UNIVERSITÀ DEGLI STUDI "ROMA TRE"

FACOLTÀ DI ECONOMIA

DOTTORATO DI RICERCA IN
ECONOMIA E METODI QUANTITATIVI

Sezione INSTITUZIONI AMBIENTE SVILUPPO

XXV CICLO

Understanding biofuels through patent data

Ylenia Curci

Relatore: Prof.ssa Valeria Costantini

Correlatori: Prof. Frabrizio De Filippis and Dott.ssa Annalisa Zezza

Contents

Introduction V

Chapter 1. Biofuels definition and policies 1

Chapter overview 1

Background and aim of the chapter 1

Classification and definition of liquid biofuels 4

Biodiesel 7

Bioethanol 13

Biofuels international political framework 17

Demand side instruments 18

Supply side instruments 22

Some issues hampering the diffusion of first generation biofuels 27

Economic aspects 28

Environmental aspects 34

Social aspects 42

Some issues hampering the diffusion of second generation biofuels 45

Chapter 2. Environmental goals and technological innovation 49

Chapter overview 49

Background and aim of the chapter 49

Innovation and eco-innovation 55

Evolutionary economics: innovation in perspective 57

Co-evolutionary processes : the carbon lock in 59

The carbon lock-in: a review of its theorization by Unruh 61

Chapter 3. Building BioPat 65

Chapter overview 65

Background and aim of the chapter 65

The IPC system and the Green Inventory database 67

The BioPat methodology 72

Database structure and preliminary descriptive statistics 80

Further characterizations of the biofuels sector through BioPat 90

Chapter 4. An adjusted text mining approach to network analysis of biofuels patent data 95

Chapter overview 95

Background and aim of the chapter 95
Biofuels as a complex technological domain 97
Patent analysis and patent quality 98
Network analysis 101
Keywords vectors 107
Performance of patent network analysis 109
Network identification 113
Visual and quantitative analysis 116
Conclusion 123
Annex I. Green Inventory biofuels classes 126
Annex II. Examples of keywords 127
References 129

List of Tables

Table 1 Conventional and advanced biofuels 7
Table 2 Biodiesel blending mandates (M) and targets (T) 11
Table 3 Bioethanol blending targets and mandates 16
Table 4 CO_2 net reduction by fuel 35
Table 5 Comparison of corn ethanol and gasoline GHGs with and without land-use change by stage of production and use 38
Table 6 NER by biofuels type 41
Table 7 NER calculated by different authorities, by raw material and country 42
Table 8 Production costs by pretreatments 46
Table 9. Data available on Thomson Innovation 76
Table 10. Information available in the BioPat database 81
Table 11. Selected countries in BioPat for descriptive statistics 85
Table 12- Count of records and share of patents by main country and patent office 85
Table 13. Validation of BioPat for EPO patents: percentage of patents related to the biofuels sector 85
Table 14. Highly cited patens 87
Table 15. Selection of new IPC classes included in BioPat/GI from 2000 92
Table 16. Keywords occurrence 108
Table 17. Matrix of association value 111
Table 18. Matrix of adjacency 112
Table 19. Stemmed keywords-topic 116
Table 20. Number of ties and keywords 119

List of Figures

Figure 1 Global biofuels production 1975-2005 3
Figure 2. Global biofuel production 2000-2011 4
Figure 3 Feedstock (thousand tonne) usage for bioethanol production, 2010 8
Figure 4 Biodiesel production by region 2000-2010 (kt) 10
Figure 5 Feedstock (thousand tonne) usage for bioethanol production, 2010 14
Figure 6 Trend (2006-2010) in global production (thousand tonnes) of grain ethanol animal feed co-products 17
Figure 7 Trends in public R&D expenditures in selected energy sectors 1990-2010 (Index number 1990=100) 26
Figure 8 Trends in biofuels patents application as % of total renewables, 1990-2008 26
Figure 9 Trend in the price of oil and sugar 1980-2012 30
Figure 10. Breakeven price for crude oil (US$) and other raw materials, 2005 31
Figure 11. Different production costs for raw materials and country (€/hl) 33
Figure 12Exemplificative alternative structures of database and classifications using keywords (case a) 79
Figure 13. Exemplificative alternative structures of database and classifications using keywords (case b) 79
Figure 14. Number of patents and number of patents weighted by forward citations by year, Japan, European countries and United States of America 88
Figure 15. Number of patents by year for food-related technologies 89
Figure 16. Number of patents by year for algae-related technologies 89
Figure 17. Variation in classgroup number in the IPC system and in BioPat 93
Figure 18. Selection of number of topics 115
Figure 19. Probability document-topic 115
Figure 20. Biofuel network, by topic and number of connections 118
Figure 21. Network by number of ties 119
Figure 22. Network by years 120

Introduction

The last decade has been a period of intense instability in oil prices and there has been growing concern about the environmental costs of carbon emissions from fossil fuels in the transport sector.

Due to high oil prices and the need to reduce GHGs (Green House Gasses) emissions, biofuels for transport use such as ethanol and biodiesel, which are the only suitable substitutes for liquid fossil fuels, have gained importance in many countries.

At the beginning of the 21th century, the European Union (EU) started fostering production and use of biofuels, and bioenergy in general, in several forms. There are various documents in place settled by the European Commission (EC) to promote the use of bioenergy such as Directives 2001/77/EC, 2003/30/EC, 2003/96/EC, the EU "Biomass Action Plan" (EC, 2005) and the "European Union Biofuel Strategy" (EC, 2006). According to the EU biofuels directive 2003/30/EC, EU Member States should ensure a minimum amount of biofuels and other renewable fuels in their total consumption of transport fuel. In the "Renewable Energy Roadmap" (EC, 2007), the EC proposed binding minimum targets of 10% for biofuels in each Member State. On 23 January 2008, the EC put forward an integrated proposal for Climate Action, including a directive that sets an overall compulsory target for the European Union of 20% renewable energy by 2020 and a 10% minimum target for the market share of biofuels by 2020, to be observed by all Member States.

Despite a specific emission target for the transport sector was already in place in 2003, as described in the "Energy, Transport and Environment Indicators" published by Eurostat (2007), in 2005 the transport sector accounted for about 31% of total energy consumption in the European Union (EU-27 Members), representing 19% of total Greenhouse Gases (GHG) emissions. In 2009 the GHG emission percentage increased from 19% to 25% (EEA, 2012). In the light of this, in subsequent years, European policy makers set specific policy instruments to abate transport sector GHGs emission.

In the US, the 2005 Energy Bill established a mandate requiring minimum levels of biofuel consumption from 11.9 million tons in 2006 up to 22.1 million tons in 2012.

Despite the fact that the US mandate had almost been reached by 2007, and despite the very recent change in petroleum consumption among OECD countries which is showing a slow decrease, the past ten years demonstrate that current European policies for a sustainable energy system are inadequate in the transport sector and highly dependent on fossil fuels, thus requiring further efforts to expand alternative energy sources.

Global production of biofuels amounted to 59,261 ktoe in 2010, which represents around 1-2% of total fuel consumption in transportation. Projections on market shares shape a huge increase reaching around 13% of global fuel consumption in 2050 (IEA, 2007). The size of such an increase will depend critically on the rate of technological change and the diffusion rate of new technologies in the biofuels sector. It is worth mentioning that the OECD-FAO (2010) projection for 2010-2019 on bioethanol and biodiesel production pointed out that the 13% growth rate is probably underestimated. In 2009, alternative energy sources to fossil fuels accounted for more than 50 % of installed capacity in US and above 60% in the EU (UNEP, 2010), remaining almost resilient against economic turbulence. Among renewable energy sources, investments in biofuel plants declined in 2009, whereas waste-to-energy investment increased from 9 to 11 billion dollars. In 2008, the biofuels sector had a total investment of 18 billion dollar whereas in 2009 it ended up with just 7 billion dollars. The UNEP Energy

Finance Initiative Report suggests that investment in first generation biofuels is declining due to the fact that most firms are not operating at full capacity: "investment in new biofuel plants declined from 2008 rates, as corn ethanol production capacity was not fully utilized in the United States and several firms went bankrupt. The Brazilian sugar ethanol industry also faced economic troubles, with no growth despite on-going expansion plans. Europe faced similar softening in biodiesel, with production capacity only half utilized." (UNEP, 2010, p. 6).

The recent evolution in the biofuels sector has been characterized by strong price volatility and a mismatch between demand and supply. Part of the responsibility for the current situation can be attributed to the confusion created by governmental policies that conflict with one another and a lack of knowledge of the biofuels production system (Costantini and Crespi, 2012). However, the increased price of fossil fuels as well as a need for environmental-friendly and cost-effective technologies for the production of clean energy, made us believe that these changes should be reflected in evolution of the sector's technological regime.

Additionally to policy incoherence, this study supports the idea that the biofuels sector suffers from severe drawbacks linked to innovative improvements of production processes and adaptability of existing technologies. In particular we refer to the existence of a transport sector carbon lock-in. A carbon lock-in is a phenomenon where the emergence of new technologies is hampered by a historic dependency on fossil energy. Above all, the adaptability of existing fleets and distribution apparatus to different bioliquids seems to play an important role in the deployment of the biofuels sector.

Chapter 1. Biofuels definition and policies

Chapter overview

This chapter provides an insight into liquid biofuels universe: product characteristics, technological perspectives, environmental, social and economic issues and policy considerations. While the biofuel industry showed a positive trend during the last decades, concerns remain on whether this industry would survive in the absence of fiscal incentives. Furthermore, biofuels are considered environmentally friendly, but still the global impact results ambiguous. Advanced generation biofuels seem to address some of the most crucial outstanding issues and technological improvements are gaining importance in the process of transition towards sustainability.

