The authors, using an extensive patent citation network as their data source, presented a comprehensive analysis of the interaction between homophily and path dependence in the field of green technologies. Homophily, the tendency for similar entities to interact or congregate together, and path dependence, the idea that the possibilities of the present are influenced by the events of the past, are both essential concepts for comprehending the dynamics of technology development. The Nomaler and Verspagen (2019) discovered that green technologies exhibited a greater degree of homophily, indicating that they are more likely to reference other green technologies in their patent applications. This finding suggests that green technologies may be evolving along a distinct trajectory compared to non-green technologies. This study simultaneously highlights the significant path dependence observed in the development of green technologies. This indicates that the evolution of these technologies is heavily influenced by previous technological decisions and advancements, further directing their development along a distinct path. According to Nomaler and Verspagen (2019), a combination of robust homophily and path dependence can facilitate the transition to a more sustainable technological paradigm.

Later on, the same authors published an additional significant work on patent analysis applied to the study of sustainable technologies (Nomaler and Verspagen, 2021), where they presented a novel methodology for patent landscaping. The new approach is based on green technological trajectories, with an emphasis on the evolution and development of green technologies over time. Specifically, through the usage of patent data, Nomaler and Verspagen (2021) showed that it is possible to trace the developmental paths of green technologies, providing a clear understanding of their evolution and possible future direction. Nomaler and Verspagen (2021) covered a broad spectrum of technologies and among them also presents the evolution of some hydrogen-related innovations. Their findings revealed distinct trajectories for every different technology, highlighting how each of these elements had a unique progression pattern, and thus, demonstrating the complex and multifaceted character of the development of green technologies.

Differently, Montresor and Quatraro (2020), analysed patent data, by specifically adopting the technology codes of the International Patent Classification, to investigate the nuanced relationship between green technologies and smart specialization strategies. Montresor and Quatraro (2020) investigated the technological relatedness of renewable technologies with the Key Enabling Technologies (KETs), within the context of European patent data, finding that the two are closely related. This correlation suggests a synergistic relationship in which advances in KETs can stimulate the development and enhancement of green technologies, and vice versa. The interdependent development of these two industries provides an essential basis for the implementation of smart specialization strategies. The smart specialization strategy concept emphasizes the need to concentrate resources on important fields where regions have a competitive advantage or potential.

Therefore, the co-evolution of green technologies and KETs can enhance regional competitive advantage in sustainable practices (Montresor and Quatraro, 2020).

Additionally, Chai et al. (2020), studying patents on green and sustainable technologies, provided an exhaustive and relevant analysis of the factors that substantially influence the citation performance of these green patents. Chai et al. (2020) focused on the examination of two crucial factors: inventor collaboration and technological relatedness, as well as their influence on the citation performance of green patents. The collaborative networks established by inventors appears to be a critical success and impact factor for green patents. These partnerships can facilitate the exchange of ideas and knowledge, thereby augmenting the potential and quality of ecological innovations (Chai et al., 2020). Similarly, the author investigated the role of technological relatedness in green patents, that is, the extent to which diverse technologies share common knowledge bases. The Chai et al. (2020) showed that a green patent has a greater chance of being recognized and cited if it closely aligns with existing technological pathways. The findings indicate that both of these factors positively influence the citation performance of green patents and therefore this highlights the importance of fostering collaborative networks and leveraging technological relatedness to enhance the impact and diffusion of green innovations.

Hydrogen-related Technologies and Patent Analysis

An extremely meaningful study that addresses the technological development involving the hydrogen economy, using patents as a tool, was conducted by Sinigaglia et al. (2018), which presented a comprehensive analysis of patents related to the hydrogen economy from 1998 to 2018. Initially, utilizing the “Questel Orbit” platform for data collection, the authors meticulously identified and catalogued patents, focusing on their geographic distribution and their correlation with technological advancements in hydrogen production, storage, and usage. Sinigaglia et al. (2018) utilized keywords as a methodology to filter patents, ranging across various stages of the hydrogen economy, coupled with International Patent Classification (IPC) codes. The data analysis provided an effective approach for understanding the technological landscape and the leading countries and companies in the hydrogen economy. Sinigaglia et al. (2018) unveiled that Japan and the United States have made significant strides in hydrogen technology, leading the pack with the highest number of patent families. Toyota Motor and Honda Motor emerged as the leading companies in terms of patent ownership, indicative of their active involvement in this technology domain. Interestingly, the Sinigaglia et al. (2018) identified an overall increase in patent publications from 2001 to 2006, followed by a notable decline between 2012 and 2017. This decline might suggest a waning interest in certain technologies related to the hydrogen economy during that period. Despite this trend, the authors underscored the importance of hydrogen as a future energy source given its high energy density and potential for zero carbon emissions. Ultimately, Sinigaglia et al. (2018) indicated that patent analysis could serve as a valuable tool in comprehensively mapping the progress and direction of technological advancements.

More recently, Yu et al. (2022) developed a significant analysis that aimed to predict the future of the hydrogen supply chain by identifying the most promising technologies. The authors examined patent databases and research paper databases from Korea, the United States, Europe, China, and Japan, which enabled them to derive the key technologies for future hydrogen supply chains. By analysing the development of storage, transport, and charging technologies, Yu et. Al, (2022) utilized text mining and Generic Topographic Map (GTM) analysis to identify emerging technologies. This approach enabled the authors to bridge the 18-month blind spot in patent analysis, thus capturing the latest advancements in hydrogen technologies. As a matter of fact, patents often have an "18-month blind spot" or "blind period," which is a term that refers to the time from when a patent application is filed until it is published (Yu et. Al, 2022). This period is typically 18 months long, and during this time, the contents of the patent are not publicly available. From the patent analysis, the Yu et. Al, (2022) identified three promising technologies: the compression, cooling, and liquefaction processes for storing hydrogen in vehicle-mounted cylinders; the use of tanks for transporting stored hydrogen fuel; and the charging of transported liquid hydrogen to vehicles in stations. Finally, Yu et. Al, (2022) underlined the potential of three promising technologies: the use of Liquid Organic Hydrogen Carriers (LOHC) for storing hydrogen fuel energy; transportation by train or truck using containers; and the remodelling of existing train stations to supply hydrogen energy fuel to vehicles.

One of the most significant studies regarding the analysis of the status of hydrogen technologies with the help of patent data was conducted by Ampah et al. (2022), who provided an in-depth investigation into the trends and key contributors to hydrogen production technologies, utilizing a combined approach of patent-life cycle and econometric analysis. The emphasis on patent data offers a unique perspective on the technological progression in this field. Through the patent-life cycle analysis, Ampah et al. (2022) determined the technological maturity of different hydrogen production technologies and the analysis revealed that about 60% of patents were filed between 2000 and 2010, predominantly by companies from Japan, the US, and China. Additionally, Ampah et al. (2022) found that fossil-based technologies have a maturity rate of around 66%, indicating limited space for further growth. In contrast, renewable-based technologies, with a maturity rate of nearly 57%, show higher potential for technological advancement. The authors also employed econometric analysis to identify the key drivers of development in the hydrogen production technologies sector and their findings suggest that research and development expenditure, along with low carbon energy consumption, are significant positive contributors to the advancement of hydrogen production technologies. Also, Ampah et al. (2022) found an ascendant trend of renewable-based technologies in recent years, despite the existing challenges like high production costs and the lack of a sustainable clean hydrogen value chain. The study demonstrates that a committed push towards innovative technologies like water electrolysis, biomass gasification, and nuclear thermal pathways can overcome these challenges.

Furthermore, Zhang et al. (2023) conducted a further study of hydrogen technologies through the analysis of patent data that filled a knowledge gap in renewable energy literature by reviewing relevant US patents on hydrogen-based renewable energy systems and energy management strategies. Zhang et al. (2023) focused on hydrogen's industrial applications, given its attributes such as high efficiency and high energy density, which position it as a crucial component in renewable energy systems. To conduct this analysis, the authors performed an extensive search on "Google Patents" and "Free Patents Online" using keywords like hydrogen, management, and power system. They then filtered the gathered data and categorised the patents into three groups: renewable energy hydrogen production technology, the application of hydrogen-containing in renewable energy systems, and energy management of power systems with hydrogen energy storage. Zhang et al. (2023) found that the trend towards hydrogen production through wind power and photovoltaic electrolysis of water will be significant in the future, owing to the continuing decrease in the cost of electricity generation from renewable sources. However, issues such as system efficiency and economic feasibility still need to be addressed. Hydrogen energy storage presents a viable solution to the limitations of traditional energy storage methods, exhibiting multiple application values both on the load side and power supply side (Zhang et al., 2019). However, the authors recognize challenges, including lack of integrated development strategy, low energy conversion efficiency, and monitoring difficulties. They also note that there is no "one-size-fits-all" energy management strategy; the best scheme depends on the specific optimization objectives. Ultimately, Zhang et al. (2023) concluded that the future of energy management systems with hydrogen storage is promising and forecasted that research will shift towards multi-energy complementary energy management systems that incorporate big data and intelligent autonomous optimization design.

To conclude the review of hydrogen publications, it is also noteworthy to cast light on the findings of Chung et al. (2023). The authors employ patent analysis to explore the life cycle, commercial readiness, and technological advancements of Hydrogen Production Technology (HPT). In particular, the analysis utilized patents from the US Patent and Trademark Office (USPTO) database. Chung et al. (2023) found that HPT has progressed from its initial hype stage to a mature phase of development, signifying stabilized expectations and readiness for commercial dissemination. Trends in patent applications have shown a global hydrogen ecosystem, involving not just the US and Japan who led the early stages, but also the EU and emerging countries, fostering greater innovation. Also, Chung et al. (2023) highlighted the shift in technological competitiveness within HPT. Initially, reforming technology dominated, but as technology accumulation grew, the superiority transitioned to electrolysis. The competitive gap between these two technologies has subsequently expanded, and electrolysis is expected to play a vital role in a sustainable hydrogen economy. Furthermore, Chung et al. (2023) pointed out a strong correlation between HPT development trends and fossil fuel prices, implying that fossil fuel costs significantly impact HPT as an alternative technology. Despite the technological maturity and strengthening competitiveness of HPTs, their commercialization largely depends on cost efficiency (Chung et al., 2023). Currently, electrolysis, though technologically and environmentally suitable, lacks economic feasibility due to the high costs of catalysts like platinum, gold, and silver. Therefore, the authors emphasized the need for the development of cost-effective electrocatalysts with improved system efficiency for large-scale hydrogen electrolysis commercialization. Chung et al. (2023) also acknowledged the limitations of their study, above all the reliance on patents from USPTO, which could not represent technology development trends globally. Second, patent data doesn't reflect completely ongoing research trends or potential technologies developed at the laboratory scale. Despite these limitations, the Chung et al. (2023) provided valuable insights into the technological trajectory and commercial readiness of HPT.

In conclusion, this third section literature review provides an extensive exploration of green technologies and the use of patent analysis in uncovering trends, advancements, and challenges in the field of hydrogen technology. These studies collectively highlight the value of patent data in offering unique insights into technological progression, maturity, and readiness for commercialization of hydrogen technologies and renewable energy systems. Furthermore, they underline the promise and potential of hydrogen as a key component of future energy systems, a prospect that is evidently supported by the global shift towards sustainable and renewable energy sources. Despite evident challenges, such as economic feasibility and technology development constraints, the advancements and research trends reviewed here inspire confidence in the potential for significant growth and innovation in the hydrogen technology sector. On the broader perspective, previous literature illustrates the increasingly crucial role of green technologies in shaping the future of our energy landscape and their capacity to address urgent global issues such as climate change and energy sustainability.

# THE HYDROGEN VALUE CHAIN

## Hydrogen Production: Brown, Gray, Blue, Green Hydrogen and Other Colours of Hydrogen

Hydrogen can be obtained in various ways, from different input feedstocks and different sources of energies can be used in the transformation process. The different hydrogen generation methods carry along different environmental impacts, as well as different costs, which also largely depend on the cost of feedstock/resources in a specific geographical area.

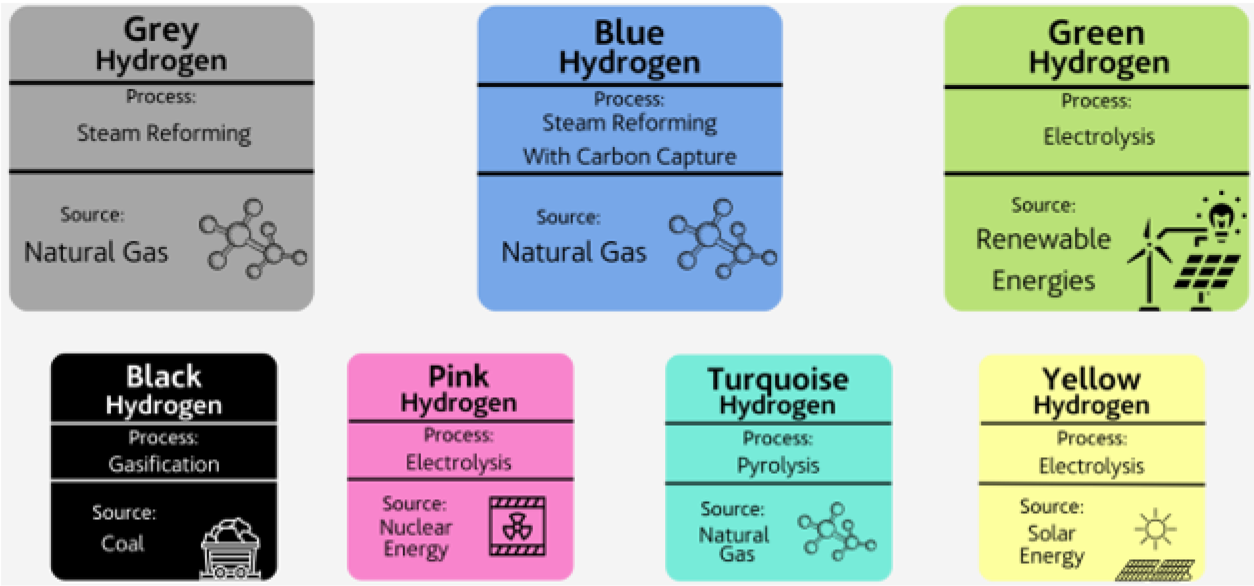



IMAGE 1: THE COLOURS OF HYDROGEN, SOURCE: ACCIONA ENERGIA

**Black or Brown Hydrogen**

The production of brown hydrogen involves utilizing coal in the gasification process. This method, which is on the opposite end of the spectrum from the green hydrogen's electrolysis process, is commonly used in various industries to convert carbon-rich materials into hydrogen and carbon dioxide. However, the emissions released during this process contribute to pollution and make it the second most harmful form of hydrogen for the environment, behind only grey hydrogen. Furthermore, brown hydrogen accounts for over 27% of global hydrogen production (World Economic Forum, 2021).

## Grey Hydrogen

Currently, grey hydrogen production is the most common and cost-effective method of producing hydrogen. It is a source of propellant and accounts for 71% of the world's hydrogen production. However, its production does produce greenhouse gas emissions. Using steam methane reforming (SMR), which extracts hydrogen from natural gas, grey hydrogen is produced from natural gas. However, this process's technology does not retain the resulting carbon emissions, which are instead released into the atmosphere (World Economic Forum, 2021).

## Blue Hydrogen

Steam reforming produces blue and grey hydrogen. However, unlike grey hydrogen, carbon emissions from the process are collected and stored using carbon capture and storage (CCS) technology, reducing atmospheric emissions but not eliminating them. Blue hydrogen is called "low-carbon hydrogen" because it stores greenhouse gasses during creation. Blue hydrogen's restrictions have limited its application. It uses limited resources, fluctuates fossil fuel costs, and doesn't provide energy security. Blue hydrogen needs CO2 storage, transport, and monitoring. Blue hydrogen can boost the hydrogen industry by upgrading current resources with CCS technology to reduce greenhouse gas emissions. CCS efficiency is predicted at 85-95%, therefore some CO2 will be discharged into the environment. Thus, blue hydrogen is a short-term solution for net-zero emissions. Green hydrogen is expected to replace blue hydrogen to accomplish carbon neutrality as renewable energy prices drop (World Economic Forum, 2021).

## Green hydrogen

Green hydrogen is a form of hydrogen that is produced from renewable energy sources, making it the optimal means of transitioning to a completely sustainable energy system. It is produced by electrolysing water with electricity derived from renewable sources such as solar, wind, and hydropower. In the future years, the proportion of hydrogen produced from renewable resources is expected to increase significantly from its current level of less than 1%. Renewable energy cost reductions and technological advances are essential for lowering the price of producing green hydrogen and making it a financially viable and sustainable option (World Economic Forum, 2021).

## Other Colours of Hydrogen

Numerous technological advancements in hydrogen production have emerged in recent years. Electrolysis powered by nuclear energy is used to produce pink hydrogen, which is one of these techniques. Another new technique, known as turquoise hydrogen, is undergoing research to determine its potential for wide-scale application. This variety of hydrogen is produced by the methane pyrolysis process, which employs heat to convert a material into hydrogen and solid carbon. Instead of being released into the atmosphere, the carbon produced is retained in solid form. In the event of success, turquoise hydrogen could be deemed a low-carbon option, provided the carbon is stored in an environmentally responsible manner. Last, yellow hydrogen is a less common term for hydrogen produced through electrolysis powered by solar energy (a subset of green hydrogen) (World Economic Forum, 2021).

## Hydrogen Distribution

Distribution of hydrogen is a crucial aspect of the hydrogen value chain, which entails transporting hydrogen from production sites to end-users. Due to its unique characteristics, such as its low density, high flammability, and need for specialized infrastructure, the distribution of hydrogen presents several difficulties. There are numerous distribution methods for hydrogen, including pipelines, vehicles, and ships.

Hydrogen's low density is one of the most significant obstacles to its distribution. Hydrogen has a lower energy density than petroleum and diesel, meaning it occupies more space per unit of energy. This results in larger storage containers and transport vessels, making hydrogen distribution more difficult and expensive. Another difficulty is the excessive flammability of hydrogen. Hydrogen is more explosive than other fuels due to its low ignition energy and vast flammability limits, necessitating additional safety measures and specialized infrastructure to ensure its safe handling and distribution. In comparison to other fuels, the present infrastructure for hydrogen distribution is relatively underdeveloped. The hydrogen pipeline network is limited, and there are fewer hydrogen fuelling stations than petroleum stations. As the demand for hydrogen rises, more extensive infrastructure development is required to support its increased use. To address these obstacles, ongoing research and development efforts are being made to enhance the safety and efficacy of hydrogen distribution. For instance, efforts are made to develop materials that are resistant to hydrogen embrittlement and to optimize the compression and storage of hydrogen for transportation purposes. In addition, it may be necessary to develop new distribution methods, such as the use of hydrogen carriers such as liquid organic hydrogen carriers (LOHC) or metal hydrides (Hydrogen Europe, 2022).

**Pipelines**

Pipelines, which currently are most widely adopted distribution method, offer several benefits, including reduced costs, greater efficiency, and a more reliable supply. Hydrogen conduits are typically constructed from specialized materials that are resistant deterioration caused by hydrogen absorption. Hydrogen can also be mixed with natural gas and transported through existing natural gas conduits. Hydrogen can be separated from natural gas at the point of use, allowing for a gradual transition to a network of purified hydrogen without requiring significant infrastructure modifications. Hydrogen integration is a cost-effective strategy for incorporating hydrogen into the current natural gas infrastructure (Hydrogen Europe, 2022).

**Trucks**

Another method of hydrogen distribution is truck distribution, which entails compressing and storing hydrogen in high-pressure cylinders on board the trucks. This method is more adaptable than pipeline distribution, allowing hydrogen to be transported to regions lacking pipeline infrastructure. Due to the limited quantity of hydrogen that can be transported at one time, it is additionally more expensive and less efficient (Hydrogen Europe, 2022).

**Ships**

Hydrogen is still transported over vast distances via ship distribution, though it is less common. For transport aboard ships, hydrogen is chilled and compressed into liquid form, and upon arrival at the destination, it is heated and vaporized back into gas form. This distribution method is appropriate for transporting large quantities of hydrogen over long distances, but it is costly and requires specialized apparatus (Hydrogen Europe, 2022).

## Hydrogen Storage & Transformation

The ability to store hydrogen is a key link in the hydrogen supply chain. Hydrogen, as a renewable energy transporter, might be essential in the shift to a low-carbon economy. However, because of its low density and high flammability, storing it can be difficult. To guarantee the secure and effective management of hydrogen, several storage technologies have been devised. There are several variables to consider while deciding on a storage technique, including the data type, the amount of space needed, and the budget. Most hydrogen is stored as compressed gas or as a liquid, although there are intriguing alternatives, such as solid-state storage and chemical storage (IEA, 2019).

**Gaseous Storage**

Compressed hydrogen gas is one of the most often used ways of long-term storage. High-pressure tanks composed of carbon fibre composite materials are commonly used to store compressed hydrogen gas because of their low weight and resistance to pressure. Tank pressures vary from around 350 bar to about 700 bar, depending on the use case. In order to guarantee the tanks' safety, they must be tested often and conform to strict safety regulations.

**Liquid Storage**

Liquid hydrogen is another way that hydrogen may be stored. The volume of hydrogen gas is reduced by a factor of around 700 when it is cooled to extremely low temperatures (-253°C) and turns into a liquid condition. After that, the liquid hydrogen is kept in insulated storage tanks. Since liquid hydrogen can carry more energy per unit volume than compressed hydrogen gas, it is a promising fuel for rockets and spacecraft that travel long distances. However, due to its low boiling point, liquid hydrogen necessitates special care when being handled and stored, which can be both time-consuming and costly.

**Solid Storage**

Solid-state materials, such as metal hydrides, can also be used to store hydrogen since they can absorb and release hydrogen through reversible chemical processes. Hydrogen may be stored in metal hydrides at low to high temperatures and pressures without harming the material. They are versatile, with several possible uses including fuel cells and portable power sources due to their large hydrogen storage capacity. However, their widespread application has been hampered by metal hydrides' slow hydrogen release rate and low capacity.

**Chemical Storage**

Chemical forms of hydrogen storage, such ammonia and methanol, are also on the rise as a means of energy conservation. Hydrogen is stored in these chemical compounds and then released as hydrogen gas as needed. Due to its high hydrogen density and its manageability, ammonia shows promise as a hydrogen transporter. The energy density of methanol is lower than that of ammonia, but it can also store hydrogen effectively. However, the value chain for hydrogen could become more expensive and complicated if chemical storage methods are used.

**Geological Storage**

Finally, geological structures like depleted oil and gas reserves or salt caverns can be used to store hydrogen. Hydrogen storage in underground geological formations (UGF) describes this technique. UGF is an attractive choice for hydrogen storage on a wide scale because of its high storage capacity and long-term storage stability. The danger of leakage and environmental damage must be carefully analysed and managed, and only certain sites should be used for UGF storage.

## Hydrogen End-Uses

Hydrogen's potential in global decarbonization is strictly connected to the versatility of this resource. As a matter of fact, hydrogen can serve as a fuel for transport applications (road, rail, aviation and maritime) and power/electricity generation, as a source of heat in hard to abate industry sectors (steel, cement, paper and pulp, food and aluminium) and in residential and commercial buildings, and last as a feedstock for chemicals production (fertilizers, fuel refining, plastics) and products such as steel, glass, food and metallurgy. Currently, Hydrogen is primarily used as a raw material in the chemical and refining industries, with over 90% of hydrogen in Europe being used in the production of ammonia, methanol, and refining. However, it is expected that hydrogen will play a significant role in reducing carbon emissions in various sectors in the future, including the power system, industries, transportation, and buildings (IEA, 2021).

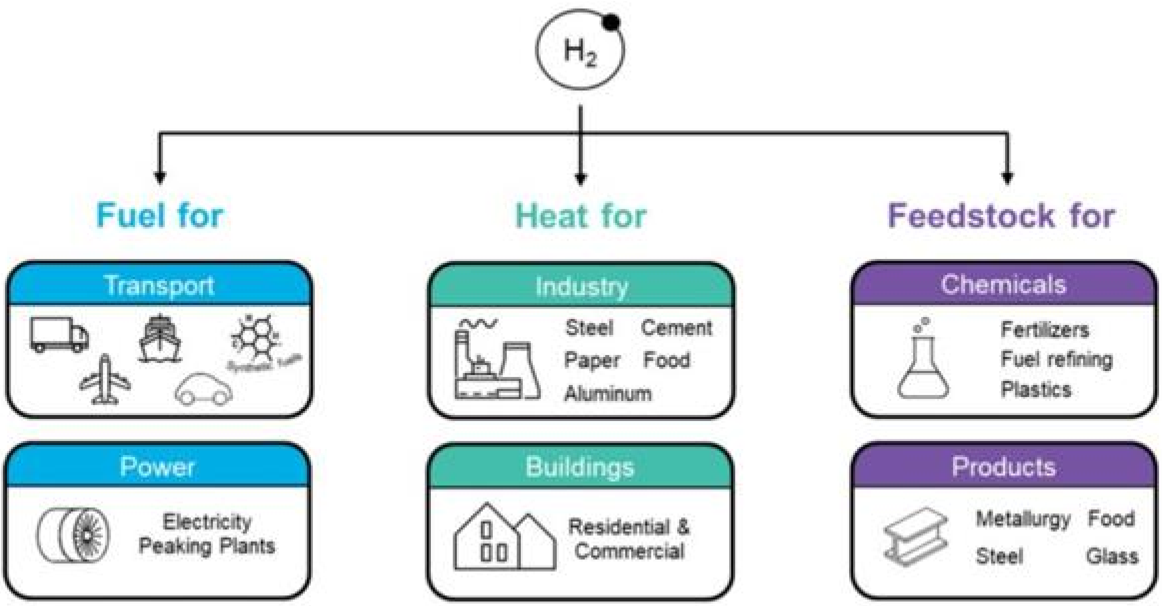



IMAGE 2: THE MANY USES OF HYDROGEN, SOURCE: BLOOMBERGNEF

## Hydrogen as a Fuel

### Road Transport

Hydrogen can be utilized as a vehicle fuel through the process of reverse electrolysis, in which hydrogen reacts with oxygen to produce electricity. Hydrogen is stored in onboard containers, while oxygen is extracted from the ambient air. This reaction generates only electricity, heat, and water vapor, rendering hydrogen-powered vehicles emission-free. Hydrogen can also be consumed in an internal combustion engine to generate power, albeit less frequently. Hydrogen-powered vehicles generate their own electricity via the fuel cell and do not require an external power source. The electricity produced by the fuel cell can either directly power the vehicle's electric motor or charge a smaller, lighter battery used for intermediate storage. Unlike the larger batteries used in pure electric vehicles, this traction battery is continuously supplied by the fuel cell. Like other electric vehicles, hydrogen cars can recuperate energy during deceleration by converting the vehicle's kinetic energy into electricity and storing it in the buffer battery (IEA, 2019).

### Rail Transport

Hydrogen has demonstrated its viability as a rail transportation fuel. Since the early 2000s, the use of hydrogen and fuel cell technology in rail transportation has been demonstrated, with applications ranging from mine locomotives to trams. Multiple nations, including Germany, Austria, the United Kingdom, and the Netherlands, have developed and tested hydrogen fuel cell passenger trains over the past few years. Hydrogen fuel cell trains offer a clean, efficient, and dependable alternative to traditional diesel-powered trains and can contribute to the decarbonization of the rail sector in regions where electrification is difficult or expensive. For long-distance and freight applications, where battery-powered trains may not be practicable, fuel cell trains are particularly advantageous. In the rail transportation industry, hydrogen fuel cell technology has the potential to substantially reduce greenhouse gas emissions and enhance sustainability. As more nations invest in this technology, it is probable that hydrogen fuel cell trains will become more prevalent worldwide (IEA 2019).

### Maritime Transport

Hydrogen is emerging as a viable fuel source for the maritime transportation sector due to its potential to provide a pure and sustainable alternative to conventional fossil fuels. Due to its low volumetric density, the direct use of hydrogen is presently restricted to short- and medium-range vessels. However, the use of hydrogen-based fuels, such as green ammonia, in larger oceangoing vessels is being investigated. Passenger ships, ferries, roll-on/roll-off ships, and tugboats are also adopting fuel cell technology at an increasing rate. These fuel cells convert hydrogen into electricity that can be used to power the ship's propulsion system, resulting in zero emissions of greenhouse gases and other contaminants. As the maritime industry continues to seek out healthier and more efficient sources of fuel, the development and use of hydrogen-based fuels and fuel cells are likely to play an important role in the propulsion of vessels of all sizes (IEA 2019).

**Aviation Transport**

To decarbonize the aviation industry, hydrogen is being considered as a potential solution. Potentially, fuel cell technology and hydrogen combustion could be used to power commercial flights and produce eco-friendly aircraft fuels. Even though hydrogen combustion could be utilized for extended flights, additional equipment would be required to reduce NOx emissions. Sustainable "drop-in" aviation fuels, such as hydrogen-based fuels and biofuels, will be required to decarbonize long-distance flights. However, additional measures may be required to address climate warming effects unrelated to CO2. The growing interest in utilizing hydrogen for aviation purposes underscores the need for sustained research and development in this area, as well as investment in the infrastructure required to support the adoption of hydrogen-based aviation technologies (IEA 2019).

**Power Generation**

Hydrogen's use in power generation is minimal at present, accounting for less than 0.2% of supply. Hydrogen can be utilized as a propellant in reciprocating and gas turbine engines. Currently, reciprocating gas engines can manage gases with a hydrogen content of up to 70% (on a volumetric basis), and several manufacturers have developed 100% hydrogen-powered engines that will soon be commercially available. Gas turbines can also operate on hydrogen-rich gases, and manufacturers are confident that by 2030, standard gas turbines will be able to operate on unadulterated hydrogen. Hydrogen can be converted into electricity and heat by fuel cells, which produce water but no direct emissions. Fuel cell systems can attain high electrical efficiencies (over 60 percent) and maintain high efficiencies even at partial load, making them suitable for flexible operations such as load balancing (IEA 2019).

## Hydrogen as a Source of Heat

### Iron & Steel Industry

The steelmaking industry is beginning to test the use of hydrogen to reduce emissions during the steelmaking process. Coal is commonly used for both high temperatures and chemical reactions in the steelmaking process, which necessitates a great deal of heat. In this process, hydrogen can be substituted for both the required heat and the chemical reactions. As steel is one of the fundamental building elements of modern structures and industrial processes, the use of pure hydrogen has the potential to reduce emissions significantly (Deloitte, 2023).

### Cement, Paper, Food and Aluminium Industries

Additionally In the cement, paper, food, and aluminium industries, hydrogen can be utilized in a variety of methods to reduce emissions. Hydrogen can be used as a fuel source in the cement industry to substitute coal, which is the main source of energy for cement furnaces. Hydrogen peroxide is used as a bleaching agent in the paper industry, and it can be produced using hydrogen as a feedstock, thereby reducing the use of chlorine-based compounds that emit hazardous contaminants. Hydrogen produced from renewable energy sources can be used for hydrogenation in the food industry to reduce emissions associated with the production of hydrogen. Hydrogen can be employed as a reducing agent in the production of aluminium from bauxite ore, thereby minimizing the quantity of energy required and emissions (IRENA, 2021).

### Heating of Residential & Commercial Buildings

The use of hydrogen as a thermal source for residential and commercial buildings is a promising decarbonization strategy for the building sector. However, it confronts obstacles due to the high efficacy of electricity-based solutions and the energy losses associated with hydrogen conversion and transportation. In addition, the expense and complication of assuring safe operations and converting gas infrastructure make it difficult to decarbonize this sector. Despite these obstacles, the localized use of hydrogen in existing building energy systems could support decarbonization in specific contexts with existing gas infrastructure. Incorporating hydrogen with other heat production technologies can increase the flexibility of the electricity grid, especially in extremely frigid regions where other storage options may not be adequate. Hydrogen reactors, fuel cells, hybrid heat pumps, and gas-driven heat pumps are the four major categories of technologies that can operate on hydrogen at the building level. Each technology has advantages and disadvantages, and the optimal solution depends on the building's particular context and needs (IRENA, 2021).

## Hydrogen As a Feedstock

### Oil Refining & Synfuel Production (Chemical Sector)

The oil refining industry, which converts crude oil into various end-use products such as transportation fuels and petrochemical feedstocks, is currently the largest consumer of hydrogen, consuming approximately 40 million metric tons of hydrogen per year (IEA, 2019). Hydrogen is primarily used for hydrotreating and hydrocracking in refineries. Hydrotreatment is used to remove impurities from crude oil, specifically sulphur, whereas hydrocracking is a refining procedure that uses hydrogen to transform heavy residual hydrocarbons into more valuable oil products. With the rising demand for light and intermediate distillates (including the upgrading of oil sands and the hydrotreatment of biofuels) and the falling demand for heavy residual oil, hydrocracking is gaining in significance. Another potential remedy is the production of low-carbon footprint synthetic hydrocarbon fuels (synfuels). These fuels are referred to as "drop-in" fuels because they can be used to replace current oil-derived fuels and the same distribution networks and end-use equipment can be utilized without modification (IRENA, 2021).

### Ammonia Production (Chemical Sector)

Ammonia (NH3) is presently the second greatest consumer of hydrogen with approximately 31 million metric tons per year (IEA, 2019). Hydrogen and nitrogen are combined through the Haber-Bosch process to produce ammonia. The remainder is used in industrial applications such as explosives, synthetic fibres, and other specialty materials. With an increasing drive toward decarbonization across industries, novel ammonia applications are emerging. Ammonia, which is well-known in the freight shipping industry as a sustainable shipping propellant, can also serve as a transport vector for hydrogen. Currently, ammonia is the preferable method for long-distance hydrogen transport. This is because the cost of energy storage is less for ammonia than for hydrogen or liquefied petroleum gas, and because ammonia can store more hydrogen by volume than either hydrogen pipelines or liquid hydrogen. In addition, ammonia is already utilized globally as a fertilizer, so its transportation and storage infrastructure are already in place. Due to its extensive use, there are already established regulations governing its production, transportation, and application (Deloitte, 2021)

### Methanol Production (Chemical Sector)

Methanol (CH3OH) is presently the third greatest consumer of hydrogen with approximately 12 million metric tons per year (IEA, 2019). Like ammonia, methanol has multiple uses as a chemical, basic material, hydrogen transport, and electronic fuel. Utilizing CO2 and hydrogen as inputs to produce methanol significantly reduces the CO2 emissions that are typically generated during the manufacturing process. Additionally, e-methanol can help reduce emissions during its use. By procuring the CO2 used as a feedstock through direct air capture, e-methanol used in internal combustion engines can be carbon neutral. Furthermore, methanol has numerous industrial applications, including as a solvent, antifreeze, and in construction materials. Methanol is also used in the production of other industrial chemicals and in the methanol-to-gasoline process, which converts natural gas and coal into gasoline and has attracted interest in regions with an abundance of coal or gas but limited or no domestic oil production. This is one of the uses of methanol as a fuel, accounting for roughly a third of its global consumption, whether in its unadulterated form or after further conversion. In addition, the development of methanol-to-olefins and methanol-to-aromatics technology has created an indirect path from methanol to high-value chemicals (HVCs) and, by extension, plastics (Deloitte, 2021).

**Products**

Hydrogen is a versatile element that can be used as a raw material in a variety of industries, including the metallurgical, food, steel, and glass industries. Hydrogen is used as a reducing agent in the metallurgy industry to extract metals from their ores, such as iron from iron oxide. Hydrogenation is a process in the food industry that entails adding hydrogen to unsaturated lipids to make them more saturated, thereby extending their shelf life. Hydrogen is used as a reducing agent in the steel industry to remove impurities from iron ore and produce clearer, stronger steel. In the glass industry, hydrogen is used as a fuel source for high-temperature furnaces that shape glass into various shapes by melting it (IEA, 2019).

# INNOVATION AND INTELLECTUAL PROPERTY

## The Importance of Innovation for Organizations and Intellectual Property Rights

The protection of innovation's value is of paramount importance to the inventor. Innovation is the propelling force behind economic expansion, competitiveness, and societal progress. It propels technological advancements, encourages innovation, and produces novel solutions to complex problems (Grossman and Helpman, 1991). By safeguarding the value of innovation, businesses can defend their investments in R&D, encourage additional innovation, and create incentives for continuous improvement (Aghion and Howitt, 1990). Thus, especially in markets characterized by a high degree of technological specialization, innovation becomes an essential element for gaining competitive advantages over competitors. Therefore, intangible value plays a significant strategic function in complex economic systems and consequently, the exploitation of intangible assets becomes essential for the creation of value.

Nonetheless, innovation needs a suitable preservation strategy. In this regard, numerous context variables (such as the nature of the investment, the presence of complementary assets, and eligibility schemes) can impact the management of innovation and can lead to significant effect on its economic outcome (Schilling, 2013). Consequently, various strategies can be used to safeguard the value of innovation:

- Industrial secrecy (practice of keeping proprietary information and trade secrets confidential).
- Intellectual property rights (i.e., patents, trademarks, copyrights, industrial designs)
- Knowledge learning curves (keeping ahead from competitors by delivering continuous innovation and retaining knowledge)
- Exploitation of complementary assets: (e.g., large-scale production capabilities, distribution channels, access to key resources…)
- Lock-in of customers: (e.g., network externalities, industry standards, high switching costs for consumers...).

It is necessary to focus on Intellectual Property Rights and, more specifically, on patents in order to comprehend the evolution of the inquired technology. Intellectual Property (IP) refers to mental creations such as inventions, literary and artistic works, commercial symbols, identities, and images. IP is divided into two main categories:

- Copyright, which encompasses artistic and creative forms, such as literary works, films, music, works of art (such as drawings, paintings, photographs, and sculptures), and architectural designs. Copyright also includes rights pertaining to live performances of artists, recordings of music producers, and radio or television programs. Therefore, this privilege is applied automatically to all unpublished works at the time of their creation.
- Industrial Property consists of patents for inventions, trademarks, industrial designs, and geographical indications; unlike Copyright, this right does not arise automatically, but there is an application and publication process for patents and a registration process for trademarks and designs.

Intellectual Property Rights, like any other property rights, enable the creators or proprietors of patents, trademarks, or patented works to profit from the investment made or the labour performed in creating the work. Particularly, the author has the right to enjoy the protection of his own moral and material interests deriving from any of his works (Art. 27 Universal Declaration of Human Rights). The Paris Convention for the Protection of Industrial Property (1883) and the Bern Convention to Protect Literature and Artistic Works (1886) acknowledged for the first time the significance of intellectual property. The World Intellectual Property Organization (WIPO) administers both agreements.

There are numerous justifications for promoting and defending intellectual property rights (IPR). First, the progress and welfare of humanity are dependent on its capacity for creativity and innovation in the domains of culture and technology. Second, the protection of new inventions encourages the allocation of additional resources for future innovations. Third, the promotion and protection of IPRs stimulates economic growth, generates new employment and industries, and enhances the quality of life. An effective and equitable system of innovation protection can benefit all nations by fostering economic growth and social and cultural prosperity. As a matter of fact, the intellectual property system serves to strike a balance between innovators' and the public's interests by providing a protected environment in which creativity and innovation can flourish for the benefit of all.

## Patents

### Characteristics of a Patent

A patent is a contract between an inventor and a state that grants an exclusive right to an invention, product, or process that offers a novel technical solution to an existing problem. A patent is also a technical-legal document that contains a detailed technical description of the subject of the patent itself as well as its claims of protection. Thus, a patent must include a summary of the prior state of the art (i.e., the technology known at the time of filing, the problem that the invention is intended to address, and a description of how to implement the invention). For an invention to be patented, it must meet the following requirements:

- Must be original/novel. The invention being submitted is not part of the current state of the art and is therefore not yet available to the public or patented. In the EU, everything that is publicly available in any form before the filing date is not considered new (i.e., any disclosure made by the inventor generates prior art). On the other hand, in other patent systems, the inventor has some additional time (i.e., the so-called grace period) to file a patent application after disclosure (e.g., 1 year in the US)
- Must be inventive/non-obvious. It is not possible to obtain a patent for an idea or invention that is considered obvious or falls within the realm of common knowledge. An invention is deemed to be inventive when it is not obvious to a person skilled in the art in view of the state of the art. The person skilled in the art is a skilled practitioner in the relevant technical field who has access to the entire state of the art, is aware of technical knowledge, and capable of routine work.
- Must be useful. The invention being submitted must serve some purpose or have some use that would be desired, such as solving a technical problem and/or having an industrial application.

The assignee, the individual or entity that has been granted the rights to a patent, can decide to whom to grant the right to use it by licensing it based on agreements between the parties or also to transfer his patent rights to third parties. A patent provides assignees with guaranteed protection for their inventions for a limited period of time, typically twenty years, to enable the return of sustained investment in research and development and to consolidate market position and competitiveness. As the inventor has the right to exclude others from making, using, selling, and importing the patented invention for a limited time, any use of a patented product or process without the owner's permission constitutes a patent infringement.

From the perspective of the state, patents are a highly efficient means of achieving several essential objectives. Patents facilitate the dissemination of new technical knowledge, to begin with. By granting exclusive rights to inventors, patents encourage them to disclose their innovations so that others can learn from and build upon them. This dissemination of knowledge is beneficial to society, as it enables further innovation and advancement. Second, patents prevent R&D efforts from being duplicated. Through the patent system, inventors receive a transitory monopoly on their inventions in exchange for their R&D investments. This exclusivity encourages inventors and companies to invest in hazardous and expensive research, knowing that if their innovations are successful, they can recoup their investments and receive the rewards. As a result, duplicate R&D efforts are reduced, optimizing resource allocation, and nurturing more effective innovation. Last, patents play a crucial role in promoting innovation. Patents encourage inventors and companies to invest in the development of ground-breaking technologies and solutions by providing legal protection and exclusive rights. Individuals and organizations are motivated to stretch the limits of knowledge and develop inventive solutions by the potential for financial gain and market advantage. This, in turn, generates a dynamic and competitive marketplace, which fuels innovation and economic expansion.

In complex industries where products and services are susceptible to intricate variations, a single patent may not provide adequate protection for an invention. Businesses may employ patent fencing strategies to maximize economic benefits and prevent imitation. These strategies entail obtaining distinct patents for each variant of the invention, with the intention of establishing a dense network of intellectual property rights. By doing so, businesses can effectively thwart imitation attempts and maximize the value of their innovations. Patent fencing enables businesses to derive greater economic benefits by assuring comprehensive protection and control over diverse aspects and variants of their inventions.

Patenting Procedure

As a patent consists of a contract between an inventor and a state, each country will have a national office where it is possible to apply for a patent. However, in today's global business environment, companies and organizations require protection beyond national borders for their innovations. Patent rights can be extended internationally or continentally by submitting applications with organizations such as the European Patent Office (EPO) or the World Intellectual Property

Organization (WIPO) to address this need. By utilizing international patent organizations and adhering to established patenting procedures, businesses can protect their innovative assets beyond national borders and ensure the proper defence and enforcement of their intellectual property rights.

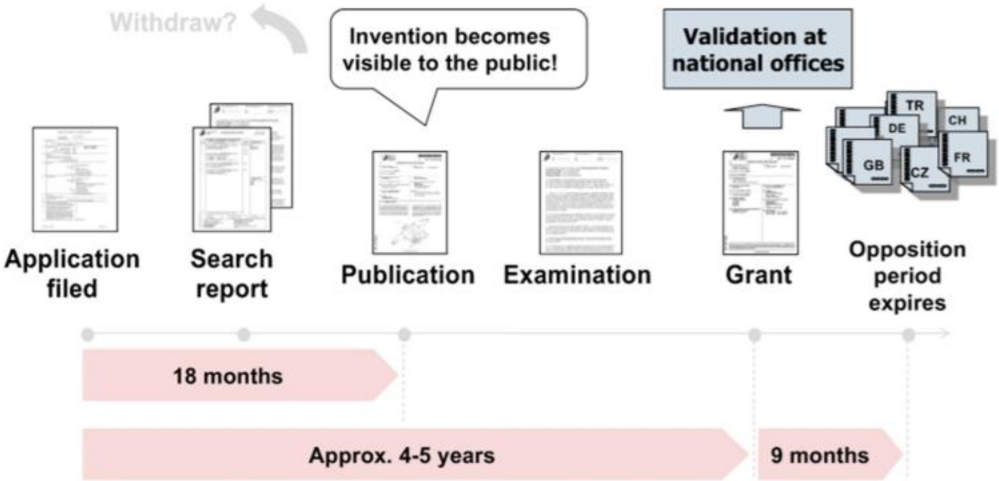
The patenting procedure begins on the Application Date, date on which the patent application is officially submitted to the competent patent office. Aside from signifying the beginning of the patenting procedure, the Application Date carries significant legal importance as it establishes the invention's priority by determining the order in which patent applications are evaluated for the granting of rights. As patent rights are typically granted based on the "first-to-file" principle in most jurisdictions, it is vital for inventors to file their applications promptly to secure an early priority date. Once the Application Date has been established, the inventor or applicant has a period of time, typically twelve months, to request territorial rights extensions. This permits them to pursue patent protection in additional countries or regions beyond the jurisdiction of the initial filing. Requests for extensions may be submitted through a variety of channels, including national patent offices, regional patent organizations (e.g., the European Patent Office), and international patent systems (e.g., the Patent Cooperation Treaty administered by the World Intellectual Property Organization).