Background and aim of the chapter

Even though concerns related to the biofuels production can be considered a new phenomenon, it worth reminding that liquid biofuel date back to the second half of the XIV century: in 1853, the scientists E. Duffy and J. Patrick discovered the so called transesterification process, a chemical process allowing the transformation of crude vegetable oil into a blend of methyl esters and fat acids burnable as fuel. In 1900 Rudolf Diesel presented an engine at the universal exposition of Paris. Diesel's engine was tested to work burning raw peanut oil. After 12 years, Diesel release a very modern-day statement:

> *«We can get fuel from fruit, from that shrub by the roadside, or from apples, weeds, saw-dust—almost anything! There is fuel in every bit of vegetable matter that can be fermented. There is enough alcohol in one year's yield of a hectare of potatoes to drive the machinery necessary to cultivate the field for a hundred years. And it remains for someone to find out how this fuel can be produced commercially—*

better fuel at a cheaper price than we know now. »

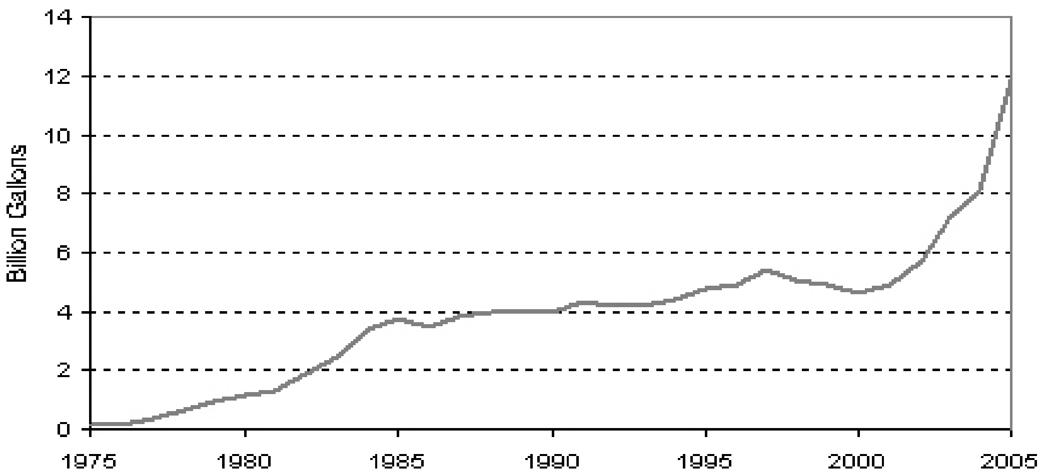
(Henry Ford, 1925)

Lately, the use of biofuels felt into disuse due to clear economic causes: it will not be reasonable to produce biofuels instead of using the vast fossil resources available at the beginning of the last century. This cheap price of fossil fuels was translated into the edification of a whole industrial paradigm supported by fossil energy and cars, that were created together with biofuels, were assembled following specific characteristics of fossil petrol viscosity and requirements (rare exceptions is represented by model T by Henry Ford, projected to work with corn derived ethanol).

The development of biofuels demand during the XX century has been determined basically by the need of alternatives for fossil fuel imports, in an energy security perspective. In fact, during the second world war, some nations used renewable fuels as energy input for military equipment (Germany, for example, overhauled a missile powered by hydrogen peroxide and ethylic alcohol obtained by potatoes' fermentation). After the end of the conflict, falling prices of Middle-Eastern fossil fuel stopped the technological research pushed by the military sector.

Biofuels went back on stage thanks to the unstable geopolitical situation and economics interest caused by fuel crisis started during the seventies (Figure 1). Cuts in production announced by OPEC, increasing energy needs of modern societies, escalation of concerns related to environmental degradability and climate changes attributed to CO2 release in the atmosphere, put biofuels in the middle of a storm of global debate. Few years after OPEC petroleum crisis, two important countries, United States of America and Brazil, initiated their probiofuel action (respectively with tax exempt for ethanol-gasoline blends in 1978 and the pro-alcohol program in 1975).

Figure 1 Global biofuels production 1975-2005



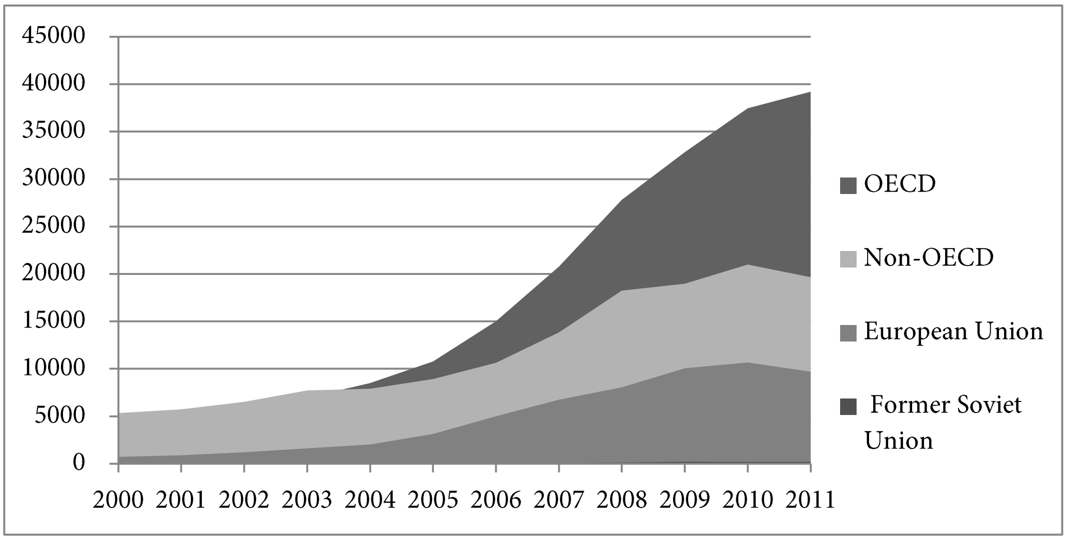
Source: EarthTrends, 2007

While environment protection and high energy prices might have drawn back attention to biofuels, the expansion of the sector has been certainly driven by fiscal incentives, subsidies, targets and mandated that governments put in place all around the word (Figure 2).

This chapter aims to provide an insight into the biofuels sector and illustrate the centrality of the debate on technological improvements. In order to achieve this, the chapter is structured in the following way: in the next section we provide definition and description on the two main biofuels products (bioethanol and biodiesel), we illustrate technological differences in terms of raw material production and conversion processes and biofuels uses in transportation. Then we discuss some environmental, social and economic issue related to biofuel production, commercialization and final use. Finally policy treatment of biofuels will be reviewed.

Figure 2. Global biofuel production 2000-2011

Thousand tonnes of oil equivalent



Source: our elaboration on BP, 2012

Classification and definition of liquid biofuels

Biofuels cover a wide range of fuels derived from biomass, which is material of biological origin excluding material embedded in geological formation and/or transformed to fossil. Biofuels include:

- liquid biofuels: fuels and bio additives such as bioethanol, biodiesel, biobutanol, biomethanol, bioETBE (ethyl tert-butyl ether), bioMTBE (methyl tert-butyl ether), bio gasoline, and combustible oils produced by plants;
- gaseous biofuels: such as biogas, mainly methane and carbon dioxide produced by the process of anaerobic digestion of biomass;
- solid biofuels: such as wood pellets, wood chips and charcoal, including char-briquettes.

In the further pages we analyse only liquid biofuels. In particular we concentrate on the two most commercially available: biodiesel and ethanol. In the next section of the chapter we will illustrate extensively biofuels' raw materials and conversion processes that can be summarized as follow:

- Biodiesel production is essentially based on vegetables oils, waste oils and animal fats extraction and conversion into fuel trough esterification by using alcohols. The most advanced production techniques include gasification and catalytic conversion of biomass to liquid fuels (BTL) and hydrogenation of fat and oil,
- Bioethanol production mainly consists in conversion of starch or sugar-rich biomass into fermentable sugar. The most advanced production techniques include hydrolysis of ligno-cellulosic biomass.

Biofuels are an energy source derived from organic material of diverse origin and they show some interesting features from an economic, environmental and social[1] point of view. Biofuels are:

- renewables
- biodegradables
- partially composed by oxygen, natural element which allow the abatement of gaseous polluting components in vehicles' emissions (in particular carbon monoxide)
- have a low sulphur level, allowing the reduction of sulphur dioxide
- have a high energy content, comparable with fossil fuels
- produced under favourable circumstances, they show a positive energy balance.

There is a large discussion on how biofuels can be classified because of the numerous and not clearly defined criteria that can be used.

According to directive 2003/30/EC of the European parliament, "biofuels' means liquid or gaseous fuel for transport produced from biomass", where "biomass' means the biodegradable fraction of products, waste and residues from

[1] While the economic and environmental value of biofuels appears evident, the social implications of biofuel consumption are trickier to deduce. For example, a very important social implication is given by the reduction in children and women lung cancer mortality due to indoor smoke (Fullerton *et al.*, 2008). Shifts from stoves fuelled by solid biomass towards stoves fuelled by bioliquids (which can be used in the so-called modern cook stoves) contribute to limit the indoor pollutant emissions. As recalled by Hilary Clinton, sponsor of the Global Alliance for Clean Cookstoves, indoor smoke mow down every year more victims than malaria.

agriculture (including vegetal and animal substances), forestry and related industries, as well as the biodegradable fraction of industrial and municipal waste"

The European directive provides a list of what can be unequivocally considered biofuel:

"At least the products listed below shall be considered biofuels:

- 'bioethanol': ethanol produced from biomass and/or the biodegradable fraction of waste, to be used as biofuel;
- 'biodiesel': a methyl-ester produced from vegetable or animal oil, of diesel quality, to be used as biofuel;
- 'biogas': a fuel gas produced from biomass and/or from the biodegradable fraction of waste, that can be purified to natural gas quality, to be used as biofuel, or woodgas;
- 'biomethanol': methanol produced from biomass, to be used as biofuel;
- 'biodimethylether': dimethylether produced from biomass, to be used as biofuel;
- 'bio-ETBE (ethyl-tertio-butyl-ether)': ETBE produced on the basis of bioethanol. The percentage by volume of bio-ETBE that is calculated as biofuel is 47 %;
- 'bio-MTBE (methyl-tertio-butyl-ether)': a fuel produced on the basis of biomethanol. The percentage by volume of bio-MTBE that is calculated as biofuel is 36 %;
- 'synthetic biofuels': synthetic hydrocarbons or mixtures of synthetic hydrocarbons, which have been produced from biomass;
- 'biohydrogen': hydrogen produced from biomass, and/or from the biodegradable fraction of waste, to be used as biofuel;
- 'pure vegetable oil': oil produced from oil plants through pressing, extraction or comparable procedures, crude or refined but chemically unmodified, when compatible with the type of engines involved and the corresponding emission requirements.