During the 18-month period preceding the publication of the patent, research is carried out. The patent office conducts a thorough search for relevant prior art. Examining existing patents, scientific literature, and technical publications, this search seeks to determine the novelty and non-obviousness of the invention. The results are compiled into a search report that contains a list of prior art documents with citations and summaries. This report aids the patent examiner in assessing the patentability of the invention and provides the applicant with an understanding of existing technologies. It assists the applicant in revising or strengthening the patent application's claims. Overall, the search period and search report play a significant role in determining the patentability of the invention and informing the subsequent examination procedure.

Following the search period, the patent application enters the publication phase, where it typically becomes accessible to the public 18 months after its filing. The publication permits the dissemination of invention details and establishes a period of provisional protection. After its publication, the patent application enters the phase of examination. The application is reviewed by a patent examiner who considers its novelty, inventive step, and industrial applicability. The examiner may ask the applicant for additional information, amendments, or clarifications.

Examining the claims, description, and prior art references cited in the search report is part of the examination procedure. A patent is granted if the examiner determines that the application fulfils the criteria for patentability. Nevertheless, if issues are identified, the applicant can resolve them via arguments or amendments. Depending on factors such as the complexity of the invention, the congestion at the patent office, and the jurisdiction's practices, the examination procedure can last anywhere from several months to several years. Once granted, the patent grants the inventor exclusive rights to the invention for a specified period, typically twenty years from the date of filing, allowing them to prohibit others from producing, using, or selling the invention without permission.

Following the grant of a patent, there is a nine-month period for appealing the decision. If there are objections, interested parties can file an appeal by presenting evidence and arguments against the patent. Objections must be pursued through civil proceedings, typically in a court of law, once the appeal period has expired. In civil proceedings, parties may file petitions to contest the validity of a patent. The outcome of these proceedings will determine whether the patent is maintained in its current form, modified, or invalidated. This assures a comprehensive evaluation of the granted patent and permits interested parties to challenge its validity or seek necessary modifications.




Technological Classifications

For the purpose of facilitating research and grouping patents according to universal criteria, technological classifications have been developed, the most prominent of which is the International Patent Classification (IPC), which was established in 1971 by the Strasbourg Agreement. The IPC code appears on the front of the patent document and divides patentable technologies based on their functional properties into eight sections, denoted by letters from A to H, until descending into ever-greater levels of detail; each patent is therefore marked by at least one code indicating the main class of belonging, followed by additional codes if the invention belonged to multiple classes. In addition, depending on the jurisdiction of deposit, additional classifications occur, such as the European Classification (ECLA) in Europe or the United States Patent Classification (USPC) in the United States.

Patenting offices

The procedure of obtaining intellectual property rights for an innovation through patenting differs depending on the country where the patent application is submitted. The majority of the time, registering a patent with a national patent office ensures that it is protected in the nation where it was filed. This indicates that only inside that specific jurisdiction are the patent rights legitimate and enforceable.

A different strategy is provided by filing a patent with international bodies like the European Patent Office (EP) or the World Intellectual Property Organization (WIPO). A single worldwide patent application may be submitted by applicants under the Patent Cooperation Treaty (PCT), which is administered by WIPO. This method makes the initial filing process simpler by enabling applicants to look for patent protection across several member nations. A "worldwide" patent is not, however, granted. Instead, a PCT application goes through an international search and preliminary review before the applicant can decide whether to move on to the national or regional phase in the nations or regions of their choosing.

Through a centralized process, applicants may request patent protection in the nations that are parties to the European Patent Convention (EPC) through the European Patent Office (EP). An approved European patent offers defence in numerous European nations chosen by the applicant.

Inventors primarily obtain protection within a particular country by filing at a national patent office. By contrast, international agencies like WIPO or regional offices like the European Patent Office simplify the process of obtaining patent protection across numerous nations or regions. It is essential for inventors to take into account the desired geographic breadth of protection and assess the best strategy when registering their patents.

# DATA & METODOLOGIES

## Selection and Description of the Research Perimeter

The objective of this master's thesis is to provide a comprehensive overview of the current state of the technologies related to the hydrogen industry, through the analysis of patent data. To conduct a complete analysis of the topic, the entire hydrogen value chain (production, transformation, storage, distribution, and end uses) was selected as a research perimeter. Thus, to identify patent classes, the report "*Hydrogen patents for a clean energy future - A global trend analysis of innovation along hydrogen value chains*" (IEA, 2023), was utilized as a benchmark to construct a taxonomy for the patent landscape. This report provides a thorough perspective of all patented technologies across the entire hydrogen technology value chain. As indicated by the report, within the research perimeter that incorporates the entire value chain, the analysis carried out will focus on the most patent-intensive technologies. Thus, it is worth to precise that not all hydrogen related worldwide patented technologies will be included in this analysis, but only the most significant ones. These identified hydrogen value-chain-related technologies are listed in *Table 1* and described into details in this section. Each technology was highlighted in yellow if it represents a *process innovation*, counter wise it was highlighted in green if it accounts for a *product innovation*.

| Hydrogen Value Chain | Category | Technology |
|---|---|---|
| 1.Hydrogen Production | 1.1 Low-carbon Technologies for Hydrogen Production (from Light Hydrocarbons) | 1.1.1 Steam Methane Reforming: with CCUS (Blue Hydrogen) |
| | | 1.1.2 Steam Methane Reforming: Electrically Heated (Grey Hydrogen) |
| | | 1.1.3 Steam Methane Reforming: Sorption-enhanced (Grey Hydrogen) |
| | | 1.1.4 Steam Methane Reforming: Plasma Reforming (Grey Hydrogen) |
| | | 1.1.5 Methane Pyrolysis (Turquoise Hydrogen) |
| | 1.2 Water Electrolysis (Green Hydrogen Production) | 1.2.1 Alkaline Electrolysers (AEL) |
| | | 1.2.2 Polymer Electrolyte Membrane (PEM) or Proton Exchange Membrane Electrolysers (PEMEL) |
| | | 1.2.3 Solid Oxide Electrolysers (SOEL) |
| | | 1.2.4 Anion Exchange Membrane Electrolysers (AEMEL) |
| | 2.1 Transformation into Hydrogen-Based Fuels | 2.1.1 Transformation of Pure Hydrogen into Low-emission Hydrogen-based Synthetic Fuels: Synthetic Methane & Others |

| | | |
|---|---|---|
| 2. Hydrogen Transformation, Storage & Distribution | (also part of Hydrogen End Use for the Chemical Sector) | 2.1.2 Transformation of Pure Hydrogen to Ammonia & Low-temperature Ammonia Cracking (from Ammonia to Pure Hydrogen) |
| | | 2.1.3 Transformation of Pure Hydrogen to Liquid Organic Hydrogen Carriers |
| | 2.2 Hydrogen Storage & Distribution | 2.2.1 Gaseous Storage (Fuel stations, Terminals or Platforms, by Burying Tanks, by Digging Cavities, by using Natural Cavities, Deep Sea, Offshore) |
| | | 2.2.2 Liquid Storage (Fuel stations, Terminals or Platforms, by Burying Tanks, by Digging Cavities, Deep Sea, Offshore) |
| | | 2.2.3 Solid storage (Hydrides/Adsorption) |
| 3. Hydrogen End-Uses | 3.1 Fuel Cells and ICE | 3.1.1 Proton-exchange membrane fuel cells (PEMFC) |
| | | 3.1.2 Alkaline Fuel Cells (AFC) |
| | | 3.1.3 Phosphoric Acid Fuel Cell (PAFC) |
| | | 3.1.4 Molten Carbonate Fuel Cell (MCFC) |
| | | 3.1.5 Solid Oxide Fuel Cell (SOFC) |
| | | 3.1.6 Direct Methanol Fuel Cell (DMFC) |
| | | 3.1.7 Internal Combustion Engine (ICE) |
| | 3.2 Iron & Steel Manufacturing | 3.2.1 Direct Reduced Iron |
| | | 3.2.2 Blending in Blast Furnaces |
| | | 3.3.3. Smelting Reduction |

TABLE 1: LIST OF HYDROGEN VALUE-CHAIN-RELATED TECHNOLOGIES FOCUS OF THE STUDY

### 1.1.1 Steam Methane Reforming: with CCUS (Blue Hydrogen)

Steam methane reforming with carbon capture and utilization (CCUS) is a process that converts natural gas into hydrogen while capturing and storing the carbon dioxide (CO2) emitted during the conversion. Using steam and methane, SMR with CCUS produces a chemical reaction that produces hydrogen gas, carbon monoxide, and carbon dioxide. The captured CO2 emissions are then redirected to other industrial operations or buried underground. This technology is acquiring popularity as a low-carbon alternative to conventional SMR, also known as grey hydrogen, which lacks carbon capture. SMR with CCUS has the potential to substantially reduce greenhouse gas emissions associated with hydrogen production by capturing and storing CO2 emissions (IEA, 2023).

### 1.1.2 Electrically Heated Steam Methane Reforming (Grey Hydrogen)

An option for reducing the two-fifths of SMR emissions that result from thermal requirements is substituting electricity for natural gas combustion. Innovation in this field has centred on the design of compact reformers that eliminate the need for a large gas furnace with an array of hundreds of reformer tubes that are each longer than 10 meters and contain a catalyst. In contrast to the gas-based heating system, which necessitates flame temperatures above the reaction temperature to account for heat transfer losses, an electrical resistance heating system can use much more precise and efficient heating, varied in real time according to the chemical reaction profile, to achieve greater methane conversion ratios. If such systems were applied to all SMRs utilizing renewable or nuclear energy, it would be possible to reduce global $CO_2$ emissions by 1%.9 Because an eSMR can be operated with some degree of flexibility, it is possible that it could be derated when renewable electricity is in limited supply if incentives are in place to incentivize "system-friendly" operation (IEA, 2023).

### 1.1.3 Sorption-enhanced Steam Methane Reforming (Grey Hydrogen)

In the SMR process, methane is initially reformed with steam to separate its carbon and hydrogen components. The resulting carbon monoxide (CO) is then reacted with additional vapor in a second step to extract additional hydrogen from the water molecules. This two-step procedure is hampered by the need for high temperature and pressure (800–1000°C and 1.53 MPa) as well as the inability to achieve extremely high conversion rates. SE-SMR combines these stages into a single phase with more moderate operating conditions and a potential output containing up to 98% $H_2$ and significantly reduced levels of CO and $CO_2$. Therefore, it requires less natural gas, less energy to purify the $H_2$ product, and inexpensive reactor materials that do not need to withstand such extreme conditions. In addition, $CO_2$ separation is considerably more straightforward with CCUS. In addition, the high-temperature, high-alloy steels required for the reforming reactor can be substituted with less expensive building materials (IEA, 2023).

### 1.1.4 Steam Methane Reforming through Plasma (Grey Hydrogen)

The creation of a plasma of heated ionized gas, in which the reaction occurs, is a more radical method for transitioning to electricity-based reforming heating. There is no need for water inputs; the equipment can be very compact; it can process biomass, heavy hydrocarbons, and natural gas to produce hydrogen; smaller amounts of catalyst can potentially be used, with the plasma's free radicals helping to achieve higher yields; the reaction conditions could potentially be modified so that the hydrogen product is converted to synthetic fuels using the same equipment (IEA, 2023).

### 1.1.5 Methane Pyrolysis (Turquoise Hydrogen)

Methane pyrolysis is the high-temperature decomposition of methane into its constituent elements, predominantly hydrogen and carbon. Typically, the process takes place in a reactor at temperatures above 800°C, with or without a catalyst. The process can also be performed in the presence of a catalyst, which can reduce the required temperature for the reaction to occur and increase the hydrogen yield. The main benefit of methane pyrolysis is that it can produce highly pure hydrogen without additional purification and $CO_2$ is not produced as a by-product. However, methane pyrolysis is still a relatively new and developing technology, and it faces challenges in terms of scalability, energy efficiency, and cost-effectiveness in comparison to other hydrogen production methods such as steam methane reforming and electrolysis (IEA, 2023).

### 1.2.1 Alkaline Electrolysers (AEL)

Alkaline electrolysis (AELs) is the most mature and widely used technology for stationary and/or continuous applications, accounting for approximately 70% of the market for green hydrogen production. It has a low cost and a long operating life, but continuous operation is required, or the apparatus may be damaged. Applications requiring flexible operation and intermittent electrical production employ AELs less frequently. The electrolyte is typically a liquid solution of KOH or NaOH that is circulated between two Ni-alloy electrodes. This method transfers $OH^-$ ions between the cathode and anode at temperatures between 60 and 80 degrees Celsius. A permeable diaphragm is used to prevent the mingling of hydrogen and oxygen and to maintain their separation on the cathodic and anodic sides. Two liquid-vapor separators receive the gas and electrolyte fluxes departing the cathode and anode. The residual electrolyte is recirculated, while the purified gases are sent for external use (IEA, 2023).

### 1.2.2 Polymer Electrolyte Membrane (PEM) or Proton Exchange Membrane Electrolysers (PEMEL)

Polymer Electrolyte Membrane (PEM) or Proton exchange membrane electrolysis (PEMEL) are also commercially available, but their industrialization and experimentation are not as advanced as alkaline electrolysers. PEM employs a polymer membrane electrolyte that facilitates the transfer of protons ($H^+$ ions) in the presence of water, producing hydrogen with a near-zero oxygen concentration. The hydrogen is stored between metal electrodes at temperatures between 50 and 70 degrees Celsius. Due to the reduced thickness and high current density, medium-high pressure operation, and rapid response to electrical power transients, the design of PEM/PEMEL electrolysers allows for the development of compact stacks. The requirement for precious materials such as catalysts (Platinum, Iridium) is a significant disadvantage of this technology, and ongoing research focuses on reducing the quantity of catalyst required and making them entirely recyclable.

PEM/PEMEL electrolysis can produce hydrogen of a higher quality and can operate intermittently, but it is more expensive and has lower production rates than alkaline electrolysis (IEA, 2023).

### 1.2.3 Solid Oxide Electrolysers (SOEL)

Solid oxide electrolysis (SOEL) is a technique that uses elevated temperatures to produce hydrogen from water vapor. The technology is in the pre-commercial stage of development at the moment. SOEL cells employ solid oxide ceramic electrolytes that permit oxygen exchange and have high electrical efficiencies, ranging from 80 to 95% depending on thermal integration. These cells are desirable for use in high-temperature industrial processes such as steel manufacturing and refining. However, SOEL cells lack operational flexibility due to their high operating temperatures and the consequent thermal inertia. They cannot withstand frequent on/off cycles due to their high operating temperatures and thermal inertia. Moreover, while increased production is anticipated to reduce investment costs, the longevity of the cells still needs to be enhanced. Solid oxide electrolysis has the potential to accomplish high efficiency at a low cost, but it still requires increased adaptability and extended component lifetimes (IEA, 2023).

### 1.2.4 Anion Exchange Membrane Electrolysers (AEMEL)

Anion exchange membrane electrolysis (AEMEL) is a relatively novel technology that operates at low temperatures (30 to 60 °C) and has recently made significant advancements. Although these cells are less well-known than other technologies, several companies are already producing at a pre-commercial level. They have several advantages over other technologies, such as the use of an alkaline environment, which reduces the need for costly materials, and solid polymer electrolyte membranes that are capable of transferring OH- ions selectively. This technology reduces the presence of corrosive fluid and has reduced membrane and material costs in comparison to PEMEL (IEA, 2023).

### 2.1.1 Transformation of Pure Hydrogen into Low-emission Hydrogen-based Synthetic Fuels: Synthetic Methane & Others

The manufacture of low-emission synthetic fuels derived from hydrogen, such as synthetic methane and others, is a crucial aspect of the transition to a sustainable energy future. Power-to-gas technology combines hydrogen with carbon dioxide captured from industrial processes or the atmosphere to produce synthetic methane. This method utilizes renewable electricity to produce hydrogen, which is then combined with carbon dioxide to produce methane. The resultant synthetic methane can be used as a low-emission transportation fuel or injected into the natural gas infrastructure for energy generation and heating. Other hydrogen-based synthetic fuels, such as methanol, may also function as carbon-neutral energy carriers and chemical feedstocks (IEA, 2023).

### 2.1.2 Transformation of Pure Hydrogen to Ammonia & Low-temperature Ammonia Cracking (from Ammonia to Pure Hydrogen)

The conversion of pure hydrogen to ammonia and its conversion back to pure hydrogen via low-temperature ammonia cracking, is essential to the development of a hydrogen economy. The Haber-Bosch process combines nitrogen from the air with hydrogen from natural gas or renewable sources, such as electrolysis, to produce ammonia. Ammonia can be used in transportation as a low-emission propellant, as a fertilizer, and as a chemical feedstock. Low-temperature ammonia cracking is the process of separating ammonia into nitrogen and hydrogen, which can be used as a source of purified hydrogen for fuel cells and other applications. This process can be conducted at substantially lower temperatures than traditional steam methane reforming, reducing energy consumption and carbon emissions (IEA, 2023).

### 2.1.3 Transformation of Pure Hydrogen to Liquid Organic Hydrogen Carriers

Important to the development of hydrogen-based energy storage systems is the transmutation of purified hydrogen into liquid organic hydrogen carriers (LOHCs). Through reversible hydrogenation and dehydrogenation reactions, LOHCs are able to absorb and release hydrogen. In the hydrogenation process, hydrogen is added to LOHC, which can be transported as a liquid at ambient temperatures, allowing for the safe and efficient storage of large quantities of hydrogen. The process of dehydrogenation releases hydrogen that can be used as fuel in a fuel cell or combustion engine. LOHCs can be produced from a variety of organic compounds, such as hydrocarbons and alcohols, and can be used to store and convey hydrogen in regions where hydrogen infrastructure is not yet complete. This technology offers a promising solution for the safe and efficient storage and transport of hydrogen, thereby facilitating the incorporation of renewable energy sources into the energy balance (IEA, 2023).

### 2.2.1 Gaseous Hydrogen Storage

High pressure storage of hydrogen is one method to increase its storage density. At 700 bar, the density of hydrogen is 42 kg/m3, allowing a 125-liter storage vessel to store up to 5 kg of hydrogen. However, high-pressure storage containers are expensive to manufacture and require special materials that can withstand the pressure. There are also safety concerns associated with high-pressure storage, as the abrupt discharge of hydrogen can result in powerful explosions. Hydrogen is stored using gaseous storage technologies at gas stations, terminals, and platforms by burying containers, excavating cavities, utilizing natural cavities, deep sea, and offshore (IEA, 2023).

### 2.2.2 Liquid Hydrogen Storage

Liquid hydrogen storage involves retaining liquid hydrogen at cryogenic temperatures to prevent its evaporation into gas. At -252.87°C and 1.013 bar, liquid hydrogen has a higher energy density than its gaseous form, with a density of approximately 71 kg/m3. However, the procedure of transforming hydrogen vapor into liquid state is costly. In addition, cryogenic liquid hydrogen storage containers and facilities must be adequately insulated to prevent evaporation caused by conduction, convection, or radiation. The energy density per unit volume of liquid hydrogen is roughly four times less than that of gasoline and other hydrocarbons. Liquid hydrogen storage can be utilized in a variety of applications, including gas stations, terminals or platforms, subterranean containers, excavated cavities, deep sea, and offshore facilities (IEA, 2023).

### 2.2.3 Solid Hydrogen Storage

Solid hydrogen storage entails the use of substances that can absorb or adsorb hydrogen via chemical reactions. By reacting hydrogen with specific metal alloys, solid metallic hydrides, such as magnesium and alanates, can be produced. Hydrogen is stored through a reversible chemical reaction with the elements of the material. Solid hydrogen storage is advantageous because it eliminates the need for cryogenic temperatures and high-pressure storage. To remain solid, hydrogen must be stored at specific temperatures and pressures (typically below -253 degrees Celsius or at high pressures, depending on the storage material). Typically, only 2% to 3% of the total weight of the storage material consists of hydrogen, which is the most significant disadvantage of this technology. Therefore, additional research is necessary to optimize critical parameters such as the efficacy of the storage material, the temperature and pressure during the hydrogen charge and discharge cycles (IEA, 2023).

### 3.1.1 Proton-exchange membrane fuel cells (PEMFC)

Proton Exchange Membrane Fuel Cells were invented in 1960 and have nowadays become the most prevalent fuel cell technology. In contrast to the direct combustion of hydrogen and oxygen gases to produce thermal energy, a proton exchange membrane fuel cell converts the chemical energy liberated during the electrochemical reaction of hydrogen and oxygen into electrical energy. PEMFCs currently are the most promising fuel cell design for transportation for a number of reasons: it operates at a relatively low temperature range of 100°–180°C; it can quickly vary its output; it is smaller in volume and size than most other types; it has a good supply of membranes (e.g., NAFION or CELTEC, produced in large quantities); and it has a simple, scalable production process. PEMFC membranes must be able to conduct hydrogen ions (protons) for them to function; however, this requires rather expensive platinum catalysts (IEA, 2023).

### 3.1.2 Alkaline Fuel Cells (AFC)

Alkaline fuel cells, also known as alkaline membrane fuel cells (AMFCs) or alkaline anion exchange membrane fuel cells (AAEMFCs), function by transporting alkaline anions – typically hydroxide (OH) – between electrodes. Initially, aqueous potassium hydroxide (KOH) was used as the electrolyte in AFCs. In the 1960s, NASA utilized AFCs for the Apollo and Space Shuttle programs. As it is responsible for the transport of OH– ions, the anion exchange membrane (AEM) has been the focus of numerous recent advancements, as it is an essential component of AFCs. This contrasts with PEM, which is an H+ conductive membrane, and is the primary reason why this type of fuel cell is less popular (IEA, 2023).

### 3.1.3 Phosphoric Acid Fuel Cell (PAFC)

Phosphoric Acid Fuel Cells are a form of fuel cell whose electrolyte is aqueous phosphoric acid. They were the first commercially available fuel cells. Developed in the mid-1960s and field-tested since the 1970s, their stability, performance, and cost have significantly increased. Due to these qualities, the PAFC was an excellent candidate for early stationary applications. Due to the risk of corrosive acid, they are utilized less frequently for transport (IEA, 2023).

### 3.1.4 Molten Carbonate Fuel Cell (MCFC)

Molten Carbonate Fuel Cells operate at temperatures above 600 degrees Celsius and are designed to directly convert natural gas or biogas. Due to the required high temperatures, less rare metals can be used as catalysts, resulting in significant cost savings compared to PAFCs. MCFCs do not require an external reformer to transmute more energy-dense fuels into hydrogen, unlike PAFCs, AFCs, and PEMFCs. Due to the high temperatures at which MCFCs operate, these hydrocarbons are converted to hydrogen within the fuel cell itself via a process known as internal reforming, which reduces costs. Before MCFCs can be used for transportation, additional investigation on the employed materials is necessary due to their still-huge size. However, they have tremendous potential due to their durability. Presently, MCFCs are discussed primarily in terms of stationary use (IEA, 2023).

### 3.1.5 Solid Oxide Fuel Cell (SOFC)

Solid Oxide Fuel Cells are distinguished by their electrolyte material, which is either a solid oxide or a ceramic. SOFC employ the simplest fuel cell design, consisting only of gas and particulates. This type of fuel cell features a high combined heat and power efficiency, long-term stability, fuel versatility, low emissions, and a relatively low cost. The greatest drawback is the high operating temperature (500–1000°C), which necessitates prolonged start-up times and causes mechanical and chemical compatibility problems. In the 1990s, SOFCs were utilized in automobiles, but have since been supplanted by PEMFCs. They are still the subject of intensive research for multiple transport applications, particularly shipping and rail (IEA, 2023).

### 3.1.6 Direct Methanol Fuel Cell (DMFC)

A Direct Methanol Fuel Cell is a variety of fuel cells that converts the chemical energy of methanol directly into electrical energy without requiring a separate reformer device. DMFCs were first developed in 1955, but their potential use in portable electronic devices and as a secondary power source for buildings has garnered significant attention in recent years. In a DMFC, two electrodes, an anode and a cathode, are separated by a polymer membrane. At the anode, methanol undergoes a chemical reaction with water to generate protons, electrons, and carbon dioxide. Protons pass through the membrane to the cathode, whereas electrons travel through an external circuit to generate electricity. Oxygen is supplied at the cathode, where it reacts with protons and electrons to form water. As methanol is a liquid fuel that can be readily stored and transported, DMFCs offer the benefit of a high energy density. However, they have several disadvantages, including low efficacy and a sluggish reaction time. To enhance the efficacy and durability of DMFCs for commercial applications, ongoing research is being conducted (IEA, 2023).

### 3.1.7 Internal Combustion Engines (ICE)

Hydrogen-powered internal combustion engine vehicles are distinct from hydrogen fuel cell vehicles, which use electrochemical hydrogen utilization as opposed to combustion. Instead, the hydrogen internal combustion engine is merely a modified variant of the conventional internal combustion engine propelled by gasoline. The absence of carbon means that no CO2 is produced, which eliminates the principal greenhouse gas emission of a conventional petroleum combustion. Carbon-based pollutants such as carbon monoxide (CO), carbon dioxide (CO2), and hydrocarbons are absent from the exhaust, as pure hydrogen is carbon-free. In an atmosphere containing nitrogen and oxygen, the combustion of hydrogen can produce oxides of nitrogen known as NOx. In this fashion, the combustion process is comparable to that of other high-temperature fuels, such as kerosene, gasoline, diesel, or natural gas. Hydrogen combustion engines are therefore not considered zero emission (IEA, 2023).

### 3.2.1 Direct Reduced Iron

Direct reduced iron (DRI) is a process that entails reducing iron oxide pellets or masses with a reducing gas, typically natural gas, hydrogen, or syngas. In a vertical shaft furnace or rotary kiln, iron ore is heated to between 800 and 1050 degrees Celsius, and reducing gas is introduced to convert iron oxide into metallic iron. The reduced iron product is commonly known as "sponge iron" and has a maximum purity of 98%. Due to its high purity and minimal residual elements, DRI is becoming an increasingly popular feedstock for electric arc furnaces (EAF). In addition, the use of hydrogen as a reducing agent in DRI production has the potential to substantially reduce carbon emissions in comparison to conventional blast furnace processes, making it an attractive option for sustainable steel production (IEA, 2023).

### 3.2.2 Blending in blast furnaces

Blending in blast furnaces is a prevalent practice in the iron and steel industry, in which various types of iron ore and additives are mixed to produce the desired chemical composition for a blast furnace feed. Traditionally, anthracite or coal is burned to produce carbon monoxide, which acts as a reducing agent and converts iron oxide in the ore to metallic iron. Nonetheless, there is a growing interest in using hydrogen as a reducing agent in blast furnaces to reduce carbon emissions. The concept is to heat the iron ore mixture to between 1200 and 1300 degrees Celsius and add hydrogen gas to reduce the iron oxide to metallic iron. Despite the potential benefits of hydrogen, this technology is still in its early phases of development, and its ubiquitous use is likely to be limited by the high cost of producing hydrogen at scale relative to coal (IEA, 2023).

### 3.3.3. Smelting Reduction

Smelting reduction is a process involving the direct reduction of iron oxide with carbon, which yields a heated metal product that is refined in an electric arc furnace or basic oxygen furnace. Various reducing agents, such as hydrogen, methane, and coal, can be utilized to complete the procedure. However, hydrogen's potential to reduce carbon emissions makes it an appealing option. In this procedure, iron ore is elevated to between 1,200 and 1,400 degrees Celsius and hydrogen gas is introduced to convert iron oxide to metallic iron. The use of hydrogen in smelting reduction has the potential to reduce carbon emissions considerably compared to traditional blast furnace processes, making it an essential technology for the transition to more sustainable steel production. However, the technology is still in its infancy, and substantial investment and innovation will be required to make it commercially viable (IEA, 2023).

## Identification of the Dataset

*Derwent Innovation*[1] was utilized to identify and download patents related to each subcategory of the taxonomy. *Derwent Innovation* is a robust database of patents that provides several advantages over other databases. First, it encompasses numerous jurisdictions, including the United States Patent and Trademark Office (USPTO), European Patent Office (EPO), and World Intellectual Property Organization (WIPO). This ensures that a diverse range of patents are included in the inquiry, thereby providing a more comprehensive view of the patent landscape. In addition, *Derwent Innovation* provides sophisticated search capabilities, such as truncation (*), logical operators (AND, OR), and proximity operators (NEAR), which permit precise and targeted searching when developing the research queries. Moreover, *Derwent Innovation* provides a vast array of search filters that enable precise and targeted searches (Title/Abstract/Claim, IPC or CPC Classification, Text Fields, Assignee/Applicant, Citations, Priority Data, etc.). Furthermore, the platform also contains some advanced tools that allow the user to create customized fields and easily export patent data in Excel format. Overall, *Derwent Innovation* is its intuitive and user-friendly interface. Throughout the search procedure, plain instructions are provided to facilitate platform navigation, making it accessible to users with varying degrees of patent searching experience and expertise.

Multiple phases and an iterative strategy were required to identify patents for each leaf of the taxonomy. The procedure began by conducting research to identify the technology's most important keywords and synonyms. Using these to refine results, a query was then modelled that accounted for all these variations. This required trial and error as various search criteria were evaluated and the query was modified as required. Various search tools, including truncation (*), logical operators (AND, OR), and proximity operators (NEAR/ADJ), were used to ensure that the search was exhaustive and accurate. Truncation permitted the incorporation of variations of a specific keyword, whereas logical operators enabled the combination of multiple search criteria to

[1] Derwent Innovation is a global patent search and analysis utility that provides access to more than 50 million patents and patent applications.

refine results. Finally, proximity operators enabled the identification of patents where two keywords appeared near together in the text, resulting in a more precise search. It is worth noting that the 'NEAR' operator allows for a bidirectional search between two parts of text. This means that the command will consider matches even if the order of the parts is inverted. On the other hand, the 'ADJ' operator is more stringent and considers the exact order as written. If the second part of text appears before the first, the result will be filtered and dropped.

As shown in *Table 2* in *Annex*, the main search fields adopted to identify each patented technology were: (“CTB” filter) and eventually the IPC codes (“IC” filter). It is important to precise that adopting Title/Abstract/Claim search filter leads to the identification of a wide range of patents, that at first sight may not seem exactly pertaining to the scope of research. For instance, a technology grouped under the “Fuel Cell” category, may not be a fuel cell itself, but most likely it will refer to components, systems and/or other complementary technology, necessary for the functioning of the “Fuel Cell”. Being the Claims section far more expanded than the Title and the Abstract of a given patent, most references will be found in this section.

The patent research was carried out at a worldwide level, with no country code restrictions, meaning that the patent could have been filed or granted anywhere. To broaden the search for patents related to specific technological fields such as “Fuel cells” and “ICE”, *Derwent Innovation* incorporates the use of International Patent Classification (IPC). IPC is a standard methodology for classifying patents according to their technical subject matter. It provides a hierarchical classification structure that facilitates the efficient organization and retrieval of patent information across countries and patent offices. By employing IPC (“IC” filter), it was possible to improve the research for patents within targeted technological areas, by considering a broader range of relevant patents and access comprehensive information for research and analysis.

*Table 2* illustrates the search queries utilized to investigate patents within a particular technological category of the hydrogen value chain. Information regarding the number of individual documents, applications, and patent families for each technology are also provided. First, the various records contain all of the various papers and entries linked with a patent application or issued patent, such as bibliographic data, claims, descriptions, drawings, legal status, and related communications. Each patent application or awarded patent may contain several unique records indicating various phases, changes, and events within the patent process. The application number, on the other hand, is a unique identifier assigned to a specific patent application at the time of filing. It is used to track and distinguish distinct patent applications by serving as a reference number. Each patent application is normally assigned a single application number. Last, Patent Families group together patents with the same priority application that are related. Priority applications are the first patent applications submitted for an invention. If a single applicant submits multiple patent applications for the same invention in different countries, those applications are regarded as members of the same patent family. The family view provides an overview of the patents that belong to the same family, such as granted patents, pending applications, and related documents.

The search queries identified a substantial number of records, applications, and families associated with hydrogen technologies. There was a total of 150,778 individual records, 119,380 applications, and 79,962 families discovered. It is essential to observe that these numbers may comprise more than one hydrogen technology. This is due to the fact that a single patent typically contains multiple claims, each of which may potentially cover various processes or technologies within the broader hydrogen domain. Consequently, these numbers reflect the overall scope and extent of hydrogen-related patents, which incorporate numerous innovations in the field. These aspects will be better explained in the next chapter of this Master thesis.

## Download, Cleaning and Refinement of the Dataset

After having identified and refined the research queries across all patent classes, a custom category was created in *Derwent Innovation* to enable the export of all patents pertaining to hydrogen technologies. Within this custom field, a list of values corresponding to the various categories of the hydrogen technologies taxonomy were inserted. Each value represents a single leaf in the taxonomy, enabling patents to be easily categorized. The distinct queries for each technology were then imported into a custom field. This populated the custom field with all hydrogen-related patents, which were now separated by value/category. Subsequently, it was possible to download the dataset in Excel in a more streamlined and effective manner, by tagging each patent to its related category. The final download of the dataset was conducted on 12/05/2023. During the download procedure from *Derwent Innovation*, the database was filtered by publication date to adhere to the 30,000-record limit. Various time periods, including until 2004, 2005-2010, 2010-2015, 2015-2020, and beyond 2020, were filtered using the Publication Year of each patent ("PY" filter). This strategy enabled retrieval of patent records within the specified time ranges, ensuring a targeted and manageable data extraction from the platform. After having united the entire dataset, all patents with a Publication Date prior to 1978 were removed. This operation was conducted as older patents lacked some important information such as technological codes (IPC). The threshold year was chosen for a precise reason: the analysis that will be represented in the next chapter of this master thesis involves the use of *PATSTAT*[2] to seek for EPO patents, of which data is available starting from 1978, thus this date is useful for a purpose of comparability. For the same reason, data only until the end of 2022 was considered. Overall, around 10% of the observations were lost in this passage. As described above, the extraction from the *Derwent Innovation* database after the cleaning and refining operations identified a final sample consisting of 101.834 applications.

[2] PATSTAT is a global patent database managed by the European Patent Office (EPO), providing comprehensive patent information for worldwide analysis and research on intellectual property, innovation, and technology trends

## DESCRIPTIVE STATISTICS OF THE DATASET

Once the new dataset was created, a wide range of descriptive analyses of the sample were ran to investigate the nature and characteristics of Hydrogen Value-Chain-Related Technologies. Among others, the following section will focus on analysing the evolution, main trends, occurrence, geography, inventors, assignees, innovators, technological areas (IPC Codes), and share of green technologies ("Y" CPC Section) among hydrogen patents.

### The Evolution of Hydrogen Patents

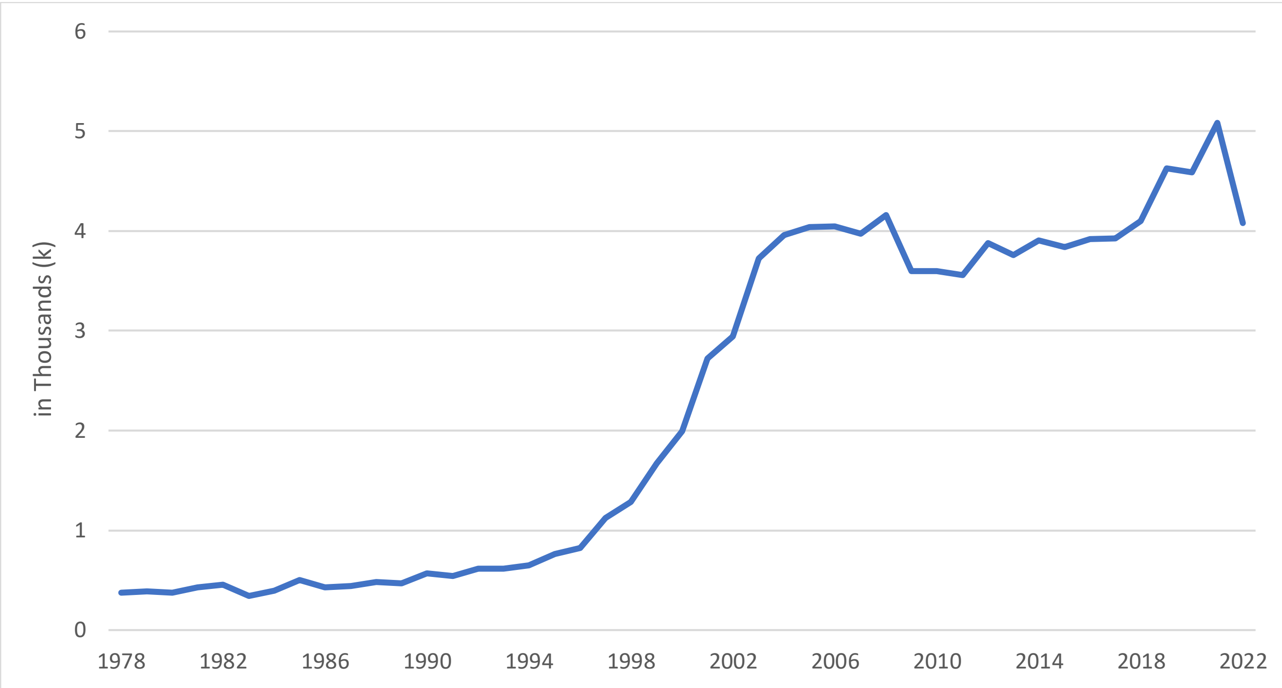



FIGURE 1: THE EVOLUTION OF HYDROGEN PATENTS BY APPLICATION YEAR

*Figure 1* depicts the number of hydrogen technology patents from 1978 to 2022, revealing a remarkable trend that reflects the increasing prominence of Hydrogen technologies in last two decades. Over the entire period, there is a discernible upward trajectory in patent counts, indicating a growing interest and investment in hydrogen-related research and innovation. However, the most remarkable aspect of the graph lies in the years ranging from 1998 to 2006, where the curve demonstrates a steep and exponential rise. Such a sharp rise highlights a significant surge in patent filings and suggests an increased focus on developing and commercializing hydrogen technologies. Following a noticeable decrease in hydrogen patent filings after the 2008 crisis, the graph displays a rather stable trend up to 2018, followed by a new steep rise up to 2021. It is imperative to notice that the apparent decline in 2022 does not reflect the true patenting activity, but it is due to the fact that patents become visible to the public only after their publication date. As explained in the section above, on average it takes 18 months from when the application of a patent is filed to when they are published. Therefore, it is likely that both 2022 and 2021 may suffer from distortion bias, as the actual number of patent applications in these years is expected to be significantly higher in 2022 and slightly higher in 2021. Overall, the data presented is consistent with the findings of Sinigaglia et al. (2018), which identified an overall increase in patent publications from 2001 to 2006, followed by a period of stagnation from 2012 to 2017.

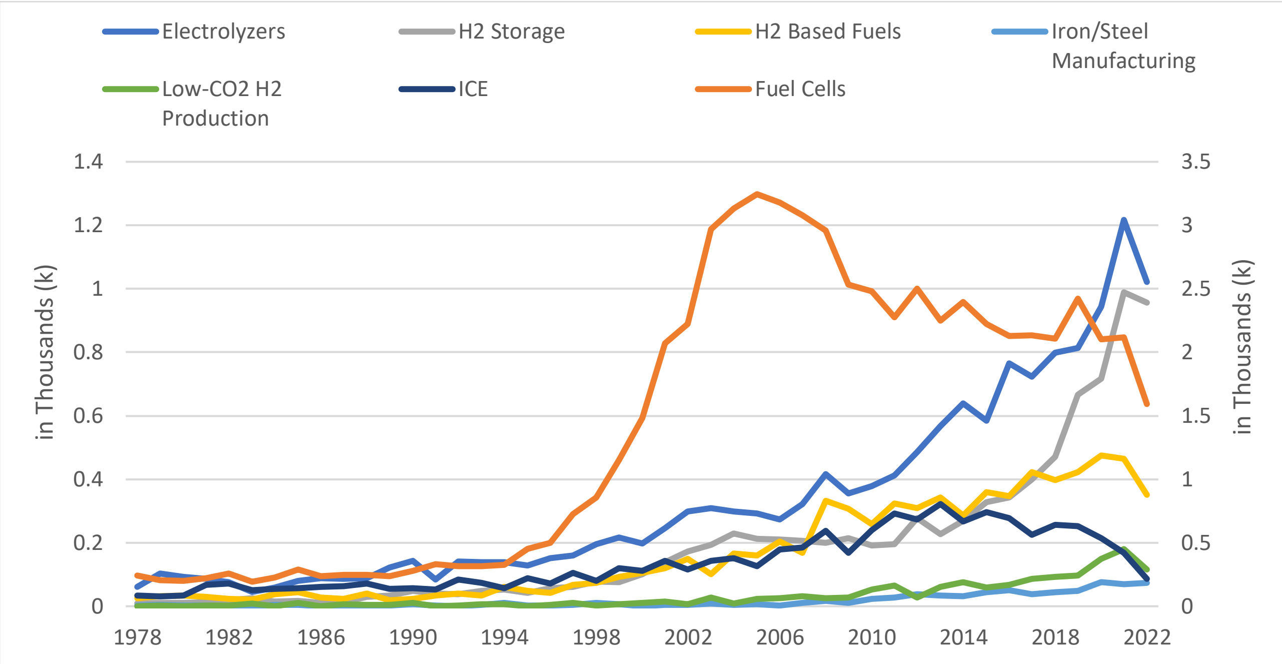



FIGURE 2: THE EVOLUTION OF HYDROGEN PATENTS BY APPLICATION YEAR PER MACRO CATEGORY OF THE HYDROGEN VALUE CHAIN (FUEL CELLS SCALED ON RIGHT AXIS, ALL OTHER CATEGORIES SCALED ON LEFT AXIS)

Above, *Figure 2* displays the number of hydrogen patents divided by macro group of the value chain, revealing important insights into patenting trends. Fuel cells stand out as the most patent-intensive technology, with 63,649 patents filed from 1978 to 2022. The graph indicates a significant rise in fuel cell patenting around 1998, but a rather declining trend after 2008. In contrast, electrolysers, storage technologies, and Hydrogen based fuels experienced a visible increase in patents throughout their entire lifetime. The graph clearly suggests that fuel cells are the most mature technology across the hydrogen value chain and that even though they remain at the forefront of innovation activity, other complementary technologies are rapidly catching up, indicating a wider expansion of interest and investment in their development. This demonstrates the expanding focus on advancing hydrogen technologies across various sections of the value chain.

## Breakdown of Hydrogen Technologies across the Value-Chain

*Figure 3* illustrates the distribution of patented technologies across the entire hydrogen value chain.