Further than product specific classification, an alternative classification criteria is represented by biofuels generation: as presented in Table 1, conventional biofuels (first generation) are produced at commercial scale following well-understood technologies, advanced biofuels (from second to fourth generation) are comparatively immature and there is large room for further deployment.

Table 1 Conventional and advanced biofuels

Conventional Biofuels
Bioalchools (especially ethanol) and biodiesel derived from food crops rich in sugar (such as sugarcane and sugar beet) or starch (such as wheat and corn) or vegetable oil (such as palm oil, soybean and rapeseed). They also include product obtained from chemical combination of biologic molecules with fossil molecules, such as ETBE.
Advanced biofuels
Bioalchool based on ligno-cellulosic biomass, HVO (Hydrotreated vegetable oil), BTL diesel and Bio-SG (bio-syntetic gas), algae-based biofuels, sugar conversion into diesel-type biofuels, artificial photosyintesys reactions (solar-to-fuel) and hydrocarbons obtained from genetically modified organism.

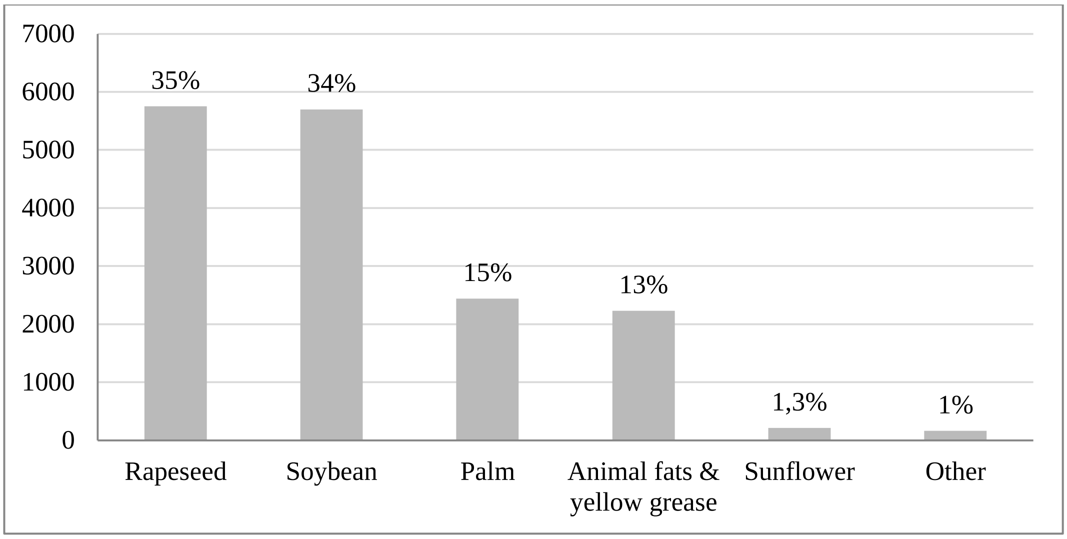
Source: IEA, 2008

Biodiesel

The term Biodiesel is commonly used as a joint name to indicate a number of liquid renewable fuels formally known as Fatty Acid Methyl Esters (FAMEs) which are characterized by a lower viscosity than the corresponding fossil fuel (diesel, obtained through fractional distillation of crude petroleum). Up to now, biodiesel is mainly derived by extensive oleaginous plants such as rapeseed and soya (Figure 3).

Using biodiesel to replace diesel oil can end up in a considerable GHG and local air-pollutant (especially PM_{10}) reduction. However it is well known that neat FAMEs exhibit a corrosive effect on metal engine components which may hamper the diffusion of its use as fossil fuel additive. In other words, the addition of biodiesel to diesel typically increases the corrosive power of the base fuel. Hence, high-level biodiesel blends can damage engines and fuel systems that previously used petroleum diesel by degrading seals and clogged fuel filters.

Figure 3 Feedstock (thousand tonne) usage for bioethanol production, 2010



Source: our elaboration on F.O. Litch 2011

Even though modern diesel engines are becoming more and more adaptable to biodiesel[2], over recent years the number of manufacturers allowing biodiesel as a

---

[2] Abstract of Volvo biodiesel statement (3rd March 2010)
Volkswagen Group does approve the use of diesel fuel containing up to 7% Biodiesel in all of their diesel vehicles. The use of this fuel requires no modifications to the vehicle or changes to the vehicles' maintenance schedule. Volkswagen Group does not approve the use of B30 Biodiesel in any of its vehicles. B30 Biodiesel is a blend of 70% Fossil fuel and 30% Biofuel, derived from Fatty Acid Methyl Ester (FAME). The use of this fuel in Volkswagen Group vehicles may invalidate the engine and exhaust system warranty. Certain Volkswagen Group vehicle models are approved to run on 100% Rapeseed Methyl Ester (RME) Biodiesel. 100% RME Biodiesel compatible parts fitted during manufacture of the vehicle is denoted by the vehicle PR code 2G0, which is found on the data sticker; however some of the older vehicles may not have this PR code. In vehicles that are 100% Biodiesel compatible, Volkswagen Group has only approved the use of Rapeseed Methyl Ester (RME) to standard DIN EN 14214. No other Biodiesel can be used. Important:

- Vehicles that do not have the factory preparation for Biodiesel cannot use 100% Biodiesel.
- Vehicles fitted with a Common Rail fuel injection system cannot use 100% Biodiesel.
- Vehicles fitted with Pumpe-Duse injectors cannot use 100% Biodiesel.

fuel beyond low blending percentages has decreased. This is especially due to the use of more advanced exhaust gas treatment technology, for example regenerating particulate filters: when regenerating the filters the exhaust gas temperature is increased by late injection of the fuel in the cylinder. As biodiesel has a higher boiling point than diesel oil, a consequence might be that biodiesel is mixed with the engine oil. This may be a minor issue with B5 or B7 blends but it might cause problems with higher biodiesel percentages.

On the other hand, heavy duty vehicles accept biodiesel in their engines demanding a contract with the consumer in which is defined what special measures the consumer has to take when using biodiesel (e.g. biodiesel has to meet certain requirements and the range of the service intervals has to be decreased). Furthermore, the engine has to be adapted to biodiesel when it comes to sealing, gaskets and rubber hoses, otherwise these parts will be affected after a little time and will start to decompose resulting in leakages (Biofuel cities, 2008).

For this and other reasons, the blends percentage varies significantly among countries. Table 2 summarizes biodiesel blending mandate and targets.

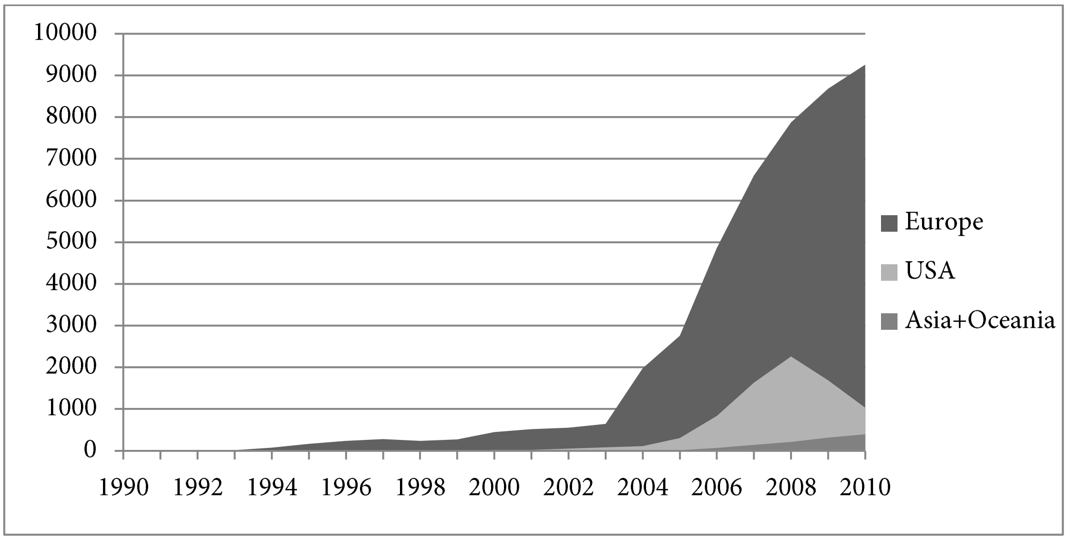
Organically derived oils such as vegetable oils (including recycled vegetable oils) and animal fats can be subjected to a transesterification process with an alcohol (typically a C1 to C5 alcohol) to form the corresponding fatty esters, typically mono-alkylated. This process, which is suitably either acid- or base-catalysed such as with the base KOH, converts the triglycerides contained in the oils into fatty acid esters and free glycerol, by separating the fatty acid components of the oils from their glycerol backbone (Crawshaw *et al.*, 2006).