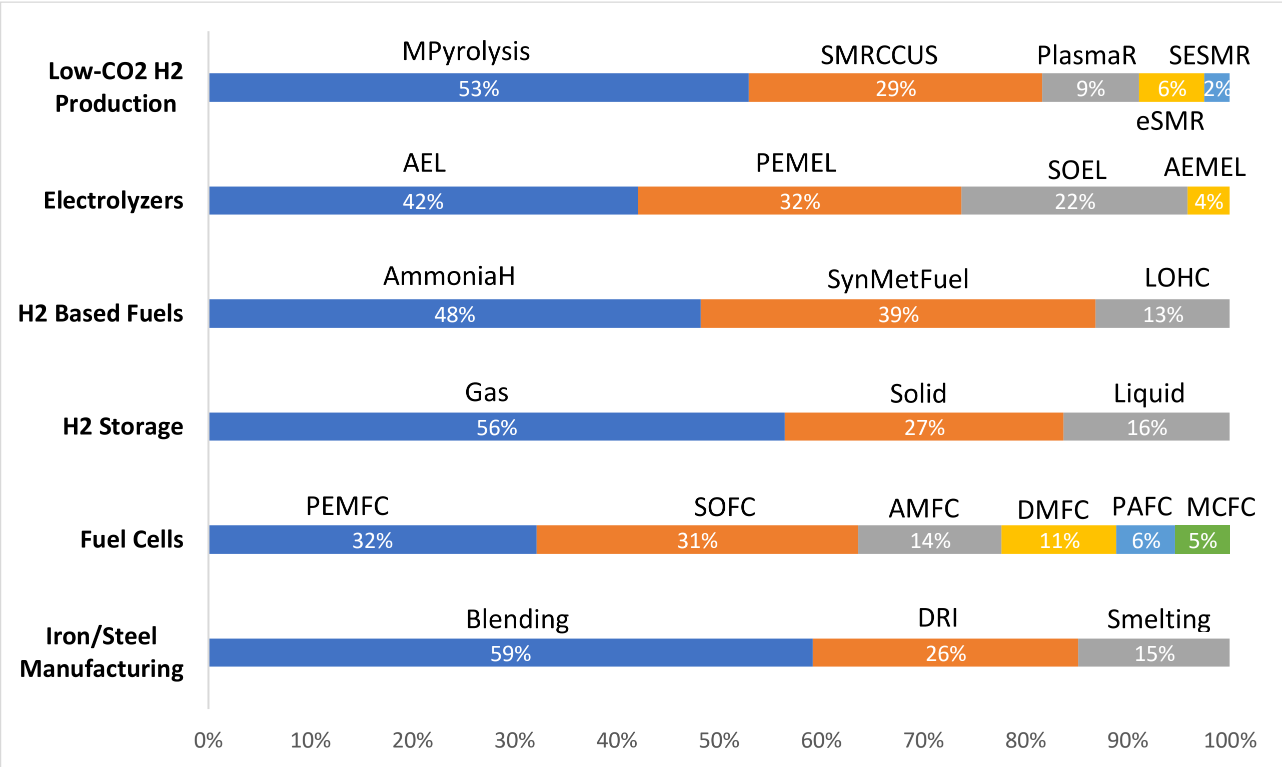



FIGURE 3: BREAKDOWN OF PATENTING ACTIVITY ACROSS THE HYDROGEN VALUE CHAIN

Regarding the patent distribution of emergent hydrogen production technologies with low carbon emissions, Methane Pyrolysis is the most prevalent technology, accounting for 52.9% of all patents. It is followed by Steam Methane Reforming with CCUS (28.7%), Electrically Heated Reforming (6.5%), Sorption-Enhanced Steam Reforming (2.5%), and Plasma Reforming (9.5%). As depicted in the graph, the distribution of emergent low carbon emission hydrogen production technologies is closely aligned with the respective technology readiness levels defined in the report "*Hydrogen Patents for a Clean Energy Future*" (IEA, 2023). The fact that Methane Pyrolysis accounts for 52.9% of all patents validates the report's claim that this technology has reached the pre-commercial stage, indicating its advanced state of development. Likewise, the prevalence of Steam Methane

Reforming with CCUS, which accounts for 28.7% of the patents, corresponds with its classification as a pre-commercial stage in the report. In addition, the classification of Electrically Heated Reforming as a significant prototype is consistent with its 6.5% patent share. In addition, a 9.5% patent share of Plasma Reforming is consistent with the report's description of the technology as being in the conceptual stage. These correlations emphasize the congruence between patent distribution and the technology readiness levels reported by the IEA, confirming the ongoing advancement of low-carbon hydrogen production methods.

Secondly, the distribution of patents depicted in *Figure 3* provides insight into the proportional distribution of electrolysers for the production of green hydrogen. AEL (Alkaline Electrolysis) holds the largest market share with 42.0%, followed by PEMEL (Polymer Electrolyte Membrane Electrolysis) with 31.7%, SOEL (Solid Oxide Electrolysis) with 22.1%, and AEMEL (Anion Exchange Membrane Electrolysis) with 4.2%. Overall, the patent distribution corresponds to the readiness levels reported in the report "*Hydrogen Patents for a Clean Energy Future*" (IEA, 2023), reinforcing the correlation between patent activity and the technological advancement of electrolysers for the production of green hydrogen. Alkaline Electrolysis's patent share dominance is consistent with its characterization as having market penetration. This demonstrates that the technology has acquired traction and is currently being utilized in the industry. Similarly, the substantial patent share of Polymer Electrolyte Membrane Electrolysis correlates with its classification as a market-accepted technology, further validating its maturity and market presence. Solid Oxide Electrolysis, which represents a significant proportion of patents, corresponds to its technology readiness level classification as a pre-commercial demonstration stage technology. This indicates that the technology is undergoing testing and demonstration for real-world applications as it advances toward commercial viability. Lastly, Anion Exchange Membrane Electrolysis's lower patent share is consistent with its classification as a large prototype, indicating that it is still undergoing development and evaluation.

In addition, *Figure 3* depicts the patent shares of various hydrogen-based fuel technologies, including 38.6% for Synthetic Methane and Other Synthetic Liquid Hydrogen-based Fuels, 13.1% for Liquid Organic Hydrogen Carriers (LOHC), and 48.2% for the Transformation of Pure Hydrogen to Ammonia & Low-temperature Ammonia Cracking (from Ammonia to Pure Hydrogen). Notably, the corresponding technology maturity levels from the report *"Hydrogen Patents for a Clean Energy Future"* (IEA, 2023) may contribute to erroneous interpretations for this category. This is due to the inherent difficulty of accurately distinguishing patents involving technologies that convert hydrogen to ammonia from those involving technologies that convert ammonia back to hydrogen. Due to the overlap in patent categories, it is difficult to assign specific readiness levels to each technology. Despite this challenge, the substantial patent share indicates ongoing research and development efforts to utilize ammonia as a hydrogen carrier. Due to its high hydrogen content and well-established infrastructure for storage and transport, ammonia has gained attention as a potential energy carrier. The significant patent share for Synthetic Methane and other liquid hydrogen-based fuels suggests that innovation is also occurring in this field.

Furthermore, *Figure 3* provides a breakdown of hydrogen storage and distribution technologies. Gaseous storage technologies account for 56.4% of the market, whereas liquid storage technologies account for only 16.3%. Hydrides and adsorption account for 27.3% of the total solid storage technologies. Reflecting the present state of the hydrogen industry, the dominance of gaseous storage indicates its broad applicability and well-established infrastructure. The extensive use of gaseous storage methods, such as fuel stations and terminals, demonstrates the existing infrastructure for hydrogen storage and distribution in its gaseous form. Although liquid storage technologies have a reduced market share, they still play a significant role, especially in applications or locations where gaseous storage may not be as practicable or effective. The existence of liquid storage options illustrates the versatility and diversity of hydrogen storage requirements. Hydrides and adsorption, which represent solid storage technologies, occupy a significant portion of the storage distribution of patents. The prospective benefits of these methods include a higher volumetric density and the capacity for long-term storage. Inclusion of solid storage technologies indicates ongoing research and development efforts to discover innovative and alternative hydrogen storage solutions.

Moreover, *Figure 3* depicts the proportion of patents developed for various fuel cell technologies. Proton Exchange Membrane Fuel Cells (PEMFC) possess the dominant share of patents with 32.1% (22,326 patents), followed by Solid Oxide Fuel Cells (SOFC) with 31.5% (21,887 patents). Alkaline Fuel Cells (AFC) account for 14.1% (9,778 patents), while Direct Methanol Fuel Cells (DMFC) account for 11.2% (7,817 patents). At 5.7% (3,974 patents) and 5.4% (3,734 patents) respectively, Phosphoric Acid Fuel Cells (PAFC) and Molten Carbonate Fuel Cells (MCFC) possess lesser shares. The provided patent data aligns with the findings illustrated in the report "Hydrogen fuel cells in transportation" (WIPO*,* 2022). Both studies classify Proton Exchange Membrane Fuel Cells (PEMFC) as the technology with the most intensive patenting activity. According to the report, the considerable patent share of Solid Oxide Fuel Cells (SOFC) in the dataset reflects its position as one of the most patent-intensive technologies. Direct Methanol Fuel Cells (DMFC) and other direct or reforming fuel cells with significant patent shares in the dataset validate the ongoing research and development efforts cited in the report. The lesser patent shares of Phosphoric Acid Fuel Cells (PAFC) and Molten Carbonate Fuel Cells (MCFC) in the dataset correspond to their representation in the report as technologies with relatively fewer patents. This correspondence between the dataset and the report provides assurance that the collected patent data are consistent and relevant.

Last, *Figure 3* depicts the patent allocation for numerous hydrogen-based iron and steel manufacturing technologies. Blending in Blast Furnaces holds the largest patent share at 59.2%, followed by Direct Reduced Iron (DRI) at 26.0% and Smelting Reduction at 14.8%. Examining the corresponding technology preparedness levels from the report *"Hydrogen Patents for a Clean Energy Future"* (IEA, 2023), we can establish a relationship between the proportion of patents and the readiness levels. With the highest patent share, Blending in Blast Furnaces is classified as pre-commercial demonstration, indicating that substantial research and development efforts have been made to advance this technology toward commercialization. The patent share of Smelting Reduction correlates with its classification as an early prototype in the readiness levels. The patent activity indicates ongoing research and innovation in this field, with the goal of optimizing the technology further. Direct Reduced Iron (DRI), despite having a lower patent share than Blending in Blast Furnaces, is a complete prototype according to the preparation levels. This indicates that substantial progress has been made in the development and demonstration of this technology, paving the way for its possible commercial application.

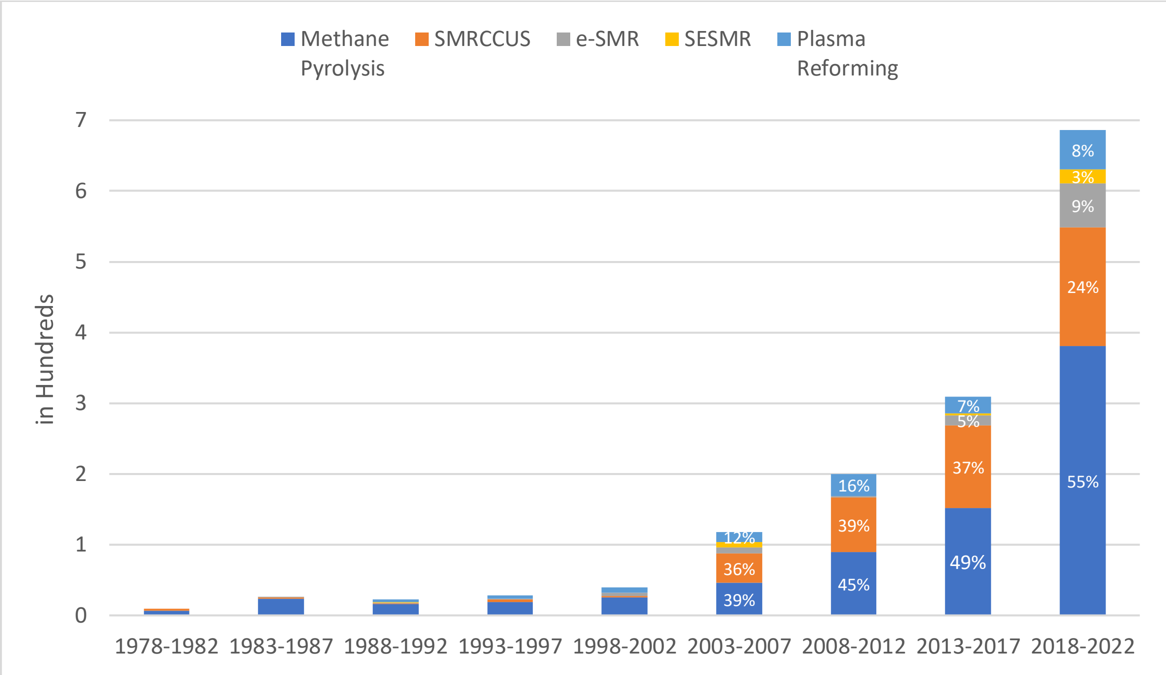



**FIGURE 4: THE EVOLUTION OF PATENTS RELATED TO HYDROGEN PRODUCTION TECHNOLOGIES WITH LOW-CARBON EMISSIONS**

*Figure 4* shows the patent counts for various low carbon emission hydrogen production technologies. Analysing the trends reveals significant growth rates and a growing interest in these technologies. For instance, Methane Pyrolysis has had a large growth in patent numbers, rising from 7 in 1978-1982 to 381 in the most recent period (2018-2022). The dominant position of this technology demonstrates a considerable industrial focus on developing methane pyrolysis as a low carbon emission hydrogen generating method. Additionally, Steam Methane Reforming with Carbon Capture and Storage methods have maintained a relatively strong share in the last 4 periods. Last, it is possible to see that Electrically Heated Steam Methane Reforming, as well as Plasma Reforming, are newer emerging technologies that are gaining increasing attention.

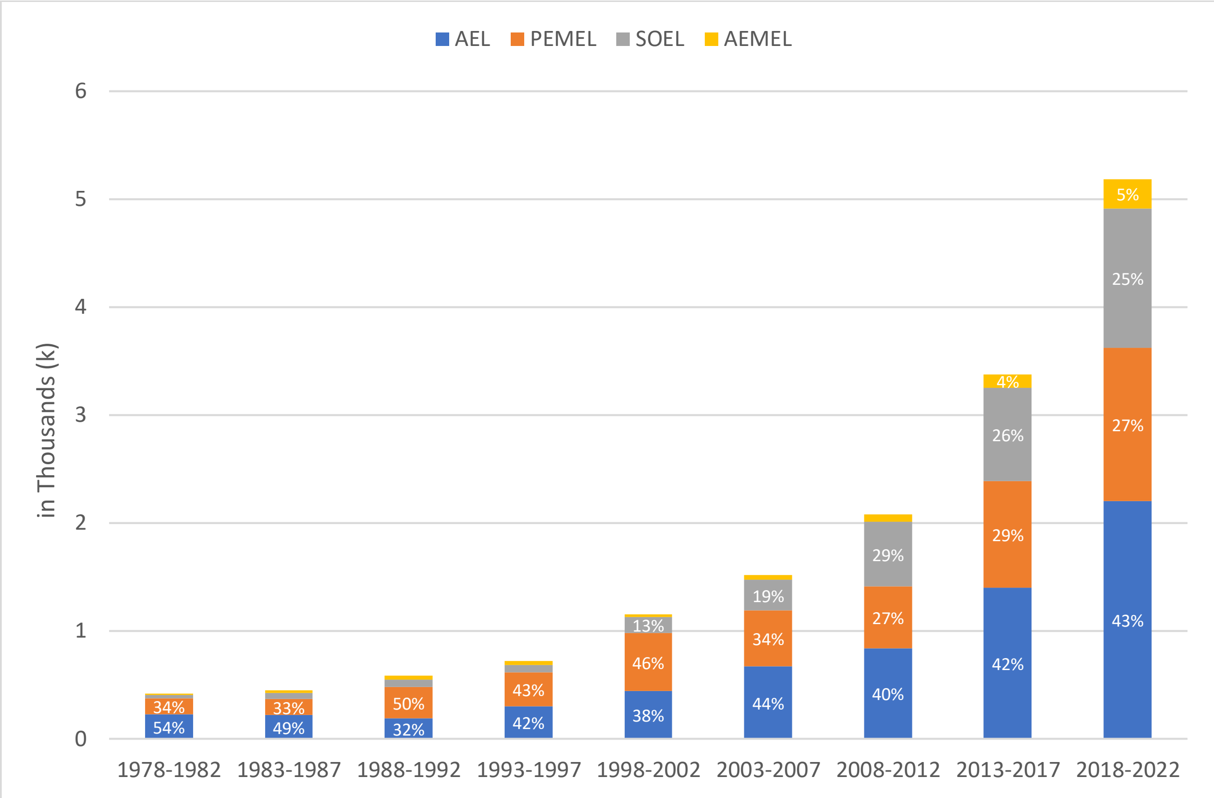



FIGURE 5: THE EVOLUTION OF PATENTS RELATED TO ELECTROLYSERS FOR THE PRODUCTION OF GREEN HYDROGEN

*Figure 5* shows the patent counts for the four types of electrolysers (AEL, PEMEL, SOEL, and AEMEL) through time. Analysing the trends reveals interesting patterns, and compound annual growth rates (CAGR) for each electrolyser type can be estimated. AEL exhibits significant growth from 1978 to 2022, with its patent count increasing from 228 to 2204, resulting in a CAGR of nearly 10.5%. PEMEL has also grown significantly, from 145 to 1.418 patents, representing a CAGR of roughly 9.6%. SOEL has increased moderately, from 34 to 1.292 patents, resulting in a CAGR of around 11.2%. AEMEL, which began with fewer patents, has had amazing growth, increasing from 14 to 270 patents with a strong CAGR of around 15.3%.

These statistics demonstrate a general upward trend in patent applications for all types of electrolysers over the selected time period. AEL and PEMEL are the most prominent technologies, with consistent and sustained growth. The faster CAGR of AEMEL in recent years indicates a growing interest in and possible developments in this specific electrolyser technology. Although SOEL has a reduced growth rate, it nevertheless reflects a consistent degree of innovation.

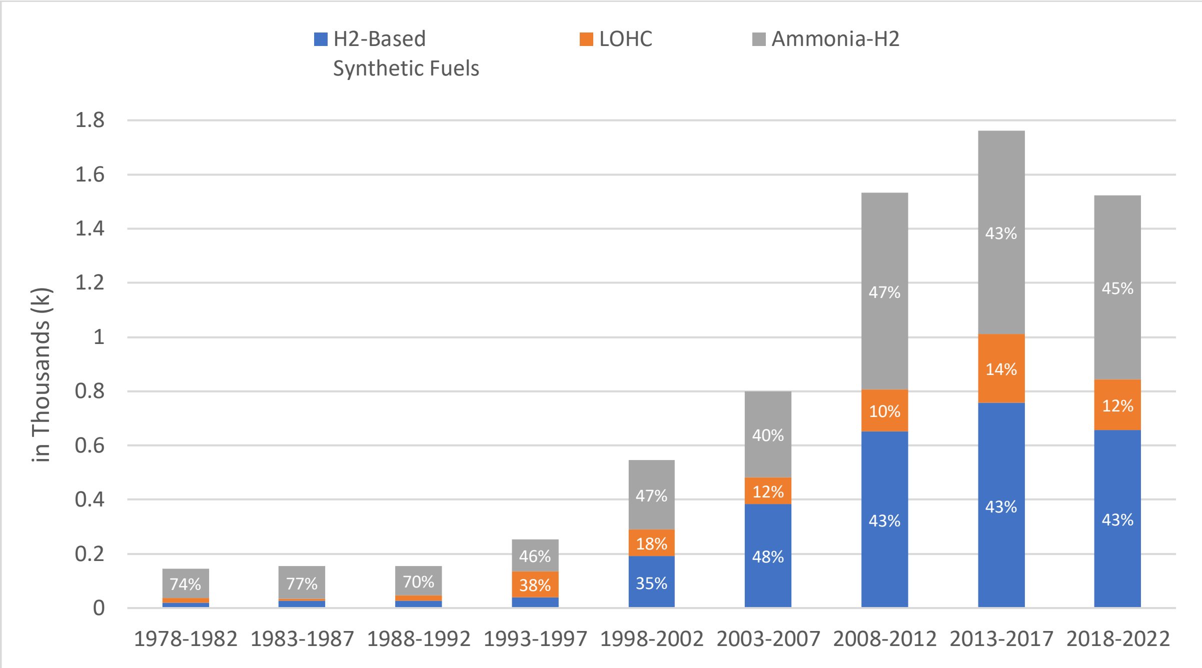



**FIGURE 6: THE EVOLUTION OF PATENTS RELATED TO HYDROGEN-BASED FUEL TECHNOLOGIES**

The temporal variations in patent activity across several hydrogen-based fuel concepts are presented in *Figure 6*. The areas of Transformation of Pure Hydrogen to Ammonia and Low-temperature Ammonia Cracking have maintained a dominant and growing position in hydrogen conversion technologies. The growth of Synthetic Methane is extremely pronounced in the past 25 years, gaining considerable attention since the 1998-2002 period. On the other hand, Other Synthetic Liquid Hydrogen-based Fuels display a much slower growth trend in terms of number of patents pertaining to that specific area.

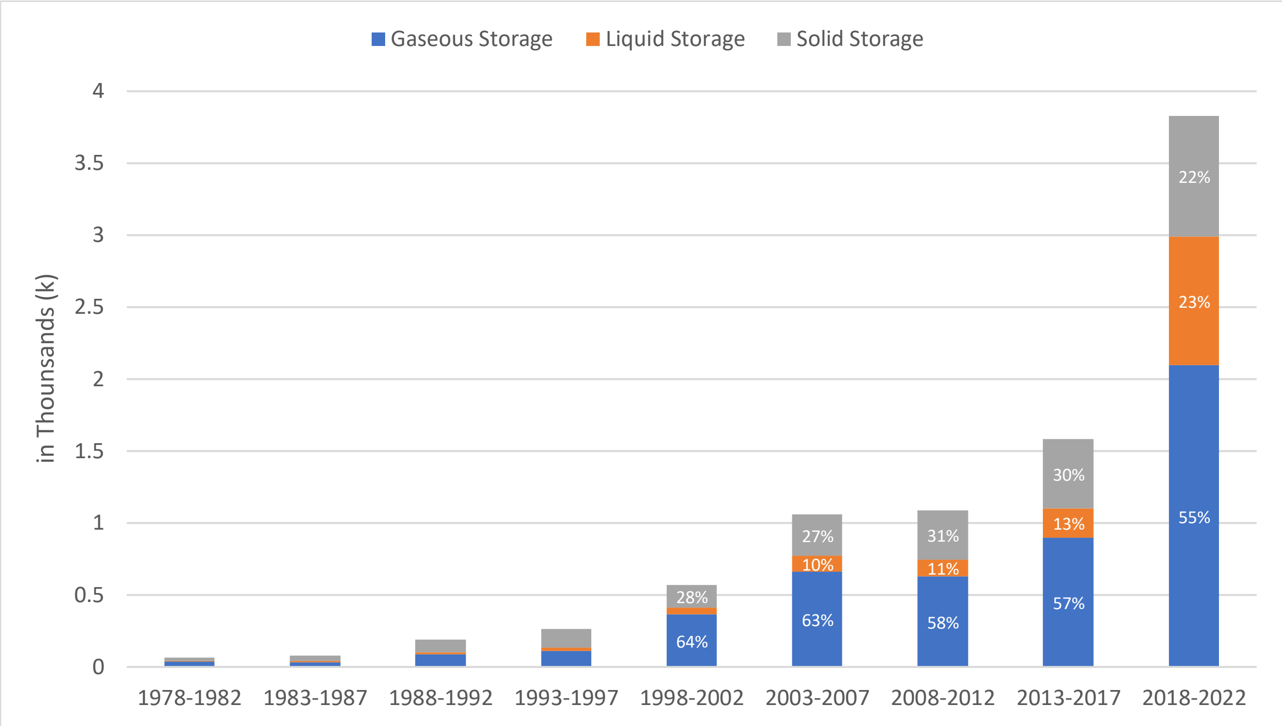



FIGURE 7: THE EVOLUTION OF PATENTS RELATED TO HYDROGEN STORAGE AND DISTRIBUTION TECHNOLOGIES

*Figure 7* represents the patent counts for three distinct categories of hydrogen storage technologies: Gaseous Storage, Liquid Storage, and Solid Storage, spanning multiple time periods. From the graph, it is possible to notice that Gaseous Storage experienced a significant surge in patent counts, soaring from 39 in the initial period (1978-1982) to 2096 in the latest period (2018-2022). This remarkable growth corresponds to an impressive CAGR of approximately 35.87%. Such a substantial increase and the dominant share of this technology suggests a continued industry focus and dedication to developing efficient and effective hydrogen storage solutions utilizing gaseous materials. Likewise, Liquid Storage experienced substantial growth, with patent counts rising from 5 to 894 over the analysed period, indicating a CAGR of approximately 40.04%. The category demonstrated steady progress, particularly in recent years, gaining a significant share of all storage technologies, which, by 2018-2022, accounted for 23% of the total storage patent counts. Solid Storage, another crucial category in the realm of hydrogen storage, witnessed a substantial rise in patent counts from 19 to 837 between 1978-1982 and 2018-2022 (45% CAGR approximately), however its trend highlights how it is losing share in favour of other storage solutions.

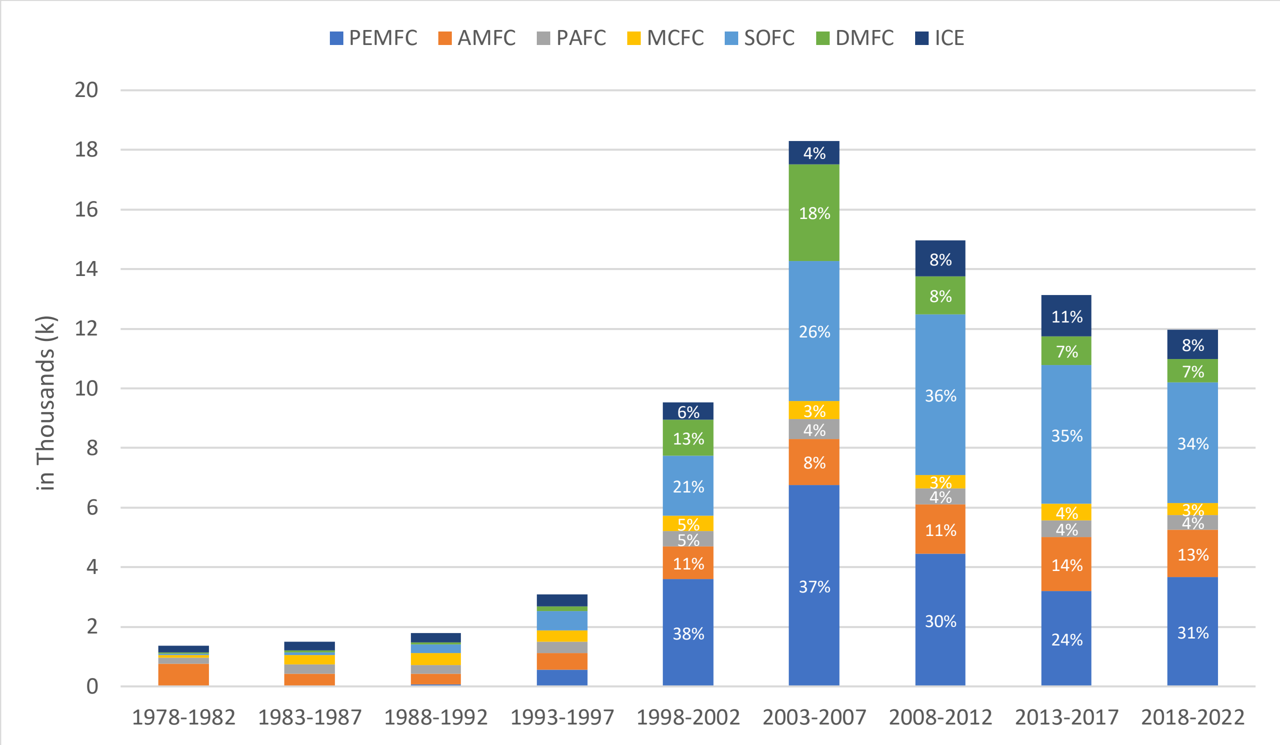



FIGURE 8: THE EVOLUTION OF PATENTS RELATED TO FUEL CELLS

A summary of the evolution of patent activity for several fuel cell technologies and the hydrogen internal combustion engine (ICE) is shown in *Figure 8*. Notably, Solid Oxide Fuel Cells (SOFC) and Proton Exchange Membrane Fuel Cells (PEMFC) both exhibit a notable increase in patent activity beginning in 1998-2002. Direct Methanol Fuel Cell (DMFC) patent activity shows an increasing interest between in 1998-2002 and 2003-2007, followed by downward trend in favour of alternative technologies. Hydrogen internal combustion engines (ICE) patenting activity presents a rather stable trend. On the other hand, and the Alkaline Membrane Fuel Cells (AMFC), Molten Carbonate Fuel Cells (MCFC), Phosphoric Acid Fuel Cells (PAFC), and display sustained but relatively sluggish development in patent activity. Overall, it is possible to infer that patenting activity related to Fuel Cells technologies has reached its peak in the 2003-2007 period.

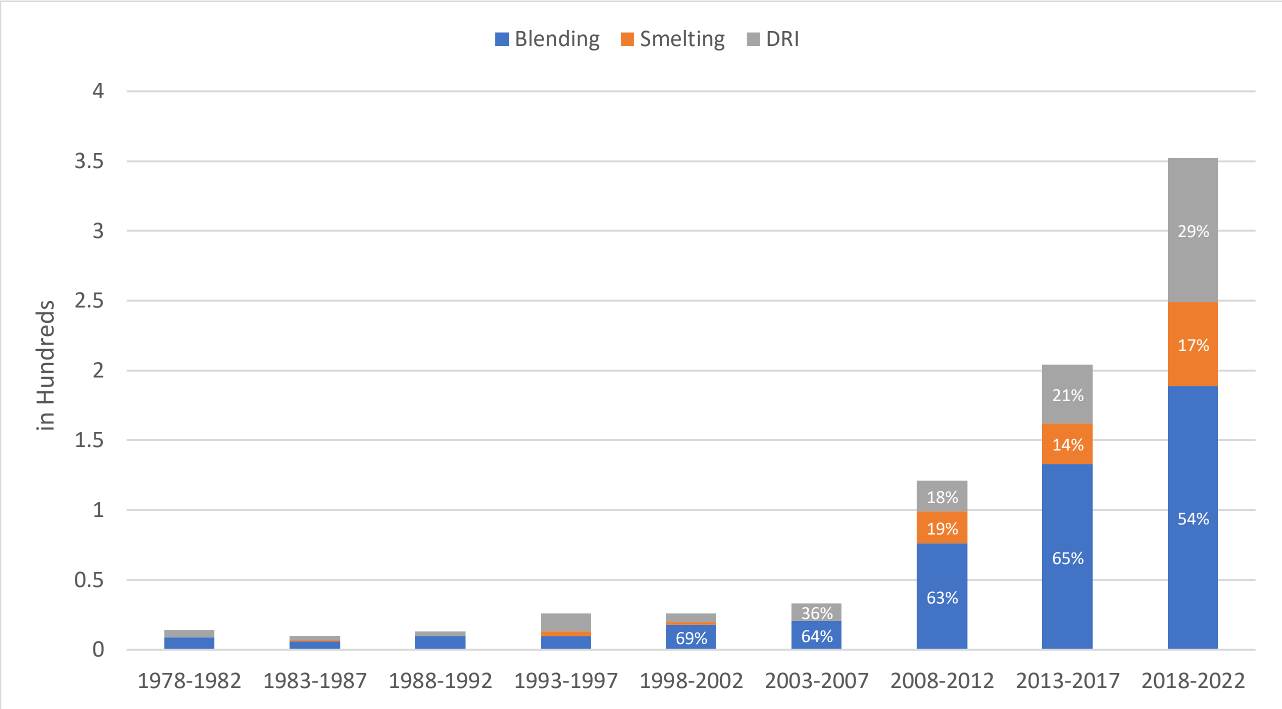



FIGURE 9: THE EVOLUTION OF PATENTS RELATED TO HYDROGEN-BASED IRON AND STEEL MANUFACTURING TECHNOLOGIES

*Figure 9* depicts the development of patent activity in hydrogen use technology in the iron and steel production sector. According to the data, there has been a noticeable increase in interest in these technologies starting from 2008. All of the subcategories have seen an uptick in patent activity since then. Particularly, a large and accelerating increase tendency may be seen in the blending of hydrogen in blast furnaces. Direct reduced iron (DRI) technologies are next, and they too show a significant increase in patent activity. Smelting reduction, on the other hand, exhibits a considerably slower growth tendency. The information demonstrates the increased interest in and innovation surrounding the use of hydrogen in the iron and steel production sector, with the integration of blast furnaces and DRI technologies leading the way.

## Geographical Distribution of Hydrogen Patents

*Figure 10* presents a compelling overview of the geographical distribution of all patented hydrogen technologies until the end of 2022 across various patent offices.

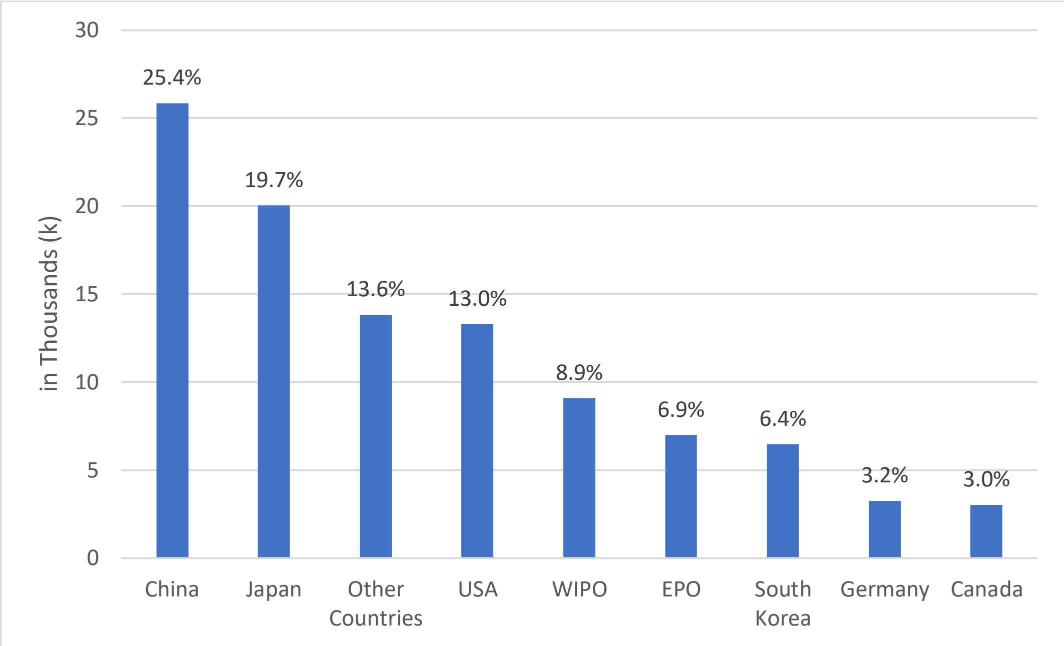



**FIGURE 10: GEOGRAPHICAL DISTRIBUTION OF HYDROGEN PATENTS BY PATENT OFFICE**

It is straightforward to notice that China, which holds a stunning 25.4% of all hydrogen technology patents, emerges as the front-runner. This sizeable proportion highlights China's continuous dedication to promoting research and innovation connected to hydrogen. With a respectable 19.7% share, Japan (JP) comes in second place, demonstrating its considerable contributions to the hydrogen industry. United States (US) patents represent 13.0% of the total hydrogen technology patents, highlighting the country's ongoing efforts and investments to advance hydrogen-related innovations. Similarly, South Korea, Germany and Canada demonstrate a notable contribution of 6.4%, 3.2% and 3.0%, respectively.

The European Patent Office (EP) region, represented by the code EP in the data, showcases a patenting activity share of 6.9% in hydrogen technology. It's important to note that the EPO Patent distribution may not fully reflect the extent of patenting activity within Europe, as many patents related to hydrogen technologies may have been filed directly at national patent office rather than through the European Patent Office. Therefore, the actual patenting activity in Europe is undoubtedly higher than the indicated percentage. Moreover, the World Intellectual Property

Organization's (WIPO) patenting activity accounts for 8.9% of the total data. Notably, many of these WIPO-registered patents may have originated from countries with a high patent intensity, such as Japan, China, or the United States. Filing with the WIPO permits inventors to pursue international protection for their inventions, ensuring broader coverage beyond the borders of a single nation. This highlights the strategic approach adopted by innovators to secure global intellectual property rights through a WIPO-facilitated centralized and efficient process. Notably, a wide variety of countries make up the "Other Countries" group, which together account for 13.6% of the total number of patents on hydrogen technology. This emphasizes that the development of hydrogen-based technologies is a worldwide effort in which many nations are actively involved.

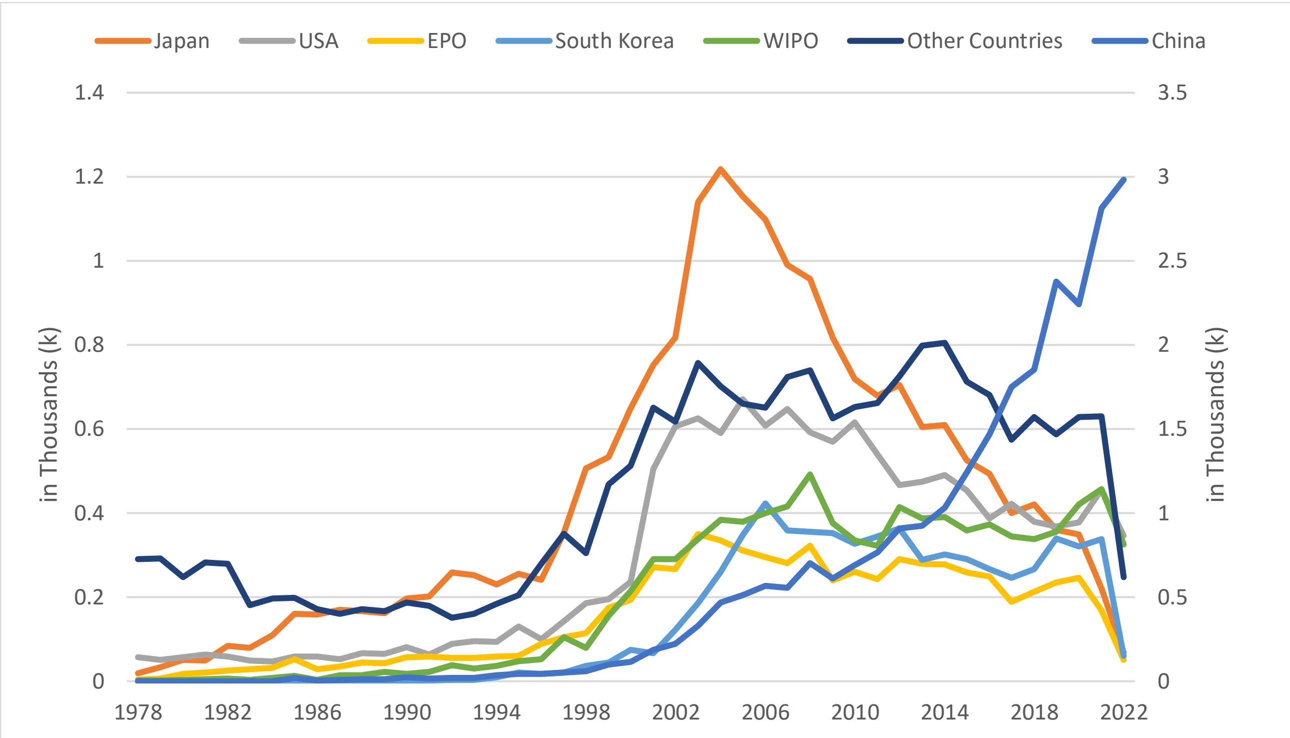



FIGURE 11: THE EVOLUTION OF HYDROGEN PATENT APPLICATIONS BY PATENT OFFICE CHAIN (CHINA SCALED ON RIGHT AXIS, ALL OTHER COUNTRIES SCALED ON LEFT AXIS)

*Figure 11* provides a representation of the evolution of hydrogen patents registered at each patent office. Notably, China experiences an exponential increase in hydrogen patents beginning in the early 2000s (CAGR approximately equal to 23% from 2000 to 2022), leading it to become the undisputed leader of hydrogen patents from 2010 onwards. On the opposite end, Japan demonstrates a mountain-shaped trend: after displaying initial exponential growth in patenting activity (23% CAGR over the 1978-2005 period) and dominating innovation in the field of hydrogen technologies, hydrogen patent count starts a rapid decline after 2006, indicating a potential shift in research focus or shifting priorities within the Japanese hydrogen industry. Concerning hydrogen patenting activity, it is straightforward that Japan, followed by the United States, which display a comparable trend, exhibited a remarkable period of growth and dominance in this field from the late 1990s up to 2010. In subsequent years, however, hydrogen patenting activity in both countries sharply declines. Again, these results align with the findings of Sinigaglia T. et al. (2018), previously cited in the Literature Review section. As a matter of fact, both studies confirm a dominant position in patenting activity of hydrogen related technologies for both countries in the first decade and a relatively declining trend in the second decade of the twenty-first century. Overall, from the conjunct patent filings across all countries, it is possible to infer that hydrogen innovation has seen an increasing trend starting from 1998, which has quite stabilized in the last years (see EPO, WIPO, South Korea and Other Countries).

It is imperative to mention that, for the same reasons previously illustrated, 2022 displays a representation patenting activity that is not free from bias and therefore should not be considered.

## The Complexity of Hydrogen Patented Technologies

The number of inventors associated with a patent is seen as an indicator of the underlying technology's complexity. In other words, the greater number of inventors in the same patent indicates a higher level of complexity and technical knowledge necessary to produce innovation (Broekel, 2019). This is due to the fact that sophisticated technologies frequently require multidisciplinary knowledge, collaboration among specialists from many fields, and elaborate problem-solving method. Analysing the number of inventors provides a more in-depth insight of the complexity nature, level and transversality of competence and knowledge needed to advance in a particular technological subject.

It is important to mention that of the 101.834 patent applications that make up the dataset, 9.675 patents (less than 10% of the total) showed no inventors. Logically, this represents a misinformation and therefore all these patents were removed from the panel of data exclusively for the purpose of the following considerations.

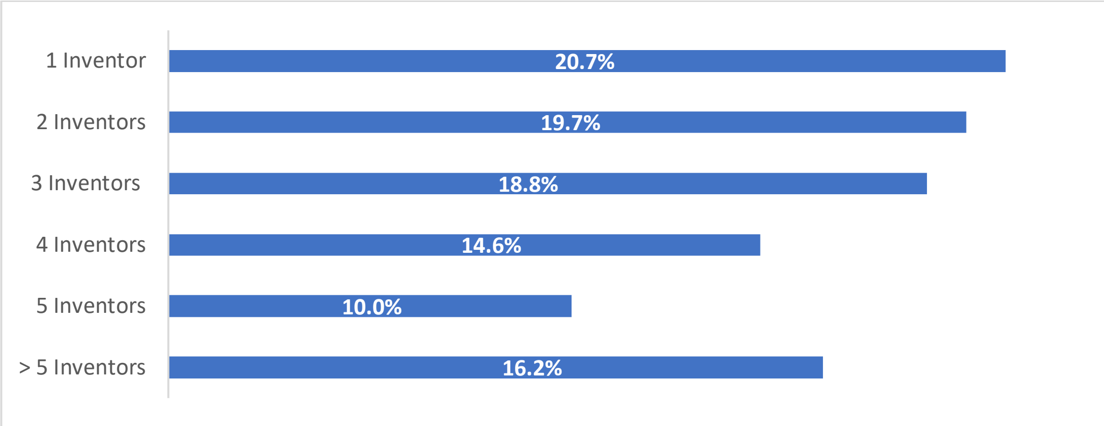



FIGURE 12: HYDROGEN PATENTS BREAKDOWN BY NUMBER OF INVENTORS

*Figure 12* provides insights regarding the complexity of hydrogen related patented technologies. The distribution of inventors in hydrogen patents demonstrates different levels of joint efforts. The majority of patents (20.7%) had only one inventor, showing that individual contributions were made. However, a sizable fraction has several inventors, with 19.7% having two and 18.8% having three. The patent share drops as the number of inventors grows, indicating larger teams or more sophisticated technology. This highlights the complexity and interdisciplinary nature of hydrogen innovation.

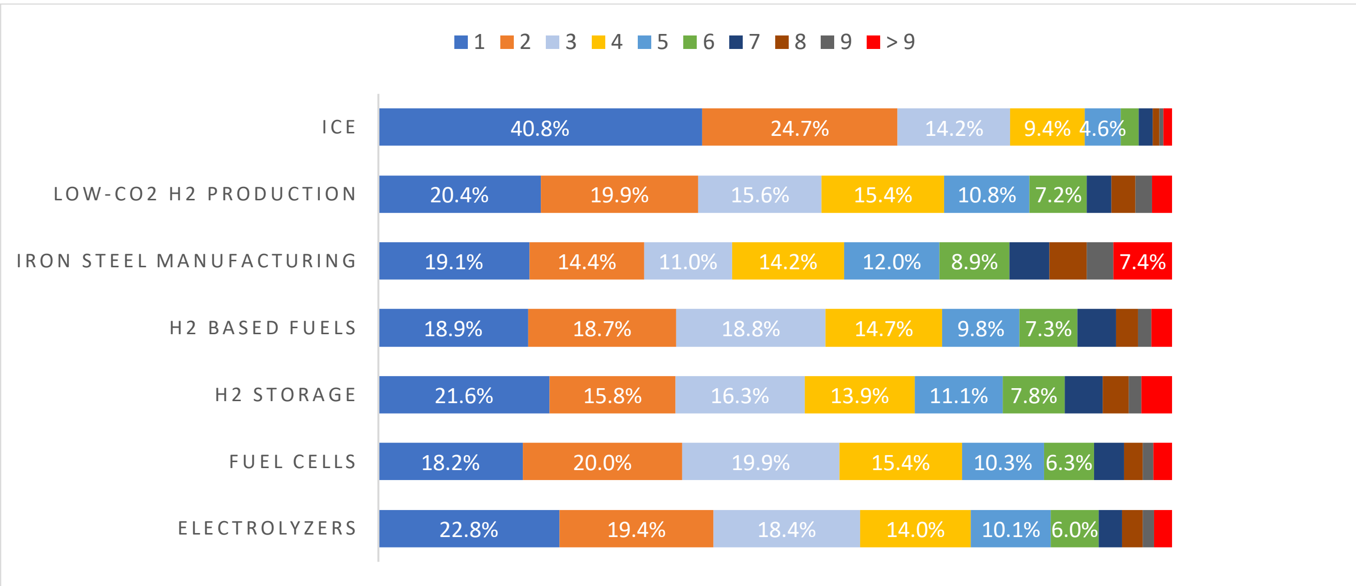



FIGURE 13: DISTRIBUTION OF INVENTORS PER PATENT BY SECTION OF THE HYDROGEN VALUE CHAIN

*Figure 13* shows that while most technological macro categories across the hydrogen value chain have similar inventor levels per patent class, hydrogen ICE-related innovations appear to be technologically easier. This discovery could be related to the existing amount of combustion engine knowledge and skill obtained from other industries, which can be easily transferred and applied to the hydrogen industry. This familiarity with combustion engines may have sped up the innovation process and contributed to the relative simplicity of hydrogen ICE-related technology.