- Vehicles fitted with a Diesel Particulate Filter (DPF) cannot use 100% Biodiesel, however EN590 containing up to 7% Biodiesel is approved for use with DPF.
- 100% Biodiesel vehicles cannot be used in temperatures below -10°C.
- The use of 100% Biodiesel may slightly reduce the driving performance of the vehicle and slightly increase the fuel consumption.
- The use of 100% Biodiesel may increase the frequency of the maintenance schedule on the vehicle.

Up to now, transesterification represents the most economically viable production process, because it is carried exploiting low temperatures and low pressures. This process allows easily both small and large scale production, making transesterificated biodiesel one of the most widespread biofuel.

According to EurObserv'ER, European countries are leading the biodiesel market: in fact, Europe feeds 70% of its road transport fuel demand trough diesel. On the contrary, diesel engines are less common in USA and they represent only 20% of the global demand.

Figure 4 Biodiesel production by region 2000-2010 (kt)



Source: Iea 2012

European countries exhibit substantial differences in term of biodiesel production. In particular Germany, after the biofuel sustainability law came into force (1 January 2011) is the first country to comply with the sustainability prescriptions of the Renewable Energy Directive (Eurobserv' Er Barometer, 2012).

Recently, world leadership in biodiesel production has been brought into question due to rapid expansion of production coming from developing countries such as Malaysia and Indonesia. I 2005 12 million of hectares were cultivated for biodiesel production from palm oil of those 4 million were in Malaysia and 5,3 in Indonesia.

Table 2 Biodiesel blending mandates (M) and targets (T)

Country	Current M/T	Future M/T	Status
Argentina	7%	10 (2015)	M
Australia (New South Wales)	2%	5% (2012)	M
Bolivia	2.5%	20% (2015)	T
Brazil	5%	n.a.	M
Canada	2-3% (3 provinces)	2% (nationwide)	M
Chile	5%	n.a.	T
Colombia	10%	20% (2012)	M
Costa Rica	20%	n.a.	M
Dominican Republic	n.a.	2% (2012)	n.a.
European Union	5.75% biofuels	10% renew. energy in transport	T
Fiji	5%	n.a.	T
India	n.a.	20% (2017)	M
Indonesia	2.5%	5% (2015), 20% (2025)	M
Jamaica	n.a.	Renewables in tra.: 11% (2012), 12.5% (2015), 20% (2030)	M
Japan	500 Ml/y (oil equival.)	800Ml/y (2018)	T
Korea	2.5%	3% (2012)	M
Malaysia	5%	n.a.	M
Mozambique	n.a.	5% (2015)	n.a.
Norway	3.5% biofuels	5%	M
Paraguay	1%	n.a.	M
Perù	5%	n.a.	M
Philippines	5%	n.a.	M
South Korea	2.5%	n.a.	M
South Africa	n.a.	2% (2013)	M
Taiwan	2%	n.a.	M
Thailand	3%	5%	M
Uruguay	2%	5%	M
United States	n.a.	1 billion gallons (2012)	T
Vietnam	n.a	50 Ml (2020)	n.a.
Zambia	10%	n.a.	n.a.

Source: Iea 2011, Biofuel digest website

Vegetable oils are mainly derived from rapeseed, soya, palm and sunflowers and from food industry waste, but it can be obtained from all oleaginous material. On earth there are more than 4000 oleaginous plants, most of which daily used at small and large scale. By this way, oleaginous plants still represent a partially underexploited source of vegetable oils and biodiesel. Researchers in Nevada noted that spent coffee grounds contain between 11 and 20 per cent oil by weight. That is about as much as traditional biodiesel feedstock such as rapeseed, palm, and soybean oil. Growers produce more than 16 billion pounds of coffee around the world each year. The used or "spent" grounds remaining from production of drinkable coffee wind up in the trash or find use as soil conditioner. The scientists estimated that spent coffee grounds can potentially add 1,2 billion litres of biodiesel to the world's fuel supply. Coffee-based fuel, which actually smells like coffee, had a major advantage in being more stable than traditional biodiesel due to coffee's high antioxidant content. Solids left over from the conversion can be converted to ethanol or used as compost (Narasimharao *et al.*, 2008). In November 2009, the flight company KLM operated the world's first demonstration flight with passengers on board using oleaginous biofuel with one engine running on a mix of 50%. On 29 June 2011, the same company operated the first commercial flight on bio kerosene (made from recycled cooking oil) with 171 passengers on board. [3]

In the biodiesel production process a number of co-products can be obtained (Cooper & Weber, 2012). For example, palm oil is usually divided in solid and liquid fractions suitable for a variety of food and non-food purposes. Biodiesel derived from animal fat has several co-products as well: feed ingredients for livestock and aquaculture and soap. In general, transesterification reaction

[3] According to KLM website, compared to 2009, KLM already strives for a 20% CO2 reduction per ton/kilometer as per 2020 of the total air transport industry reduction target (by 50% in 2050). A further example is given by some Scottish bus companies that power their fleet using cooking oil and reduce bus fares passengers who supply exhausted oils.

produces glycerin, a co-product generated by the treatment of intermediate high fatty acid/ester streams. Glycerin is an odourless, viscous liquid with a sweet taste, it is widely used as chemical component and finds many applications in pharmaceutical and cosmetics. The amount of crude glycerin generated in the transesterification process is approximately 0,4Kg per 1 litre of biodiesel. In 2010, a survey of the National Biodiesel Board (NBB) reported that 48% of their member sold glycerin for a variety of purposes.

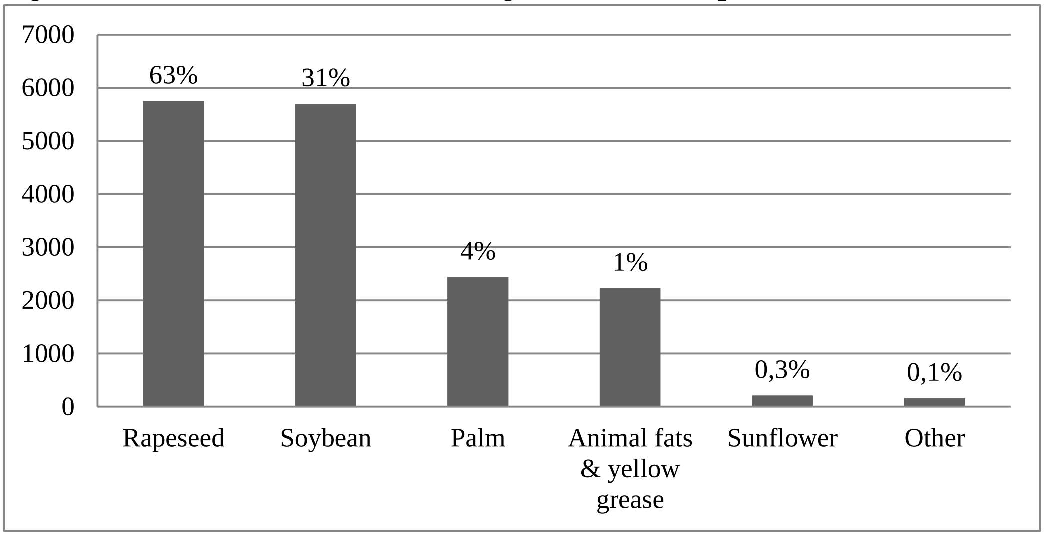
Another biodiesel co-product is fatty acids which are used to produce numerous goods such as candles, lubricants, cosmetics animal feed and asphalt. Moreover, vegetable oil meal (in particular derived from soybean) can be considered as an “indirect” co-product of biodiesel. As explained by Cooper and Weber, 2012), crushed oilseed crops generate both crude oil and oilseed meal, which is used as protein feed component in the livestock sector, Considering that only 20% of the oilseed is actually oil (while the remainder is oil meal) and considering that oilseed meal has historically been the main economic driver of oilseed crushing and it would be produced regardless of the oil demand, oilseed meal is not acknowledged as a biodiesel co-product.

Bioethanol

Bioethanol (or ethanol) is a liquid colourless biofuel obtained by sugar-rich or starch-rich biomass sources such as sugar beet, sugarcane, corn, wheat, barley and other cereals (Figure 5).

After a decade during the 1960’s and 1970’s, when methanol was tested as a vehicle fuel, bioethanol has become the main substitute for fossil fuels in transport. Bioethanol is mainly produced following well-established conversion processes: it can be derived from raw material fermentation as long as it contains sugar or starch that can be hydrolysed to sugar. A less common way to produce bioethanol involves biomass gasification, conversion process that increase the total supply of biomass compared to fermentation, since the only demand on/requirement for the biomass in this case is that it has to contain carbon. (Biofuel cities, 2008).

Figure 5 Feedstock (thousand tonne) usage for bioethanol production, 2010



Source: our elaboration on F.O. Litch 2011

The main disadvantage of conventional generation bioethanol is that, with the exception of sugarcane ethanol, it is produced using crops that are a fundamental part of worldwide diets and there are big concerns related to bioethanol adverse impact on food supply. These concerns brought growing attention to advanced generation bioethanol produced from lingo-cellulosic feedstock. Although there is a lot of research going on to overcome the technical and economic challenge posed by advanced generation bioethanol, the road to full commercial deployment seems to be very long.

Bioethanol is commonly used as vehicle fuel blended in low amounts in fossil gasoline, mostly 5%. In such a small percentages, bioethanol-gasoline fuel does not endanger engines and other components and do not require any modification (except for very old vehicles). Higher blend, up to 20%, can be used in very modern fossil gasoline engines with no adaptation, achieving full engine performance and fulfilling the EU directives on exhaust gas emissions.

Flexi Fuel Vehicles (FFV), which can be powered by any blend of gasoline and bioethanol from 0 to 100 % bioethanol constitute the majority of Brazilian car fleet. Although FFV's engine and the fuel injection system automatically adapts to any blend, 100% bioethanol fuel could be difficult to use especially with low external temperatures. In the light of this, the highest mix of gasoline and bioethanol commercially available in Europe is E85.

The profitability of the bioethanol production is strongly influenced by the exploitation of the co-products. Grains such as maize, wheat, barley and sorghum are common feedstock for both the bioethanol and feed industry. Grains transformation into biofuel basically occurs by mean of dry or wet milling processes. These processes have as co-product goods that can be directly sold into the animal feed market (as CDS – condensed distillers soluble) or processed and then sold (DDGS – distillers dried grain with soluble, starch-gluten suspensions).

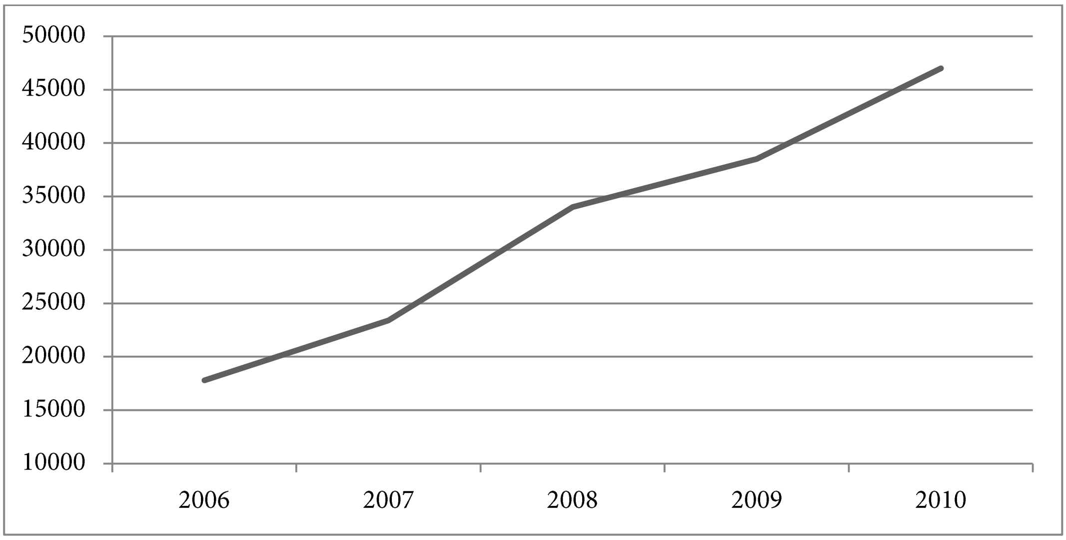
F.O. Licht (2011) estimated that in 2010 142,5 million tonnes of grain was absorbed by the ethanol industry, the 6.3% of the total grain use. Considering that one third of the grain processed to ethanol was actually used to produce animal feed, this volume can be split into 95 million tonnes for fuel and 47,5 million tonne for feed (Figure 6). Even though grains are the main sources of bioethanol co-products, other feedstock provides co-produce marketable goods. The fibrous component of the sugar beet is sold as an animal feed ingredient. Maize oil can be removed from the stillage at the back end of the process in the dry-mill maize ethanol production process.(Cooper & Weber, 2012 and used in feed or biodiesel production.

Table 3 Bioethanol blending targets and mandates

Country	Current M/T	Future M/T	Status
Argentina	5%	10 (2015)	M
Australia (New South Wales)	6%	5% (Queensland on hold until end 2011)	M
Bolivia	10%	n.a.	T
Brazil	18-20%	n.a.	M
Canada	5% (up to 8% in 4 provinces)	n.a.	M
Chile	5%	n.a.	T
China	10% (9provinces)	n.a.	T
Colombia	10%	n.a.	M
Costa Rica	7%	n.a.	M
Dominican Republic	n.a.	15% (2012)	n.a.
Ethiopia	10% (Addis Abeba)	n.a.	M
European Union	5.75% biofuels	10% renewables in transport	T
Fiji	10%	n.a.	T
India	5%	20% (2017)	M
Indonesia	3%	5% (2015), 15% (2025)	M
Jamaica	10%	Ren. in transp.:12.5% ('15), 20% ('30)	M
Japan	500 Ml/y (oil equivalent)	800Ml/y (2018)	T
Kenya	10% (Kisumu)	n.a.	M
Malawi	10%	n.a.	M
Mexico	2% (3 cities)	n.a.	M
Mozambique	n.a.	10% (2015)	n.a.
Norway	3.5% biofuels	5%	M
Nigeria	10%	n.a.	T
Panama	2%	5% (2014), 7% (2015), 10% (2016)	M
Paraguay	24%	n.a.	M
Perù	7.8%	n.a.	M
Philippines	10%	n.a.	M
South Africa	n.a.	8% (2013)	M
Taiwan	3%	n.a.	M
Thailand	3 Ml/day	9 Ml/d (2017)	M
Uruguay	n.a.	5% (2015)	M
United States	n.a.	Advanced biofuels (2.0 billion gal) Cellulosic biofuels (3.45 – 12.9 mgal)	T
Vietnam	n.a	50 Ml (2020)	n.a.
Zambia	5%	n.a.	n.a.

Source: Iea 2011, Biofuel digest website

Figure 6 Trend (2006-2010) in global production (thousand tonnes) of grain ethanol animal feed co-products



Source: F.O. Licht, 2011

Biofuels international political framework

In this part of the chapter we illustrate the current worldwide biofuel policy instruments, giving particular emphasis to those applied in the three leading biofuels markets: United States, European Union and Brazil.

Biofuel production increased exponentially under the pressure of national governments demand for energy security and energy sources that would allow the abatement of Co2 emissions. A further boost to biofuel policies came both in the US and Eu by the attempt to secure new source of income to the rural sector. While in favourable circumstances producing energy from biofuels is cost competitive, in many cases economic incentives are needed to off-set cost differences between biofuels and fossil fuel production costs.

In order to achieve environmental and energy-security goals, several countries are imposing blending targets and mandates (see section above) that impacts on biofuel production and consumption as well on biofuels-associated sectors, and primarily the raw material industry. The primary sector is then affected through direct and indirect land use change, agricultural practices adjustments, water and other inputs use, up to final consumption both in developed and developing countries.

In the following section, biofuels policies are analysed, having been classified into demand side and supply side instruments.

Demand side instruments

Targets

In the European biofuel directive, guiding targets are set for the transport sector aiming to increase the share of biofuels used in transportation. The directive states that biofuels should constitute 2% of total energy use in the sector by 2005 and 5.75% in 2010. In the Renewable Energy directive (RES-directives) there is now a unique binding target (equal for all member states) saying that the use of renewable fuels in the transport sector should be 10% in 2020. In addition, the fuel quality directive (FQD) strongly influences the potential use of biofuels by setting limits to how much blending of biofuels into diesel and gasoline that is allowed. According to this directive (which was to be implemented by 2010-12-31) it should be allowed to blend 10 % into gasoline and 7% vol. into diesel (previously the limits for both fuels were 5% vol.). Further, according to the FQD, fuel suppliers in the union should reduce GHG emissions by 6% per energy unit until 2020. For the latter target only the selling of biofuels fulfilling the sustainability criteria set in the RES directive could be accounted for.

Brazilian government started to support ethanol in 1975, following the petroleum shortage caused by the first OPEC oil cut. With the Proalcohol program, Brazilian government opened the door to a series of actions aiming at addressing energy security and sustainability issues. This program consisted in a variety of instruments organised under three approaches: global, regional and bilateral. Global by adopting international standards to facilitate the creation of a world market, regional by strengthening south American energy integration and sources diversification, and bilateral, by fostering scientific and academic exchanges (GBEP , 2007).

US biofuel policies were initiated already in 1970 motivated by environmental and energy security issues. In 2005, the US Energy Bill established a mandate requiring minimum levels of biofuel consumption from 11.9 million tons in 2006 up to 22.1 million tons in 2012. The Energy Independence and Security Act of 2007 expanded the 2005 program by adding a biodiesel mandate and expanding the specific total mandated quantity of renewable fuel to be blended into transport fuel to 9 billion. The target has been extended to year 2022 specifying that at least 21 billion gallons of the new mandate have to come from second or higher generation biofuels (Janda et al, 2012).

Tax exempt and other fiscal incentives for biofuels and flex-fuel vehicles

In addition to the Biofuels directive, European countries adopted an Energy taxation directive (2003/96/EC) which allows member states to grant tax reductions or exemption on biofuels, in order to contribute to technological development of more environmentally friendly fuels. In Europe the procedure of granting tax exempts is regulated by EU state aid rules which prohibit overcompensation. Overcompensation takes place when a tax reduction or exempt makes the biofuel cheaper than the fossil (conventional) fuel it replaces[4].

[4] In 2011 the Swedish National Audit Office performed an audit of the tax exemption for biofuels in Sweden and according to it, the tax reduction (of CO_2 and energy tax) for biofuels is the single most important instrument for increasing the amount of biofuels in Sweden. However, the main conclusion drawn by the SNAO is that the tax exempt for biofuels contributes to reaching the climate objective set by the Swedish parliament but not at a reasonable cost. It has been necessary to increase the use of biofuels but it is a relatively expensive way to reduce GHG emissions. SNAO

European countries are implementing both carbon tax and energy tax on fossil energy. In some countries the energy tax is determined individually for each fuel and is not proportional to the energy content of the fuel, in other countries the tax increase proportionally with fuel market price. Carbon tax is commonly proportional to the carbon content of the fuel.

Brazilian ethanol taxation varies from region to region. On average, tax level on ethanol with respect to gasoline is substantially lower (~50% of consumers price for fossil fuel vs. 15-20% for ethanol). In the light of social inclusion concerns in Brazil development goals, a system of tax incentives and subsidies has been established for the supply of biodiesel raw materials from small farms in the North and North-East regions of Brazil, especially in the semi-arid areas ("Social Fuel" Label-Decree 5.297/04 – 5.457/05). To facilitate emissions reduction of pollutants from automobiles sources (Law 8723 – National Council for the Environment which establishes mandatory actions from automobile and fuels producers to reduce emissions levels of CO, NOx and other pollutants according to Brazilian environmental policy) Brazilian government established differentiated IPI (Tax on industrialized good) rates for vehicles running on anhydrous ethanol and gasoline (GBEP, 2007).