Additionally, it is interesting to track the evolution of complexity of hydrogen patented technologies over time. Thus, *Figure 14* illustrates the average number of inventors per patent year on year from 1978 to 2022 for the entire population of hydrogen-related patents.

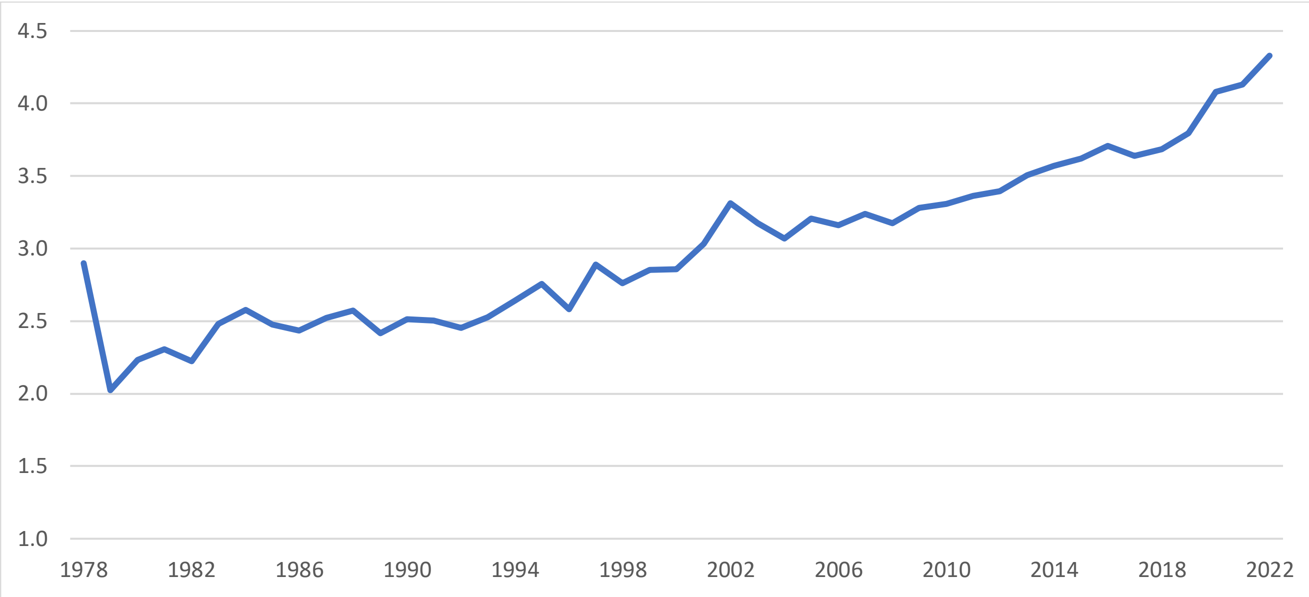



**FIGURE 14: THE EVOLUTION OF THE COMPLEXITY IN HYDROGEN TECHNOLOGIES: AVERAGE NUMBER OF INVENTORS PER HYDROGEN PATENT**

It is easy to notice that, over the years, the average number of inventors per patent has shown a consistent upward trend, indicating a complexity in hydrogen-related research and development. In 1979, the average number of inventors per patent was 2.02, and by 2022, it had increased to 4.33. As a result, excluding 1978, which appears as an outlier year, the CAGR for the average number of inventors per patent from 1979 to 2022 is 1.8%. The rising average number of inventors per patent suggests that hydrogen-related innovation has become more complex and interdisciplinary over time, necessitating the knowledge and participation of several innovators. Similar results are observable also when accounting exclusively for three of the main macro technology groups of the hydrogen value chain (electrolysers, fuel cells and hydrogen storage technologies). From *Figure 15* it is interesting to note that even though each of the three technologies follow comparable trends, hydrogen storage solutions have had the steepest increase in average number of inventors per patent over the last fifteen years, surpassing even fuel cells, which have long been the most complex technology of the hydrogen value chain.

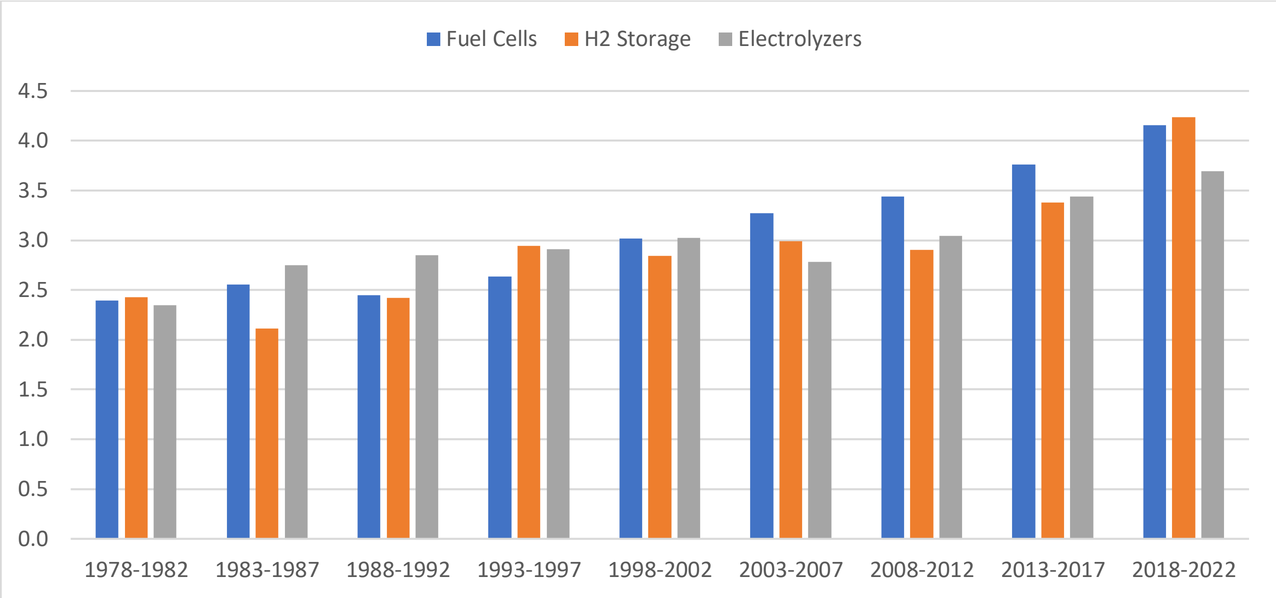



**Figure 15: The Evolution of the Complexity in Hydrogen Technologies: Average Number of Inventors per Hydrogen Patent across the Main Categories of the Value Chain**

## The Collaborativity in Developing Hydrogen Patented Technologies

The number of patent assignees is an indicator of collaborativity in the technology sector (Geum et al., 2021). In other words, the higher the number of assignees in the same patent, the higher propensity of dividing investments, sharing resources, and developing economies of scale in producing the innovation. Investigating the number of assignees that support the development of a specific technology can provide significant insights into the collaborative dynamics and collective accomplishments within a certain technological domain. According to WIPO, a collaborative patent is one that incorporates more than one assignee. It denotes the collaborativity of numerous entities, such as businesses, research institutions, or people, to contribute their knowledge, resources, or intellectual property to the development, invention, or implementation of the patented technology. Vice versa, a non-collaborative patent, on the other hand, is one that has only one assignee. In this situation, a single entity, such as a firm or an individual, is entirely responsible for the patented technology's discovery, development, and ownership. Non-collaborative patents indicate that intellectual property rights and control over technologies are not shared.

*Figure 16* provides insights regarding the collaborativity of hydrogen related patented technologies across the entire value chain. It is important to mention that of the 101.834 patent applications that make up the dataset, 3.972 patents (less than 4% of the total) showed no assignees. Logically, this represents a misinformation and therefore all these patents were removed from the panel of data exclusively for the purpose of the following considerations.

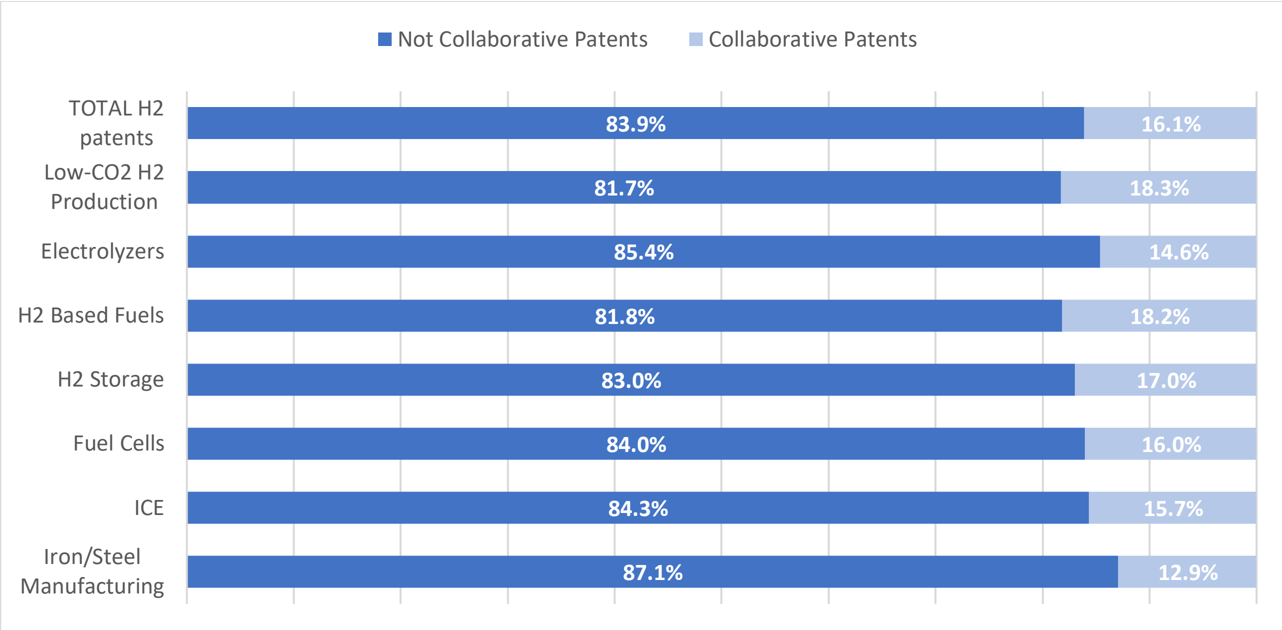



**FIGURE 16: DISTRIBUTION OF COLLABORATIVE PATENTS ACROSS HYDROGEN PATENTS**

The distribution of assignees in relation to hydrogen technology patents reveals interesting insights. Overall, 83.9%, have a single assignee, indicating individual ownership. However, 16.1% of the patents incorporate two or more assignees, indicating also the presence of a collaborative approach to hydrogen technology creation. This implies that a sizable part of hydrogen patents is the result of collaborative efforts between numerous firms. The prevalence of multi-assignee patents highlights the importance of collaborativity and collaborative efforts in advancing hydrogen technology.

Additionally, according to the data, low CO2 emission hydrogen production technologies have a higher degree of collaborativity, with around 18.3% of the patents in this category being collaborative. This implies that the development of sustainable hydrogen generation systems necessitates greater investments and resources, resulting in increased collaborativity and cost sharing. On the other hand, the adoption of hydrogen in the iron and steel manufacturing industry process appears to be a less collaborative field, with only 12.9% of patents classed as collaborative. This lower degree of collaborativity may be linked to variables such as the iron and steel industry's distinct needs and characteristics, which may necessitate less comprehensive collaborativity or entail fewer resource-intensive technologies.

Additionally, it is interesting to track the evolution of collaborativity of assignees in developing hydrogen patented technologies over time. Thus, *Figure 17* shows the average number of assignees per patent over time.

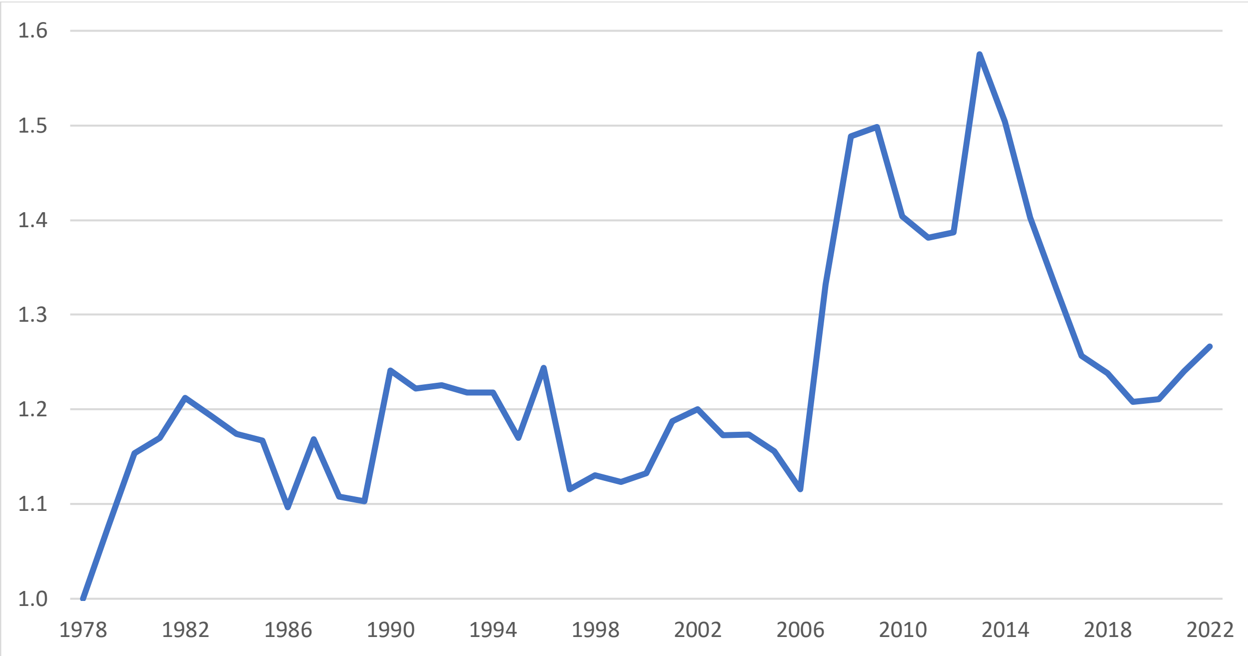



FIGURE 17: THE EVOLUTION OF THE COLLABORATIVITY IN HYDROGEN TECHNOLOGIES: AVERAGE NUMBER OF ASSIGNEES PER HYDROGEN PATENT

The average number of assignees per patent remained relatively steady from 1978 until the early 1990s, ranging between 1.0 and 1.2. However, there are certain swings during this time frame. The average number of assignees per patent appears to have increased from the late 1990s to the early 2000s, peaking at roughly 1.575 in 2013. This shows that collaborative patent activity increased throughout that time period. This rise could be attributed to the emergence of public-private partnerships and international collaborations in hydrogen research, particularly between 2005 and 2010. During this time, governments worldwide increased funding and support, establishing consortia and programs to accelerate hydrogen technology development. Additionally, an interdisciplinary approach involving experts from various fields gained momentum, leading to collaborative patent filings. After 2013, the average number of assignees per patent gradually declines, with slight volatility but generally remaining over 1.2.

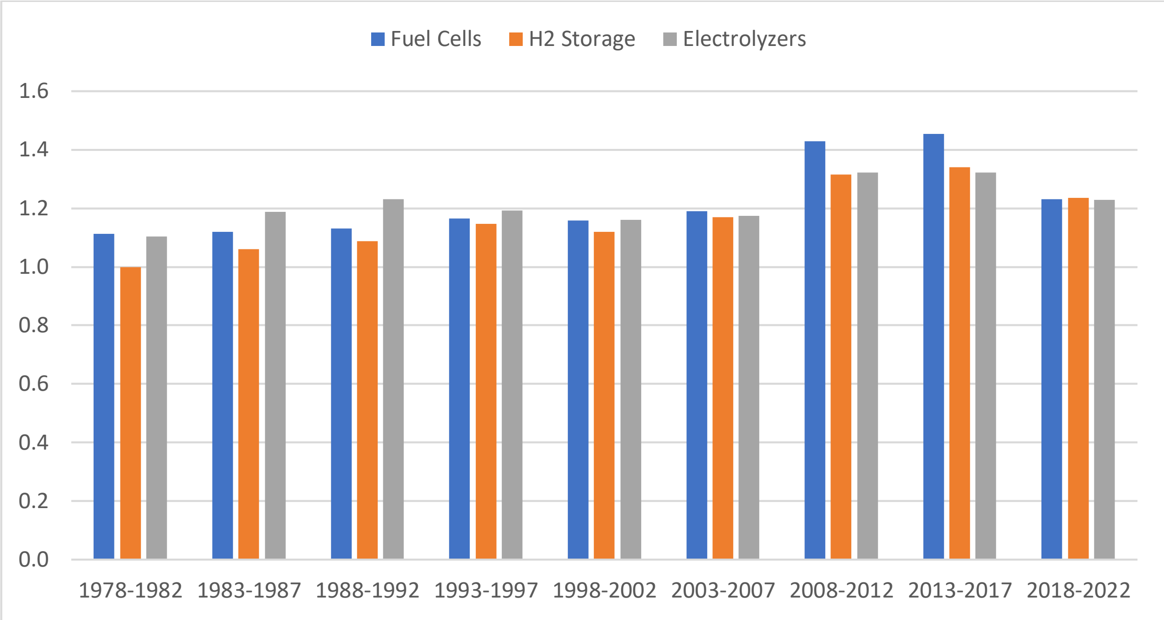



FIGURE 18: THE EVOLUTION OF THE COLLABORATIVITY IN HYDROGEN TECHNOLOGIES: AVERAGE NUMBER OF ASSIGNEES PER HYDROGEN PATENT ACROSS THE MAIN CATEGORIES OF THE VALUE CHAIN

*Figure 18* depicts the evolution of collaborativity in developing hydrogen patents among assignees, exclusively for three of the main macro technology groups of the hydrogen value chain (electrolysers, fuel cells and hydrogen storage technologies). From both graphs it is possible to evince that collaborativity for developing hydrogen patents is subject to fluctuations and cycles, which could be justified by a variety of factors. Technological improvements play a role, with discoveries necessitating greater collaborativity at first and then resulting in fewer assignees as the technology matures. Changes in research funding, government initiatives, and industry agendas can all have an impact on partnership patterns. Market dynamics and rivalry may influence fluctuations, as collaborativity may be viewed as a means to obtain an advantage or share costs.

## Top Innovators in the Hydrogen Sector

*Table 3* displays a list of the 50 innovators that are assignees of the hydrogen patents identified in the dataset.

| Innovator | Industry | Country | Number of H2 Patents |
|---|---|---|---|
| PANASONIC | Consumer Electronics | Japan | 1,376 |
| TOYOTA | Automotive | Japan | 1,326 |
| MITSUBISHI | Conglomerate | Japan | 1,209 |
| HONDA | Automotive and motorcycles | Japan | 882 |
| SAMSUNG | Conglomerate | South Korea | 845 |
| SIEMENS | Engineering and electronics | Germany | 816 |
| NISSAN | Automotive | Japan | 808 |
| HITACHI | Conglomerate | Japan | 784 |
| GENERAL MOTORS | Automotive | United States | 779 |
| TOSHIBA | Conglomerate | Japan | 741 |
| SUMITOMO | Conglomerate | Japan | 615 |
| COMMISSARIAT ENERGIE ATOMIQUE | Research and development in nuclear energy | France | 590 |
| AGC INC | Glass manufacturing | Japan | 559 |
| MERCEDES-BENZ | Automotive | Germany | 504 |
| TOPSOE | Catalysts and sustainable energy solutions | Denmark | 503 |
| GENERAL ELECTRIC | Conglomerate | United States | 495 |
| TOTO LTD | Sanitary ware and plumbing fixtures | Japan | 491 |
| BOSCH | Engineering and electronics | Germany | 488 |
| FUJI ELECTRIC | Electrical equipment and systems | Japan | 487 |
| DALIAN INSTITUTE OF CHEMICAL PHYSICS | Research and development in chemical physics | China | 483 |
| LG | Conglomerate | South Korea | 477 |
| BLOOM ENERGY | Fuel cells and clean energy solutions | United States | 444 |
| EXXONMOBIL | Oil and gas | United States | 382 |
| KOREA INSTITUTE OF SCIENCE AND TECHNOLOGY | Higher education and research | South Korea | 341 |
| GUANGDONG HYDROGEN ENERGY SCIENCE AND TECHNOLOGY | Hydrogen energy solutions | China | 334 |
| SANYO ELECTRIC | Electronics and appliances | Japan | 330 |
| UNIVERSITY OF CALIFORNIA | Higher education and research | United States | 300 |

| | | | |
|---|---|---|---|
| ZHEJIANG UNIVERSITY | Higher education and research | China | 288 |
| CERES POWER | Fuel cells and energy systems | United Kingdom | 286 |
| PEKING UNIVERSITY | Higher education and research | China | 284 |
| UNITED TECHNOLOGIES CORPORATION | Conglomerate | United States | 282 |
| OSAKA GAS | Energy and utilities | Japan | 281 |
| BALLARD POWER SYSTEMS | Fuel cells and clean energy solutions | Canada | 281 |
| ZAHNRADFABRIK FRIEDRICHSHAFEN | Automotive components and systems | Germany | 269 |
| NGK INSULATORS | Ceramic products and electrical components | Japan | 267 |
| TORAY INDUSTRIES | Advanced materials and chemicals | Japan | 262 |
| HUANENG CLEAN ENERGY RESEARCH INSTITUTE (CERI) | Clean energy research and development | China | 262 |
| TOPPAN PRINTING | Printing and packaging solutions | Japan | 258 |
| HYUNDAI | Automotive | South Korea | 253 |
| FUELCELL ENERGY | Fuel cells and clean energy solutions | United States | 251 |
| DELPHI TECHNOLOGIES | Automotive components and systems | United Kingdom | 245 |
| TECHNICAL UNIVERSITY OF DENMARK | Higher education and research | Denmark | 227 |
| AIR LIQUIDE | Industrial gases and services | France | 227 |
| NIPPON CATALYTIC CHEMICALS | Chemicals and catalysts | Japan | 222 |
| KANSAI ELECTRIC POWER | Electric utility company | Japan | 204 |
| KOREA INSTITUTE OF ENERGY RESEARCH | Energy research and development | South Korea | 204 |
| TSINGHUA UNIVERSITY | Higher education and research | China | 203 |
| TIANJIN UNIVERSITY | Higher education and research | China | 202 |
| FORSCHUNGSZENTRUM JUELICH GMBH | Higher education and research | Germany | 198 |
| SHELL INTERNATIONALE RESEARCH MAATSCHAPPIJ | Oil and gas research and development | Netherlands | 194 |
| JOHNSON MATTHEY | Sustainable technologies and materials | United Kingdom | 193 |
| BRIDGESTONE | Automotive and rubber products | Japan | 193 |
| DUPONT DE NEMOURS | Chemicals and materials | United States | 192 |

TABLE 3: LIST OF THE TOP 50 INNOVATORS BY PATENTING OF HYDROGEN VALUE-CHAIN-RELATED TECHNOLOGIES

It is crucial to note that the following geographical considerations are only estimates and may not be completely correct, as the country of the patenting organization may differ from the publication office where the innovation was patented.

| Country | Total H2 Patents | Patents from Recurring Assignees | Share of top 50 innovators | Count of Recurring Assignees |
|---|---|---|---|---|
| Japan | 20,028 | 11,295 | 56% | 19 |
| China | 25,861 | 2,056 | 7.95% | 7 |
| United States | 13,287 | 3,125 | 24% | 8 |
| South Korea | 6,470 | 2,120 | 33% | 5 |

TABLE 4: GEOGRAPHICAL DISTRIBUTION OF THE TOP 50 INNOVATORS BY PATENTING OF HYDROGEN VALUE-CHAIN-RELATED TECHNOLOGIES

Surprisingly, innovation in hydrogen technologies appears to be densely concentrated in Japan. In fact, 19 of the most recurrent assignees present in *Table 3* are from Japan, accounting for 11,295 patent applications out of the total 20,028 patents identified in Japan. This means that these organizations are responsible for more than 56% of all hydrogen patents in Japan. In contrast, hydrogen technology innovation in China presents to be much more widely spread. As a matter of fact, only 7 assignees are mentioned as the most frequent assignees of hydrogen patents, accounting for fewer than 8% of all hydrogen patents in China. Instead, the United States and South Korea have an intermediate situation. The US contains 8 organizations with considerable hydrogen patenting activity, accounting for 24% of all hydrogen patents found in the country. Similarly, South Korea has 5 organizations that account for 33% of the country's hydrogen patents.

It should be emphasized that drawing comparable conclusions for European countries would be deceptive because most assignees may have patented their innovations at the European Patent Office (EPO), making it impossible to assign these to single counties.

| Most Recurring Industry/Sector of Recurring Assignees | Count of Assignees |
|---|---|
| Automotive | 9 |
| Education / Universities | 8 |
| Conglomerate | 7 |
| Energy / Utilities | 6 |
| Hydrogen Focus | 5 |
| Research Centers | 5 |
| Electronics | 4 |
| Oil & Gas | 2 |
| Engineering | 2 |

TABLE 5: MOST RECURRING INDUSTRIES/SECTORS OF THE TOP 50 INNOVATORS BY PATENTING OF HYDROGEN VALUE-CHAIN-RELATED TECHNOLOGIES

As shown in *Table 5*, the organizations that have the most intense patenting activity across hydrogen technologies belong to the Automotive (9), Conglomerates (7) and Energy or Utility (6) industries. Additionally, much of the innovation activity in the field of hydrogen is carried out by universities (8) and or Public Research Centres (5).

It is interesting to note also that the actors driving hydrogen innovation in China, Japan, the United States, and Europe range significantly, with varying degrees of involvement from private enterprises to universities and national research organizations. The results displayed suggest that the hydrogen innovation landscape in China is characterized by the presence of numerous National Research Centres. This evidence could indicate that the Chinese government has a noteworthy inclination to exert control over innovation, particularly in strategic technologies that have the potential to shape the future, such as hydrogen. Universities typically focus on lower Technology Readiness Levels (TRLs), which represent the maturity of a technology, compared to private companies. This is because universities often engage in early-stage research and development, aiming to explore new concepts and prove their feasibility. Private companies, on the other hand, prioritize commercialization and market-ready technologies, targeting higher TRLs for immediate deployment. Consequently, this suggests that the innovation conducted in China, where universities play a significant role, may be more concentrated at an embryonic stage compared to other regions with a stronger presence of private companies. In the context of hydrogen related technologies, it is clear that private corporations, notably huge multinationals from the automotive industry and conglomerates, are the primary innovators in Japan and the United States. This demonstrates a definite tendency in these countries for private sector innovation in the hydrogen sector. Private parties own and control the majority of hydrogen-related innovation, emphasizing the importance of market-driven techniques and the role of competition in pushing breakthroughs in hydrogen technologies. While cooperation with universities is possible, the emphasis is on the contributions of private sector entities in pushing hydrogen innovation, distinguishing it from China's state-led model.

### Hydrogen Focused Corporations

This section will provide a brief overview of five most notable innovators that are exclusively involved in developing hydrogen technologies.

***Bloom Energy*** (444 hydrogen patents) is a company established in the United States in 2001 that offers solid oxide fuel cell (SOFC) technology solutions. Their fuel cell technologies convert a variety of fuels, including natural gas, into lower-emission energy.

***Guangdong Hydrogen Energy Science and Technology*** (334 hydrogen patents) is a Chinese company, founded in 2013, that specializes in the research, development, and application of hydrogen energy technologies. Their main area of expertise involves the manufacture of hydrogen fuel cells and related products for a variety of industries, helping to expand China's hydrogen energy sector.

***Ceres Power*** (286 hydrogen patents), founded in 2001 and based in the United Kingdom, is a pioneer in solid oxide fuel cell (SOFC) technology. They design and build fuel cell systems that generate energy effectively from a variety of fuels, including natural gas and hydrogen. The fuel cell technology developed by Ceres Power has applications in distributed power production, combined heat and power (CHP) systems, and electric vehicle (EV) charging infrastructure.

***Ballard Power Systems*** (281 hydrogen patents) is a Canadian firm, founded in 1979, that specializes in the design, development, and manufacture of proton exchange membrane (PEM) fuel cells and related hydrogen fuel cell products. They offer renewable energy solutions for a variety of applications such as transportation, backup power, and material handling.

***FuelCell Energy*** (251 hydrogen patents), founded in 1969 and based in the United States, is a global pioneer in the development, production, and operation of fuel cell power plants. They specialize in the development of clean, efficient, and dependable fuel cell technologies, which generate energy via the electrochemical reaction of hydrogen and oxygen and have applications ranging from stationary power generation to carbon capture and utilization.

## The Technological Domain of Hydrogen Patents: IPC Section

The International Patent Classification (IPC) is a hierarchical system for categorizing patents according to the technical subject matter they cover (*WIPO*). It is critical in organizing and retrieving patent documents all around the world. Each patent is assigned a unique code by the IPC categorization, which represents the specific technology or field to which it belongs. This classification system is significant for a number of reasons. To begin, it makes efficient patent searches possible, allowing patent examiners, inventors, and researchers to find relevant previous art and assess the uniqueness of an invention. The IPC classification is also useful for tracking technological changes and mapping the intellectual property landscape across businesses. It aids in the identification of developing technologies, possible areas for collaboration or licensing, and the distribution of technical information. Furthermore, the IPC classification is internationally recognized, providing patent offices with a uniform language and assuring worldwide consistency in patent documentation and inspection procedures.

*Figure 19* illustrates the breakdown of hydrogen related patented technologies per IPC section. It is important to mention that of the 101.834 patent applications that make up the dataset, 701 patents (less than 1% of the total) showed no IPC code. Logically, this represents a misinformation and therefore all these patents were removed from the panel of data exclusively for the purpose of the following considerations.

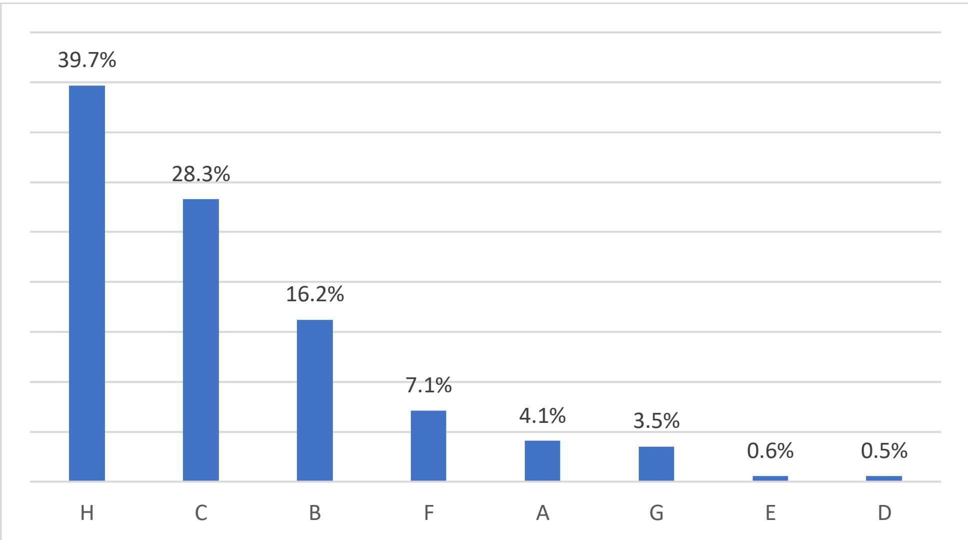



FIGURE 19: HYDROGEN PATENTS BREAKDOWN BY IPC SECTION

As depicted by the graph, the most recurring IPC Sections in the dataset are "H" (40%), "C" (28%) and "B" (16%).

IPC Section "H" (Electricity) is focuses on patents relating to electricity (*WIPO, IPC Section*). It encompasses a wide range of inventions and technologies related to the generation, distribution, and consumption of energy. This section contains several subcategories that deal with various aspects of electrical systems and equipment. IPC Section "H" is especially relevant in the context of hydrogen-related patents because it includes inventions related to hydrogen generation, storage, and utilization methods. Section "H" discusses hydrogen-related subjects such as fuel cells, electrolysis devices, hydrogen generators, hydrogen infrastructure, and hydrogen-powered vehicles. Patents relating to these inventions would be covered by IPC Section “H”.

IPC Section “C” (Chemistry; Metallurgy) is dedicated to patents relating to chemistry and chemical processes (*WIPO, IPC Section*). It includes chemical compositions, reactions, and procedures as inventions. IPC Section “C” is relevant in the context of hydrogen because it encompasses patents on hydrogen synthesis, purification, and chemical processes linked with hydrogen generation or consumption. Patents for catalysts, materials, and techniques utilized in hydrogen-related applications may be included.

IPC Section “B” (Performing Operations; Transporting) includes a variety of technological topics, such as engineering and industrial processes (*WIPO, IPC Section*). It includes inventions involving machines, apparatus, and technological systems. IPC Section “B” is relevant in the context of hydrogen since it comprises patents for engineering and manufacturing techniques that are expressly related to hydrogen-related technology. Inventions relating to hydrogen storage tanks, hydrogen transportation systems, hydrogen infrastructure, and other engineering features relating to hydrogen generation, distribution, or consumption may be included.

### IPC sections "H" (Electricity) and "C" (Chemistry; Metallurgy) Breakdown

As mentioned earlier, being the IPC a hierarchical system, it is possible to assess the technological classification of a patent into different levels of granularity. By adding two digits after the letter (Section identifier), the IPC Class is obtained.

Within IPC Section “H”, Class “H01” occurs 94.9% of the time. IPC Class “H01” specifically focuses on basic electric elements and electric power supplies. It covers inventions related to electrical components, circuits, and systems. IPC Class H01 encompasses electrical elements and components, therefore patents pertaining to the design, configuration, or improvement of electrical components used in fuel cells are contained within this class. Innovations in fuel cell electrodes, catalysts, membrane electrode assemblies (MEAs), bipolar plates, or current collectors fall under this category. Additionally, Class “H01” also includes ideas relating to electrical components and circuits utilized in electrolysis devices for hydrogen production. This includes patents for electrolyser designs, electrode materials, current distribution systems, or electrolysis-specific control circuits. Last, Class “H01” includes patents relating to the control and regulation of electrical systems in hydrogen technologies. Control circuits, power management systems, or monitoring devices specifically built for hydrogen fuel cells, electrolysers, or other hydrogen-related systems are included.

With respect to IPC "Section C", the IPC Classes that are most recurring are IPC Class "C01" (21.5%), "C25" (20.1%), "C08" (13.5%) and "C07" (11.7%).

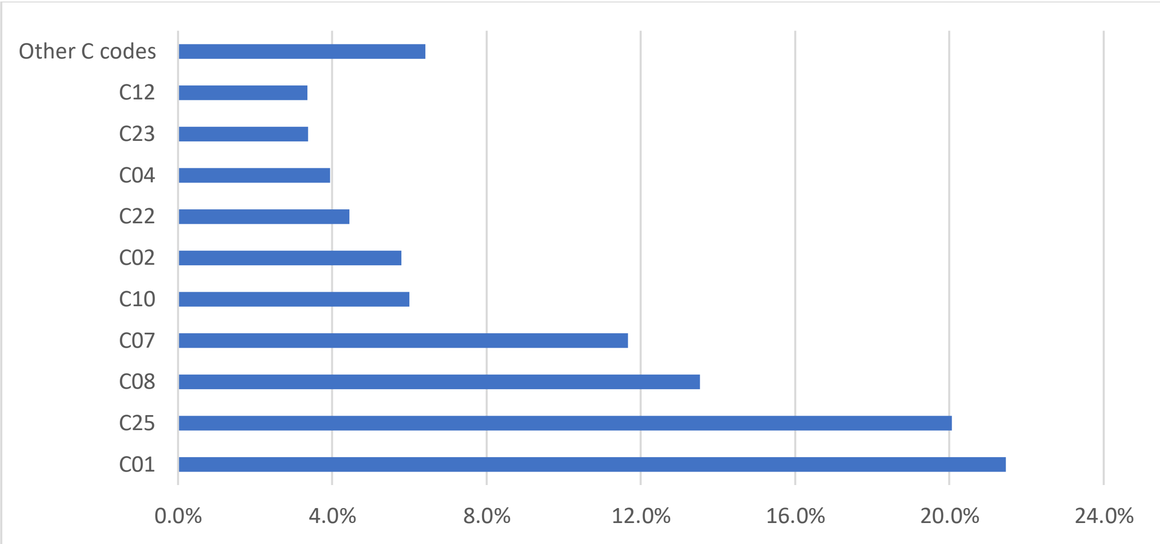



FIGURE 20: IPC SECTION "C" HYDROGEN PATENTS BREAKDOWN BY IPC CLASS

IPC Class "C01" focuses on inorganic chemistry, which includes inventions pertaining to hydrogen synthesis, purification, and chemical processes linked with hydrogen production or consumption. This category includes hydrogen-related technologies such as catalysts, materials, and procedures for hydrogen production, storage, and conversion.

IPC Class "C25" is connected to electrochemistry and covers patents relating to electrochemical processes, including those related to hydrogen technologies such as fuel cells and electrolysis devices. This category includes inventions relating to hydrogen fuel cells, electrolysers, and associated components.

IPC Class "C08" is concerned with organic macromolecular substances, such as polymers. While it is not entirely focused on hydrogen, it may include patents on hydrogen-related materials such as polymer electrolyte membranes used in fuel cells or hydrogen gas barrier materials.

IPC Class "C07" is dedicated to organic chemistry, and while it is less directly related to hydrogen, it may include patents related to organic compounds utilized in hydrogen-related applications, such as hydrogen storage materials or organic catalysts for hydrogen reactions.

## Hydrogen Green Patents: CPC Code Y

The European Patent Office (EPO), in collaboration with the United Nations Environment Programme (UNEP) and the International Centre for Trade and Sustainable Development (ICTSD), created a dedicated tagging scheme to identify low-carbon, sustainable, and climate change mitigation technologies (CCMTs) (Veefkind et al., 2012; Favot et al., 2023). These technologies are targeted by specific classes (Y02 and Y04S) target those technologies directly. The tagging activity is defined in algorithms created by a group of experts and they are re-run on a regular basis, so that new documents that meet the search requirements are automatically detected and tagged (Angelucci et al., 2018). Thus, the Y codes are added to the classification that already exists. This methodology is used by the European Commission's Joint Research Centre (JRC) to discriminate between green and non-green patents (Bellucci et al., 2021; Pasimeni et al., 2019).

Overall, the Cooperative Patent Classification system (CPC) is an extension of the International Patent Classification (IPC) system, which has been in use since January 2013 by the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO). Subsequently, it has also been adopted by the State Intellectual Property Office of China (SIPO), the Korean Intellectual Property Office (KIPO) and the Japan Patent Office (JPO) as of 2013, 2015 and 2016, respectively. The presence of a specific "Y" technology segment in the CPC is one noticeable variation between the CPC and IPC classifications. This "Y" section comprehends the “General Tagging of New Technological Developments; General Tagging of Cross-sectional Technologies Spanning over Several Sections of the IPC; Technical Subjects Covered by Former USPC Cross-reference Art Collections [XRACs] and Digests” (USPTO, EPO). Within the "Y" section of the Cooperative Patent Classification (CPC), the "Y02" subclass specifically focuses on 'green' patents related to sustainable energy technologies: “Technologies or Applications for Mitigation or Adaptation Against Climate Change” (USPTO, EPO) and the “Y04S” subclass focuses on “Smart grid technologies, including hybrid vehicles interoperability” (USPTO, EPO).

The EPO produces a new list with all codes multiple times a year, making it simple to create a list of CPC Green codes. The "Y" CPC-section contains all Y02/Y04S codes. The final May 2022 list includes 381 CPC Green codes. The key components of the Tagging scheme are shown in *Table 6*. The Y02 and Y04S tags associated with climate change mitigation technologies (CCMT) are used in this methodology to detect patents in green technology (Favot et al. 2023).

| CPC Section | Description |
|---|---|
| Y02 | Technologies or applications for mitigation or adaptation against climate change |
| Y02A | Technologies for adaptation of climate change |
| Y02B | Climate change mitigation technologies related to buildings, e.g., housing, house appliances or related end-user applications |
| Y02C | Capture, storage, sequestration or disposal of greenhouse gases [GHG] |
| Y02D | Climate change mitigation technologies in information and communication technologies [ICT], i.e. information and communication technologies aiming at the reduction of their own energy use |
| Y02E | Reduction of greenhouse gas [GHG] emissions, related to energy generation, transmission or distribution |
| Y02P | Climate change mitigation technologies in the production or processing of goods |
| Y02T | Climate change mitigation technologies related to transportation |
| Y02W | Climate change mitigation technologies related to wastewater treatment or waste management |
| Y04 | Information or communication technologies having an impact on other technology areas |
| Y04S | Systems integrating technologies related to power network operation, communication or information technologies for improving the electrical power generation, transmission, distribution, management or usage, i.e. smart grids |

TABLE 6: LIST AND DESCRIPTION OF CPC "GREEN CODES" OF PATENTED TECHNOLOGIES

Therefore, for the purpose of the considerations addressed in this master thesis, coherently with the European Commission's Joint Research Centre (JRC) methodologies, the presence of the "Y02" or "Y04S" CPC code indicates whether a patent related to hydrogen technologies is categorized as "green" or sustainable. This classification provides valuable insights into the environmental alignment of patented hydrogen technologies, allowing for a better understanding of their sustainability implications. Overall, considering the entire dataset, as depicted in *Figure 21*, 78.1% of all hydrogen patented technologies have been labelled as "green" or "sustainable".

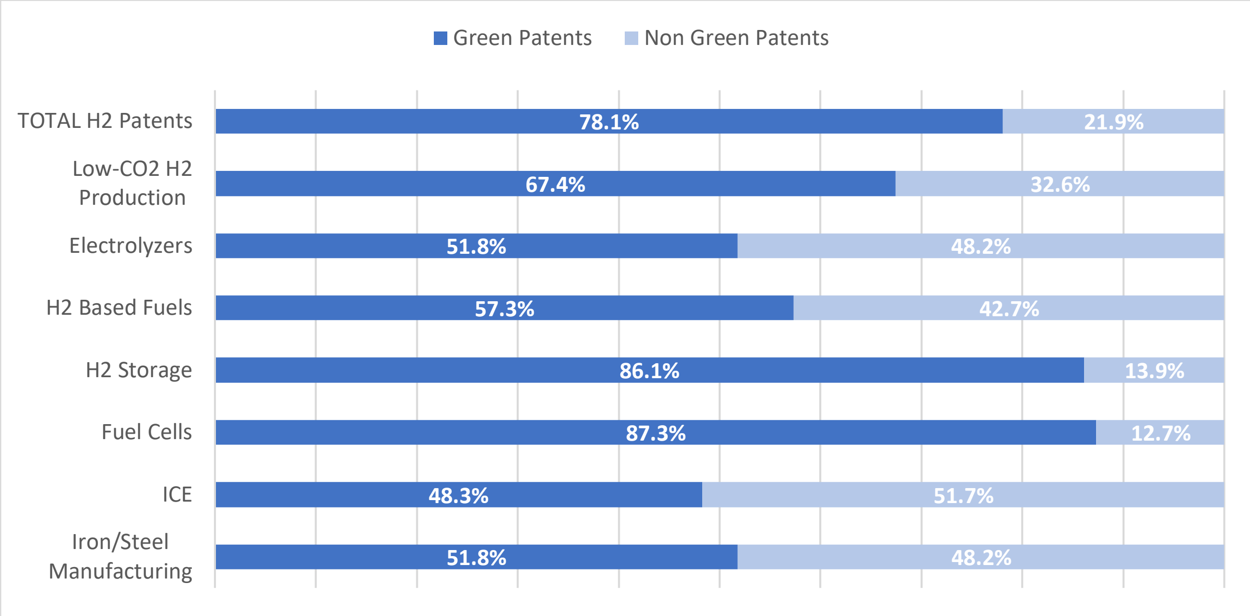



**FIGURE 21: THE DISTRIBUTION OF GREEN PATENTS AMONG HYDROGEN PATENTED TECHNOLOGIES ACROSS THE ENTIRE VALUE CHAIN**

The distribution of patents labelled with the “Y02” or “Y04S” CPC code across different macro categories of the hydrogen value chain, represented in *Figure 21*, reveals an interesting finding: only around 50% of certain macro technologies have been designated as "green."