US fiscal incentives differ depending on the state and they are completed by those at the federal level. The Energy Tax Act of 1978 introduced tax exemptions

---

state that the tax exemption has been necessary to create the market for low-blend of biofuels in gasoline and diesel. The tax exempt has not been a general exempt but decided upon based on individual applications sent in by fuel suppliers (who are the ones that pay the tax). The decisions have mainly been short term (one- two years at a time) and not always granted on the same reasons. Further SNAO claims that the tax exempt has not had a significant effect on technology development nor can it be claimed as technology neutral, due to low transparency and inequity in treatment of companies. In order to be allowed to grant tax exempts Sweden has agreed to submit annual reports to EU and thereby the tax exempts are approved. Currently there is an approval from the EU that Sweden can continue with this exempt until 2013, an approval that possibly could be prolonged but not longer than until 2020.

and subsidies for the ethanol blended gasoline. Biodiesel subsidies were introduced with the Conservation Reauthorization Act in 1998.

Consumers benefit from the blender's tax credit only when the tax credit is combined with a binding blend or consumption mandate. However, consumers do not benefit when the tax credit is the only binding policy as in this case, final price does not depend on the level of the tax credit (De Gorter & Just, 2008). The Volumetric Ethanol Excise Tax Credit and the Volumetric Biodiesel Excise Tax Credit provide the largest subsidies to biofuels, while there are some smaller additional subsidies connected to biofuel outputs both on the states and federal levels (Janda *et al.*, 2012).

Mandatory blending rates

European countries have a common mandatory blend rate that represents the minimum blending level. Each country can establish a higher blend level, accordingly to fuel quality requirements, which can depend on regional climate conditions.

Brazilian blending mandates are different according to filling stations. In the ethanol-gasoline blend (anhydrous ethanol mixed whit conventional gasoline) ethanol goes from a minimum of 18% to a maximum of 25% (Law 737-1938, MAPA). On the contrary, blends for flex-fuel vehicles are imposed by cars' apparatus, which can run up to 85% of anhydrous ethanol or on up to 100% hydrous ethanol (Janda *et al.*, 2012). In 2008, the Ministry of Mines and Energy *National Program for The Production and Use of Biodiesel* has introduced *The Biodiesel Law* (Law 11.097/05) which established biodiesel blending targets at 2% that has been raised to 5% in 2013.

Other instruments

On the demand side, the following instruments indirectly affect the demand for biofuels by increasing the demand for vehicles that use alternative fuels. Those instruments are variably implemented in all developed countries and in some big developing countries:

- CO2-based vehicle tax or green vehicles tax-reduction
- Environmental car premium
- Reduced taxable value for fringe benefits
- Free or reduced parking fees
- Congestion tax - exempt for environmental cars
- Standards for car producers

Supply side instruments

Obligations for fuel suppliers

In April 2009 the European parliament adopted biofuels sustainability criteria in the 2009/28/EC directive "On the promotion of the use of energy from renewable sources". The directive sets out sustainability criteria for biofuels, irrespective of whether raw materials are cultivated inside or outside European territory. Article 17, 18, and 19 present criteria related to greenhouse gasses savings, land with high biodiversity value, land witch high carbon stock and agro-environmental practices. Articles 17 states that biofuels and bioliquids shall be taken into account for the compliance with the directive obligations[5] only if they fulfil the following sustainability criteria:

[5] Measuring compliance with the requirements of the directive concerning national targets; measuring compliance with renewable energy obligations; eligibility for financial support for the consumption of biofuels and bioliquids.

- The greenhouses gasses emission saving from the use of biofuels shall be at least 35%[6],
- Biofuels shall not be made from raw material obtained from land with high biodiversity value,
- Biofuels shall not be made from raw material obtained from land with high carbon stock (such as wetlands and continuously forested areas) or that was grassland or peatland before 2008.

Article 18 clarifies the raw material and biofuels with high sustainability features can compensate for less sustainable products. In the fuel quality directive there is a requirement for fuel suppliers to report and to reduce greenhouse gas emissions by up to 10% on a life cycle basis by 2020 (compared to 2010 level). There are three possible ways of fulfilment:

- Increase the use of biofuels, alternative fuels and reduction in flaring and rejection at site of production
- To reduce emissions by applying CCS (carbon capture and storage) or electrical vehicles
- To buy CDM credits

Import/Export tariff and duty for biofuels

Ethanol produced in a third country and imported into the EU is associated with a tariff of 10.2 €/hl (denatured) and 19.2 €/hl (un-denatured). This tariff is kept since the product is seen as an agricultural product. Few countries are managing to get an exempt for the import of Brazilian ethanol, where the product instead is seen as a chemical product and the tariff to be paid is only ~0.25€/hl. However, in order not to violate the EU state aid rules of overcompensation,

[6] With effect from 1 January 2017, the greenhouse gas emission saving shall be at least 50%. From 1 January 2018 the greenhouse gas emission saving shall be at least 60% for biofuels and bioliquids produced in installations in which production started on or after 1 January 2017.

ethanol has to be used for low-blending at the higher tariff rate in order to get the tax exempt. In the EU, under the Generalized System of Preferences (GSP), EBA agreement and agreement with ACP countries, alcohol import from these countries occur at zero or at reduced tariffs. Biodiesel tariffs differ depending on the oil: from 4.99% for soybean oil, 9.6% for rapeseed oil.

The Brazilian sugar and ethanol market is protected from external competition by a tariff respectively of 20% and 30% (Excluding Mercosur countries).

The US apply a 1,9% ad valorem tariff on denatured ethyl alcohol imports and a 2,5% ad valorem tariff on ethanol. Moreover, US apply a specific duty on imports of ethyl alcohol for the production of biofuel, which has been recently reduced from 54 to 45c/gallon. This rate has been fixed in such a way as to withdraw benefits arising from taxation on imported products (45 c / gal). The duty affects predominantly imports from Brazil, which so far produce the most environmentally friendly ethanol, and has been highly criticized because of its contradiction with the aim of reducing emissions. Some US trade partner countries enjoy preferential treatment such as Mexico and Canada that, under NAFTA (North American Free Trade Agreement), can export ethanol in the U.S. at zero duty. Another key instrument of trade policy is the Caribbean Basin Economic Recovery Act (CBERA) which brings together Central America and Caribbean countries. According to CBERA agreement, when ethanol is produced in one of these countries using at least 50% of local raw material United States duty is zero (Zezza, 2011).

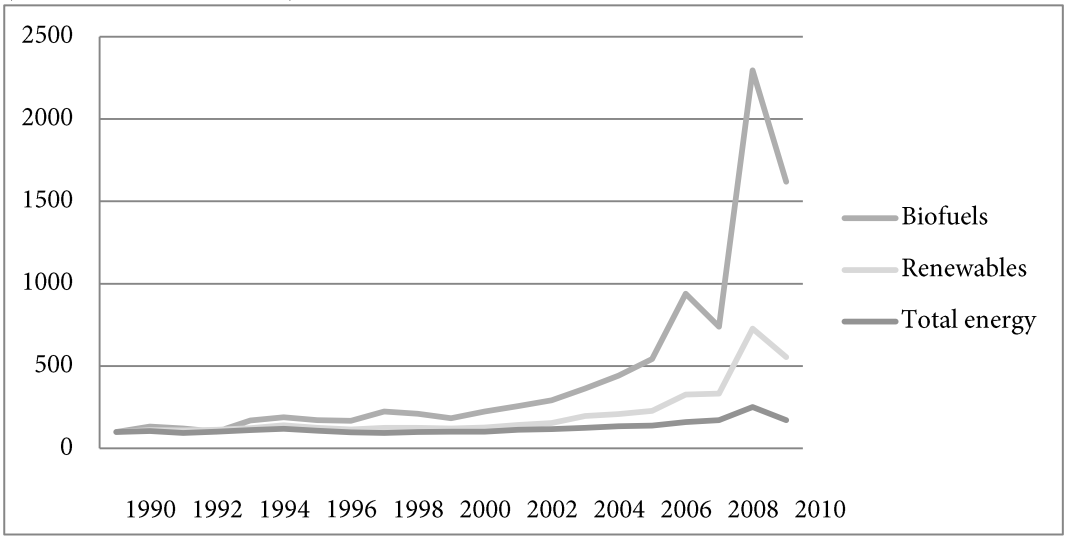
R&D expenditure

By looking at some stylized facts, as revealed by R&D expenditures by the public sector (Figure 7), during the period 1990-2008 efforts in the biofuels sector increased consistently more than in the other energy sectors, including all other renewable energy sources. At the same time, it is possible to highlight how sensitive the sector is to agricultural shocks by looking at the effect of the world food price crisis of 2007-2008, as a positive peak in 2006-2007 is suddenly followed by a dramatic reduction in R&D expenditure in 2008. It is worth noting that while the 2009 peak and the dramatic fall in 2010 is a common trend for renewables also, the rapid decrease occurred in 2007-2008 is specific for biofuels, thus revealing that it is a particularly fragile sector with respect to the overall international context.

Similar trends occurred also in the private sector, as patent applications in the same period reveal that new applications between 2007 and 2008 are spare with respect to the previous years (Figure 8).

The fact that a decrease in patent applications is substantially lower than the fall in public R&D expenditures might be explained by distinguished factors. The first and quite obvious one is the time lag between the input (R&D expenditures) and the output (patents) variable here compared. The second one is that public support to the biofuels sector is more influenced by public opinion than private sector investments. Firms investing in new technologies for producing biofuels would be less favourable to changing investment profiles in order to satisfy public opinion rather than national governments. Third, and probably more intriguingly for our purpose, the food crisis has influenced only a portion of the biofuels sector, namely the conventional generation side, while efforts in advanced generation technologies have continued and maybe expanded as a response to the food price increase and the food versus fuel debate.

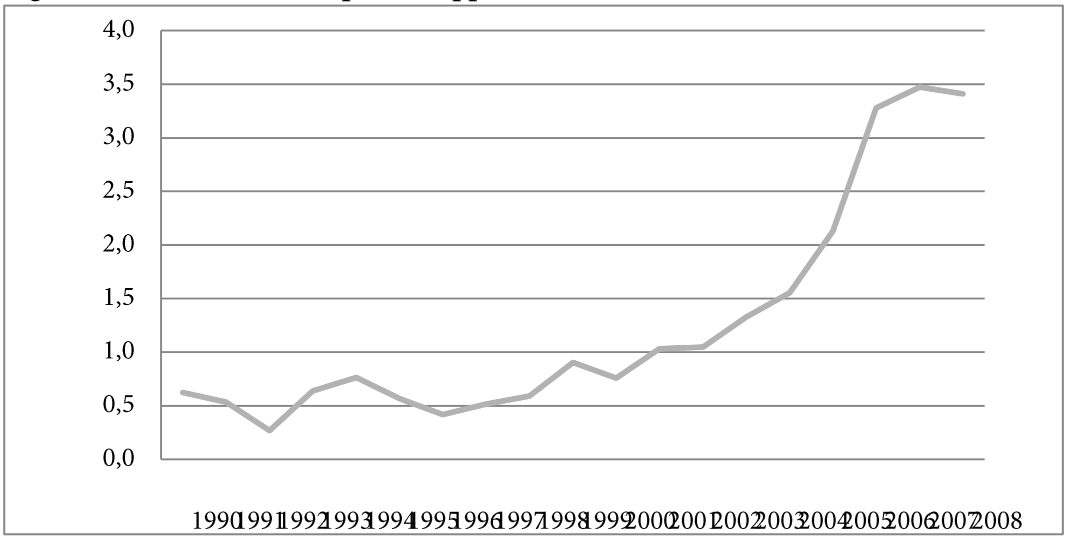
Figure 7 Trends in public R&D expenditures in selected energy sectors 1990-2010 (Index number 1990=100)



Source: our elaboration on IEA, 2011

While specific data on R&D expenditures in biofuels are available only for the public sector and for aggregated items, working with patents allows considering innovative propensity by both public and private firms as well as distinguishing theme by more specific technological domains.

Figure 8 Trends in biofuels patents application as % of total renewables, 1990-2008



Source: our elaboration on EPO applications by OECD-PATSTAT, 2012

In order to better identify the evolution of technologies in this sector it is necessary to disentangle innovative activities according to a punctual classification.

Other instruments

A premium for energy crops grown outside set-aside land is acknowledged according to the CAP (European Common Agricultural Policy during the period 2003-2008). In these years the total area used for energy crops more than doubled but the premium was abolished in 2009. The reason for the abolishment was that there was no longer considered to be a need for a specific support for energy crops, mainly due to the strong demand for these products on the international markets and the establishment of binding targets for bio-energy in total fuel by 2020. Before the premium was introduced in 2003 there had been some establishment of energy crops on the mandatory set-aside land, which allowed for production of non-food crops.

Some issues hampering the diffusion of first generation biofuels

The development and deployment of first generation biofuels started off whit the need for alternative sources of energy from an environmental and national security point of view. Moreover, bioenergy policies are considered as an instrument for creating employment in agriculture, improving farmers' resilience and preventing rural areas depopulation. While the expansion of biofuels production has been significant, the real contribution of biofuels in achieving these goals appears to be controversial. Concerns regard the environmental and social impact of large scale biofuels production and high tech input use such us fertilizers and harvest machinery. For example, the International Union for the Conservation of Nature affirms that monocultures for biofuels production are worsening biodiversity loss, pesticide dependency and plantations resilience to infestations (IUCN, 2008).

Some highly debated issues arising are reported in the following paragraphs regarding respectively the economic, social and environmental aspects of large scale biofuel production.

Economic aspects

Up to recent years, in absence of support policy[7], biofuels profitability mainly depended on three factors: feedstock availability and price (first generation biofuels are made through agricultural raw material transformation of sugar, starch or oil crops), feedstock production costs (note that transformation process costs are relatively constant among countries) and petrol availability and price (as production input and as substitute good).

It is worth remarking that biofuels supply chain is becoming more and more complex and interconnected with co/by-product industries. At the same time, feedstock availability and profitability opportunities are becoming greater. For example molasses, which is a by-product of the raw sugar production, is used as feedstock in the ethanol production. In Brazil most of the ethanol plants are equipped to produce both sugar and ethanol from sugar molasses, using their co-products as (free) feedstock for a different production. (According to F.O. Licht, in 2010 Brazil processed into ethanol more than 6,2 million tonne of molasses). Furthermore, by/co-products opens a way toward diversified end markets.

[7] So far, the main role in biofuel profitability has been played by government incentives and mandatory blends basically all over the world. In 2008, under the leadership of Dominique Strauss-Khan, former managing director, the IMF stressed the importance of biofuels subsidies in driving a whole sector that survive thanks to increasing oil prices and public generous support (IMF, 2008).

A rather new opportunity for biofuels is given by fuel/electricity co-generation which is profusely exploited by Brazil with sugar bagasse ethanol/electricity co-production.

We discuss in turn the three aspects mentioned above.

Biofuels and the food production chain

The energy and food production chain interact in different ways. On the one hand, biofuel production competes for feedstock with the food industry and contributes to agricultural prices volatility. On the other hand, biofuels could make livestock production less expensive and may provide a cheaper and more stable energy source.

Historically, agricultural prices and energy prices have always been connected[8], especially in poor countries, where the primary sector represents the pillar of the whole economy.

With the development of the biofuel industry, the link between energy and agricultural prices has become even stronger (Figure 9). An obvious example of this relationship is given by the Brazilian sugarcane industry, which in turn can produce sugar for the food industry or ethanol fuel. In Brazil these two markets are fully integrated: this is possible thanks to the flex-fuel fleet, which can run from 100% fossil gasoline to 100% bioethanol. Ethanol is also available on the entire national distribution network and therefore it can be bought with the same easiness of gasoline[9]. Given the absolute indifference for consumers in selecting

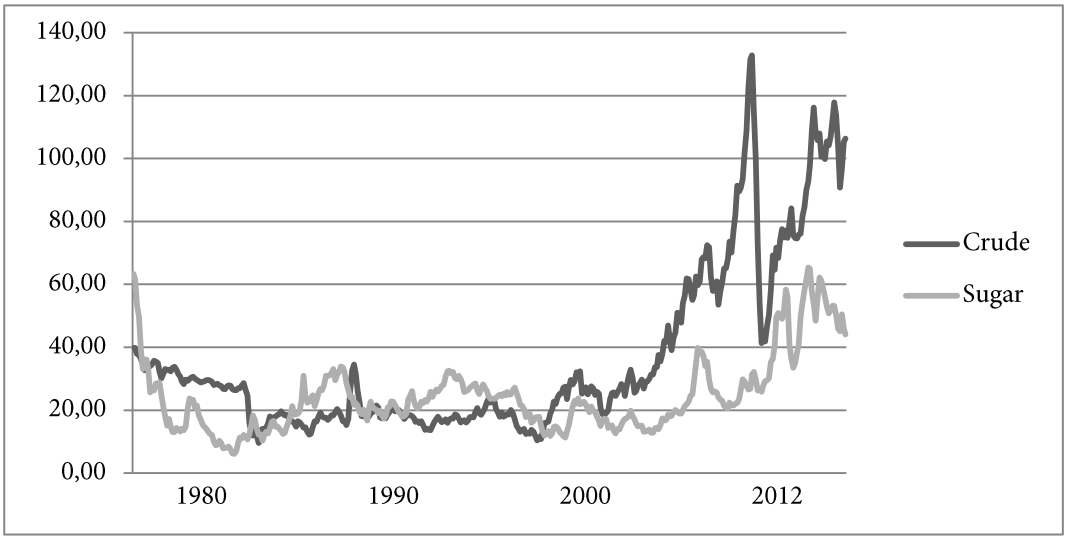
[8] Fossil energy is one of the main inputs of agricultural production. Energy sector provides fuels for the propulsion of harvest and other agricultural machinery. Moreover, the cost of energy stands for the great part of inputs prices, such as fertilizers and pesticides.

[9] Fuel must be easy to find. The lack of refill stations along the roads and the necessity to plan refuelling makes people refrain from using methane vehicles and other alternative fuels (Steenberghen & Lòpez, 2006)

gasoline or ethanol, their choice depends entirely on the fuel price. Similarly, from the production side, Brazilian sugarcane mills, which are usually able to indiscriminately supply both sugar and ethanol, take their decisions on the base of market prices. In Brazil, therefore, the increase in oil prices determines a decline in gasoline demand and an increase in the production and consumption of bioethanol and vice versa (Zezza, 2008). On the contrary, an increase of international sugar price results in a decrease of bioethanol production.