However, it is important to consider several factors when interpreting these results. Firstly, the inclusion of the Y section in the CPC system began in 2013, and while efforts have been made to retroactively label older patents, there is a conservative tendency in assigning the "green" label to pre-2013 patents. Furthermore, the assignment of CPC codes relies on the interpretation and classification choices of patent examiners or applicants, leading to potential variations in labelling. Additionally, some patents may focus on technical advancements, efficiency improvements, or other aspects of hydrogen technology without explicitly highlighting their environmental or sustainability benefits. The extensive nature of the CPC classification system, encompassing various technological domains, adds further complexity to the labelling process. It is worth noting that the absence of the “Y02” or “Y04S” code in CPC classifications for hydrogen patents does not necessarily indicate a lack of sustainability. It may be a result of factors such as the patent's age, differing interpretations by examiners or applicants, specific focus, or limitations in the classification system's coverage.

# ANALYSIS OF THE TECHNOLOGICHAL SPECIALISATION IN HYDROGEN IN EUROPEAN REGIONS

Having completed the descriptive analysis of the sample of patents on hydrogen technologies, in this chapter some statistical and econometric analysis will be presented and their results will be discussed.

The purpose of these further studies, is to analyze the geographical distribution of hydrogen patents issued at the European Patent Office (EPO) for each European region. Specifically, the analysis will be carried out based on the second level of the Nomenclature of Statistical Territorial Units (NUTS2). As discussed in more detail below, the NUTS classification represents the standard for geographic classification of the European territory for statistical purposes. Through the use of statistical indices constructed from patent data, the hydrogen technology specialization of each NUTS2 region will be investigated. Then, two different research questions will be addressed. In the first strand, this thesis will study the spatial autocorrelation of the technological specialization in hydrogen of the European regions. In the second strand, on the hand, the characteristics that affect the technological specialization in hydrogen will be investigated through the implementation of two econometric models.

This chapter will be structured in four separate sections.

In the first part, the process of sample construction and the calculation of the hydrogen technological specialization will be presented. In the second section, the structure and characteristics of the sample will be described, showing which regions and countries have the highest concentration of hydrogen patents. In the third part, the first research question will be presented. Here, the spatial autocorrelation will be studied at the global and local level through some statistical indicators. In the fourth and last part, the second research question will be introduced. In particular, the construction of the econometric models will be shown and the variables used will be presented in detail. As will be described, the independent variables of the regressions were selected from the *Regional Innovation Scoreboard* which is a report produced biennially by the European Commission to assess innovation and innovative activities on different dimensions at the NUTS2 level. In particular, a static econometric model and a dynamic econometric model will be developed. For the static model, patents will be clustered regionally regardless of time while for the dynamic model, patents will be clustered by region while also taking into consideration the patent filing year (namely the application year). In the static model, variables from the *Regional*

*Innovation Scoreboard 2021* (*European Commission, 2021*) will be used as independent variables while for the dynamic model, all versions of the Regional Innovation Scoreboard will be considered for the independent variables.

## Data & Methodologies

**Patent Sample Construction**

The testing sample was constructed from the sample downloaded from Derwent Innovation, which, as seen in the previous chapter, contains 109,945 patents on hydrogen. Only hydrogen patents published at the EPO were extrapolated from this sample. In total, there are 6,994 EPO hydrogen patents, accounting for 6.36% of total hydrogen patents.

Through the use of STATA, the hydrogen total patents sample was then matched with the OECD's REGPAT database.

REGPAT is a database containing the total universe of patents filed with the EPO (this justifies using only hydrogen patents filed with the EPO) since 1977 (priority date). In this thesis, the February 2022 edition of REGPAT was used. As a result, all patents issued at the EPO with application year up to 2021 are contained within the database. As already explained, the patent becomes publicly accessible from the date of publication. There is usually an average period of 18 months between the application date and the patent publication date. Therefore, although the February 2022 edition is used, records up to 2021 (application date) are included inside the database. In total, almost 4 million patents (precisely 3,836,758) are contained within REGPAT. The choice to use REGPAT and not PATSTAT is due to the specific characteristics of REGPAT. As a matter of fact, REGPAT is a patent database where information on the region address of residence of the patents' inventors and assignees was added and linked to each patent (*Maraut, S. et al., 2008*). The address of patents' inventors, in particular, is the crucial and essential information for developing the analysis presented in this research study. The geographical information on the address of residence of inventors and assignees, within the database, is given both in text form and in NUTS code (at least for European addresses). Indeed, in REGPAT, regional breakdowns are based on the 2013 version of the nomenclature of territorial statistical units.

Below, the NUTS standard is briefly described.

The Nomenclature of Statistical Territorial Units (NUTS) is a geographical classification system introduced by the European Union (EU) since the 1970s and formally entered into force in 2003[3]. This system was designed with the intention of offering a unified and consistent framework for the collection, processing and analysis of regional statistical data in Europe, providing a solid basis for comparisons between different regions. The NUTS standard is also a key tool because it enables socio-economic studies by facilitating the definition of structural plans and policies at regional level. The NUTS structure is based on three distinct levels of territorial units. The first level, NUTS1, comprises large socio-economic regions within each country. These may include, for example, groupings of administrative regions or distinct geographic areas with significant populations. The second level, NUTS2, represents a more detailed breakdown of each member state and often refers to a nation's division into administrative regions. Finally, the third level, NUTS3, refers to the smaller and more detailed divisions of each region. The NUTS3 units often correspond to provinces or other types of administrative divisions at the local level.

Each territorial unit contained in the NUTS classification, is uniquely identified by an alphanumeric code. The codes follow a specific pattern based on the hierarchical level and the country to which the region considered belongs. The first two characters of the code are capital letters and identify the corresponding country according to the *ISO 3166-1 alpha-2 standard*[4]. For example, the country code for Italy is "IT," for France "FR," and for Germany "DE." Then, depending on the level considered, there can be from 1 to 3 characters. The first character following the country code represents the identifying character of the first level of the NUTS classification and can be either a number or a letter. The second and third characters, identifying characters of the second and third levels of the nomenclature respectively, on the other hand, are always numeric characters. Accordingly, a NUTS1 code will consist of 3 total characters, a NUTS2 code of 4 total characters, and a NUTS3 code of 5 total characters. For example, the "ITC" code (NUTS1 code) is the identifier for Northwest Italy. The "ITC1" code (NUTS2 code) identifies Piedmont and the "ITC11" code (NUTS3 code) identifies the province of Turin. In some cases, there may be additional subdivisions below the NUTS3 level that are identified by additional digits. These subdivisions are identified as local territorial units.

[3] The NUTS classification currently in force and its historical evolution can be found at the following link: https://ec.europa.eu/eurostat/web/nuts/background
[4] The ISO 3166 Country Codes standards are available at the following link: https://www.iso.org/iso-3166-country-codes.html

As preannounced in the introduction of the chapter, the analysis proposed by this master's thesis aims to study the European geography of hydrogen patents at the NUTS2 level, and consequently we will focus only on these identifying codes. There is a twofold justification for this choice. On the one hand, the NUTS2 level is the reference for the planning of European structural funds because NUTS2 regional areas are often used as a reference for the planning and implementation of regional and structural policies. On the other hand, the choice to use the NUTS2 level, is also dictated, as will be seen below, by the limited size of the analysis patent sample and thus by the difficulty in collecting meaningful information at a more detailed geographic level.

Going back to the construction of the testing sample, the database of total hydrogen patents filed with the EPO was matched with REGPAT through the application number, which is a unique identifying code for each patent. Specifically, the REGPAT dataset containing the inventors' residence address was used.

In fact, the study of the geographical distribution of hydrogen patents was conducted by analyzing the geographical location of the inventors of those patents. Studying the location of those who invented or introduced a new technology is an approach frequently used in the literature to study innovation and the geographic spread of technologies (e.g., *Jaffe et al., 1993; Breschi and Lissoni, 2005*).

In the match between the two databases, of the 6,994 hydrogen patents, 1,014 patents (the 14.5% of the total data) were lost. These 1,014 patents, represent hydrogen patents with a priority date before 1977 and an application date after 2021, which as mentioned, represent the cut-off dates outside of which there is no data in REGPAT. The overwhelming majority of this set falls into the second category, patents with application dates after 2021. Indeed, the sample of hydrogen patents was downloaded from Derwent Innovation in April 2023 and consequently includes very recently published patents. Thus, the initial sample is represented only by the 5980 hydrogen patents that matched with the data stored in REGPAT.

This patent group has a total of 19,411 inventors. Hence, on average, a hydrogen patent filed with the EPO has 3.24 inventors. This figure is in line with the analysis of the total sample of hydrogen patents (see *Figure 14*). It is also interesting to note that 19,411 inventors of hydrogen patents represent 0.19% of the total inventors of all patents filed with the EPO. There are 10,230,273 inventors in REGPAT.

Next, the amount of hydrogen inventors based in Europe was also studied. REGPAT contains all patents issued at the EPO but many inventors of these patents are non-European. This analysis showed that only 6,090 inventors, or 31.5% of the inventors of patented hydrogen technologies are from Europe. In contrast, 61.5% of inventors (13,229) reside outside Europe. Despite the fact that at this level the number of inventors and not the number of patents actually filed at the EPO are being considered, it can be said that the majority of the innovative activity on hydrogen technologies is performed outside Europe and then extended into Europe. According to the data collected, it can therefore be said that Europe for the majority of hydrogen-based technologies represents an extension market.

As stated, the presented study approach suggests that the position of the inventor of a patent can be used as an indicator of the location of the innovation and the patent itself. Up to now inventors were counted but from now on, patents will be counted directly. To move from the number of inventors to the number of patents, a common expedient employed in the literature was used. If a patent has more than one inventor and these are located in the same NUTS2 then the following patent is counted only once in the respective NUTS2. However, if a patent has more than one inventor and the inventors are located in different geographic regions, the patent is counted separately for each of these locations. In other words, the patent is duplicated for each of the different geographical areas where the inventors reside. For example, if a patent has two inventors, one residing in NUTS2 ITC4 (Lombardy) and one residing in NUTS2 ITC1 (Piedmont), the patent is counted once in both ITC4 and ITC1. While this method provides a detailed view of the geographic distribution of innovation, it may overestimate the presence of innovation in areas that host only a smaller proportion of the inventor team. In fact, a patent that has 10 inventors of which 9 are located in the same region and 1 located in another region is counted in both regions. This is an important aspect that will be considered when interpreting the results.

By the application of the approach described above, from more than 19 thousand inventors 8,172 patents were obtained, 2,192 more than the number of patents that was found previously (5,980 patents). These 2,192 patents represent all those patents with at least two inventors who are resident in two different geographic regions, and consequently for the purpose of the analysis proposed in this master thesis they were duplicated and counted for each different region of residence. Among the 8,172 patents constituting the analysis sample, 39.61% (3,237 patents) are located in Europe while 60.39% (4,935 patents) are located outside Europe. Just as with inventors even when considering the patents granted at the EPO, it can still be said that most of the innovative activity on hydrogen is carried out outside Europe. This result is consistent with what was seen in the previous chapter, in the descriptive analysis of the entire patent sample where it was shown that the most significant concentration of hydrogen patents is found in China, Japan and the United States (respectively 25.4%, 19.7% and 13.6% of the total sample). For the purpose of this study, only patents located in Europe are considered. Consequently, only 3,237 patents are analyzed from now on. Moreover, we should also point out that a total of 2,429,846 patents filed with the EPO are located throughout Europe (in the time period defined in the beginning of this chapter), of which 2,426,609 patents are non-hydrogen patents. Therefore we can say that just the 0.13% of the patents located in Europe and filed with the EPO are related to hydrogen or hydrogen-based technologies.

Afterwards, the 3,237 European H2 patents were broken down for each NUTS2 region. Obviously, for the purposes of the econometric models, this breakdown applies only to the static model. for the dynamic model, the 3237 patents were grouped by both region and application year.

**Technological specialization in hydrogen at NUTS2 level**

After having break down the 3,237 European H2 patents for every NUTS2 region, the specialization index in hydrogen for each European subregion was calculated. The hydrogen specialization degree for a particular area, within this master thesis, depends just on the H2 patents located in that specific region. In particular two very common indexes in the literature were used: the *Revealed Technology Advantage* (RTA) and the *Normalized Revealed Technology Advantage* (NRTA) indices. These indicators provide the same information but they have different statistical properties that will be exploited in the econometric model proposed in the second section of the chapter. Their characteristics are presented in detail below.

The *Revealed Competitive Advantage* (hereinafter RTA), also known as "Specialisation Index" (OECD, 2009) is the most common metric used in the realm of patents to investigate the technological advantages in a certain field. In particular it was first proposed by Balassa (*1965*) in his seminal work on Revealed Comparative Advantage (RCA) in the context of trade. After Balassa, the index was adopted by several scholars to study the technological specialization of different firms or countries in specific technology areas. The RTA index is defined as the ratio between the share of patents related to a particular technology field in a region and the share of patents in all technology fields (i.e., the patent universe in its entirety) related to the same region. The formula for the RTA indicator is the following (*Balassa, 1965; Khramova et al., 2013*):

$$RTA_{ij} = \frac{Po_{ij}/\sum_j Po_{ij}}{\sum_i Po_{ij}/\sum_i \sum_j Po_{ij}}$$

*$RTA_{ij}$* is the Revealed Technology Advantage index for the technological domain *i* (within this thesis with the subscript *i*, we refer to the hydrogen-based technologies) for the geographical region *j* (within this thesis with *j* we refer to the European NUTS2 regions). *$Po_{ij}$* indicates the number of hydrogen-related patents located in the region *j*. $\sum_j Po_{ij}$, instead, refers to the total number of hydrogen related patents located in the whole Europe. As mentioned several times, this amount is equal to 3,237 i.e. the European hydrogen patents filed at the EPO. Looking at the RTA denominator, we have the ratio between the sum of all patents from all the different technological domains released in a specific subregion *j* ($\sum_i Po_{ij}$) and the total number of patents located in the whole Europe ($\sum_i \sum_j Po_{ij}$). Again, this quantity has already been commented already and it is equal this to 2,429,846 patents.

The RTA as it is defined can take values in the range between zero and +∞.

The indicator assumes the following values if the following conditions are met (*Caviggioli et al., 2023*):

- ***$RTA_{ij}$* = 0** : if in region *j* there is no patents of technology domain *i*;
- **0 < *$RTA_{ij}$* < 1** : if in region *j* the share of patents of technology domain *i* is in percentage less than the share of patents of all technology domains;
- ***$RTA_{ij}$* = 1** : if in region *j* the share of patents of technology domain *i* is in percentage equal to the share of patents of all technology domains;
- ***$RTA_{ij}$* > 1** : if in region *j* the share of patents of technology domain *i* is greater than the share of patents of all technology domains.

As a result, if for a specific region the *RTA* index is less than one, no specialization is observed in the same region with respect to the considered technology domain. Conversely, if the *RTA* index for the selected region is greater in value than one then specialization is observed in the technological field analyzed. Thus, in this master's thesis, NUTS2 regions associated with an *RTA* value greater than one are considered to be specialized in the hydrogen technological domain, while regions with an *RTA* value between zero and one are considered to be non-specialized in the hydrogen technological domain.

Below, both the advantages and the drawbacks of adopting this kind of measure are presented.

The RTA possesses several strengths. First, it represents one of the most widely used indices in the literature. Among the several reasons behind its popularity, there is definitely the *RTA* intrinsic straightforwardness. The *RTA* represents a simple statistical measure both in terms of calculation and interpretation, making it easily accessible even to those people without advanced statistical knowledge.

Moreover, the RTA is a quantitative measure on the technological specialization of a region allowing direct comparison between different regions. The index is flexible and can be adapted to different dimensions of analysis. Indeed, meaningful studies can be conducted, for example, both at the firm level and at the country level. In the latter case, through the use of the RTA, international comparative studies regarding innovation on specific technological fields can be conducted.

The RTA is employed to identify strengths in the technological sphere of regions by providing information that can be used for the development of effective innovation policies and investment plans by national and international authorities.

Nevertheless, this metric also has some limitations. First, the RTA only takes into account the quantity of patents issued in a region/country without considering the patents quality in any way. This could result in an overestimation of technological specialization where there is a large amount of patents but of low quality.

The RTA also does not take into account external factors that may influence patent activity in a given technological field. For example, a region might be found to be highly specialized because it has policies that encourage a high patenting rate.

In addition, RTA should only be used to analyze developed countries with a large number of patents. Indeed, the analysis of regions/countries with a small number of patents can lead to distortions of that regions/countries specialization advantages. In fact, by measuring the ratio between the patent share of a specific technology domain and the patent share of all domains for a region/country, the RTA will tend to define a region/country with few patents of the analyzed domain and few patents in general as specialized and, at the same time, a region with many patents of the analyzed domain and many patents in general as unspecialized. The RTA turns out to be very sensitive to even small changes in the number of patents for countries or regions with low patenting rates.

The second indicator used to study the hydrogen specialization of European regions is the *Normalized Revealed Technology Advantage* (hereinafter *NRTA*). This index is also known as *Revealed Patent Advantage* (RPA). Basically the *NRTA* is the standardized version of the *RTA*.

The *NRTA* is computed with the following formula (*Laursen, 2015*):

$$NRTA_{ij} = \frac{({RTA_{ij}}^2 - 1)}{({RTA_{ij}}^2 + 1)}$$

The *NRTA* has been introduced to avoid the *RTA* uneven distribution of values. In fact, as seen before, the *RTA* is defined in the range [0, 1) in the absence of specialization in a technology domain and in the range (1, +∞) in the presence of a competitive advantage in the same technology domain. The *NRTA*, which is the normalized form of the *RTA*, provides, on the other hand, a symmetric distribution for detecting the technological advantage of a region/country within the analyzed scope of study. More precisely, the *NRTA* has a symmetric distribution around the value of zero and it assumes values in the range (-1,+1). In particular the *NRTA* takes the following values based on the occurrence of the following conditions:

- **$-1 < NRTA_{ij} < 0$** : no specialization within the technology field *i* is observed in the region *j*;
- **$NRTA_{ij} = 0$** : in the region *j* the specialization within the technology domain *i* is equal to the specialization of all technology domains together;
- ***$0 < NRTAij < 1$*** : specialization within the technology domain *i* is observed in the region *j.*

Thanks to its symmetry, the *NRTA* is easily interpreted and offers a more balanced measure for the technological specialization: a positive value identifies a specialization whereas a negative value indicates a sub specialization.

However, like the *RTA* index, the *NRTA* index is also subject to limitations. It does not consider global trends and does not reflect the intrinsic qualitative value of patents.

In conclusion, while the *RTA* and the *NRTA* provide valuable insights into the technological specialization and patenting activity, given the limitations of these two indices, their use, although being outside of the scope of this thesis, should be complemented by other indicators and considerations to provide a more comprehensive and accurate picture of technological advantage. These additional indicators/considerations should even incorporate non-patent data analysis, thus providing a more comprehensive view of technological specialization. In this way it is possible to mitigate the limitations of relying solely on patent data, thereby enhancing the robustness and reliability of the analysis.

The indicators described so far, as will be seen below were used for the study of spatial autocorrelation of technological specialization in hydrogen and for the static econometric model. For the dynamic econometric model, however, compared to the previous formulas, the time variable that is given in the case of patents by the application year must also be considered. The *RTA* and *NRTA* indicator formulas for the dynamic model are given below. The same considerations made above regarding interpretation, advantages and disadvantages also apply to the time-variant indicators.

$$RTA_{ijt} = \frac{Po_{ijt}/\sum_j Po_{ijt}}{\sum_i Po_{ijt}/\sum_i \sum_j Po_{ijt}}$$

$$NRTA_{ijt} = \frac{({RTA_{ijt}}^2 - 1)}{({RTA_{ijt}}^2 + 1)}$$

t represents the application year of the patent i.e., the year in which the patent is filed with the relevant patent office (see subchapter *Patents – Patenting Procedure*), in this thesis the European Patent Office (EPO).

## Descriptive Analysis of the Sample and Spatial Autocorrelation Analysis

The goal of this chapter is to provide an answer to the first research question in this thesis: *is there spatial autocorrelation among European hydrogen specialized regions? Do clusters of specialization in hydrogen exist in Europe?*

In this regard, this chapter is divided into different parts. The first part describes the database of patents filed with the EPO. The description of the database in turn is separated into two sections. The first one studies the geography of hydrogen patents granted at the EPO and having inventors located in Europe. In the second one, on the other hand, the geographic distribution and concentration of European regions specializing in hydrogen is described. In order to investigate the hydrogen specialization across Europe, both the NRTA index (the time-independent version) and the Boolean variable constructed from this indicator (and from the RTA index) are analyzed.

In the second section of the chapter, instead, an answer to the above research question is sought by studying the spatial autocorrelation across Europe of the regional hydrogen specialization. As described later, spatial autocorrelation is analyzed at both the global and local levels. In order to do that, the software GeoDa[5] was used, and the Moran scatterplot and LISA map were calculated.

## European hydrogen patents distribution and concentration analysis

The description of the dataset starts with the breakdown of European-based hydrogen related patents filed at the EPO by European country.

| Country | Total Patents | Total H2 Patents | H2 - Total patents Ratio |
|---|---|---|---|
| Germany | 1,003,957 | 1,468 | 0.15% |
| France | 330,466 | 408 | 0.12% |
| United Kingdom | 24,458 | 307 | 0.13% |
| Italy | 155,382 | 167 | 0.11% |
| Sweden | 98,198 | 162 | 0.16% |
| Denmark | 41,327 | 154 | 0.37% |
| Switzerland | 139,781 | 129 | 0.09% |
| Netherlands | 128,636 | 118 | 0.09% |
| Austria | 62,754 | 85 | 0.14% |
| Belgium | 60,193 | 64 | 0.11% |
| Norway | 16,620 | 43 | 0.26% |
| Finland | 46,457 | 40 | 0.09% |
| Spain | 39,857 | 29 | 0.07% |
| Other Countries | 65,760 | 63 | 0.10% |

TABLE 7: BREAKDOWN OF TOTAL PATENTS AND HYDROGEN PATENTS GRANTED AT EPO BY COUNTRY

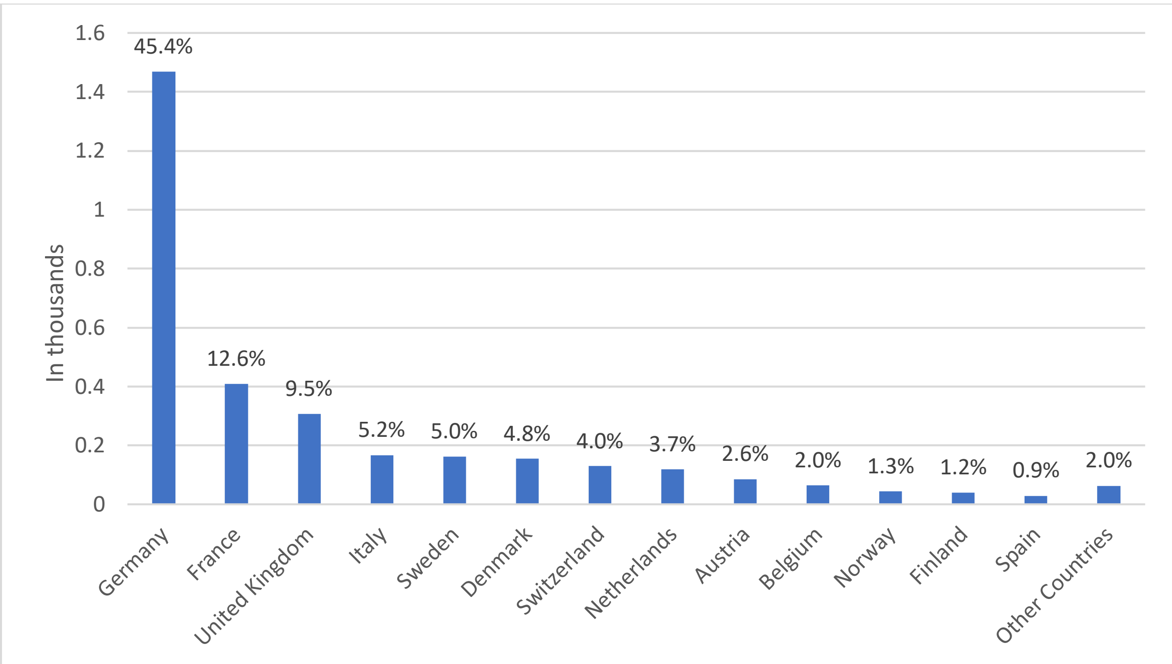



FIGURE 22: GEOGRAFICHAL DISTRIBUTION OF HYDROGEN PATENTS GRANTED AT EPO

[5] GeoDa is an open source software tool that develops spatial data analysis, geovisualization, spatial autocorrelation analysis and spatial modelling; GeoDa is available at the following link: http://geodacenter.github.io/

From *Table 7* and *Figure 22*, the German dominance is clear. Germany is the undisputed leader in this field with 1,468 hydrogen patents, representing 45.35 % of total hydrogen patents in Europe. Even in absolute terms, Germany has the highest number of total patents (1,003,957 patents, 41.3 % of the total patents located in Europe), suggesting a strong ecosystem of innovation in general and in the hydrogen field in particular. This result is quite consistent with the analysis of the total patent sample (see *Geographical Distribution of Hydrogen Patents*, *Figure 10*) where Germany stood out as the only European country with a significant share of patents on hydrogen technologies. The other prominent countries following Germany are France, the United Kingdom, and Italy, which account for 12.60 %, 9.48 %, and 5.16 % of total hydrogen patents, respectively. These countries, together with Germany, represent the main innovation hubs in Europe for hydrogen-related technologies.

Another interesting finding from *Table 7* is the country concentration of hydrogen patents relative to total patents. Looking at this figure, which provides insights into where the main hydrogen expertise and specializations might lie in Europe, we see that there is a clear diversity among countries in terms of volume and intensity of innovative hydrogen activities. The ratio of hydrogen patents to total patents, as the RTA and NRTA indices are defined, is very significant for hydrogen specialization. Denmark has a concentration of 0.37 %, the highest among the countries listed, despite having a relatively low number of total patents. This suggests that, proportionally, Denmark has a strong propensity toward innovation in hydrogen. After Denmark, the second highest country in hydrogen patent concentration is Norway (0.26 %).

Looking at the cumulative percentage of hydrogen patents, the top five countries (Germany, France, UK, Italy and Sweden) hold as much as 77.60 % of the total hydrogen patents in Europe. This indicates a strong concentration of innovative activity in these countries.

The breakdown analysis of the number of hydrogen patents in Europe was also conducted at the NUTS2 level, which as mentioned is the dimension of analysis of the two research questions addressed in this work.

| Region (NUTS code) | Country | Total Patents | Total H2 Patents | % H2 patents | H2 – Total patents ratio |
|---|---|---|---|---|---|
| Stuttgart (DE11) | Germany | 84,574 | 160 | 4.9% | 0.2% |
| Rhône-Alpes (FRK2) | France | 56,090 | 144 | 4.4% | 0.3% |
| Darmstadt (DE71) | Germany | 70,900 | 124 | 3.8% | 0.2% |
| South East (UKJ) | United Kingdom | 63,095 | 122 | 3.8% | 0.2% |
| Tübingen (DE14) | Germany | 34,382 | 120 | 3.7% | 0.3% |
| Oberbayern (DE21) | Germany | 95,672 | 120 | 3.7% | 0.1% |
| Île de France (FR10) | France | 119,616 | 119 | 3.7% | 0.1% |
| Köln (DEA2) | Germany | 62,788 | 113 | 3.5% | 0.2% |
| Mittelfranken (DE25) | Germany | 34,350 | 101 | 3.1% | 0.3% |
| Hovedstaden (DK01) | Denmark | 20,125 | 90 | 2.8% | 0.4% |
| Düsseldorf (DEA1) | Germany | 69,523 | 63 | 1.9% | 0.1% |
| Rheinhessen-Pfalz (DEB3) | Germany | 44,566 | 61 | 1.9% | 0.1% |
| Karlsruhe (DE12) | Germany | 62,229 | 58 | 1.8% | 0.1% |
| Schwaben (DE27) | Germany | 24,986 | 53 | 1.6% | 0.2% |
| Lombardy (ITC4) | Italy | 47,923 | 49 | 1.5% | 0.1% |
| Stockholm (SE11) | Sweden | 31,861 | 48 | 1.5% | 0.2% |
| Västsverige (SE23) | Sweden | 18,212 | 39 | 1.2% | 0.2% |
| East Midlands (UKF) | United Kingdom | 16,124 | 39 | 1.2% | 0.2% |
| Nordwestschweiz (CH03) | Switzerland | 31,534 | 37 | 1.1% | 0.1% |
| Schleswig-Holstein (DEF0) | Germany | 18,085 | 37 | 1.1% | 0.2% |

TABLE 8: BREAKDOWN OF TOTAL PATENTS AND H2 PATENTS BY TOP 20 EUROPEAN REGIONS

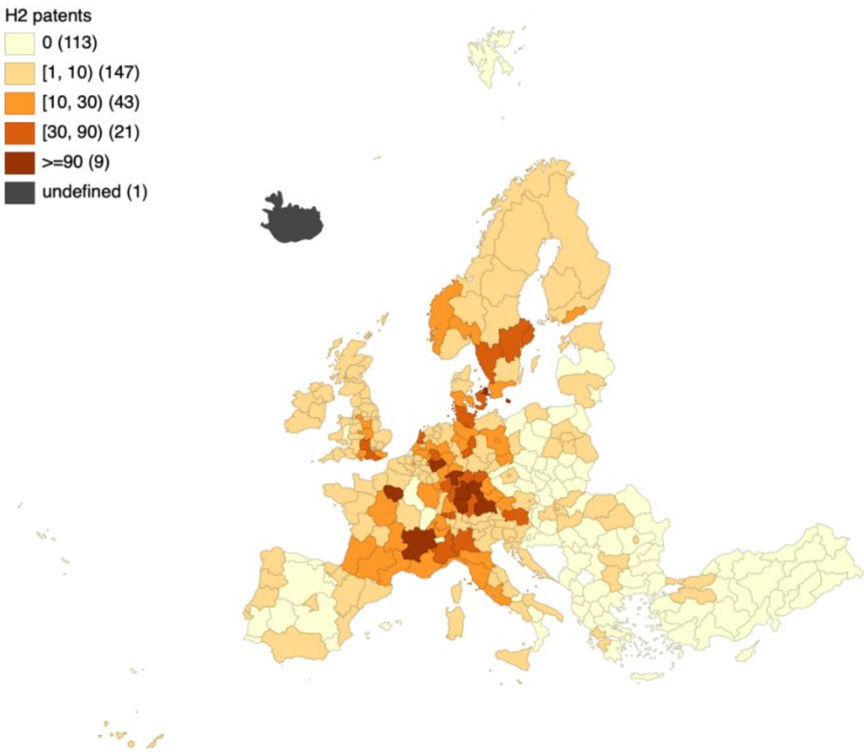



FIGURE 13 BREAKDOWN OF HYDROGEN PATENTS IN EUROPE

*Table 8* shows data for the 20 European regions with the highest number of hydrogen patents. In *Figure 23*, on the other hand, the European regions, as suggested by the legend, have been separated into 5 categories based on the number of hydrogen patents, and a different color has been assigned to each category.

Looking first at the data of *table 8*, the dominance of Germany once again becomes very evident. German regions dominate the list, with as many as 11 of the 20 regions listed. This is in line with the previous country-level analysis, which highlighted Germany as a leader in the sector. Stuttgart, Darmstadt and Tübingen are among the regions with the highest number of hydrogen-related patents.

Regarding France, although there were a significant number of patents in the country-level analysis, in *Table 8* we see only two French regions, but with a significant contribution: Rhône-Alpes and Île de France. There are 263 hydrogen patents in these two regions, accounting for 64.5% of all French hydrogen patents. This result suggests that hydrogen innovation in France is concentrated in a few key areas.

The 20 regions with the most hydrogen patents also include two English regions (South West and East Midlands), two Swedish regions (Stockholm and Vastsverige), one Danish region (Hovedstaden), one Italian region (Lombardy), and one Swiss region (Nordwestschweiz). Although these regions contribute less than some German and French regions, they emphasize the importance of hydrogen in their local innovation ecosystems and play a key role in driving hydrogen-related innovation in Europe.

Globally in the top 20 regions are concentrated 52.4% of all European hydrogen patents.

Regarding the intensity of innovation in hydrogen, the Danish region Hovedstaden has a hydrogen patents-total patents ratio of 0.4%, the highest among the listed regions. This underscores the region's strong inclination toward hydrogen technologies compared to other areas of innovation.

If we look at *Figure 23*, we have a clear indication of the distribution and concentration of hydrogen-related innovation in Europe at a regional level.

Several European regions, the 33.8% of the total number, do not even have a hydrogen patent. As shown in *Figure 23*, these regions, in some areas of Europe, are grouped together resulting in technologically backward clusters in the field of hydrogen. Such clusters are evident in parts of central and southern Spain and in the majority of Eastern Europe. The largest category is between 1 and 10 hydrogen patents. It includes 44% of European regions. These results highlight that despite the growing importance of hydrogen technologies, many regions in Europe are still starting or are in the early stages of adoption and of innovation in this area.

In contrast, only 9 regions across Europe (2.7%) have more than 90 hydrogen-related patents, suggesting that hydrogen innovation in Europe, according to patent data (in terms of volume), appears to be highly concentrated. The 43 regions with between 10 and 30 patents and the 21 regions with between 30 and 90 patents represent areas where hydrogen innovation is maturing. These regions could become leaders in the future, as long as they continue to invest in and support innovation.

In general, looking at *Figure 23*, the difference between Western and Eastern Europe is clear. The former, despite a few special cases, shows a clear superiority in terms of the number of hydrogen patents, demonstrating a significant advantage and a more developed ecosystem in the hydrogen technology domain.

Up to this point, the sample of hydrogen patents has been described at a general level. However, the previous analyses were also conducted at the level of each macro-category of the hydrogen technology taxonomy. For each hydrogen macro-technology, quite similar results to those just discussed were obtained. Before delving into these results, let us quickly analyze the breakdown of H2 patents by technologies.

*Figure 24* shows how hydrogen patents are distributed across the different macro-technologies presented in this master thesis.

In the sample of European patents filed with the EPO, the dominant macro-technology is Fuel Cells. Indeed, there are 1841 patents on Fuel Cells in Europe, accounting for 56.5% of all European hydrogen patents. After Fuel Cells, the technologies with the largest number of patents are hydrogen internal combustion engine (ICE), electrolyzers, H2 based fuels and H2 storage with 650 (19.9%), 414 (12.7%), 209 (6.4%) and 99 patents (3.0%) respectively. The other two technologies have a much smaller number of patents: 36 patents for Low-CO2 H2 production and 10 patents for Iron & Steel Manufacturing. These results are consistent with those presented earlier regarding the total sample of hydrogen patents.

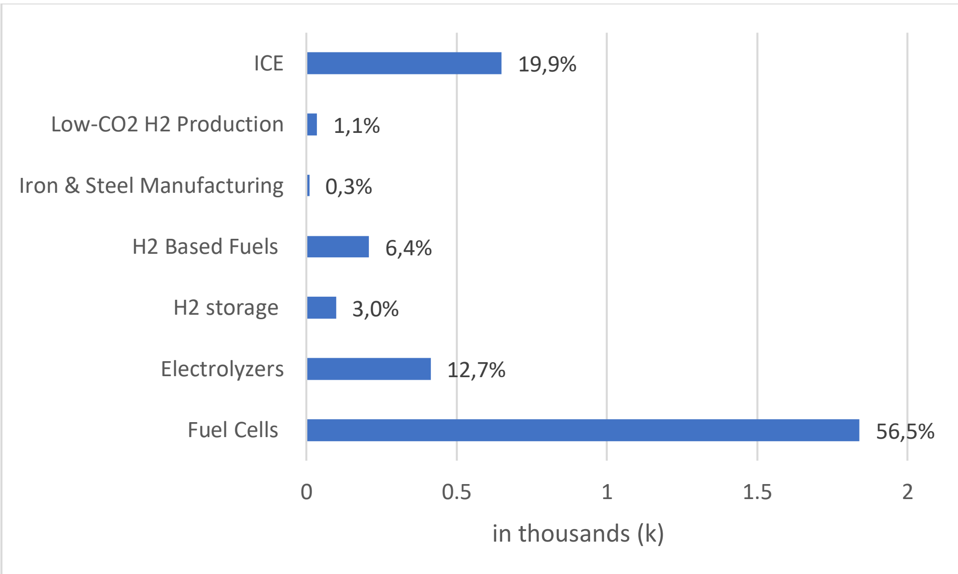



FIGURE 24: BREAKDOWN OF EUROPEAN PATENTS FILED AT THE EPO BY HYDROGEN MACRO-TECHNOLOGY

Returning to the geographical breakdown analysis of patents, *Table 9* presents the top 3 countries in terms of number of patents for each hydrogen technology.

| TOP 3 countries | Specific Hydrogen Technology patents | Specific technology patents - Total H2 patents ratio | % of specific hydrogen technology patents |
|---|---|---|---|
| | **Fuel Cells** | | |
| Germany | 830 | 56.5% | 45.1% |
| United Kingdom | 243 | 79.2% | 13.2% |
| France | 230 | 56.4% | 12.5% |
| | **Electrolyzers** | | |
| Germany | 160 | 10.9% | 38.7% |
| France | 62 | 15.2% | 15.0% |
| Denmark | 32 | 20.8% | 7.7% |
| | **H2 Storage** | | |
| Germany | 43 | 2.93% | 43.4% |
| France | 20 | 4.90% | 20.2% |
| United Kingdom | 10 | 3.26% | 10.1% |
| | **H2 Based Fuels** | | |
| Germany | 59 | 4.0% | 28.2% |
| France | 41 | 10.0% | 19.6% |
| United Kingdom | 28 | 9.1% | 13.4% |
| | **Iron & Steel Manufacturing** | | |
| Austria | 4 | 4.71% | 40.0% |
| Germany | 2 | 0.14% | 20.0% |
| France | 2 | 0.49% | 20.0% |
| | **Low-CO2 H2 Production** | | |
| France | 10 | 2.5% | 2.8% |
| Norway | 9 | 20.9% | 25.0% |
| Denmark | 5 | 3.2% | 13.9% |
| | **ICE** | | |
| Germany | 370 | 25.2% | 56.9% |
| Sweden | 77 | 47.5% | 11.9% |
| Austria | 52 | 61.2% | 8.0% |

TABLE 9: BREAKDOWN OF MACRO-TECHNOLOGY HYDROGEN PATENTS BY TOP COUNTRIES

Once again German leadership is confirmed. Germany clearly emerges as the European leader in almost all categories of hydrogen-related technologies (first in five categories and second in one category), except for low-carbon hydrogen production. This German lead is in line with previous analyses, which had highlighted Germany as the European country with the largest number of hydrogen patents.

Regarding fuel cells, which as seen are the most popular technology in terms of patents, holding more than 56% of all hydrogen patents, although Germany has the lead in absolute terms, it is interesting to note that in the United Kingdom fuel cells account for nearly 80% of all hydrogen patents, suggesting a strong specialization in this area.

Regarding electrolyzers, Germany holds nearly 40% of all European patents in this field, with France and Denmark following. The high percentage of electrolyzer-related patents compared to total hydrogen patents in Denmark is remarkable, indicating a particular interest in the country in this technology.

In hydrogen combustion engines, Germany has a dominant share, with 56.9% of European patents. However, it is also interesting to note here that Sweden and Austria show very high percentages of ICE patents compared to their total hydrogen patents, suggesting considerable specialization in these countries for hydrogen-fueled internal combustion engines.

Focusing on less common technologies, despite counting a very small number of patents, offer interesting insights. For example, Austria emerges as a leader in the iron & steel category, while France seems to have a particular interest in low-carbon hydrogen production.

In particular, France, having a diversified technology portfolio (it appears in almost all categories among the top three countries) would seem to embrace a broader diversification strategy than other countries in the hydrogen sector.

It is very interesting to note that for all technologies, the top three countries by number of patents always concentrate more than 50% of the patents for that specific technology.

In conclusion, this analysis testifies in addition to the undisputed German leadership, how there are different regional specializations in hydrogen-related technologies across Europe.

A similar table for the 5 regions with the most patents for each hydrogen-based technology is shown in the Annex (see *Table 10*).

This analysis concludes the description of the breakdown of European hydrogen patents.

## Geographical Analysis on Technological Specialization in Hydrogen

This section presents the geographical distribution across Europe of the specialization related to hydrogen and hydrogen-based technologies. In order to accomplish this purpose both the Boolean variable built on the NRTA index and the NRTA index itself are investigated.

As already said, when the NRTA index, for a given region, is above 0 the selected region is considered specialized in hydrogen. At the same time when the NRTA index, for a given region, is below 0, the selected region is not considered specialized. As discussed before (see *Technological specialization in hydrogen at NUTS2 level*) the usage of the NRTA to evaluate the regional specialization in a particular technology domain could lead to distorted results. As a matter of fact, regions with a significant share of patents on the technology of interest and a significant share of patents on all technologies at the same time might turn out to be unspecialized in the analyzed perimeter. Conversely, regions with far fewer patents in the technology of interest and few patents in general might turn out to be specialized in the analyzed perimeter. This aspect should be considered in the subsequent investigation.

In *Figure 25*, European regions at the NUTS2 level have been divided according to the NRTA index value into four categories (see the legend for the breakdown). Two different colour scales were chosen to visualise the difference between H2-specialised regions (positive NRTA indicator) and non-H2-specialised regions (negative NRTA indicator).

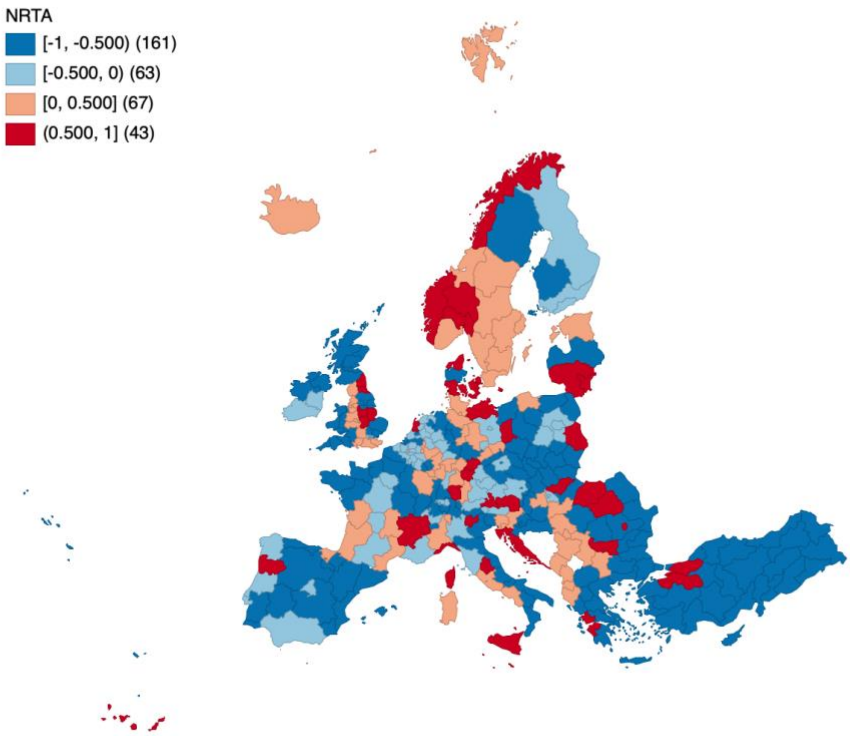



FIGURE 25: HYDROGEN SPECIALIZATION DISTRIBUTION ACROSS EUROPE BASED ON THE NRTA INDEX

Since in correspondence of a positive NRTA index value there is specialization in the relevant technology domain, from the map above we can conclude that in Europe, 110 regions (33%) are specialized in hydrogen while 224 regions (67%) are not specialized in hydrogen. In addition, 161 regions (48%) have a deeply negative NRTA index while "only" 43 regions (13%) have a strongly positive NRTA index. The large number of non-specialized regions seems to suggest that despite the growing importance of hydrogen as an energy resource, many European regions may still not have recognized it as a priority sector or may have other areas of specialization.

*Table 11* shows all European countries (in descending order by number of hydrogen-specialized regions) with at least one specialized region.

| Country | # of NUTS per Country | # of NUTS specialized in hydrogen | Hydrogen specialized NUTS per country - NUTS per country ratio |
|---|---|---|---|
| Germany | 38 | 19 | 50% |
| Italy | 21 | 8 | 38% |
| Sweden | 8 | 7 | 88% |
| Norway | 7 | 6 | 86% |
| France | 14 | 5 | 36% |
| United Kingdom | 12 | 5 | 42% |
| Denmark | 5 | 4 | 80% |
| Poland | 16 | 3 | 19% |
| Romania | 8 | 3 | 38% |
| Hungary | 7 | 3 | 43% |
| Slovenia | 2 | 2 | 100% |
| Bulgaria | 6 | 2 | 33% |
| Netherlands | 12 | 2 | 17% |
| Spain | 19 | 2 | 11% |
| Turkey | 26 | 2 | 8% |
| Portugal | 7 | 1 | 14% |
| Switzerland | 7 | 1 | 14% |
| Iceland | 2 | 1 | 50% |
| Greece | 13 | 1 | 8% |
| Austria | 3 | 1 | 33% |
| Belgium | 3 | 1 | 33% |
| Estonia | 1 | 1 | 100% |
| Malta | 1 | 1 | 100% |
| Croatia | 2 | 1 | 50% |
| Lithuania | 1 | 1 | 100% |

TABLE 11: LIST OF EUROPEAN COUNTRIES WITH AT LEAST ONE NUTS REGION SPECIALIZED IN HYDROGEN

It is immediately evident by looking at the data in *Table 11* and *Figure 25* how there is high specialization in hydrogen in some Nordic countries. Sweden, Norway, and Denmark have a very high percentage of regions specializing in H2 (88 %, 88%, and 80%, respectively). This suggests that these countries are putting a significant strategic focus on hydrogen technology, and probably possess the infrastructure and resources to support this specialization.

For other countries, however, there is diversity in specialization. Countries such as Germany and Italy, while having more specialized regions in absolute terms, show a lower percentage of specialization than Nordic countries. This may reflect a more diversified national strategy, where hydrogen energy is only one of many priority sectors.

For some small countries, the ratio of specialized regions to total regions is observed to be 100 %. Obviously, this figure does not appear very significant since these countries have only one region that encompasses the entire national land area.

When these data are viewed in light of those analyzed in the previous section, a discrepancy emerges between the absolute number of hydrogen patents and specialization in this field. For example, Germany, which was the absolute leader in the number of patents, has only 50 percent of its regions specialized in hydrogen. This difference could be attributed to the definition of the NRTA index, which assesses specialization only in relative terms. Germany, as we have seen, contains not only the majority of hydrogen patents but also the majority of all European patents, and this could negatively impact specialization in hydrogen. However, this difference could also suggest that although Germany as a whole represents a hub for hydrogen research, only a few of its regions turn out to be truly specialized in this particular technology.

Overall, however, there is a certain consistency with the previous data regarding the distribution and concentration of H2 patents.

This concludes the description of the database of European patents filed with the EPO.

## Spatial Autocorrelation Analysis

In this section we will try to give an answer to the first research question of this thesis: *is there spatial autocorrelation among European hydrogen specialized regions? Do clusters of specialization in hydrogen exist in Europe?*

To find an answer, global and local spatial autocorrelation analyses were conducted using GeoDa software.

Before proceeding with the study, let us try to make the concept of spatial autocorrelation clearer. By spatial autocorrelation, we refer to the tendency of neighboring observations in a geographical space to have similar values. In other words, if there is spatial autocorrelation in a dataset, it means that the value of a variable at a given location is influenced or correlated with the value of the same variable at surrounding locations. The study of spatial autocorrelation can be done either globally or locally. Global spatial autocorrelation assesses whether a pattern of data is dispersed, clustered or random across the entire perimeter of investigation. Essentially, global analysis provides a single measure that describes the entire data set. Local spatial autocorrelation, on the other hand, indicates where exactly any clusters or outliers are located. In this way, it is possible to identify which specific areas or regions show significant spatial autocorrelation regarding the variable under consideration. Within this thesis, the Moran's I index was used to assess global spatial autocorrelation while the Local Indicator of Spatial Association (LISA) map was used to assess local spatial autocorrelation.

Below, first the global analysis and then the local analysis are presented.

**Global Spatial Autocorrelation Analysis – Moran Scatter Plot**

To measure global spatial autocorrelation, Moran's I is probably the most widely used statistic. This indicator was first introduced by Moran (1948) but has become widely used over the years due to the spatial autocorrelation research of Cliff and Ord (1981).

Moran's I is calculated through the following formula:

$$I = \frac{n}{S_0} * \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x})(x_j - \bar{x})w_{ij}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Where:

- *n* is the total number of observations. In this thesis *n* refers to the number of European regions at the NUTS2 level which are 334;
- $w_{ij}$ is the component of the spatial weights matrix that represents the spatial relationship between the observation *i* and the observation *j*. The spatial weight can be based on various logics, such as contiguity (whether two areas share a boundary) or distance. These are called *Contiguity-Based Spatial Weights* and *Distance-Based Spatial Weights*, respectively;
- $S_0 = \sum_i^n \sum_j^n w_{ij}$ is the sum of all the spatial weights of the matrix;
- $x_i$ refers to the value of the variable of interest (in this thesis the NRTA index) for the observation *i*;
- $x_j$ refers to the value of the variable of interest for the observation *j*;

- $\bar{x}$ is the average value of the variable for all the observations. In this thesis, $\bar{x}$ is the average value of the NRTA index across all the European NUTS2 regions.

With regard to spatial weights, in this thesis, the logic of contiguity was embraced. In particular, among the *Contiguity-Based Spatial Weights*, the so-called *Queen Contiguity* was chosen, which constitutes a very common option that generally works well. The *Queen Contiguity* creates a matrix of spatial weights in which for each element a weight of 1 is given to the neighbouring elements while a weight of 0 is given to the remaining elements. In particular, *Queen Contiguity* counts two elements that share a boundary or vertex as neighbours.

The interpretation of Moran's I statistic is proposed below.

- If I is close to 0, it suggests that there is no spatial autocorrelation between the data;
- If I is positive, it indicates a tendency towards positive spatial autocorrelation, meaning that similar values are clustered in space;
- If I is negative, it indicates a tendency towards negative spatial autocorrelation, meaning that high values are close to low values and vice versa.

It is important to note that Moran's I is a measure of correlation, so its values can vary between -1 and 1, where -1 indicates perfectly dispersed spatial correlation, 1 indicates perfectly clustered spatial correlation, and 0 indicates no spatial autocorrelation.

However, the presence of a value significantly different from zero does not guarantee that spatial autocorrelation is statistically significant. Therefore, statistical tests, such as the pseudo p-value test, are used to determine the significance of the Moran Index obtained.

Moran's I for the NRTA index resulted in a value of 0.268. In light of the comments above, this results would seem to suggest that in Europe exists a moderate positive spatial autocorrelation in the hydrogen specialization. This in practical terms would mean that European regions showing strong specialisation in hydrogen technologies tend to be geographically close to other regions with similar levels of specialisation.

In an attempt to better interpret this result, the Moran scatter plot for the variable of interest (the NRTA index on H2 technology specialisation) is presented in *Figure 26*. Indeed, the Moran scatter plot (Anselin, 2019) is a common tool to graphically assess the global spatial autocorrelation of a variable.

The graph on the x-axis presents the variable of interest (from which the mean has been subtracted) while the y-axis presents the same variable (from which the mean has been subtracted) but lagged or spatially weighted. The lagged value for each spatial unit represents a kind of average of the values of the surrounding spatial units, weighted according to the spatial weights matrix. The Moran scatter plot is centred in the zero and is divided into four quadrants.

1. *Upper left quadrant (high-high)*: shows regions with high values in the variable of interest that are surrounded by regions with also high values in the same variable. In our case, it contains regions with high NRTA values surrounded by regions with high NRTA values.
2. *Lower right quadrant (Low-Low)*: shows regions with low values surrounded by regions also with low values. In our case, regions with low NRTA values surrounded by regions with low NRTA values.
3. *Upper right quadrant (Low-High)*: shows regions with low values that are surrounded by regions with high values. In our case, regions with low NRTA values surrounded by regions with high NRTA values.
4. *Lower left quadrant (High-Low)*: shows regions with high values surrounded by regions with low values. In our case, regions with high NRTA values surrounded by regions with low NRTA values.

The high-high and low-low quadrants correspond consequently to positive spatial autocorrelation while the low-high and high-low quadrants correspond to negative spatial autocorrelation.

This separation of spatial autocorrelation into four categories (high-high, low-low, low-high and high-low) represents a common point between global and local spatial autocorrelation. This discussion continues in the next section.

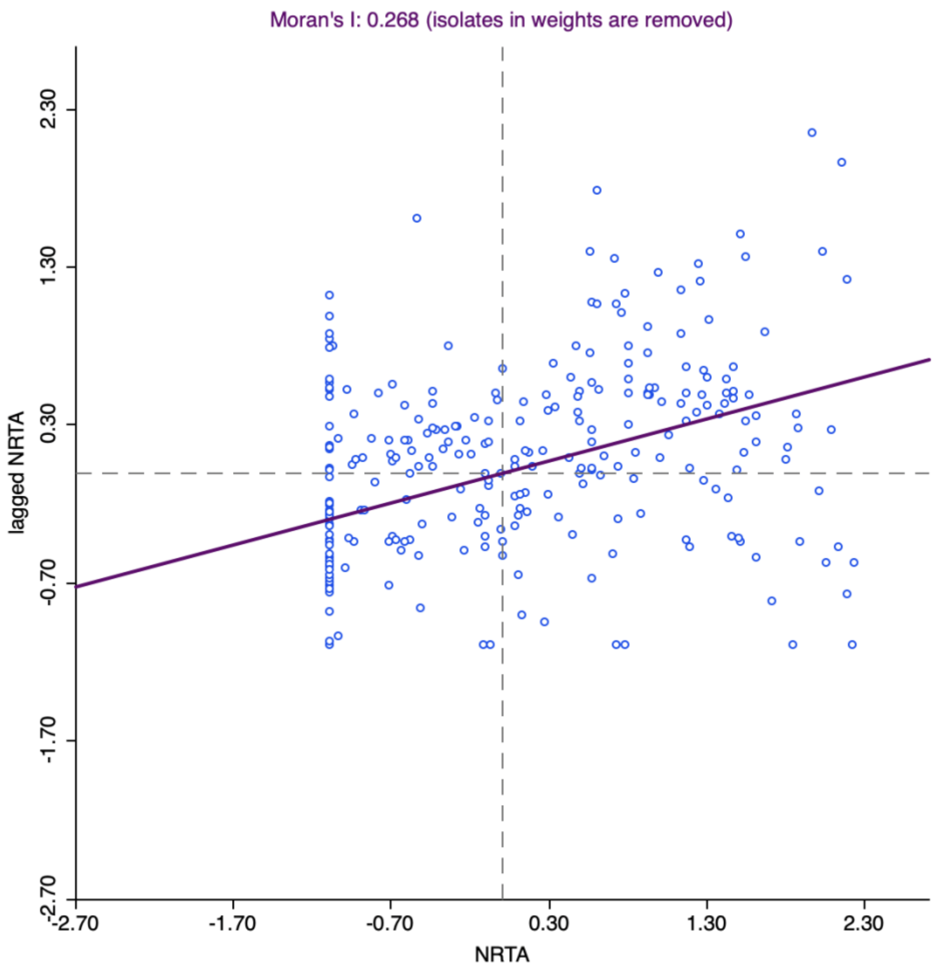



FIGURE 26: MORAN SCATTER PLOT ON NRTA SPECIALISATION INDEX

Returning to the interpretation of the Moran scatter plot, the regression line in the graph shows the relationship between the NRTA index and the lagged NRTA index. The slope of this line represents the Moran index for this variable. *Figure 26* shows a positive inclination of the line with a value of 0.268. Although a large proportion of the data is clustered around the regression line, there is nevertheless a good percentage of observations arranged in a more anomalous manner. From the graph, however, it can be seen that the majority of the data is concentrated in the high-high and low-low quadrants.

This scatter plot therefore suggests that despite some variability in the data, there is a moderate positive autocorrelation across European regions on hydrogen technology specialisation. Furthermore, this scatter plot shows that in Europe, the hydrogen specialised regions are fairly close to each other but also that the non-specialised regions are close to each other.

So far, we have analysed the value of the Moran's I statistic by making various considerations but have not verified its statistical significance. Indeed, the Moran index itself provides a measure of spatial autocorrelation, but it does not say whether this measure is "random" or actually significant.

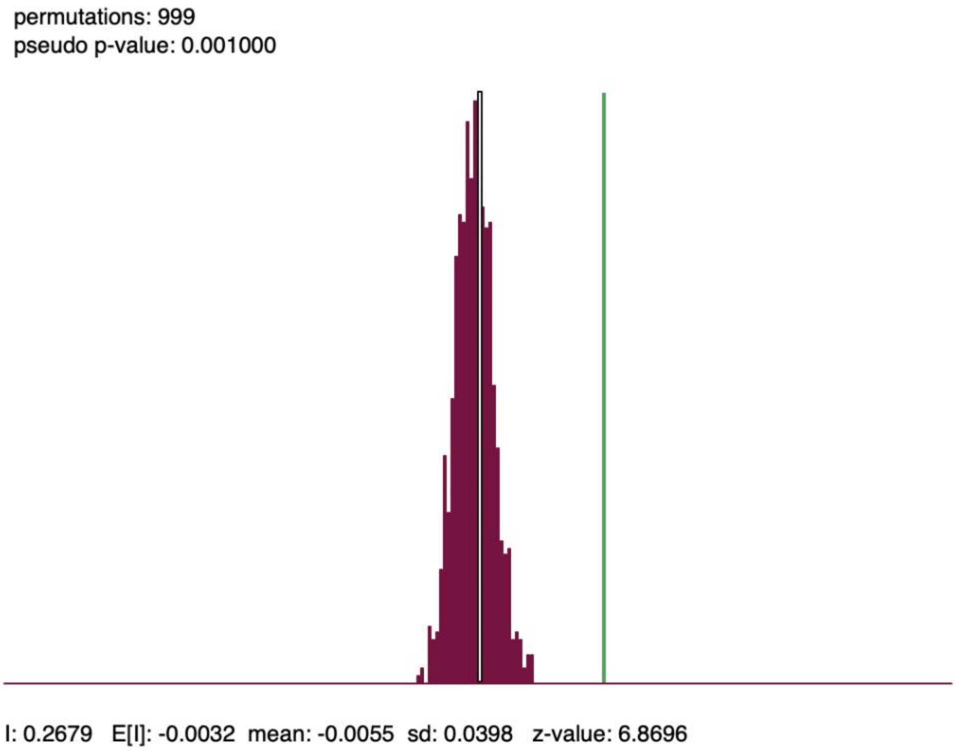



FIGURE 27: EVALUATION OF THE STATISTICAL SIGNIFICANCE OF MORAN'S I ON THE NRTA INDEX

In this regard, a hypothesis test was performed using the GeoDa software. The null hypothesis ($H_0$) states that there is no spatial autocorrelation in the variable of interest (the NRTA index), which means that the spatial distribution of the observed values is random.

The approach used by the software to test this hypothesis is randomization. In particular, 999 permutations were chosen in order to obtain an indication of statistical significance at the 1% level. *Figure 27* presents the outlook of the hypothesis test. *Figure 27* in particular compares the actual Moran index (equal to 0.268) with a distribution of Moran indices based on random configurations obtained with 999 permutations. From the figure we can clearly see how the 'real' Moran index (calculated from actual data), falls deeply outside the distribution of indices obtained through the randomization. Consequently, we can strongly reject the null hypothesis ($H_0$) of spatial randomness and state that Moran's I statistic for the NRTA variable is statistically significant at the 1% level. This result allows us to state that the positive autocorrelation observed in the specialisation in hydrogen technologies within European regions is statistically significant and not random.

This concludes the global spatial autocorrelation analysis of hydrogen specialisation in Europe.

**<u>Local Spatial Autocorrelation Analysis – LISA map</u>**

With the global spatial autocorrelation analysis, we have answered the first part of the research question shown at the beginning of this chapter, demonstrating that there is a moderate positive spatial autocorrelation in hydrogen technology specialization across Europe. However, at the moment, we have not yet identified where are the European clusters specialized in hydrogen. The purpose of this section is precisely to localize these areas.

To measure local spatial autocorrelation, Local Indicators of Spatial Association (LISA) are used. To remain consistent with the statistic used in the global spatial analysis, the Local Moran Index (Anselin, 1995) was used for the local analysis. The Local Moran, in particular, has the following formula for each observation/region i:

$$I_i = \frac{\sum_{j=1}^{n}(x_i - \bar{x})(x_j - \bar{x})w_{ij}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

The terms of the formula are the same as the general Moran's I statistic shown in the previous section.

This index, which evaluates the spatial autocorrelation at the local level, makes it possible to identify possible clusters and spatial outliers. In more detail, based on the local autocorrelation, five different scenarios emerge:

1. The so-called *Hot Spots*, regions with high values in the reference variable surrounded by regions with similarly high values (*high-high spatial clusters*);
2. The so-called *Cold Spots*, regions with low values in the reference variable surrounded by regions with similarly low values (*low-low spatial clusters*);
3. Regions with high values in the reference variable surrounded by regions with low values. These areas are referred as *high-low spatial outliers*;
4. Regions with low values in the reference variable surrounded by regions with high values. These areas are referred as *low-high spatial outliers*;
5. Regions with no statistically significant spatial autocorrelations.

*Figure 27* and *Figure 28* respectively show the Cluster Map and the Significance Map developed by GeoDa on the NRTA hydrogen specialisation variable.

The Cluster Map, in particular, splits, according to the value of the spatial autocorrelation, the 311 European NUTS2 regions into the five categories described above. The significance map, on the other hand, shows for each region the level of statistical significance of the NRTA spatial autocorrelation value. The statistical significance for each region is assessed using a randomization approach similar to the global Moran's I statistic

We should note that the number of regions (the number of observations) in this analysis is equal to 311. However, in Europe, as seen in previous analyses, there are 334 regions. This difference can be traced back to the logic of the spatial weights adopted for the calculation of the autocorrelation statistics. The use of the contiguity approach, in fact, resulted in the exclusion of all those regions that do not have any boundaries in common with other regions. In *Figure 27* and *Figure 28*, for example, islands such as Sicily, Sardinia, Corsica, Cyprus, Crete and the Canary Islands were not represented.

Below *Figure 28* there are comments on its results.

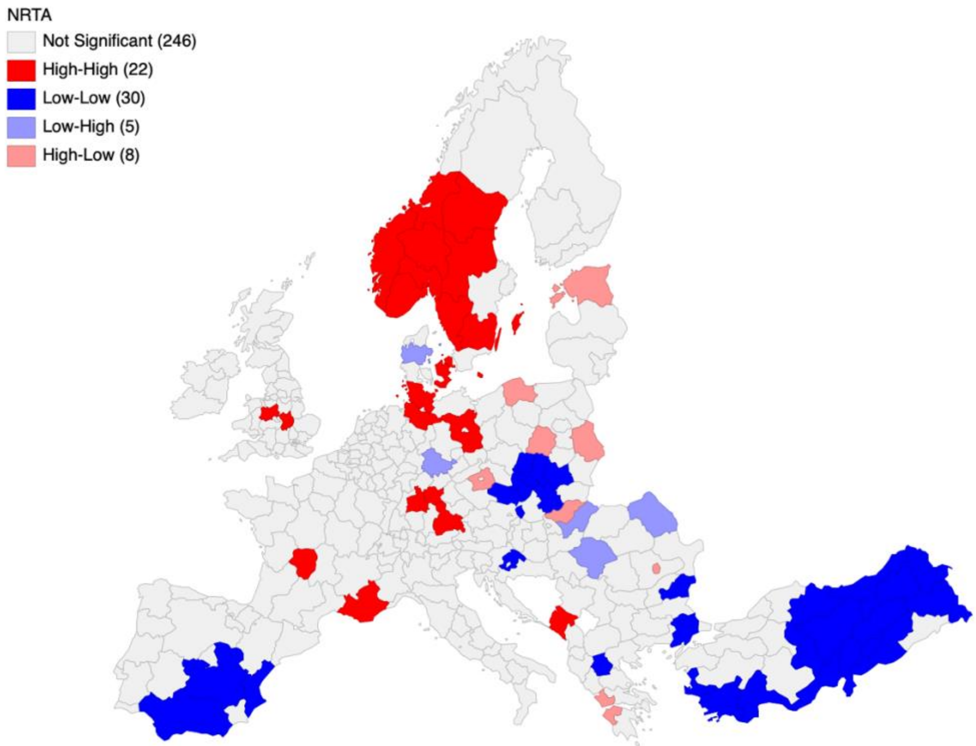



FIGURE 28: LISA CLUSTER MAP ON HYDROGEN TECHNOLOGY SPECIALISATION

The local spatial autocorrelation analysis on hydrogen specialisation revealed that in Europe there are 22 high-high regions (7% of the total), 30 low-low regions (10% of the total), 5 low-high regions (2% of the total) and 8 high-low regions (3% of the total). The remaining 246 regions do not show a clear pattern of spatial autocorrelation (79% of the total). Focusing on the geographical distribution of the different groups of regions, the cluster of high-high regions in Northern Europe immediately stands out. This area with a high and significant technological specialisation within the hydrogen sector spans almost the entire territory of Sweden and Norway and a large part of Denmark (regions of South Denmark, Zealand, Region of the Capital) and Northern Germany (regions of Brandenburg, Hamburg and Schleswig-Holstein). In the southern part of Germany, there is another high-high cluster in hydrogen specialisation consisting of the regions of Oberbayern, Mittelfranken and Stuttgart. The United Kingdom also has two contiguous high-high regions: Shropshire and Staffordshire and Leicestershire, Rutland and Northamptonshire. The remaining high-high regions in Europe are geographically isolated (these regions corresponds to smaller cluster technologically specialized in hydrogen). They include the French regions Provence-Alpes-Côte d'Azur and Limousin and the territory of Montenegro. *Figure 28* also clearly shows very large clusters of low-low regions, namely geographical areas that have no specialisation in hydrogen technologies. These include southern Spain, large parts of Turkey (Anatolia, the southern region on the Mediterranean and the Black Sea region) and in Eastern Europe parts of southern Poland, Slovakia and the Czech Republic.

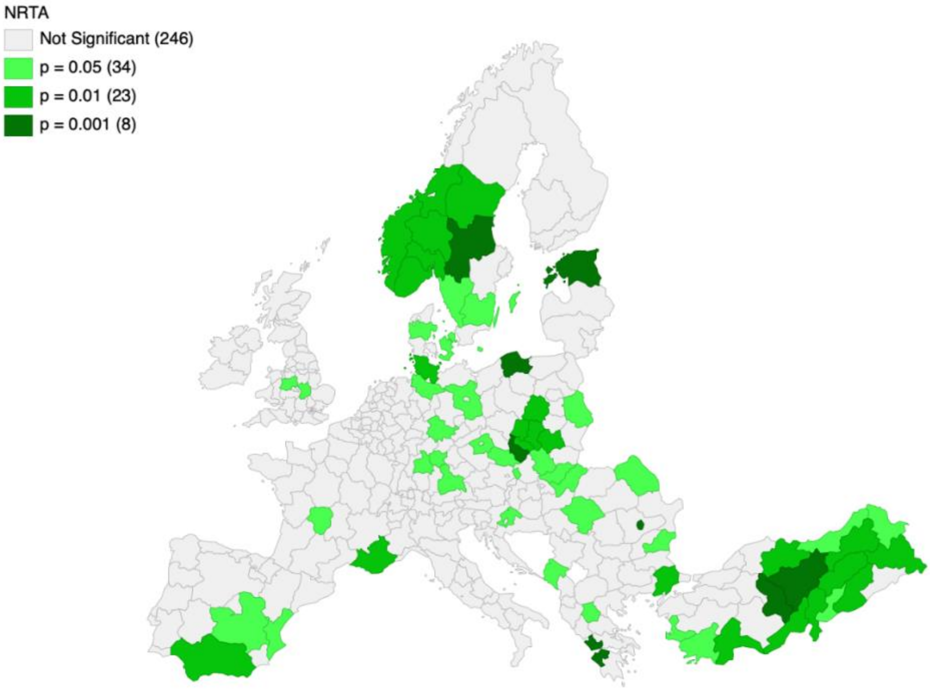



**FIGURE 29: LISA SIGNIFICANCE MAP ON HYDROGEN TECHNOLOGY SPECIALISATION**

*Figure 29* illustrates the levels of statistical significance of spatial autocorrelation in European regions. Specifically, 8 regions are significant at 0.1%, 23 regions are significant at 1% and 34 regions are significant at 10%.

This concludes the section of the analysis of local spatial autocorrelation on hydrogen specialisation and concludes the chapter on the description of the patents' sample.

## Econometric Model

As mentioned, the second research question proposed by this master's thesis is: *What are the characteristics that affect hydrogen technology specialisation in European regions?*

To answer the above question, some econometric models were developed. In particular, as mentioned above, a static time-independent econometric model and a dynamic panel econometric model were developed. Two models are presented for both static and dynamic versions.

The NRTA index was used as the dependent variable in all the models. The NRTA indeed, is a metric employed to measure a region's degree of specialisation in hydrogen.

As independent variables, on the other hand, some of the variables from the *Regional Innovation Scoreboard* (*European Commission, 2021, 2019, 2017, 2016, 2014, 2012, 2009*) were adopted. For the static model, indicators from RIS 2021 were used as independent variables, while for the dynamic model, all versions of RIS from 2009 to 2021 were used. All these variables will be presented below in detail.

In addition, some control variables were also added. Within an econometric model, the control variables help to isolate the "pure" effect of the independent variables on the dependent variable by controlling for additional factors that might have an effect on the same dependent variable.

In order to conduct a more complete and precise analysis, two separate models will be presented both for the static and dynamic models. An OLS regression model and a TOBIT regression model will be presented for the static model. For the dynamic model, on the other hand, a random effects panel model (controlling for year fixed effects and country fixed effects) and a TOBIT panel model will be presented.

OLS (Ordinary Least Squares) is the most commonly used method in statistics to estimate the parameters of a linear regression model. Precisely, The OLS aims to minimize the sum of the squares of the differences (or "residuals") between the observed values and those predicted by the model.

The TOBIT regression model is similar to the OLS but unlike the OLS, it is employed when the dependent variable is continuous but is restricted in an interval. As already shown, the NRTA (the dependent variable of the model) is a statistical index defined in the interval (-1;+1).

The random effects panel model is a statistical method used for analyzing panel data, which consists of multiple observations over time for the same units (e.g., individuals, firms, countries). This model, which differs from the fixed effects model, assumes that individual-specific effects are uncorrelated with the independent variables and allows for variations both across units and over time. It combines features of both time-series and cross-sectional data analysis.

In this thesis, as shown later on in the dedicated chapters, for the static model the baseline will be the TOBIT regression. The OLS model will used as a tool to control for the robustness of the results of the baseline model. For the dynamic model, instead, the baseline will be the random effects regression while the TOBIT will be used as robustness test.

Before describing the models and commenting on their results, the next section will describe the *Regional Innovation Scoreboard*. After the description of this report, the model variables will be described in detail and their statistics will be presented.

## Regional Innovation Scoreboard

Some of the variables defined by the *Regional Innovation Scoreboard* (hereafter RIS) were considered as independent variables in the model. As mentioned above for the static model only the 2021 version of RIS was used while for the dynamic model all published versions of RIS were used (RIS 2021, RIS 2019, RIS 2017, RIS 2016, RIS 2014, RIS 2012, RIS 2009).

The RIS is presented in detail below, for simplicity the most recent version, namely the 2021 version, will be described. In fact, the report overall structure has remained constant although the number of indicators and the definition of some variables have been changed over the years.

The RIS is developed and published every two years by the European Commission, specifically by the Directorate General for Research and Innovation. The RIS is a regional extension of the *European Innovation Scoreboard* (EIS). The EIS provides an assessment of the research and innovation performance of states in the European Union and other European states (Norway, Serbia, Switzerland, and the United Kingdom). In particular, the EIS maps the strengths and weaknesses of each country's innovation ecosystems and identifies the challenges that need to be addressed to ensure successful development.

The RIS extends the analysis dimension of the EIS from the country level to the regional level. In fact, one of the goals of this report is to monitor and promote innovation at the regional level. Precisely, through particular metrics, called within the report "indicators," the 2021 RIS provides a comparative assessment of the innovation systems of 239 European regions. Based on regional availability of data and information, 46 regions are considered at the NUTS1 level and 193 regions are considered at the NUTS2 level. The regions of France and the United Kingdom for example are at the NUTS1 level while the regions of Italy and Germany are at the NUTS2 level. Cyprus, Estonia, Latvia, Luxembourg and Malta, on the other hand, are analyzed at the country level because NUTS1 and NUTS2 coincide with the entire national extent. For the latter states, RIS indicators are not defined and consequently will not be included in the model.

Based on the methodology proposed by the EIS, the RIS calculates the *Regional Innovation Index* (RII) for each region. The RII is a composite metric that measures the state of innovation and innovative activities for a region or country. The RII is calculated by taking the average of the normalized values of the indicators considered in the RIS. Within the RIS, unlike the EIS where 32 indicators are defined, only 21 indicators are used. These indicators measure different dimensions of innovation. To give a few examples, metrics on human resources, the state of attractiveness of the research system, the state of digitalization, private and public investment in innovation, the number of innovators and existing relationships between actors in innovation ecosystems, intellectual assets, employment and environmental sustainability are considered in the report.

Based on the value of RII, RIS classifies European regions into four categories:

- *Innovation Leaders*: includes regions with RII greater than 125 % of the European Union average;
- *Strong Innovators*: includes regions with RII between 100 % and 125% of the European Union average;
- *Moderate Innovators*: includes regions with RII between 70% and 100% of the European Union average;
- *Emerging Innovators*: includes regions with RII less than 70% of the European Union average.

In the RIS of 2021 there are 38 regions among Innovation Leaders (15.8% of the total), 67 regions among Strong Innovators (27.9% of the total), 67 regions among Moderate Innovators (27.9% of the total), and 67 regions among Emerging Innovators (27.9% of the total).

Each category in turn is divided into 3 other subgroups. Thus, in total, there are the 12 subgroups: + Innovation Leaders, Innovation Leaders, - Innovation Leaders, + Strong Innovators, Strong Innovators, - Strong Innovators, + Moderate Innovators, Moderate Innovators, - Moderate Innovators, + Emerging Innovators, Emerging Innovators, - Emerging Innovators.

In RIS 2021, the regions with the highest RII value (compared to the 2014 European Union value) are:

- Stockholm - SE11 (Sweden) - RII = 177.5
- Helsinki-Uusimaa - FI1B (Finland) - RII = 174.2
- Oberbayern - DE21 (Germany) - RII = 173.5
- Hovedstaden - DK01 (Denmark) - RII = 171.1
- Zurich - CH04 (Switzerland) - RII = 168.2

It should be noted that within the RIS database of 2021, for a few regions, data for some indicators are not available. For example, data for Swiss regions are not present in the database for 4 different indicators (Non-R&D innovation expenditures, Innovation expenditures per person employed, Employment in innovative SMEs, Sales of new-to-market and new-to-firm innovations).

However, the lack of data represents a small and not significant portion compared to the whole sample (less than 1%).

Nevertheless, in order to run the regression, it was necessary to estimate the missing data. Unavailable values were estimated using the data mean imputation method. This procedure, involves replacing the missing data for a variable with the average of the available values for that specific variable. Specifically, to obtain a more precise estimate, the value of the missing variable was calculated with the average between the values of the same variable for regions belonging to the same innovation subcategory as the region under consideration. For example, for the Zurich region (CH04), the values of the unavailable indicators were estimated by averaging the values of the same indicators for regions classified as + Innovation Leaders (the same category of the Zurich region).

Finally, before moving on to the description of the variables, it is worth noting that although in the RIS 2021 there are data for 239 European regions, only 229 observations are considered in the model developed in this thesis. The observations reduction is partly due to the elimination of some regions and partly due to adjustments necessary to match at the regional NUTS code level the RIS database and the patent database described in the previous chapter. In the 239 European regions analyzed by RIS, please note that the countries of Cyprus (CY00), Estonia (EE00), Latvia (LV00), Luxembourg (LU00) and Malta (MT00) are not counted because the data for calculating the indicators were not collected.

Below are the changes made to match the two databases:

- Deletion of Serbia NUTS codes (RS11, RS12, RS21, RS22) and of Spanish region of Ceuta NUTS code (ES63) because there is no patent data in REGPAT;
- For the Ireland NUTS codes IE05 and IE06, the indicators values were averaged and only one NUTS named IE02 was created for consistency with the patent database;
- For the Croatia NUTS codes HR02, HR05 and HR06, the indicator values were averaged and only one NUTS named HR04 was created for consistency with the patent database;
- For the Lithuania NUTS codes LT01 and LT02, the indicator values were averaged and only one NUTS named LT00 was created for consistency with the patent database;
- For the Hungary NUTS codes HU11 and HU12, the indicator values were averaged and only one NUTS named HU10 was created for consistency with the patent database;
- For the Poland NUTS codes PL91 and PL92, the indicator values were averaged and only one NUTS named PL12 was created for consistency with the patent database

The previous changes therefore reduced the total observations from 239 to 229.

As mentioned several times before, some of the metrics used to measure the regions' innovation activity, the so-called "indicators," were selected as independent variables of the econometric model. Although the RIS defines 21 indicators, only eight variables were selected within the model. The reasons behind this choice are several. On the one hand, since the RIS variables are highly correlated with each other, in an attempt to comply with commonly accepted constraints in econometric models, only those indicators with lower correlation were included. In fact, in econometrics, the common norm is to build models with a correlation between variables of less than 30/40% (this aspect will be taken up later in the analysis of the correlation matrix). On the other hand, however, having a small analysis sample (only 229 observations), the degrees of freedom of the model appear significantly limited. For this reason, a restricted number of variables had to be selected to meet the model's degrees of freedom.

Furthermore, some variables were not found to be significant in explaining the variability of the dependent variable. Among these for example, metrics on the state of regions digitization are not considered because no clear relationship is seen regarding the regional specialization in hydrogen. Finally other variables, instead, were not considered to avoid the reverse causality or endogeneity problem. Reverse causality occurs when there is a bidirectional relationship between an independent variable and one or more dependent variable. This means that when there is reverse causality, the effect of the independent variable on the dependent variable is affected by the effect of the dependent variable on the independent variable. To avoid this problem, variables on intellectual assets such as the number of PCT patents and trademarks were not considered.

The model's variables, the formula by which they are calculated and the reasons why they are considered are described in detail below.

## Model Variables

This section describes the models variables.

The dependent variable of the econometric models is the NRTA index. As seen, the NRTA index measures the degree of hydrogen specialization of each region and can take values between -1 and +1. If the value of NRTA for a region is positive then the region is specialized in the relevant technology perimeter namely hydrogen.

All independent variables used in the models are described below. As indicated in the title of each variable, some were used only in the dynamic models or in the static models while others in both. These choices were made after doing the correlation analyses between the variables. In fact, if the correlations were too high, the variables were excluded. In addition, the reason for including different variables in the models is that since the RIS 2021 version, some indicators have been changed by the European Commission. As a matter of fact, in the static model, the variables of RIS 2021 were used while in the dynamic models, all versions of RIS from 2009 to 2021 were used. In the first version of the RIS (the one from 2009), the indicators from 2004 onwards are present.

- **R&D expenditures per GDP** (used only in static models)

  This variable was calculated by making the average of the following RIS variables: *R&D expenditures in the public sector as percentage of GDP* and *R&D expenditures in the business sector as percentage of GDP*.

  The variable is calculated by doing the ratio of the sum of all R&D expenditures both by the public sector (by the government and university system) and by the private sector (business expenditures) to regional gross domestic product.

  This variable on the one hand captures the creation of new knowledge within firms. In the business sector, the R&D voice is very significant for all those industries where knowledge is created in or near research laboratories. Among them for example there are the pharmaceutical industry, the chemical industry and the electronics industry. On the other hand, public sector R&D investment is one of the main drivers of economic growth in the industrialized regions. Trends in R&D provide a very significant indication of a region's future economic well-being and competitive advantage.

Globally, R&D investment is critical in fostering a transition to a knowledge-based economy and promoting the improvement of a region's technology ecosystem.

- **Public Sector R&D Spending Relative to Regional GDP** (Used only in dynamic models)
  If in the static models, the investments in R&D both in the public and business sector are condensed into the same variable by taking the average, in dynamic models the two components are kept separate. In particular, the variable related to the R&D expenditures in the public sector is calculated by taking the total value of the R&D investments within both the government sector (often referred to as *GOVERD*) and the higher education sector (known as *HERD*). This cumulative value is then divided by the Regional Gross Domestic Product to determine the proportion of the GDP that is invested in R&D activities within the public domain. The essence of this metric lies in its reflection of a region's commitment to fostering a knowledge-driven economy. R&D investments are pivotal indicators of a region's potential for future economic competitiveness and prosperity. By channeling funds into research and development, regions not only pave the way for advancements in production technologies but also catalyze economic growth.
- **Business Sector's R&D Investment Relative to Regional GDP** (Used only in dynamic models)
  This metric is derived by taking into account the entirety of R&D expenditures within the business domain, commonly referred to as BERD. This total is then juxtaposed against the Regional Gross Domestic Product to ascertain the fraction of the GDP dedicated to R&D endeavors within the corporate landscape.

  The significance of this indicator is rooted in its ability to gauge the intensity of knowledge generation within corporate entities. Especially in sectors that are heavily reliant on scientific advancements, the creation of new knowledge predominantly occurs in proximity to R&D hubs. This suggests that a higher percentage indicates a region's proactive approach to fostering innovation, staying at the forefront of technological advancements, and ensuring a competitive edge in the global market. Moreover, a robust investment in R&D by businesses can be a testament to a region's conducive environment for research, which can attract further investments and skilled talent.

- **SMEs' non-R&D innovation expenditures relative to total revenue** (used in both static and dynamic models)

  This indicator quantifies the ratio of small and medium-sized enterprise (SME) non-research and development (non-R&D) innovation expenditures to total revenue. It is calculated by summing all innovation expenditures for SMEs, excluding both internal and external R&D expenditures, and dividing this sum by the total turnover of SMEs. This indicator is significant due to its ability to capture diverse components of innovation expenditures, such as investments in apparatus and machinery and the purchase of patents and licenses.. These elements reflect the dissemination of new production technologies and innovative ideas. In other words, this indicator considers other forms of innovation investment beyond R&D, that can be crucial for the growth and competitiveness of SMEs. This can be particularly relevant in sectors where innovation is not strictly tied to R&D, but can also stem from improvements in production processes, work organization, or marketing strategies.

- **Sales of Innovative Products in SMEs Relative to Total Turnover** (used in both static and dynamic models)

  This variable is calculated by doing the ratio between the combined revenue generated from newly introduced or substantially enhanced products within SMEs and the SMEs' overall turnover. This ratio, expressed as a percentage, represents the proportion of an SME's revenue that is derived from innovative products. These innovative products can be categorized into two main types: those that are pioneering and novel to the market, representing cutting-edge technologies, and those that, while perhaps already existing in the market, are new additions to the company's product portfolio. In essence, this metric not only highlights an SME's capacity to lead with groundbreaking technologies but also its adaptability in assimilating and leveraging existing innovative solutions for growth and competitiveness. The indicator is built on the SMEs but It can be easily interpreted as an estimate for the all firms.

- **SMEs Launching Innovative Products as a Proportion of All SMEs** (Used only in static models)

  In order to determine this metric, the RIS evaluated the number of SMEs located in a particular region that have launched at least one innovative product (a product that is either entirely new or has undergone significant improvements in terms of its features, user experience or components) and compares it to the total count of SMEs within the same region. As a result this metric is presented as a percentage. The essence of this indicator lies in its ability to measure the innovative strength within the business sector on a regional basis. Introducing product innovations is pivotal for companies, as such advancements can carve out new market niches and bolster their competitive advantage. A higher percentage of SMEs engaged in product innovation signifies a vibrant and proactive innovation landscape, indicating that a substantial portion of these enterprises are actively pushing the boundaries and contributing to market dynamism.

- **SMEs presenting Products or Process Innovation as a Proportion of All SMEs** (Used only in dynamic models)

  Compared to the static model where only the variable that takes into account the number of product innovations is considered, in the dynamic model the variable that also takes into account process innovations is employed. This is due to the fact that in the prior versions of RIS 2021, there is only one variable that takes into account both product and process innovations, whereas in RIS 2021, there are two separate variables. The numerator of this ratio is given by the number of SMEs that have successfully launched a novel product or implemented a new process within their operational markets. This figure is then compared to the total count of SMEs to determine the percentage of enterprises actively engaged in technological innovation. The significance of this metric is its ability to gauge the technological innovation vigor within the SME sector. The introduction of new products or the adoption of advanced processes is a cornerstone of innovation, especially in the manufacturing domain. Therefore this variable is indicative of a region's or sector's commitment to staying ahead of the curve, ensuring that SMEs, often considered the backbone of many economies, remain competitive, adaptive, and at the forefront of technological advancements.

- **In-house Innovation Activities Among SMEs as a Proportion of Total SMEs** (used only in dynamic models)

  This variable takes into account the number of SMEs that have undertaken in-house innovation endeavors. Specifically, these are enterprises that have either internally introduced a new product or process or have done so in collaboration with other firms, excluding innovations developed externally by other entities. This count is then juxtaposed against the comprehensive number of SMEs to ascertain the percentage of SMEs actively innovating within their own confines. This indicator highlights the self-reliance and internal innovation capabilities of SMEs. By introducing new or substantially improved products or processes, these SMEs demonstrate their intrinsic capacity to innovate without solely depending on external sources or collaborations. This is particularly noteworthy in the context of SMEs, as larger corporations typically have the resources and infrastructure to innovate consistently. A higher percentage in this metric underscores the entrepreneurial spirit and self-sufficiency of SMEs in a given region or sector.

- **SMEs collaborating with others as percentage of all SMEs** (used only in dynamic models) This metric depends on the number of SMEs actively engaged in collaborative innovation endeavors. Specifically, these are enterprises that have established any form of cooperative agreements centered around innovation activities with other businesses or institutions. This count is then set against the overall number of SMEs to derive the percentage of SMEs participating in such collaborative efforts. The importance of this indicator lies in its ability to shed light on the extent of collaborative innovation within the companies landscape. In the realm of intricate innovations, a firm's capacity to tap into a diverse array of information sources and knowledge pools, or to jointly spearhead the innovation development process, becomes paramount. This metric, therefore, offers insights into the knowledge exchange dynamics, be it between academic research entities and businesses or inter-business collaborations. It's noteworthy that the focus is primarily on SMEs, as the vast majority of larger corporations are invariably engaged in such cooperative innovation activities. This emphasis underscores the importance of fostering a collaborative spirit among smaller enterprises, which can be pivotal for regional innovation ecosystems and for ensuring that SMEs remain competitive and adaptive in rapidly evolving markets.
- **Percentage of population with tertiary education** (used in both static and dynamic models) The indicator is obtained through the analysis of the number of individuals within a given age range (the range of age varies between the different versions of the RIS, in the RIS of 2021 for example the range investigated is between 25 and 34 years old) who have attained some form of tertiary education, such as a bachelor's degree, master's degree, doctoral degree or other post-secondary qualifications. This figure is then divided by the total population within the same age bracket, resulting in a percentage. The significance of this indicator is its ability to provide insights into the availability of advanced skills within a specific age group. While it encompasses a broad spectrum of educational fields beyond just science and technology, its importance is underscored by the understanding that innovation

adoption across various sectors, including services, requires diverse skill sets. In the 2021 version, by focusing on the 25 to 34 age group, this metric offers a timely reflection of the impact of educational policies and shifts, showcasing the evolving landscape of tertiary education attainment.

- **Percentage of Employment in Knowledge-Intensive Activities** (used in both static and dynamic models)

  This proportion is defined by counting the number of individuals employed in sectors deemed as knowledge-intensive. These sectors, are identified following the EU Labour Force Survey guidelines. The above count of individuals employed in knowledge-intensive activities is then compared to the total employment figure, yielding a percentage. The essence of this indicator is its capacity to gauge the concentration of knowledge-based roles within the broader employment landscape. The importance of this variable lies in its reflection of the modern economy's shift towards knowledge-based roles. Knowledge-intensive activities not only offer direct services to consumers, but they also play a pivotal role in fueling innovation across various sectors by providing essential ideas. By emphasizing roles that require advanced education and specialized skills, this metric underscores the evolving nature of the workforce and the increasing demand for expertise in today's dynamic economic landscape.

- **Particulate Matter Emissions in the Manufacturing Sector** (used in both static and dynamic models)

  This metric evaluates the emissions of fine particulate matter (PM2.5) within the manufacturing sector, quantified in tonnes. These emissions are then contextualized against the economic output of the sector, specifically the value added in chain-linked volumes (with 2010 as the reference year) measured in million euros. The significance of this indicator stems from the profound health and environmental implications of air pollution. While pollution can be both man-made and natural, certain pollutants, notably PM2.5, nitrogen dioxide, and ground-level ozone, are particularly hazardous to human health. Prolonged and peak exposure to these pollutants can lead to severe health conditions, including cardiovascular and respiratory problems, and even increase the risk of cancer. The focus on PM2.5 is especially pertinent given its minute size, allowing it to penetrate deep into the respiratory system.

  This variable is used in both static and dynamic models. However, within the dynamic models, the variable was transformed into a Boolean variable with a value of 1 for polluted regions and a value of 0 for unpolluted regions. In particular, the threshold for determining whether a region is or is not polluted was set at the value of 0.5. Consequently, regions with values greater than or equal to 0.5 are considered polluted, while those with values less than 0.5 are considered unpolluted. This variable has only been defined since the RIS of 2021 and is consequently constant over the years in the dynamic models.

As described within the report of the RIS, the values of the published indicators (namely the values used in the models proposed in this thesis) were modified to some extent and were normalized. Specifically, for each indicator, observations –the outliers- with a value higher (lower) than the mean value of all regions plus two times the standard deviation (minus two times the standard deviation) were replaced with the maximums (the minimums) observed in all regions. In addition, all indicators with skewness degree greater than 1 were modified with a square root transformation to lower the skewness below the threshold value of 1. Finally, the data for each indicator were normalized through the minimum-maximum method. As a result, as also shown by the following matrix of descriptive statistics, all RIS variables vary in the range [0, 1] and take the value of 1 at the highest value among all observations and 0 at the lowest value among all observations.

The use within an econometric model of normalized variables is a constrain because this reduces the variability of the data and consequently the statistical analysis on the dependent variable will be worse. This aspect will be taken up later in the chapter *Limitations and Possible Future Developments*.

Along with the previous variables, control variables have also been considered to develop more reliable models. Control variables are always used in econometric models to account for all those effects that might influence the relationship between the dependent variable and the independent variables of interest. Thus, control variables are factors that are included in the model to control for the effect of external variables that could influence both the dependent variable and the independent variables and that if excluded could generate biased estimates. In other words, control variables serve to isolate the pure effect of the independent variables on the regressed variable. In the literature, this aspect is called Omitted Variable Bias.

The data regarding these variables were downloaded from the Eurostat databases, precisely from the *General and regional statistics* section.

The control variables are presented below.

- **Population in millions of people** (used only in static models)
  The reference year considered is 2018. This choice is due to the greater availability on a regional basis of the population data in the Eurostat databases in 2018. Moreover, the 2018 was selected as the baseline year to be coherent with the data used for the subsequent control variable (the GDP per capita). The unit of measure in millions of people was selected to have data with an order of magnitude comparable to the RIS indicators presented earlier. This variable allows to control for size effects on the dependent variable and the independent variables.
- **Gross Domestic Product (GDP) per capita in thousands of EURO** (used both in static and dynamic models)
  GDP per capita was chosen as the second control variable. The unit of measurement (thousands of EURO), as in the previous control variable, was chosen to maintain an order of magnitude comparable with the values of the other variables in the model. As with the population, the reference year selected is 2018. It should be noted that the reference year of 2018 was set only for the static model, while for the dynamic model, data from 2004 to 2021 were considered. For each region, GDP per capita was calculated by doing the ratio of GDP to population. This variable is included in the model to control for the effects of wealth and economic well-being on the model variables.
- **Population (in thousands) per square kilometer** (used only in dynamic models)
  In the dynamic models, to control for size effects on the dependent variable and the independent variables, the population density is considered instead of population by region. Using population density as a control variable versus population in absolute terms allows a more precise measure of the concentration of people in a specific area. Also for these data, the source used was Eurostat databases, and in particular data from 2004 to 2021 were used. The unit of measurement considered for population density is thousands of people per square kilometer.

Only in dynamic models have additional controls been added for so-called *year fixed effects* and *country fixed effects*.

**Year fixed effects** control for events, shocks, or trends that occur in a particular year and affect all regions equally. For example, an economic crisis or a new policy that has uniform impacts in all regions.

**Country fixed effects**, on the other hand, control for time-invariant characteristics that are specific to each country. For example, cultural, institutional, or historical aspects that affect regions in a particular country differently than regions in another country.

Country fixed effects, being constant over time, could also have been considered in the static model. However, given the limited number of observations and limited degrees of freedom, only population and GDP per capita were considered in the static model as controls.

### Static Econometric Models

Static models are presented in this section. The temporal dimension is not taken into account in these models. The indices of technological specialization in hydrogen, were calculated in aggregate and not for each year. In fact, the patents were summed without considering the year of application. For the independent variables, on the other hand, as mentioned, only the indicators from RIS 2021 were used. Below, the descriptive statistics, the correlation matrix, the regression equation and the results of TOBIT (used as baseline model) and OLS (used as robustness check) models are presented.

#### Variables Descriptive Statistics

*Table 10* below shows some descriptive statistics of the variables considered in the econometric model. Precisely, the statistics reported are: the number of observations, mean value, median, standard deviation (SD), minimum value, and maximum value.

Note that the hydrogen specialization variable (which will not be included in the subsequent regressions) is a Boolean variable constructed from the NRTA or RTA index values. Hydrogen specialization is equal to 0 if the region is not hydrogen specialized (when the NRTA value is negative) and is equal to 1 if the region is hydrogen specialized (when the NRTA value is positive).

| Variable | Count | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|
| 1 - Hydrogen specialization | 229 | 0.341 | 0.000 | 0.475 | 0.000 | 1.000 |
| 2 - NRTA | 229 | -0.308 | -0.412 | 0.640 | -1.000 | 0.995 |
| 3 - R&D expenditures per GDP | 229 | 0.383 | 0.336 | 0.228 | 0.021 | 1.000 |
| 4 - Firms' non-R&D innovation expenditures relative to total revenue | 229 | 0.419 | 0.399 | 0.164 | 0.000 | 1.000 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5 - Sales of innovative products in firms relative to total revenue | 229 | 0.588 | 0.589 | 0.176 | 0.115 | 1.000 |
| 6 - % of firms who introduced a product innovation | 229 | 0.588 | 0.614 | 0.251 | 0.025 | 1.000 |
| 7 - % of population with tertiary education | 229 | 0.536 | 0.540 | 0.242 | 0.042 | 1.000 |
| 8 - % of employment in knowledge-intensive activities | 229 | 0.532 | 0.540 | 0.243 | 0.000 | 1.000 |
| 9 - PM2.5 emissions in the manufacturing sector | 229 | 0.529 | 0.570 | 0.219 | 0.000 | 0.969 |
| 10 - population (millions of people) | 229 | 2.275 | 1.630 | 1.986 | 0.029 | 12.213 |
| 11 - GDP per capita (thousands of €) | 229 | 30.490 | 29.019 | 17.073 | 5.209 | 96.094 |

TABLE 12: DESCRIPTIVE STATISTICS STATIC MODEL

The dependent variable, the regional specialization index NRTA varies in the range of -1 and +1. NRTA Positive values indicates a specialization whereas negative values indicates a non-specialization. Because the NRTA median takes a negative value (as well as the mean), the number of non-specialized NUTS regions in hydrogen is greater than the number of specialized NUTS regions. This result is consistent with what is shown in the database description section (see *Geographical Analysis on Technological Specialization in Hydrogen*, *Figure 25*).

As mentioned above, all the variables selected from the RIS take values between 0 and 1. This is due to the fact that these variables were normalized by the min-max method.

The two control variables (population and GDP per capita) have an order of magnitude greater than one unit compared to the other variables and even a greater variability of the data.

**Correlation Matrix**

As mentioned in the RIS description, in econometric models it is usually common to include variables that are weakly correlated with each other (often orthogonalized variables are used i.e., completely uncorrelated with each other) to avoid multicollinearity problems. Multicollinearity occurs when two or more independent variables in a regression model are highly correlated with each other. This can make it difficult, to accurately estimate the partial effects of each independent variable on the dependent variable. In addition, it can lead to unstable estimates and very wide confidence intervals.

Within this thesis, given the high correlation between the RIS variables, it was decided to keep 40 percent as the threshold limit for correlation. In some cases, as notable in the following table, the correlation between the variables is greater than 40%.

The imposition of a maximum correlation threshold led to the exclusion of some RIS variables that could be relevant in explaining how the hydrogen specialization of European regions is affected. For example, the following indicators were excluded: the number of scientific publications among the 10 percent most cited, the percentage of firms introducing process innovations, the percentage of innovative firms collaborating with each other, the percentage of employment in innovative firms, and the volume of innovation expenditures per employee in innovative firms.

The correlation matrix of the variables included in the model is shown below in *Table 11*. Unlike the matrix of descriptive statistics, the Boolean variable on hydrogen specialization was not considered.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** - NRTA | 1.000 | | | | | | | | | |
| **2** - R&D expenditures per GDP | 0.333 | 1.000 | | | | | | | | |
| **3** - Firms' non-R&D innovation expenditures relative to total revenue | 0.042 | 0.144 | 1.000 | | | | | | | |
| **4** - Sales of innovative products in firms relative to total revenue | 0.032 | 0.245 | 0.442 | 1.000 | | | | | | |
| **5** - % of firms who introduced a product innovation | 0.287 | 0.589 | 0.359 | 0.529 | 1.000 | | | | | |
| **6** - % of population with tertiary education | 0.046 | 0.422 | -0.040 | 0.052 | 0.097 | 1.000 | | | | |
| **7** - % of employment in knowledge-intensive activities | 0.212 | 0.571 | 0.013 | 0.081 | 0.304 | 0.282 | 1.000 | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **8** - PM2.5 emissions in the manufacturing sector | 0.222 | 0.361 | 0.134 | 0.242 | 0.379 | 0.246 | -0.008 | 1.000 | | |
| **9** - population (millions of people) | 0.080 | 0.137 | 0.029 | -0.056 | -0.0004 | 0.230 | 0.243 | -0.027 | 1.000 | |
| **10** - GDP per capita (thousands of €) | 0.324 | 0.698 | 0.063 | 0.258 | 0.538 | 0.479 | 0.499 | 0.524 | 0.100 | 1.000 |

TABLE 13: CORRELATION MATRIX STATIC MODEL

The control variable GDP per capita appears to be highly correlated with several variables. The highest correlation is with R&D expenditures (69.8%). This level of correlation is not surprising because it is obvious that richer regions spend more money on R&D projects. Similarly, the most economically developed regions are also usually the regions that offer the most innovative products (53.8% correlation with variable 5) and where there is the highest degree of post-secondary education due to the high concentration of universities and other professional educational institutes (47.9% correlation with variable 6). The positive correlation between GDP per capita and PM2.5 emissions is also not surprising. In fact, a wealthy region very often is a highly industrialized region and consequently turns out to be a polluted region.

The other variable, besides GDP per capita, that is highly correlated with the remaining variables is R&D expenditure per GDP. Not surprisingly, there is a high correlation with the percentage of companies that have introduced product innovation (58.9 percent). In fact, the higher is the R&D expenditure in a region, the higher will be the concentration of firms in the same region that have introduced at least one product innovation. The high correlation of R&D expenditures with the degree of population with tertiary education (42.2 percent) and the percentage of employment in knowledge-intensive activities (57.1 percent) is quite obvious.

Finally apart from a few cases of high levels of correlation (variable 3 with variable 4 and variable 4 with variable 5), the remaining variables are weakly correlated.

For highly correlated variables, as seen in the subsequent model results, regressions were implemented both considering these variables and not considering them. When having highly correlated variables in fact in the econometric models it is usual to show how the estimates vary when these variables are included and when they are not included.

**Regression formula**

Before delving into the output of the econometric model on the study of regional specialization in hydrogen, the regression analysis formula is shown below. As mentioned above, the model used as a baseline in this thesis is a TOBIT model. An OLS model will also be presented as a robustness analysis. For both TOBIT and OLS, batteries of regressions were implemented by inserting a new variable each time to study how at each iteration the significance, the sign and the magnitude of each variable varied.

$$y_i = \alpha_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \beta_3 * x_{3i} + \beta_4 * x_{4i} + \beta_5 * x_{5i} + \beta_6 * x_{6i} + \beta_7 * x_{7i} +$$
$$+ \beta_8 * x_{8i} + \beta_9 * x_{9i} + \varepsilon_i$$

Where:

- $y_i$ = NRTA index for region i;
- $\alpha_0$ is the general intercept of the model;
- $x1_i$ = R&D expenditures over GDP for region i;
- $x2_i$ = SMEs' non-R&D innovation expenditures relative to total revenue for region i;
- $x3_i$ = Sales of Innovative Products in SMEs relative to total turnover for region i;
- $x4_i$ = SMEs launching Innovative Products as a proportion of all SMEs for region i;
- $x5_i$ = Percentage of population with tertiary education for region i;
- $x6_i$ = Percentage of employment in knowledge-intensive activities for region i;
- $x7_i$ = Particulate matter emissions in the manufacturing sector for region i;
- $x8_i$ = Population in millions of people for region i;
- $x9_i$ = GDP per capita in thousands of EURO for region i;
- $\varepsilon_i$ is the error term.

**TOBIT Model**

As stated several times before, the TOBIT model is considered the baseline model in this thesis for the static econometric models. TOBIT is an econometric model similar to OLS that is used when the dependent variable varies within a limited range. The dependent variable is the NRTA index that takes values within the range [-1;+1].

Specifically, the TOBIT model shown in the *Table 14* proposes 8 different models.

For each variable, in each model, the coefficient value and in parentheses the standard error of the estimate are given.

The stars indicate statistical significance. Respectively: * indicates 10% significance level (10% < p-value < 5%), ** indicates 5% significance level (5% < p-value < 1%) and *** indicates 1% significance level (p-value<1%).

All variables related to innovation expenditures (R&D and non-R&D) and related to innovative products (% of revenues from innovative products and % of firms with innovative products) were included in Model (1). In fact, this set of variables appears to be quite similar and therefore was considered in bulk.

From model (2) to model (6), on the other hand, all remaining variables were included one at a time. These indicators measure different characteristics and so they were added one by one. Model (6) contains all the variables of interest. In Model (7) and (8), on the other hand, the % of population with tertiary education and Sales of innovative products in SMEs relative to total revenue were eliminated, respectively.

Statistical significance, sign and magnitude are analyzed below for each variable. Sign and magnitude will be commented on only for the statistically significant variables.

Model (6) is taken as the reference model, because, as mentioned contains all independent variables and controls.

R&D expenditure per GDP is statistically significant in all models except model (7) where the tertiary education variable was excluded. In model (6), the variable is statistically significant at the 10% level.

Notably, there is a partial positive correlation between R&D expenditures per GDP and hydrogen specialization. As regional R&D expenditure per GDP increases by one unit, regional specialization in hydrogen increases by 0.648 (model (6)).

A positive correlation suggests that regions that invest more in research and development relative to their GDP tend to have higher hydrogen specialization. This makes sense, as greater investment in R&D could translate into a greater ability to develop and adopt new technologies, such as those related to hydrogen.

Firms' non-R&D innovation expenditures relative to total revenue is not statistically significant in any of the proposed models.

Sales of innovative products in SMEs relative to total turnover is statistically significant in all models. In model (6), it is statistically significant at 5% level. Surprisingly, there is a partial negative correlation with the dependent variable. As sales of innovative products in SMEs relative to total turnover increases by one unit, specialization in hydrogen decreases by 0.754. This result might suggest that firms with a higher proportion of their revenues derived from innovative products might be more diversified in their innovation activities. Consequently, this could mean that while they may have a strong presence in other innovative sectors, they may not be as specialized in hydrogen. Another possible explanation behind this negative correlation could lie in the presence of particular competitive dynamics or barriers to entry in the hydrogen sector in some regions. Indeed, SMEs that are highly innovative in other sectors might find the hydrogen industry too competitive or dominated by large companies making it complex for them to specialize in this sector. Moreover, regulations may exist in some European regions that make it more difficult for SMEs to enter the hydrogen market.

In Model (8), where this variable was excluded, no major changes were reported except the loss of statistical significance of the percentage of firms that introduced a product innovation.

<u>The percentage of companies that have introduced product innovations</u> is statistically significant in all models (except model (8)), and in model (6) it is statistically significant at 10% level. There is a partial positive correlation with specialization in hydrogen. As the value of the percentage of firms that introduced product innovations increases by one unit, specialization in hydrogen increases by 0.547. This impact is not surprising because if more companies in a region introduce product innovations, this could indicate a dynamic and innovative business environment. This environment could foster specialization in cutting-edge fields such as hydrogen.

<u>The degree of population having tertiary education</u> becomes statistically significant in Model (6) (at 5% level) when GDP per capita is introduced. The significance in this case seems to be due to the high correlation with GDP per capita. In fact, as seen in the correlation matrix, the two variables have a correlation level of 47.9%. Because of this reason, Model (7) was also presented where the variable on tertiary education was removed.

Despite the previous clarification, returning to the analysis of model (6), there is surprisingly a partial negative correlation with the dependent variable. Specifically as the value of the percentage of the population having tertiary education increases by one unit, hydrogen specialization decreases by 0.574.

One possible interpretation of this result could be that while tertiary education is generally associated with higher innovation capacity, it could also be that regions with a high percentage of population with tertiary education have more diversified economies and, therefore, less specialization in a single sector such as hydrogen.

The percentage of employment in knowledge-intensive activities is not statistically significant except in model (5) where it is weakly significant (at 10% level).

PM2.5 emissions in the manufacturing sector, on the other hand, is always statistically significant and in model (6) it is significant at 5%. There is a partial positive correlation with the dependent variable. As PM2.5 emissions from the manufacturing sector increase by one unit, specialization in hydrogen technologies increases by 0.671. A positive correlation here could suggest that regions with higher emission intensity in manufacturing are actively seeking cleaner and more sustainable solutions, such as hydrogen technologies. However, it should also be noted that this variable could be partly endogenous and thus suffer from the so-called reverse causality problem. In fact, it could also be argued that regions are specialized in hydrogen because they are industrialized and consequently polluted.

Coming to the control variables, population appears to be weakly significant (10% level) except in model (7). There is a partial positive correlation with the regressed variable. If population increases by one unit then hydrogen specialization increases by 0.054.

GDP per capita is also found to be, albeit weakly, statistically significant (except in model (7)). Here again there is a partial positive correlation with the dependent variable. When GDP per capita increases by one unit, hydrogen specialization increases by 0.09.

| Tobit: NRTA | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| R&D expenditures per GDP | 1.076*** | 1.168*** | 1.012*** | 0.784** | 0.808** | 0.648* | 0.475 | 0.715* |
| | (0.291) | (0.324) | (0.359) | (0.364) | (0.363) | (0.372) | (0.367) | (0.374) |
| Firms' non-R&D innovation expenditures relative to total revenue | 0.056 | 0.032 | 0.067 | 0.083 | 0.008 | 0.076 | 0.131 | -0.159 |
| | (0.371) | (0.372) | (0.374) | (0.371) | (0.372) | (0.372) | (0.376) | (0.354) |
| Sales of innovative products in firms relative to total revenue | -0.752* | -0.731* | -0.729* | -0.783** | -0.737* | -0.754** | -0.812** | |
| | (0.383) | (0.383) | (0.384) | (0.382) | (0.382) | (0.379) | (0.384) | |
| % of firms who introduced a product innovation | 0.919*** | 0.876*** | 0.870*** | 0.654** | 0.677** | 0.547* | 0.753** | 0.319 |
| | (0.309) | (0.315) | (0.315) | (0.321) | (0.321) | (0.327) | (0.316) | (0.308) |
| % of population with tertiary education | | -0.160 | -0.176 | -0.334 | -0.429 | -0.574** | | -0.609** |
| | | (0.251) | (0.251) | (0.255) | (0.260) | (0.271) | | (0.272) |
| % of employed persons in knowledge-intensive activities | | | 0.279 | 0.545* | 0.454 | 0.287 | 0.298 | 0.283 |
| | | | (0.278) | (0.292) | (0.296) | (0.309) | (0.313) | (0.311) |
| PM2.5 emissions in the manufacturing sector | | | | 0.865*** | 0.890*** | 0.671** | 0.629* | 0.655** |
| | | | | (0.297) | (0.298) | (0.321) | (0.322) | (0.322) |
| population (millions of people) | | | | | 0.051* | 0.054* | 0.043 | 0.058** |
| | | | | | (0.027) | (0.027) | (0.027) | (0.027) |
| GDP per capita (thousands of €) | | | | | | 0.009* | 0.006 | 0.009* |
| | | | | | | (0.005) | (0.005) | (0.005) |
| Constant | -1.000*** | -0.925*** | -1.019*** | -1.301*** | -1.352*** | -1.231*** | -1.438*** | -1.437*** |
| | (0.213) | (0.242) | (0.260) | (0.279) | (0.280) | (0.285) | (0.274) | (0.270) |
| Observations | 229 | 229 | 229 | 229 | 229 | 229 | 229 | 229 |
| Log-likelihood | -241.6 | -241.4 | -240.9 | -236.5 | -234.8 | -233.3 | -235.5 | -235.2 |
| Chi-squared | 49.2 | 49.7 | 50.7 | 59.3 | 62.8 | 65.9 | 61.4 | 61.9 |

**TABLE 14: TOBIT MODEL (STATIC) ON THE REGIONAL HYDROGEN SPECIALISATION**

**OLS Model**

To verify the robustness and reliability of the TOBIT model and the previously described results, it was decided to present an OLS model with the same variables.

As with TOBIT, several batteries of models were implemented. In particular, 7 different models are presented in *Table 15*. The sequence of entering the regressors and control variables is the same as that of the TOBIT model. The only difference from the previous analysis is that for the OLS, the model, without the variable "sales of innovative products in SMEs relative to total turnover", was not implemented.

Again for the OLS, coefficients and standard error are given in the table for each variable. The latter is indicated in parentheses. Stars indicate the level of statistical significance of each variable and are interpreted as in the previous case.

Below, the statistical significance, sign and magnitude for each variable are commented on (the last two only for significant variables). As with TOBIT, the model taken as a reference is model (6) which contains all variables.

R&D expenditures per GDP is always statistically significant except in model (7) where the variable on the level of tertiary education is excluded. Specifically, in model (6) R&D expenditures per GDP is statistically significant at the 10% level. As in TOBIT, a partial positive correlation with the regressed variable is observed. As R&D expenditures increases by one unit, hydrogen specialization increases by 0.503.

These results confirm what was observed in the TOBIT model.

Firms' non-R&D innovation expenditures relative to total revenue and sales of innovative products in SMEs relative to total revenue are not statistically significant in any model. The result of the second variable (sales of innovative products in SMEs relative to total turnover) differs from what was previously observed in TOBIT where the indicator was always significant.

Percentage of firms that introduced a product innovation remains statistically significant until the control variable of GDP per capita is added (model (6)). However, in model (7) when tertiary education is excluded, the variable returns to 5% statistically significance. When the variable is significant, a partial positive correlation with hydrogen specialization always emerges.

The results for this indicator, excluding model (6), are consistent with what was seen in TOBIT.

The percentage of population with tertiary education, just as in TOBIT, assumes statistical significance only when GDP per capita is considered (significant at 5 %). Even in the OLS, there is a partial negative correlation with the dependent variable. As the percentage of the population with tertiary education increases by one unit, the specialization in hydrogen technologies decreases by 0.429. This result probably turns out to be influenced by the correlation with GDP per capita.

The percentage of employment in knowledge intensive activities is never found to be statistically significant. Again, this result aligns with what was seen in TOBIT.

PM2.5 emission in the manufacturing sector is statistically significant (at 5 percent) only in models (4) and (5). When GDP per capita is considered, this variable loses significance. This result is different from the TOBIT result.

The population control variable is never statistically significant.

GDP per capita remains statistically significant only in model (6) although weakly (10%).

Overall, the results are consistent with what was shown in the TOBIT model and thus ensure its robustness and reliability.

The only significant deviation lies in the total loss of significance of sales of innovative products in SMEs relative to total turnover. In TOBIT this variable showed unexpected and surprising results. Consequently, the previous estimate, even in light of the results provided by the OLS, appears even more uncertain.

| OLS: NRTA | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| R&D expenditures per GDP | 0.661*** | 0.798*** | 0.743*** | 0.621** | 0.625** | 0.503* | 0.365 |
| | (0.213) | (0.238) | (0.260) | (0.262) | (0.264) | (0.258) | (0.252) |
| innovation expenditures (not R&D) per turnover | -0.045 | -0.067 | -0.060 | -0.052 | -0.074 | -0.013 | 0.004 |
| | (0.299) | (0.300) | (0.300) | (0.303) | (0.304) | (0.305) | (0.308) |
| sum of total turnover of new products over sum of total turnover | -0.508 | -0.484 | -0.479 | -0.508 | -0.489 | -0.505 | -0.540 |
| | (0.329) | (0.330) | (0.334) | (0.330) | (0.334) | (0.325) | (0.331) |
| % of firms who introduced a product innovation | 0.575** | 0.520** | 0.519** | 0.427* | 0.434* | 0.348 | 0.484** |
| | (0.223) | (0.234) | (0.235) | (0.227) | (0.228) | (0.238) | (0.223) |
| % population with tertiary education | | -0.231 | -0.235 | -0.302 | -0.333* | -0.429** | |
| | | (0.188) | (0.189) | (0.195) | (0.199) | (0.213) | |
| % of employed persons in knowledge-intensive activities | | | 0.091 | 0.208 | 0.174 | 0.062 | 0.081 |
| | | | (0.205) | (0.211) | (0.214) | (0.230) | (0.229) |
| PM2.5 emissions in manufacturing over value added in manufacturing sector | | | | 0.418** | 0.425** | 0.265 | 0.248 |
| | | | | (0.200) | (0.201) | (0.213) | (0.216) |
| population (millions of people) | | | | | 0.019 | 0.021 | 0.012 |
| | | | | | (0.017) | (0.017) | (0.017) |
| GDP Pro Capite (thousands of €) | | | | | | 0.007* | 0.004 |
| | | | | | | (0.004) | (0.003) |
| Constant | -0.582*** | -0.483** | -0.514** | -0.647*** | -0.666*** | -0.592** | -0.739*** |
| | (0.176) | (0.201) | (0.212) | (0.219) | (0.221) | (0.231) | (0.213) |
| Observations | 229 | 229 | 229 | 229 | 229 | 229 | 229 |
| Adjusted R-squared | 0.123 | 0.125 | 0.122 | 0.133 | 0.133 | 0.139 | 0.125 |

**TABLE 15: OLS MODEL (STATIC) ON THE REGIONAL HYDROGEN SPECIALISATION (NRTA)**

## Dynamic Econometric Models

The models described above are static econometric models and as such do not allow causal inference between the variables investigated. Static models are observational models that only allow the detection of partial correlations between variables.

Models of this type may suffer from endogeneity and reverse causality problems. Furthermore, static models, by neglecting the temporal evolution of variables, are unable to capture any transformations that have occurred over time. To partially solve this problem, as already mentioned, dynamic econometric models have also been developed considering the evolution of the technological specialization in hydrogen over time.  Through dynamic models, as well as moving in the direction of causal interpretation among variables and observing transformations that have occurred over the years, the number of observations and degrees of freedom of the models are increased significantly. In the static models, as can be seen from *Table 14* and *Table 15* there are only 229 observations while in the dynamic models as seen in *Table 18* and *Table 19* there are 3,777 observations. The increase in degrees of freedom in the dynamic models also allowed us to control for year fixed effects and country fixed effects and consequently to obtain more sophisticated estimates.

To develop the database of dynamic models, the technological specialization variable in hydrogen, the NRTA index, was calculated year by year. Patents were grouped by each region, as in the construction of the static model database, but were separated by application year. The RTA index was calculated by taking the ratio of hydrogen patent shares to the total patent shares year by year. Consequently, for each European region, the RTA and NRTA indicators vary from year to year.

Regarding the independent variables, all reports of the *Regional Innovation Scoreboard* (RIS 2021, RIS 2019, RIS 2017, RIS 2016, RIS 2014, RIS 2012, and RIS 2009) were used. It should be noted that in the 2009 RIS, the indicators are measured starting from 2004.

For this reason, in the dynamic models, patents with Application Year from 2004 to 2021 are considered.

Since the RIS variables are defined every two years and not year by year (the report is indeed published biennially by the European Commission), the RIS variables of a particular year were associated with both patents with the same application year and patents with the subsequent application year. In other words, the 2019 RIS variables, for example, were associated with both patents filed in 2019 and those filed in 2020.

This one-year time lag of the independent variables compared to the technological specialization variable in hydrogen allows for addressing the issue of reverse causality and endogeneity.

Also for the panel model, as for the static model, two models are presented: a random-effects model used as a baseline and a TOBIT model used as a robustness test for the results.

Below, the descriptive statistics of the variables, the correlation matrix, the regression formula and the results of the random-effects model and TOBIT are presented.

**Variables Descriptive Statistics**

*Table 16* shows the descriptive statistics of the dynamic model.

| Variable | Count | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|
| 1 - NRTA index | 3,777 | -0.654 | -1.000 | 0.651 | -1.000 | 1.000 |
| 2 - R&D expenditures (public sector) over GDP | 3,777 | 0.411 | 0.410 | 0.201 | 0.000 | 1.000 |
| 3 - R&D expenditures (business sector) over GDP | 3,777 | 0.421 | 0.410 | 0.214 | 0.000 | 1.000 |
| 4 - innovation expenditures (not R&D) per turnover | 3,777 | 0.429 | 0.415 | 0.160 | 0.000 | 1.000 |
| 5 -firms' sum of total turnover of new/innovative products over sum of total turnover | 3,777 | 0.454 | 0.460 | 0.176 | 0.000 | 1.000 |
| 6 - % of firms who introduced a product or process innovation | 3,777 | 0.477 | 0.483 | 0.227 | 0.000 | 1.000 |
| 7 - % of firms with in-house innovation activities | 3,777 | 0.449 | 0.465 | 0.208 | 0.000 | 1.000 |
| 8 - % of firms with innovation cooperation activity | 3,777 | 0.380 | 0.371 | 0.207 | 0.000 | 1.000 |
| 9 - % population with tertiary education | 3,777 | 0.487 | 0.490 | 0.191 | 0.000 | 1.000 |
| 10 - % of employed persons in knowledge-intensive activities | 3,777 | 0.498 | 0.490 | 0.195 | 0.000 | 1.000 |
| 11 - Pollution dummy variable | 3,777 | 0.678 | 1.000 | 0.467 | 0.000 | 1.000 |
| 12 - Population density (thousands of people per km2) | 3777 | 0.298 | 0.122 | 0.709 | 0.003 | 7.527 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 13 - GDP per capita (thousands of €) | 3777 | 27.641 | 27.300 | 15.473 | 1.900 | 92.700 |

TABELLA 16: DESCRIPTIVE STATISTICS DYNAMIC MODEL

*Table 16* shows the descriptive statistics of the dynamic model. With regard to descriptive statistics, what has been said and commented on in the static model also applies here (see *Static Econometric Models - Variables Descriptive Statistics*). All variables in the RIS vary between 0 and 1 while the dependent variable, the technology specialisation index NRTA varies between -1 and +1. Just as in the static model, also in the dynamic model the mean and median value of the NRTA are negative, indicating that most regions in Europe are not specialised in hydrogen.

## Correlation Matrix

The correlation matrix is shown below in *Table 17*. The same considerations made in the static model also apply here (see *Static Econometric Models - Correlation Matrix*).

| **Variables** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** - NRTA index | 1.00 | | | | | | | | | | | | |
| **2** - R&D expenditures (public sector) over GDP | 0.22 | 1.00 | | | | | | | | | | | |
| **3** - R&D expenditures (business sector) per GDP | 0.35 | 0.51 | 1.00 | | | | | | | | | | |
| **4** - innovation expenditures (not R&D) per turnover | 0.00 | 0.15 | 0.18 | 1.00 | | | | | | | | | |
| **5** – firms' sum of total turnover of new/innovative products over sum of total turnover | 0.07 | 0.28 | 0.31 | 0.31 | 1.00 | | | | | | | | |
| **6** - % of firms who introduced a product or process innovation | 0.25 | 0.43 | 0.57 | 0.21 | 0.48 | 1.00 | | | | | | | |
| **7** - % of firms with in-house innovation activities | 0.28 | 0.43 | 0.58 | 0.22 | 0.40 | 0.89 | 1.00 | | | | | | |
| **8** - % of firms with innovation cooperation activity | 0.19 | 0.39 | 0.50 | 0.21 | 0.33 | 0.55 | 0.56 | 1.00 | | | | | |
| **9** - % population with tertiary education | 0.17 | 0.33 | 0.37 | -0.16 | 0.08 | 0.15 | 0.15 | 0.33 | 1.00 | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **10** - % of employed persons in knowledge-intensive activities | 0.26 | 0.22 | 0.59 | -0.02 | 0.19 | 0.44 | 0.40 | 0.25 | 0.31 | 1.00 | | | |
| **11** - Pollution dummy variable | 0.21 | 0.30 | 0.42 | -0.05 | 0.25 | 0.46 | 0.47 | 0.36 | 0.33 | 0.16 | 1.00 | | |
| **12** - Population density (thousands of people per km2) | 0.11 | 0.19 | 0.08 | -0.06 | 0.12 | 0.14 | 0.14 | 0.16 | 0.28 | 0.23 | 0.01 | 1.00 | |
| **13** - GDP per capita (thousands of €) | 0.30 | 0.45 | 0.58 | -0.02 | 0.30 | 0.56 | 0.55 | 0.40 | 0.54 | 0.49 | 0.56 | 0.33 | 1.00 |

Tabella 17: Correlation Matrix dynamic model

The RIS variables are highly correlated with each other. Variables with high levels of correlation with many other variables have been eliminated and not considered. Among the model variables, as can be seen in *Table 17*, the percentage of firms who introduced a product or process innovation and the percentage of firms with in-house innovation activities are highly correlated with each other (0.89 correlation). In light of this result, in the regressions presented below, these two variables are never entered simultaneously. When one is entered, the other is left out.

**Regression formula**

As with the static model below is the regression formula for the variant panel models over time. The baseline is a random effects model and therefore it was also possible to include variables that do not vary over time such as the dummy variable on pollution and control for country fixed effects. As a robustness test, TOBIT is used, which in a panel model can only be done random effects. Thus, even for TOBIT, variables that are constant over the years can be considered.

$$y_{it} = \alpha_0 + \beta_1 * x_{1it} + \beta_2 * x_{2it} + \beta_3 * x_{3it} + \beta_4 * x_{4it} + \beta_5 * x_{5it} + \beta_6 * x_{6it} + \beta_7 * x_{7it} +$$
$$+ \beta_8 * x_{8it} + \beta_9 * x_{9it} + \beta_{10} * x_{10it} + \beta_{11} * x_{11it} + \gamma * Z_i + \lambda_t + \mu_i + \varepsilon_{it}$$

Where:

- $y_{it}$ = NRTA index for region i in year t;
- $\alpha_0$ is the general intercept of the model;
- $x1_{it}$ = R&D expenditures in the public sector over GDP for region i in year t;
- $x2_{it}$ = R&D expenditures in the business sector over GDP for region i in year t;
- $x3_{it}$ = SMEs' non-R&D innovation expenditures relative to total revenue for region i in year t;

- $x4_i$ = Sum of total turnover of new or innovative products over sum of total turnover in SMEs for region i in year t;
- $x5_{it}$ = SMEs presenting Products or Process Innovation as a Proportion of All SMEs for region i in year t;
- $x6_{it}$ = In-house Innovation Activities Among SMEs as a Proportion of Total SMEs for region i in year t;
- $x7_{it}$ = SMEs collaborating with others as percentage of all SMEs for region i in year t;
- $x8_{it}$ = Percentage of population with tertiary education for region i in year t;
- $x9_{it}$ = Percentage of employment in knowledge-intensive activities for region i in year t;
- $x10_{it}$ = Population density: population (in thousands) per square kilometer for region i in year t;
- $x11_{it}$ = GDP per capita in thousands of EURO for region i in year t;
- $Z_i$ = Pollution (dummy variable) by region i , constant over time;
- $\lambda_t$ = represents year fixed effects (a series of dummy variables for each year, except one which is omitted as a reference category);
- $\mu_i$ = represents country fixed effects (a set of dummy variables for each country, except one which is omitted as a reference category);
- $\varepsilon_{it}$ is the error term.

**Random Effects Model**

The panel model used as a baseline is a random effects model. The random effects model was preferred to the fixed effects model after the *Hausman test* was conducted (Hausman, 1978). The *Hausman test* is a statistical method for determining whether a fixed-effects model or a random-effects model is more appropriate for an analysis on panel data. The test resulted in a non-significant p-value of 0.9 and thus did not allow for a rejection of the null hypothesis. For this reason, the random-effects model was chosen. For the sake of completeness, however, the results of the fixed-effects model are shown in the *Annex* in *Table 20*, although not commented on. Please note that a fixed-effects model does not allow for variables that remain constant over time. For this reason, the Boolean variable on particulate emissions and the country fixed effects are not included in the fixed effects regression.

In the random effects model, and also in the model used for robustness analysis (the TOBIT model), several sets of regressions were developed, as in the static models. The variables were inserted in blocks. Initially, only variables related to monetary investments in innovation (R&D investments and other expenses for developing innovative activities) were considered. Subsequently, all variables related to different types of innovations (product/process innovation, in-house innovation, and cooperative innovation) were included. Later on, variables related to the education of the workforce (tertiary education and employment in knowledge-intensive activities) were added. The control variables GDP per capita and population density, and the boolean variable on pollution, were considered in all models. Finally, for each model two different versions are presented: one version where fixed effects for year and country are controlled for by inserting dummy variables, and a version where these effects are not controlled for.

In *Table 18*, the results of the random effects panel model are reported. The same considerations made for the previous models also apply to this model. Below, each variable is discussed in terms of statistical significance, sign, and magnitude. In particular, sign and magnitude are commented only for statistically significant variables.

<u>R&D expenditures in the public sector over the GDP</u> are not statistically significant in any model.

In contrast, <u>R&D expenditures in the business sector over the GDP</u> are statistically significant at the 1% level in all 10 models. In particular, there is a partial positive correlation with the dependent variable. As R&D spending in the business sector per GDP increases by 1€, the indicator on technological specialization on hydrogen increases by 0.299. The coefficient for model (8), which is considered the reference model because it contains almost all variables, was commented here. This result shows that R&D investment by firms at the regional scale tends to increase regional technology specialization in hydrogen. Even in the static model, as seen previously, there was a significant and positive partial correlation between general expenditures (both those in the public and business sectors) in R&D per GDP and specialization on hydrogen technologies. Considering what this model shows, the component of R&D investment that influences the development of the hydrogen innovation ecosystem is that in the business sector rather than that in the public sector.

<u>Firms' non-R&D innovation expenditures relative to firms' turnover</u> is statistically significant (at 5% level) only in model (1). There is a surprising partial negative correlation in model (1). However, by adding the remaining variables the significance of this variable is lost and consequently this negative correlation is insignificant.

<u>Sum of total turnover of new/innovative products over the sum of total turnover</u> is statistically significant in all models. The level of statistical significance varies depending on whether the dummy

variables year and country are included to control for year and country fixed effects. Without the dummy variables, the variable is statistically significant at 1% while with the dummy variables the variable is statistically significant at 10%. Just as in the static model, there is a partial negative correlation with the dependent variable. Considering model (8), as turnover from innovative products in the market increases by €1 while keeping total turnover unchanged, the index of technological specialization on hydrogen decreases by 0.15. Here again, the same considerations made in the commentary on the static model apply. Furthermore, based on this result, one could comment that companies with a large and varied portfolio of innovative products already on the market are less inclined to invest large amounts of resources in new technologies such as hydrogen. In other words, these companies, having already brought their innovations to market and thus having previously invested in research and development could maximize their economic return by investing limited amounts of money. This means that such companies follow a model of incremental innovation that is less resource-intensive than a model of radical innovation that seems to be more suitable for the development of a still immature ecosystem such as the hydrogen ecosystem.

The percentage of companies that introduced product or process innovations is not statistically significant in any model. In the static model, on the other hand, where we considered only the percentage of firms that had introduced product innovations, this variable was statistically significant (albeit weakly) in all models. It should be noted that this variable is included only when the percentage of firms with in-house innovation activities is excluded. As mentioned, this choice has to be attributed to the high level of correlation between these two variables.

The percentage of firms with in-house innovation activities is statistically significant in all models in which it is included. When there are year and country dummies, the variable is significant at 5% while when there are not, the variable is significant at 1%. There is a partial positive correlation with the dependent variable. If the number of firms with in-house innovative activities increases by one unit in a region with the same number of total firms, the hydrogen technology specialization index increases by 0.204 (coefficient of model (8)).

The percentage of companies with innovation cooperation activity is not statistically significant in any model.

In light of the results of the last two variables, it is more important to develop in-house innovation activities rather than a collaborative open innovation model to increase technology specialization in hydrogen. These findings thus suggest that the hydrogen technology supply chain is based on the ability of firms to develop in-house expertise rather than on the use of external expertise.

The percentage of population with tertiary education is not statistically significant in any model.

The percentage of people employed in knowledge-intensive activities is statistically significant in all models in which it is included. Specifically, in models (7), (8) and (10) it is significant at 5% while in model (9) it is significant at 1%. A partial positive correlation with the dependent variable is observed. In a region, with the same number of people classified as labor force, as the number of people employed in knowledge intensive activities increases by one unit, the index of technological specialization on hydrogen in the same region increases by 0.230 (coefficient of model (8)). Consequently, as the number of resources employed in knowledge-intensive activities increases, the technological specialization in hydrogen also increases. In fact, the development of the hydrogen supply chain, given its complexity, requires the employment of massive skilled and specialized human resources.

PM2.5 emissions are statistically significant only in models in which country fixed effects and year fixed effects are not considered. Statistical significance is also rather weak: at 1% in models (1), (3) and (7) and 5% in model (9). When the variable is significant, a positive partial correlation with the dependent variable is observed. The same considerations made in the static model apply to this variable.

Population density is statistically significant only in models (3), (5) and (9) when year and country fixed effects are not considered. In all these models, population density is statistically significant at the 10% level. A partial positive correlation with the dependent variable is also observed in this case.

GDP per capita is statistically significant in all models. From model (1) to model (6), GDP per capita is statistically significant at 1%. In the remaining models, when the variables related to the level of education of the population are included, GDP per capita is statistically significant at 5%. There is a partial positive correlation between GDP per capita and the dependent variable. As GDP per capita in a region increases by €1,000, the index of technological specialization in hydrogen in that region increases by 0.005. This result is also not surprising since, net of all other variables, the richest regions are also the most technologically advanced in the hydrogen domain.

| Random Effects - NRTA | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| R&D expenditures in the public sector per GDP | 0.003 | -0.005 | 0.038 | -0.008 | 0.028 | -0.013 | 0.068 | -0.041 | 0.083 | -0.041 |
| | (0.076) | (0.089) | (0.076) | (0.089) | (0.076) | (0.089) | (0.078) | (0.092) | (0.078) | (0.092) |
| R&D expenditures in the business sector per GDP | 0.641*** | 0.447*** | 0.647*** | 0.457*** | 0.618*** | 0.435*** | 0.526*** | 0.299*** | 0.550*** | 0.314*** |
| | (0.084) | (0.099) | (0.086) | (0.102) | (0.086) | (0.102) | (0.094) | (0.111) | (0.094) | (0.111) |
| Innovation expenditures (not R&D) per turnover | -0.160** | -0.131 | -0.078 | -0.102 | -0.105 | -0.128 | -0.071 | -0.105 | -0.047 | -0.078 |
| | (0.067) | (0.081) | (0.071) | (0.086) | (0.071) | (0.086) | (0.073) | (0.087) | (0.073) | (0.086) |
| Firms' sum of total turnover of new/innovative products over sum of total turnover | | | -0.251*** | -0.134* | -0.243*** | -0.141* | -0.244*** | -0.150* | -0.247*** | -0.143* |
| | | | (0.071) | (0.080) | (0.068) | (0.079) | (0.069) | (0.079) | (0.071) | (0.080) |
| % of firms who introduced a product or process innovation | | | 0.099 | 0.024 | | | | | 0.065 | 0.014 |
| | | | (0.075) | (0.090) | | | | | (0.077) | (0.090) |
| % of firms with in-house innovation activities | | | | | 0.305*** | 0.224** | 0.286*** | 0.204** | | |
| | | | | | (0.080) | (0.094) | (0.082) | (0.095) | | |
| % of firms with innovation cooperation activity | | | -0.063 | 0.040 | -0.110 | -0.007 | -0.114 | -0.013 | -0.062 | 0.031 |
| | | | (0.069) | (0.084) | (0.069) | (0.084) | (0.070) | (0.084) | (0.071) | (0.084) |
| % population with tertiary education | | | | | | | -0.020 | 0.182 | -0.060 | 0.192 |
| | | | | | | | (0.087) | (0.117) | (0.087) | (0.117) |
| % of employed persons in knowledge-intensive activities | | | | | | | 0.200** | 0.230** | 0.222*** | 0.232** |
| | | | | | | | (0.085) | (0.103) | (0.086) | (0.103) |
| PM2.5 Emissions (dummy) | 0.075* | -0.057 | 0.086* | -0.054 | 0.069 | -0.046 | 0.080* | -0.040 | 0.099** | -0.046 |
| | (0.044) | (0.062) | (0.045) | (0.063) | (0.045) | (0.063) | (0.046) | (0.063) | (0.046) | (0.063) |
| Population density (thousands of people per km2) | 0.035 | -0.004 | 0.042* | -0.001 | 0.042* | 0.003 | 0.040 | -0.006 | 0.042* | -0.011 |
| | (0.025) | (0.029) | (0.025) | (0.029) | (0.025) | (0.029) | (0.025) | (0.029) | (0.025) | (0.029) |
| GDP per capita (thousands of €) | 0.005*** | 0.008*** | 0.005*** | 0.008*** | 0.005*** | 0.007*** | 0.003** | 0.005** | 0.004** | 0.006** |
| | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.002) | (0.002) | (0.003) | (0.002) | (0.003) |
| Application Year dummies | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Country dummies | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 |

**TABLE 18: RANDOM EFFECTS MODEL ON THE REGIONAL HYDROGEN SPECIALISATION (NRTA)**

**TOBIT Model**

The TOBIT time-variant model, as mentioned above, is used as a robustness test for the previously illustrated model. As can be seen from *Table 19*, the results are consistent with the random-effects panel model.

The R&D expenditures in the business sector, the percentage of people employed in knowledge-intensive activities, and the GDP per capita are statistically significant and positively correlated with the NRTA index on hydrogen even in this model.

Compared with the previous model, however, in the econometric models where controls for year fixed effects and country fixed effects are also considered, the statistical significance of the following variables is lost: sum of total turnover of new/innovative products over the sum of total turnover and percentage of firms with in-house innovation activities.

For the remaining variables, results are similar to one of the random effects model.

Although in correspondence with some models, the level of statistical significance of some variables may change, it is important to note that the sign of the partial correlation between the significant variables and the dependent variable is always maintained.

Consequently, the robustness of the panel random effects model results commented earlier can be confirmed.

| TOBIT - NRTA | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| R&D expenditures in the public sector per GDP | -0.061 | 0.247 | 0.056 | 0.232 | 0.021 | 0.225 | 0.336 | 0.144 | 0.384 | 0.144 |
| | (0.299) | (0.333) | (0.299) | (0.333) | (0.297) | (0.333) | (0.307) | (0.343) | (0.309) | (0.343) |
| R&D expenditures in the business sector per GDP | 2.755*** | 2.218*** | 2.814*** | 2.205*** | 2.708*** | 2.155*** | 2.163*** | 1.474*** | 2.254*** | 1.510*** |
| | (0.359) | (0.391) | (0.365) | (0.400) | (0.361) | (0.398) | (0.385) | (0.440) | (0.387) | (0.441) |
| innovation expenditures (not R&D) per turnover | -0.753*** | -0.719** | -0.415 | -0.636* | -0.515* | -0.682* | -0.354 | -0.568 | -0.275 | -0.521 |
| | (0.288) | (0.364) | (0.303) | (0.379) | (0.303) | (0.380) | (0.308) | (0.383) | (0.309) | (0.382) |
| firms' sum of total turnover of new/innovative products over sum of total turnover | | | -0.922*** | -0.387 | -0.817*** | -0.380 | -0.905*** | -0.500 | -0.955*** | -0.493 |
| | | | (0.286) | (0.323) | (0.276) | (0.316) | (0.280) | (0.319) | (0.289) | (0.326) |
| % of firms who introduced a product or process innovation | | | 0.447 | 0.079 | | | | | 0.219 | 0.002 |
| | | | (0.288) | (0.332) | | | | | (0.293) | (0.334) |
| % of firms with in-house innovation activities | | | | | 1.155*** | 0.502 | 0.988*** | 0.403 | | |
| | | | | | (0.306) | (0.342) | (0.311) | (0.349) | | |
| % of firms with innovation cooperation activity | | | -0.363 | 0.327 | -0.490* | 0.233 | -0.453* | 0.213 | -0.306 | 0.302 |
| | | | (0.262) | (0.311) | (0.258) | (0.313) | (0.263) | (0.314) | (0.267) | (0.311) |
| % population with tertiary education | | | | | | | -0.369 | 0.788 | -0.534 | 0.790 |
| | | | | | | | (0.368) | (0.483) | (0.370) | (0.483) |
| % of employed persons in knowledge-intensive activities | | | | | | | 1.381*** | 1.228*** | 1.485*** | 1.242*** |
| | | | | | | | (0.371) | (0.445) | (0.378) | (0.446) |
| PM2.5 Emissions (dummy) | 0.720*** | -0.160 | 0.761*** | -0.146 | 0.684*** | -0.125 | 0.758*** | -0.073 | 0.831*** | -0.094 |
| | (0.205) | (0.238) | (0.207) | (0.238) | (0.203) | (0.238) | (0.205) | (0.239) | (0.208) | (0.239) |
| Population density (thousands of people per km2) | 0.195* | -0.001 | 0.228** | 0.004 | 0.222** | 0.015 | 0.214** | -0.026 | 0.225** | -0.038 |
| | (0.101) | (0.101) | (0.101) | (0.102) | (0.099) | (0.102) | (0.099) | (0.104) | (0.101) | (0.104) |
| GDP per capita (thousands of €) | 0.024*** | 0.034*** | 0.023*** | 0.035*** | 0.021*** | 0.033*** | 0.014** | 0.022** | 0.016*** | 0.023** |
| | (0.006) | (0.009) | (0.006) | (0.009) | (0.005) | (0.009) | (0.006) | (0.009) | (0.006) | (0.009) |
| Application Year dummies | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Country dummies | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 | 3,777 |
| Log-likelihood | -2,708.6 | -2,611.4 | -2,702.0 | -2,610.3 | -2,696.1 | -2,609.2 | -2,688.6 | -2,602.7 | -2,693.4 | -2,603.4 |
| Chi-squared | 195.8 | 360.8 | 209.1 | 361.7 | 223.8 | 364.0 | 237.4 | 368.6 | 225.0 | 367.2 |

**TABLE 19: TOBIT MODEL (DYNAMIC) ON THE REGIONAL HYDROGEN SPECIALISATION (NRTA)**

# CONCLUSIONS

The aim of this thesis was to investigate the hydrogen technology ecosystem through patent data. Specifically, in the first part of the work, the main hydrogen-related technologies were identified and the process of constructing the patent landscape was illustrated. In the second part of the thesis, on the other hand, various analyses were conducted on the data sample to seek answers to the two research questions that emerged within the thesis. The first question addressed is: Is there spatial autocorrelation among European hydrogen specialised regions? Do clusters of specialisation in hydrogen exist in Europe? To answer this question, an analysis of spatial autocorrelation was carried out, both on a global level, by calculating the Moran's I statistic, and on a local level, by calculating the LISA indicator. The second research question addressed is: What are the characteristics that affect hydrogen technology specialisation in European regions? To answer it, different regression models were developed: a static model independent of the time variable and a dynamic time variant model, namely a panel model. For both the static model and the dynamic model, two models were produced, one used as a baseline and one used as a robustness check. For the static model, a TOBIT (baseline model) and an OLS (robustness check) were developed, while for the dynamic model, a random effects model (baseline) and a dynamic TOBIT model (robustness check) were developed. Obviously, the dynamic model given the larger number of observations and the larger number of variables and factors considered is more robust and accurate than the static model. In both strands of analysis, the NRTA indicator (the normalised version of the RTA indicator proposed by Balassa) was used to calculate the technological specialisation in hydrogen.

The first strand of analysis showed that there is a moderate positive and statistically significant spatial autocorrelation in hydrogen technology specialisation in Europe. This result suggests that regions with similar technology specialisation tend to cluster. Consequently, policy makers should encourage cross-border collaboration within these clusters to maximise knowledge spillovers. Given the moderate spatial autocorrelation, policies should be designed to facilitate knowledge spillovers within a certain geographical proximity, but not beyond, as spatial propagation effects are not infinite.

In the local autocorrelation analysis, the LISA map showed the presence of several hot spots or high-high clusters: one encompassing almost the whole of Sweden and Norway, a good part of Denmark and northern Germany; one in southern Germany; one in the centre of England and two smaller clusters in France. Some cold spots or low-low clusters were also identified: one in southern Spain, one in Turkey and one in Eastern Europe between Poland, Slovakia and the Czech Republic. This result might suggest the need for targeted investments to foster technological development in the field of hydrogen. In hot spots, such as northern Europe, investments in hydrogen technologies would be more effective. Conversely, cold spots could be targeted for development policies to close the gap. In this regard, inter-regional partnerships between high-high and low-low clusters could be established to facilitate the sharing of knowledge and resources.

Concerning the second strand of analysis, in these conclusions we refer to the dynamic model which, as mentioned above, is much more robust than the static model. The econometric model showed a partial positive correlation between technological specialisation in hydrogen and the following variables: R&D expenditures in the business sector over the GDP, The percentage of firms with in-house innovation activities, The percentage of people employed in knowledge-intensive activities and GDP per capita. In addition, the model also showed a partial negative correlation between the NRTA index and Sum of total turnover of new/innovative products over the sum of total turnover. The sign of this correlation was fully explained in the commentary on the random effects model.

The positive correlation between R&D expenditures per GDP in the business sector together with the lack of significance of public sector R&D expenditures suggest that, at least in the context of European regions, private sector R&D expenditures might have a greater impact on hydrogen technology specialisation. Policies could therefore incentivise firms to invest in R&D through tax breaks, subsidies or targeted funding, rather than directly increasing public R&D expenditures.

The correlation between firms having in-house innovation activities and specialisation in hydrogen indicates the importance of in-house innovation. Policies could support training and development programmes for small and medium-sized enterprises, promoting in-house innovation as a means to achieve greater technological development in H2.

The correlation between the percentage of people employed in knowledge-intensive activities and specialisation in hydrogen underlines the importance of specialised skills. Education and training policies should aim to develop specialised skills in the hydrogen sector by promoting training and specialisation programmes.

The correlation between regional GDP per capita and H2 specialisation suggests that more prosperous regions tend to have higher specialisation. Policies should consider how to balance investments between regions with different levels of GDP per capita to ensure an equal spread of specialisation.

Finally, in the static model there is a partial positive correlation between hydrogen technology specialisation and PM2.5 emissions. This correlation also holds in the dynamic model, although in the latter it is not always statistically significant (it loses significance when controlling for year and country fixed effects). Consequently, in light of this aspect, policy makers could consider introducing more stringent environmental regulations to ensure that this specialisation is sustainable.

# LIMITATIONS AND POSSIBLE FUTURE DEVELOPMENTS

The study presented in this thesis is not exempt from some limitations. The weaknesses of the study and possible further developments to make it more precise and effective are outlined below.

Concerning the methodology, a first limitation lies in the identification of hydrogen technologies for the collection of patent data. Technologies were identified by means of keywords and technological codes. Although this approach represents the state of the art in this type of analysis, it can lead to some errors such as: exclusion of relevant technologies and inclusion of irrelevant technologies.

Regarding the construction of the data sample, the database for the analysis was developed considering only hydrogen patents filed at the European Patent Office. This choice was dictated by the need to match the patent sample downloaded from Derwent Innovation with REGPAT's database containing inventors' geographical information. The EPO sample, as seen in the chapter *Descriptive Statistics of the Dataset* (section *Geographical Distribution of Hydrogen Patents*) represents a limited part of the sample of all hydrogen patents (6.9%). The statistical analysis on hydrogen technology specialisation could consequently be extended to other geographies, also considering data from China, Japan and the United States. Furthermore, the decision to use only EPO data excluded all patents filed at the national patent offices of European countries. Thus, only the highest quality European patents were considered. Patenting an invention at the EPO has in fact a higher cost for companies than patenting at the national office. The decision to use this treatment sample, despite the exclusion of less qualitative patents, was necessary in order to be able to compare data from different countries.

Limitations also exist in the statistical analysis. First of all, the NRTA indicator (and the RTA indicator) used to study technological specialisation in hydrogen suffers from limitations in its construction and interpretation. The NRTA index, as it is defined, can lead to distortions. Regions possessing a very small number of total patents might turn out to be specialised in the reference area even with a very small number of patents in that area. Therefore, to avoid distortions in the study of technology specialisation, a minimum size threshold in terms of absolute patent count should be imposed. This would exclude all regions with few patents and consider only technologically developed areas. The use of the NRTA index in econometric models also entails an interpretative limitation. The interpretation of the NRTA is not linear but varies between regions. For example, for a region just below the specialisation threshold (and thus with a slightly negative NRTA), a small increase in the NRTA is enough to make it specialised. Conversely, for a strongly under-specialised region (deeply negative NRTA), even a significant increase in the indicator may not be enough to make it cross the specialisation threshold.

Concerning the spatial autocorrelation analysis, in the definition of the spatial weights, the queen-type contiguity logic was chosen (see section *Global Spatial Autocorrelation Analysis - Moran Scatter Plot*). This resulted in the exclusion of all regions without boundaries with other regions such as islands. To include these regions in the autocorrelation analysis, artificial spatial contiguities could be created at the most frequented port routes. For example, contiguities could be created between Sicily and Calabria or between Sardinia and Liguria.

Regarding econometric models, there are limitations both on the variables used and on the types of models themselves. The variables of the RIS are highly correlated with each other and then they are all normalised between 0 and 1. The normalisation leads to a reduction of the variability in the distribution and consequently to a less precise estimation of the dependent variable.

Both static and dynamic econometric models were presented in this thesis. While the former are merely observational models that allow to study partial correlations, the latter also allow for causal inference. However, dynamic models may also suffer from reverse causality and endogeneity problems although these effects are limited compared to static models.

# BIBLIOGRAPHY

## SCIENTIFIC PAPERS

Ball, M., & Wietschel, M. (2009). The future of hydrogen–opportunities and challenges. International journal of hydrogen energy, 34(2), 615-627.

Schumpeter, J. A. (1911). The Theory of Economic Development. Harvard University Press, Cambridge, MA.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, *70*(1), 65-94.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, *98*(5, Part 2), S71-S102.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The quarterly journal of economics*, *107*(2), 407-437.

Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European economic review*, *35*(2-3), 517-526.

Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. Econometrica, 60(2), 323-351.

Fagerberg, J. (1987). A technology gap approach to why growth rates differ. *Research policy*, *16*(2-4), 87-99.

Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The quarterly journal of economics*, *114*(1), 83-116.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. American economic review, 91(5), 1369-1401.

Lundvall, B. A. (1992). National systems of innovation: towards a theory of innovation and interactive learning. In Glaeser, E. L. (Ed.), Agglomeration Economies, pages. 85-104, University of Chicago Press, Chicago, IL.

Nelson, R. R. (Ed.). (1993). National innovation systems: A comparative analysis. Oxford University Press, New York, NY.

Arrow, K. J. (1962). The Rate and Direction of Inventive Activity: Economic and Social Factors, Chapter: Economic Welfare and the Allocation of Resources for Invention. Universities-National Bureau Committee for Economic Research, Committee on Economic Growth of the Social Science Research Council (Ed.), pages 609-626, Princeton University Press, Princeton, NJ.

Mazzucato, M. (2013). The Entrepreneurial State: Debunking Public vs. Private Myths in Risk and Innovation, pages 36-56. Anthem Press, London, UK

Griliches, Z. (1998). Patent statistics as economic indicators: a survey. In R&D and productivity: the econometric evidence (pp. 287-343). University of Chicago Press.

Hall, B., Jaffe, A., & Trajtenberg, M. (2001). *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools* (No. 3094). CEPR Discussion Papers.

Narin, F., & Noma, E. (1985). Is technology becoming science? Scientometrics, 7(3-6), 369-381.

Jaffe, A. B., & Trajtenberg, M. (2002). Patents, citations, and innovations: A window on the knowledge economy. MIT Press, Cambridge, MA.

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. Research policy, 32(1), 69-87.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. Research policy, 30(7), 1019-1039.

Verspagen, B. (2007). Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research. Advances in complex systems, 10(01), 93-115.

Newman, M. (2010). Networks: An Introduction, pages 10-19 Oxford University Press, New York, NY.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. Brookings papers on economic activity, 1987(3), 783-831.

Cohen, W. M., Nelson, R., & Walsh, J. P. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not) (No. 7552). National Bureau of Economic Research, Inc.

Barbieri, N., Marzucchi, A., & Rizzo, U. (2020). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones?. Research Policy, 49(2), 103901.

Aldieri, L., Kotsemir, M., & Vinci, C. P. (2020). The role of environmental innovation through the technological proximity in the implementation of the sustainable development. Business strategy and the environment, 29(2), 493-502.

Nomaler, Ö., & Verspagen, B. (2019). Greentech homophily and path dependence in a large patent citation network. Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT).

Nomaler, Ö., & Verspagen, B. (2021). Patent Landscaping Using" green" Technological Trajectories. Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT).

Montresor, S., & Quatraro, F. (2020). Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies. Regional Studies, 54(10), 1354-1365.

Chai, K. C., Yang, Y., Sui, Z., & Chang, K. C. (2020). Determinants of highly-cited green patents: The perspective of network characteristics. Plos one, 15(10), e0240679.

Yu, J., Han, Y. J., Yang, H., Lee, S., Kim, G., & Lee, C. (2022). Promising Technology Analysis and Patent Roadmap Development in the Hydrogen Supply Chain. Sustainability, 14(21), 14210.

Ampah, J. D., Jin, C., Fattah, I. M. R., Appiah-Otoo, I., Afrane, S., Geng, Z., ... & Liu, H. (2022). Investigating the evolutionary trends and key enablers of hydrogen production technologies: A patent-life cycle and econometric analysis. International Journal of Hydrogen Energy.

Zhang, S., Wei, Y., Guo, X., Li, Z., Song, X., & Blaabjerg, F. (2023). Overview of US patents for energy management of renewable energy systems with hydrogen. International Journal of Hydrogen Energy, 48(26), 9574-9591.

Chung, D., Kwon, Y., & Kim, Z. (2023). Technology life cycle and commercialization readiness of hydrogen production technology using patent analysis. International Journal of Hydrogen Energy.

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A new EPO classification scheme for climate change mitigation technologies. World Patent Information, 34(2), 106-111.

Favot, M., Vesnic, L., Priore, R., Bincoletto, A., & Morea, F. (2023). Green patents and green codes: How different methodologies lead to different results. Resources, Conservation & Recycling Advances, 18, 200132.

Angelucci, S., Hurtado-Albir, F. J., & Volpe, A. (2018). Supporting global initiatives on climate change: The EPO's "Y02-Y04S" tagging scheme. World Patent Information, 54, S85-S92.

Bellucci, A., Fatica, S., Georgakaki, A., Gucciardi, G., Letout, S., & Pasimeni, F. (2021). Venture capital financing and green patenting (No. 2021/11). JRC Working Papers in Economics and Finance.

Pasimeni, F., Fiorini, A., & Georgakaki, A. (2019). Assessing private R&D spending in Europe for climate change mitigation technologies via patent data. World Patent Information, 59, 101927.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., & Guellec, D. (2008). *The OECD REGPAT Database: A Presentation* (No. 2008/2). OECD Publishing.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, *108*(3), 577-598.

Breschi, S., & Lissoni, F. (2005). KNOWLEDGE NETWORKS FROM PATENT DATA. *Handbook of Quantitative Science and Technology Research*, 613.

Khramova, E., Meissner, D., & Sagieva, G. (2013). Statistical patent analysis indicators as a means of determining country technological specialisation. *Higher School of Economics Research Paper No. WP BRP*, *9*.

Caviggioli, F., Colombelli, A., De Marco, A., Scellato, G., & Ughetto, E. (2023). Co-evolution patterns of university patenting and technological specialization in European regions. *The Journal of Technology Transfer,* 48(1), 216-239.

Laursen, K. (2015). Revealed comparative advantage and the alternatives as measures of international specialization. Eurasian Business Review, 5(1), 99–115.

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, *10*(2), 243-251.

Cliff, A. D., & Ord, J. K. (1981). Spatial processes: models & applications. *(No Title)*.

Anselin, L. (2019). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In *Spatial analytical perspectives on GIS* (pp. 111-126). Routledge.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica: Journal of the econometric society, 1251-1271.

**REPORTS**

IPCC. (2020). Climate Change 2020: The Physical Science Basis. Cambridge University Press. United Kingdom.

World Intellectual Property Organization. (2022). Annual Report WIPO 2022.

International Energy Agency. (2023). Hydrogen Patents for a Clean Energy Future, IEA 2023.

Organisation for Economic Co-operation and Development. (2009). OECD Patent Statistics Manual.

United Nations Framework Convention on Climate Change. (2015). Paris Agreement.

European Commission. (2020). A hydrogen strategy for a climate-neutral Europe.

Organisation for Economic Co-operation and Development. (2020). Patents and innovation: Trends and policy challenges.

European Patent Office. (2020). What is a patent?

Guidelines for Preparing Patent Landscape Reports, Guidelines prepared for the World Intellectual Property Organization (WIPO) by Anthony Trippe, 2015

Patent Landscape Report - Hydrogen fuel cells in transportation, World Intellectual Property Organization 2022

Hydrogen patents for a clean energy future - A global trend analysis of innovation along hydrogen value chains, International Energy Agency 2023

Global Hydrogen Review, IEA 2021

Clean Hydrogen Monitor, Hydrogen Europe 2022

The potential of hydrogen for the chemical industry, Deloitte 2021

Communication From the Commission to The European Parliament, The Council, The European Economic and Social Committee and the Committee Of The Regions: A Hydrogen Strategy For A Climate-neutral Europe, European Commission, 2020

Hydrogen: A Renewable Energy Perspective, IRENA 2019

The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs, Committee on Alternatives and Strategies for Future Hydrogen Production and Use, Board on Energy and Environmental Systems, Division on Engineering and Physical Sciences, National Research Council and National Academy of Engineering Of The National Academies, 2004

European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, (2021, 2019, 2017, 2014, 2012, 2009). *Regional innovation scoreboard*, Publications Office.

# ANNEX

| Hydrogen Technology | Query – Derwent Innovation | Individual Records | Applications | Families |
|---|---|---|---|---|
| Steam Methane Reforming: with CCUS (Blue Hydrogen) | CTB=(((((methan* OR (natural* adj gas*) OR CH4) OR (therm* NEAR1 plasma) OR solar OR (molten NEAR1 metal*)) NEAR3 pyrolys*) OR ((methan* OR (natural* adj gas*) OR CH4) NEAR3 (cracking or thermolys*)) OR (solid ADJ carbon NEAR1 produc* NEAR2 (methan* OR (natural* adj gas*) OR CH4) NEAR1 (dissociat* OR decompos*))) AND ((hydrogen* OR H2 OR H) NEAR2 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*))) | 560 | 442 | 197 |
| Steam Methane Reforming: Electrically Heated (Grey Hydrogen) | CTB=((((electr* or (electr* NEAR2 steam*)) NEAR3 (methan* OR (natural* ADJ1 gas*) OR CH4) NEAR1 reform*) OR (eSMR* OR e-SMR* OR (electr* NEAR2 SMR*))) AND ((hydrogen* OR H2 OR H) NEAR2 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*))) | 112 | 100 | 69 |
| Steam Methane Reforming: Sorption-enhanced (Grey Hydrogen) | CTB=((((sorpt* NEAR2 reform*) AND (methan* OR (natural* ADJ1 gas*) OR CH4*)) OR (SESMR* OR (sorpt* NEAR2 SMR*)) OR (SMR* NEAR2 (in-situ OR integrat*) NEAR2 (CO2* OR (carbonic* NEAR1 anhydrid*) OR (carbon NEAR1 dioxid*)) NEAR2 (captur* OR sorpt*)))) | 49 | 41 | 21 |
| Steam Methane Reforming: Plasma Reforming (Grey Hydrogen) | CTB=((((plasma* OR (electr* NEAR1 discharg*)) NEAR3 (reform* OR pyroly*)) AND (methan* OR (natural* ADJ1 gas*) OR CH4*)) AND ((hydrogen* OR H2 OR H) NEAR2 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*))) | 173 | 143 | 106 |
| Methane Pyrolysis (Turquoise Hydrogen) | CTB=(((((methan* OR (natural* adj gas*) OR CH4) OR (therm* NEAR1 plasma) OR solar OR (molten NEAR1 metal*)) NEAR3 pyrolys*) OR ((methan* OR (natural* adj gas*) OR CH4) NEAR3 (cracking or thermolys*)) OR (solid ADJ carbon NEAR1 produc* NEAR2 (methan* OR (natural* adj gas*) OR CH4) NEAR1 (dissociat* OR decompos*))) AND ((hydrogen* OR H2 OR H) NEAR2 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*))) | 984 | 819 | 658 |
| Alkaline Electrolysers (AEL) | CTB=(((alkalin* OR bacon OR ((potassium OR sodium) NEAR2 hydroxid*) OR AEL* OR AEC*) NEAR3 (electrolyz* OR electrolys* OR electrodialys* OR (electrolytic* ADJ cell*) OR ((Hydrogen* OR H2 OR H) NEAR3 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*)) OR (water NEAR1 (split* OR decompos* OR (oxidat* NEAR3 reduct*) OR electrodissociat*))))) | 8.211 | 6.851 | 5.257 |
| Polymer Electrolyte Membrane (PEM) or Proton Exchange Membrane Electrolysers (PEMEL) | CTB=(((proton* NEAR1 exchang* NEAR1 membran*) OR (proton-exchang* NEAR1 membran*) OR (Polymer* NEAR1 Electrolyt* NEAR1 Membran*) OR Polymer* OR PEM* OR SPE) NEAR3 (electrolyz* OR electrolys* OR electrodialys* OR (electrolytic* ADJ cell*) OR | 6.264 | 5.106 | 3.468 |

| | | | | |
|---|---|---|---|---|
| | ((Hydrogen* OR H2 OR H) NEAR3 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*)) OR (water NEAR1 (split* OR decompos* OR (oxidat* NEAR3 reduct*) OR electrodissociat*))) NOT (FUEL ADJ CELL*)); | | | |
| Solid Oxide Electrolysers (SOEL) | CTB=(((Solid ADJ Oxide) OR SOE* or CERAMIC* OR HIGH-TEMPERATUR* OR (HIGH NEAR1 TEMPERATUR*)) NEAR3 (electrolyz* OR electrolys* OR electrodialys* OR (electrolytic* ADJ cell*) OR ((Hydrogen* OR H2 OR H) NEAR3 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*)) OR (water NEAR1 (split* OR decompos* OR (oxidat* NEAR3 reduct*) OR electrodissociat*))) NOT (FUEL ADJ CELL*)); | 4.244 | 3.488 | 2.550 |
| Anion Exchange Membrane Electrolysers (AEMEL) | CTB=(((Anion* NEAR1 exchang*) OR AEM*) NEAR3 (electrolyz* OR electrolys* OR electrodialys* OR (electrolytic* ADJ cell*) OR ((Hydrogen* OR H2 OR H) NEAR3 (produc* OR generat* OR synthesis* OR creat* OR manufact* OR fabric*)) OR (water NEAR1 (split* OR decompos* OR (oxidat* NEAR3 reduct*) OR electrodissociat*)))); | 763 | 658 | 486 |
| Transformation of Pure Hydrogen into Low-emission Hydrogen-based Synthetic Fuels: Synthetic Methane & Others | CTB=((((((low-emission* OR (low* ADJ emission*) OR low-carbon* OR (low* ADJ carbon*) OR sustainabl* OR clean* OR green* OR (climat* ADJ friendl*) OR CO2-neutral OR (CO2 ADJ neutral*) OR (carbon ADJ neutral) OR synthetic* OR bio*) NEAR3 (fuel* OR gas* OR methan* OR CH4 OR kerosen* OR diesel OR hydrocarbon*)) OR (synfuel* OR biofuel* OR syngas* OR synthetic ADJ oil* OR e-methan* OR e-diesel OR e-kerosen*)) NEAR3 (produc* OR generat* OR transform* OR reduc* OR conver* OR process* OR synthes*)) NEAR3 (hydrogen* OR H2 OR H)) NOT coal); | 3.660 | 2.925 | 1.951 |
| Transformation of Pure Hydrogen to Ammonia & Low-temperature Ammonia Cracking (from Ammonia to Pure Hydrogen) | CTB=((((low-temperatur* OR low NEAR1 temperatur*) NEAR3 ((Haber-Bosch OR Haber ADJ Bosch)) OR ((Ammonia OR NH3) NEAR3 (produc* OR generat* OR transform* OR reduc* OR conver* OR crack* OR decompos* OR process* or synthes*) NEAR3 (hydrogen* OR H2 OR H)))) NOT (Methan* OR synthetic* OR CH4 OR (Natural ADJ Gas) OR Coal OR Electrol* OR (FUEL ADJ CELL*))); | 4.724 | 3.951 | 3.019 |
| Transformation of Pure Hydrogen to Liquid Organic Hydrogen Carriers | CTB=(((LOHC OR liquid* ADJ hydrogen ADJ carrier* OR (liquid NEAR1 organic* NEAR3 hydrogen) OR Hydrogenat* ADJ (Liquid* ADJ Organic ADJ Compound* OR Organic* ADJ Liquid*) OR *cyclohexan* OR *Dibenzyltoluen* OR *ethylcarbazol* OR methyldecalin* OR *naphthalene) NEAR3 (produc* OR generat* OR transform* OR reduc* OR conver* OR crack* OR decompos* OR process*) NEAR3 (hydrogen* OR H2 OR H)) NOT (synthetic* OR Electrol* OR (FUEL ADJ CELL*) OR coal)); | 1.396 | 1.122 | 838 |
| Gaseous Storage (Fuel stations, Terminals or Platforms, by Burying Tanks, by Digging Cavities, by using | CTB=((((hydrogen* OR H2 OR H) NEAR1 (stor* OR stock* OR deposit*)) NEAR3 (gas* OR compress* OR pressur* OR high-pressur*)) NOT ((solid* OR liquid* OR liquef* OR hydrid* OR cryogen* OR cryonic*) NEAR3 (hydrogen OR H2 OR H))); | 5.972 | 4.987 | 4.038 |

| | | | | |
|---|---|---|---|---|
| Natural Cavities, Deep Sea, Offshore) | | | | |
| Liquid Storage (Fuel stations, Terminals or Platforms, by Burying Tanks, by Digging Cavities, Deep Sea, Offshore) | CTB=((((hydrogen* OR H2 OR H) NEAR1 (stor* OR stock* OR deposit*)) NEAR3 (liquid* OR liquef* OR cryogen* OR cryonic* OR dewar* OR (insulated ADJ1 tank*))) NOT ((solid* OR hydrid* OR gaseous* OR compress*) NEAR3 (hydrogen OR H2 OR H))); | 1.727 | 1.427 | 1.242 |
| Solid storage (Hydrides/Adsorption) | CTB=((((hydrogen* OR H2 OR H) NEAR1 (stor* OR stock* OR deposit*)) NEAR3 (solid* OR (metal* NEAR1 (hydrogen OR H2 OR compound*)) OR hydride* OR ((AB5 OR AB2) NEAR1 alloy*) OR (magnesium* NEAR1 hydride*) OR MgH2 OR (sodium* NEAR1 borohydride*) OR NaBH4 OR adsorpti* OR physisorpti* OR chemisorpti* OR (activated* NEAR1 carbon*) OR (metal-organic* NEAR1 framework*) OR (metal* NEAR1 organic* NEAR1 framework*) OR MOF OR MOFs OR (covalent-organic* NEAR1 framework*) OR (covalent* NEAR1 organic* NEAR1 framework*) OR COF OR COFs OR (absorb* NEAR1 (material* OR matter* OR compound* OR mixtur* OR substanc* OR component* OR constituent* OR element* OR structur*)) OR clathrat* OR (cage-like NEAR1 structur*) OR ((cage ADJ like) NEAR1 structur*) OR (crystal* NEAR1 solid*) OR (water ADJ cage))) NOT ((liquid* OR liquef* OR cryogen* OR cryonic* OR gaseous* OR compress* OR (gas* NEAR1 compress*)) NEAR3 (hydrogen OR H2 OR H))); | 2.983 | 2.390 | 1.854 |
| Proton-exchange membrane fuel cells (PEMFC) | CTB=((PEM* OR (proton* ADJ exchang* ADJ membran*) OR (proton-exchang* ADJ membran*) OR (proton-conduct* ADJ membran*) OR (proton* ADJ conduct* ADJ membran*) OR (polymer* ADJ electrolyt*)) NEAR1 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*))) OR (*PEMFC* AND (Hydrogen* OR H2 OR H))) OR (CTB=((PEM* OR (proton* ADJ exchang* ADJ membran*) OR (proton-exchang* ADJ membran*) OR (proton-conduct* ADJ membran*) OR (proton* ADJ conduct* ADJ membran*) OR (polymer* ADJ electrolyt*))) AND AIC=(H01M000800)); | 30.647 | 22.404 | 13.749 |
| Alkaline Fuel Cells (AFC) | CTB=((bacon* OR alkalin* OR (anion-exchang*) OR (anion* NEAR1 exchang*) OR (hydroxid* NEAR1 exchang*) OR (basic* NEAR1 membran*) OR (basic* NEAR1 polymer*)) NEAR1 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*))) OR ((*AFC* OR AEMFC* OR AMFC* OR HEMFC* OR SAFC*) AND (Hydrogen* OR H2 OR H))) OR (CTB=((bacon* OR alkalin* OR (anion-exchang*) OR (anion* NEAR1 exchang*) OR (hydroxid* NEAR1 exchang*) OR (basic* NEAR1 membran*) OR (basic* NEAR1 polymer*))) AND AIC=(H01M000800)); | 15.315 | 13.810 | 10.255 |

| | | | | |
|---|---|---|---|---|
| Phosphoric Acid Fuel Cell (PAFC) | CTB=(((phosphoric* near2 acid*) near3 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*)))) OR (phosphoric* near1 acid* near3 electrolyt*) OR (*PACF* AND (Hydrogen* OR H2 OR H))) OR (CTB=(phosphoric* near2 acid*) AND AIC=(H01M000800)); | 5.186 | 4.325 | 2.965 |
| Molten Carbonate Fuel Cell (MCFC) | CTB=(((((molten* OR liquid*) near2 carbonat*) or molten-carbonat* OR liquid-carbonat*) near3 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*)))) OR (MCFC* AND (Hydrogen* OR H2 OR H))) OR (CTB=(((molten* OR liquid*) near2 carbonat*) or molten-carbonat* OR liquid-carbonat*) AND AIC=(H01M000800)); | 4.897 | 3.802 | 2.108 |
| Solid Oxide Fuel Cell (SOFC) | CTB=((((solid ADJ oxide) or ceramic) NEAR1 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*)))) OR (*SOFC* AND (Hydrogen* OR H2 OR H))) OR (CTB=((solid ADJ oxide) or ceramic) AND AIC=(H01M000800)); | 31.548 | 22.169 | 12.937 |
| Direct Methanol Fuel Cell (DMFC) | CTB=((methanol* NEAR3 (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR (Fuel* ADJ Power* ADJ System*) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*)))) or (DMFC* AND (Hydrogen* OR H2 OR H))) OR (CTB=(methanol) and AIC=(H01M000800)); | 10.589 | 7.922 | 4.910 |
| Internal Combustion Engine (ICE) | CTB=(((((combust* OR piston* OR reciprocating* OR otto OR spark-ignition* OR (spark* NEAR1 ignition*) OR compression-ignition* OR (compression* NEAR1 ignition*)) NEAR3 engin*) OR ICE) AND (hydrogen* OR H2 OR H)) NOT (((Fuel* OR Electrochemical*) ADJ (cell* OR Batter*)) OR Electrochemical* NEAR1 ((Power OR Energy) ADJ (Conver* OR Generat* OR Sourc*)))) AND (AIC=(B60) OR AIC=(B62) OR AIC=(B63)); | 9.746 | 9.621 | 6.512 |
| Direct Reduced Iron | CTB=(((direct* NEAR1 reduc* NEAR3 (iron* OR Fe OR Fe2O3 OR (iron* ADJ ore*) OR steel* OR ferrous* OR ferrum*)) OR DRI ADJ H-DRI OR H2-DRI OR sponge-iron* OR (spong* NEAR1 iron*)) NEAR5 (hydrogen* OR H2 OR H)); | 282 | 244 | 188 |
| Blending in Blast Furnaces | CTB=(((hydrogen* OR H2 OR H) NEAR3 (mix* OR blend* OR reduc* OR inject* OR introduc* OR incorporat*)) NEAR3 (iron* OR Fe OR Fe2O3 OR (iron* ADJ ore*) OR steel* OR ferrous* OR ferrum*) AND (furnac* OR BF OR blast)); | 598 | 511 | 464 |
| Smelting Reduction | CTB=(((smelt* NEAR3 (hydrogen* OR H2 OR H)) OR H-smelt* OR H2-smelt* OR hydrogen-smelt* OR HIsarna* OR Hyl-SR OR (Hyl ADJ1 SR) OR HylSr) NEAR5 (iron* OR Fe OR Fe2O3 OR (iron* NEAR1 ore*) OR ferrous* OR ferrum* OR steel*)); | 147 | 122 | 120 |

TABLE 2: SEARCH QUERIES HYDROGEN VALUE-CHAIN-RELATED TECHNOLOGIES

| TOP 5 Regions (NUTS code) | Country | Specific Hydrogen Technology patents | Specific technology patents - Total H2 patents ratio | % of specific hydrogen technology patents |
|---|---|---|---|---|
| **Fuel Cells** | | | | |
| South East (UKJ) | United Kingdom | 102 | 83,6% | 5,5% |
| Darmstadt (DE71) | Germany | 92 | 74,2% | 5,0% |
| Rhône-Alpes (FRK2) | France | 91 | 63,2% | 4,9% |
| Stuttgart (DE11) | Germany | 82 | 51,3% | 4,5% |
| Mittelfranken (DE25) | Germany | 73 | 72,3% | 4,0% |
| **Electrolyzers** | | | | |
| Rhône-Alpes (FRK2) | France | 28 | 19,4% | 6,8% |
| Hovedstaden (DK01) | Denmark | 22 | 24,4% | 5,3% |
| Mittelfranken (DE25) | Germany | 18 | 17,8% | 4,3% |
| Köln (DEA2) | Germany | 16 | 14,2% | 3,9% |
| Île de France (FR10) | France | 15 | 12,6% | 3,6% |
| **H2 Storage** | | | | |
| Rhône-Alpes (FRK2) | France | 9 | 6,3% | 9,1% |
| Île de France (FR10) | France | 7 | 5,9% | 7,1% |
| South East (UK) (UKJ) | United Kingdom | 6 | 4,9% | 6,1% |
| Düsseldorf (DEA1) | Germany | 5 | 7,9% | 5,1% |
| Darmstadt (DE71) | Germany | 4 | 3,2% | 4,0% |
| **H2 Based Fuels** | | | | |
| Île de France (FR10) | France | 15 | 12,6% | 7,2% |
| Rhône-Alpes (FRK2) | France | 12 | 8,3% | 5,7% |
| Darmstadt (DE71) | Germany | 9 | 7,3% | 4,3% |
| Rheinhessen-Pfalz (DEB3) | Germany | 9 | 14,8% | 4,3% |
| North East (UKC) | United Kingdom | 8 | 50,0% | 3,8% |
| **Iron & Steel Manufacturing** | | | | |
| Niederösterreich (AT12) | Austria | 2 | 25,0% | 20,0% |
| Oberösterreich (AT31) | Austria | 2 | 12,5% | 20,0% |
| Darmstadt (DE71) | Germany | 2 | 1,6% | 20,0% |
| Île de France (FR10) | France | 2 | 1,7% | 20,0% |
| Stockholm (SE11) | Sweden | 1 | 2,1% | 10,0% |
| **Low-CO2 H2 Production** | | | | |
| Île de France (FR10) | France | 6 | 5,0% | 16,7% |
| Hovedstaden (DK01) | Denmark | 3 | 3,3% | 8,3% |
| Vestlandet (NO05) | Norway | 3 | 27,3% | 8,3% |
| Sjælland (DK02) | Denmark | 2 | 5,9% | 5,6% |
| Rhône-Alpes (FRK2) | France | 2 | 1,4% | 5,6% |
| **ICE** | | | | |
| Tübingen (DE14) | Germany | 65 | 54,2% | 10,0% |
| Stuttgart (DE11) | Germany | 64 | 40,0% | 9,8% |
| Oberbayern (DE21) | Germany | 51 | 42,5% | 7,8% |
| Stockholm (SE11) | Sweden | 31 | 64,6% | 4,8% |
| Steiermark (AT22) | Austria | 27 | 77,1% | 4,2% |

TABLE 10: BREAKDOWN OF MACRO-TECHNOLOGY HYDROGEN PATENTS BY TOP COUNTRIES

| Fixed Effects - NRTA | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| R&D expenditures in the public sector per GDP | -0.066 | -0.117 | 0.003 | -0.110 | -0.003 | -0.123 | 0.021 | -0.142 | 0.029 | -0.133 |
| | (0.095) | (0.124) | (0.097) | (0.125) | (0.097) | (0.125) | (0.100) | (0.125) | (0.100) | (0.125) |
| R&D expenditures in the business sector per GDP | 0.161 | -0.008 | 0.159 | 0.008 | 0.157 | 0.005 | 0.135 | -0.120 | 0.136 | -0.122 |
| | (0.126) | (0.138) | (0.126) | (0.139) | (0.126) | (0.138) | (0.128) | (0.145) | (0.128) | (0.145) |
| innovation expenditures (not R&D) per turnover | -0.080 | -0.134 | 0.011 | -0.093 | -0.006 | -0.115 | -0.004 | -0.105 | 0.017 | -0.082 |
| | (0.072) | (0.083) | (0.076) | (0.088) | (0.076) | (0.089) | (0.078) | (0.089) | (0.078) | (0.089) |
| firms' sum of total turnover of new/innovative products over sum of total turnover | | | -0.187** | -0.147* | -0.199*** | -0.150* | -0.206*** | -0.148* | -0.192** | -0.144* |
| | | | (0.076) | (0.083) | (0.074) | (0.081) | (0.075) | (0.081) | (0.077) | (0.083) |
| % of firms who introduced a product or process innovation | | | -0.057 | 0.022 | | | | | -0.061 | 0.020 |
| | | | (0.089) | (0.092) | | | | | (0.090) | (0.093) |
| % of firms with in-house innovation activities | | | | | 0.181* | 0.195** | 0.186* | 0.185* | | |
| | | | | | (0.096) | (0.097) | (0.096) | (0.098) | | |
| % of firms with innovation cooperation activity | | | -0.128 | -0.002 | -0.186** | -0.041 | -0.187** | -0.033 | -0.128 | 0.005 |
| | | | (0.081) | (0.087) | (0.079) | (0.087) | (0.080) | (0.087) | (0.081) | (0.087) |
| % population with tertiary education | | | | | | | 0.027 | 0.228* | 0.018 | 0.241* |
| | | | | | | | (0.115) | (0.133) | (0.115) | (0.133) |
| % of employed persons in knowledge-intensive activities | | | | | | | 0.105 | 0.243** | 0.114 | 0.244** |
| | | | | | | | (0.109) | (0.115) | (0.110) | (0.115) |
| Population density (thousands of people per km2) | -0.343 | -0.281 | -0.265 | -0.243 | -0.214 | -0.208 | -0.225 | -0.205 | -0.273 | -0.233 |
| | (0.215) | (0.215) | (0.218) | (0.218) | (0.217) | (0.217) | (0.218) | (0.218) | (0.219) | (0.218) |
| GDP per capita (thousands of €) | -0.001 | 0.004 | -0.002 | 0.004 | -0.003 | 0.003 | -0.004 | 0.004 | -0.004 | 0.004 |
| | (0.003) | (0.004) | (0.003) | (0.004) | (0.003) | (0.004) | (0.003) | (0.004) | (0.003) | (0.004) |
| Application Year dummies | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Constant | -0.524*** | -0.553*** | -0.420*** | -0.515*** | -0.493*** | -0.570*** | -0.473*** | -0.676*** | -0.427** | -0.650*** |
| | (0.115) | (0.143) | (0.120) | (0.151) | (0.122) | (0.152) | (0.167) | (0.178) | (0.166) | (0.177) |
| Observations | 3777 | 3777 | 3777 | 3777 | 3777 | 3777 | 3777 | 3777 | 3777 | 3777 |
| Adjusted R-squared | -0.064 | -0.054 | -0.061 | -0.054 | -0.060 | -0.053 | -0.060 | -0.051 | -0.061 | -0.052 |

**TABLE 20: FIXED EFFECTS MODEL ON THE REGIONAL HYDROGEN SPECIALISATION (NRTA)**