Figure 9 Trend in the price of oil and sugar 1980-2012

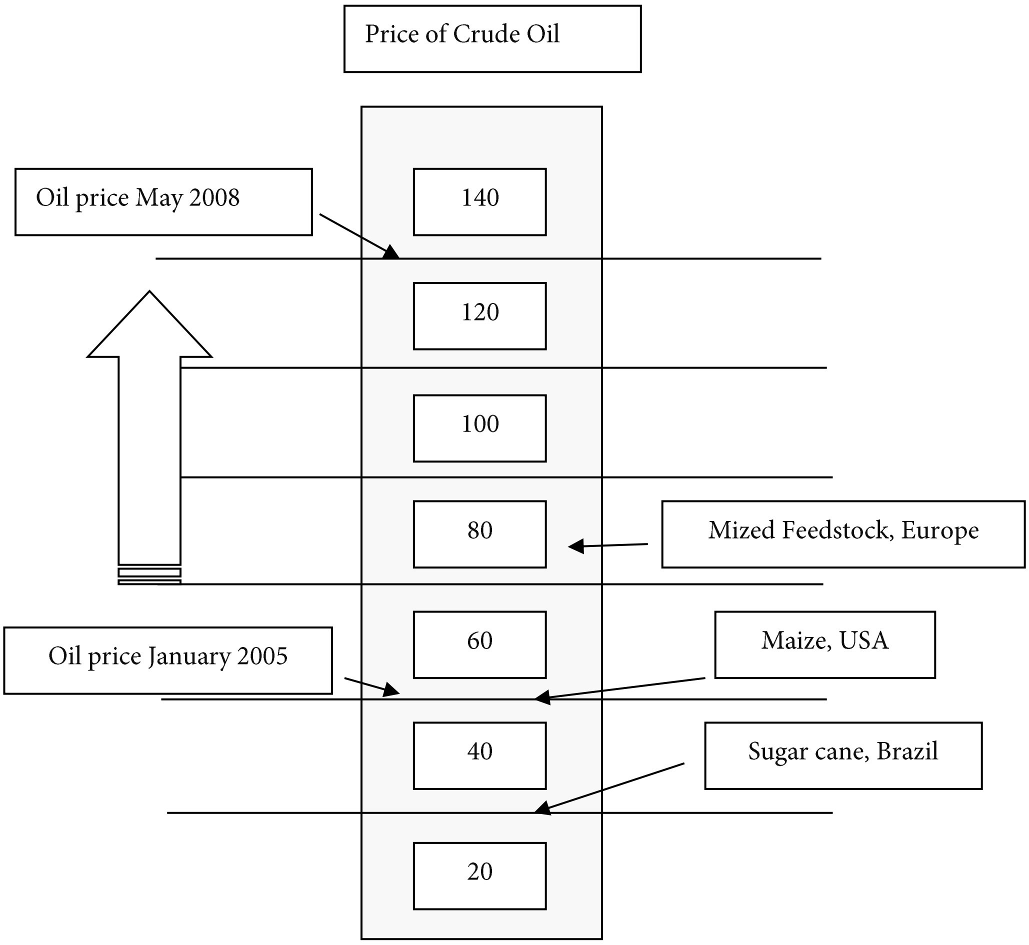
US$ per barrel (Crude oil) and tonne (Sugar)



Source: our elaboration on World bank 2013

As explained by Schmidhuber (2008), energy-feedstock connection allows energy prices to determine a minimum price for agricultural feedstock and, also, a maximum price. Through an analysis on the price of sugar and oil in Brazil from 2000 to 2008, Schmidhuber observed that the demand for bioethanol has set a minimum price for sugar equal to $ 35 per barrel. When the price of oil exceeds this threshold, producing bioethanol becomes profitable. The advantage in bioethanol production pushes up demand for sugar, influencing sugar prices (note that sugar price accounts for basically 80% of production costs). But, in the long run, if the price of sugar continues to rise, the cost of the raw material would make the production of bioethanol not more profitable, reducing the production of biofuel and thereby pushing down the demand for sugar (and, consequently, its price.)

Figure 10. Breakeven price for crude oil (US$) and other raw materials, 2005



Source: Sofa 2008

In the short run, this mechanism should operate up to expel from the market obsolete production facilities, which are not profitable with a sugar price above a certain threshold. It is important to emphasize how this mechanism stops operating in the presence of public support policies. Incentives and other policy instruments can introduce significant distortions in the biofuel market and thus keep alive economically inefficient production, letting agricultural raw material prices growing indiscriminately (Perry et al, 2004; IFPRI IMPACT 2050 model, CAWMA model, IAASTD model in the low agricultural investments scenario).

Production costs

Biofuels production costs are not the same for each country. Beyond the cost of processing, which seems to be very similar regardless production location, the variables involved in the production process differ greatly.

Producing biodiesel from vegetable oils in 2005 in Canada cost US$ 0.455 per litre. The same production cost in Brazil amounted to US$ 0.568 per litre (OECD). Again in Brazil, in 2005, bioethanol from sugarcane was produced at US$ 0.219 per litre while in the United States the cost was US$ 0.648. These profound differences can be explained through the evaluation of various factors affecting production costs (Figure 11).

For the majority of biofuels producers, the competitive advantage in the production of biofuels depends on the amount of public subsidies they receive. Without these incentives, the economic balance of most of the production would be negative.

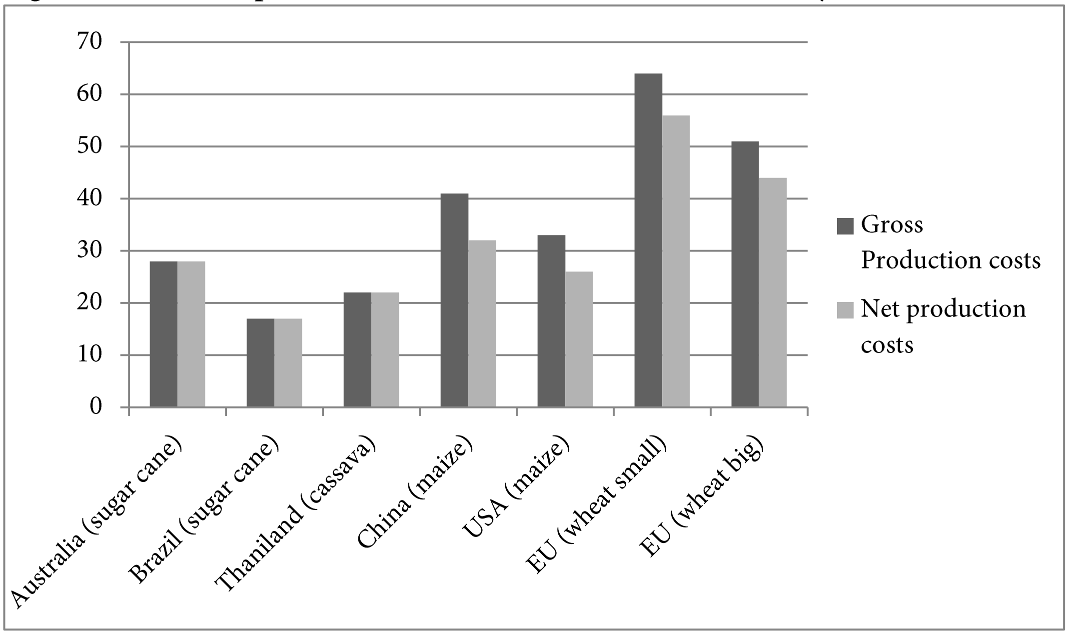
In addition to raw materials prices and public support, other factors influencing economic profitability in different countries (Figure 11) include the local cost of labour, production standards, product standards, local cost of energy and, by/co-products market.

Labour cost, production standards and product standards will be discuss respectively in the social aspect and environmental aspect sections.

Despite the standardization of transformation processes (for example biochemical or thermo-chemical processes), the quantity of energy used in the different stages of the production chain can vary significantly, especially in the fermentation and distillation phase (Zezza, 2008). Hence, depending on the local cost of energy and the level of production mechanization, the transformation of raw materials into biofuel may result more or less expensive.

Regarding by/co-products market and prices, a clear example comes from feed-fuel co-production: biofuels are made by exploiting sugar, starch or fat components of raw materials. This allow for the left overs (which can have a rich protein content) to be used in the production of feeding material for livestock industry (Esposti, 2008). Several studies have analysed the potential effect of animal feed supplies and prices and the consequent impact on numerous products of the agricultural sector (Elobeid *et al.*, 2007; USDA, 2007; Taheripour *et al.*, 2010).

Figure 11. Different production costs for raw materials and country (€/hl)



Source: IEA 2008b

On one hand, most of these analyses suggest that large scale biofuels production have a long-term impact on certain commodity prices indicating a slight reduction in production and consumption of milk, meat and eggs. On the other hand many of the studies do not take into account properly the recent increased supply of biofuels co-products. Recent analysis improved co-products substitution (Taheripour *et al.*, 2010) but still have failed in taking into consideration improvements in feed conversion efficiency nor the ability of DDGS to displace more than an equivalent mass of maize or soybean meal (Hoffman & Baker, 2011).

Fossil fuel price

Without public intervention, oil price influence on biofuels is twofold.

First, it represents a limit to biofuels market value, because biofuels are gasoline substitutes and their price conforms to fossil oil one. As long as biofuels price is higher than fossil fuels price, biofuels demand comes exclusively from blending obligations. This link also works in reverse: when the price of oil is high, the production of biofuels increase. Second, oil price is one of main part of biofuels production costs: oil is a basic input of large scale raw material production. When the price of oil increases, agricultural prices increase, creating a sort of counter balance in biofuels supply.

Consequently, despite oil price volatility introduce a source of uncertainty in the biofuels market, it also make it more attractive. In this respect, Vedenov (2006) explained how volatility of gasoline prices is able to contribute to foster substitute goods demand when they are characterized by a more stable price trend.

Environmental aspects

Many of the policies for climate change mitigation include measures to promote the production and use of biofuels. Biofuels, in fact, have a lower carbon content with respect to for fossil fuels. In terms GHGs emission reduction, first-generation liquid biofuels performs differently depending on raw materials type, origins and cultivation practices and, transformation process. So far, only few production chains show a favourable emission balance.

If we look at the big picture by taking into account fundamental phenomena such as deforestation, the expansion of cultivated land and other environmental impacts, it is questionable whether, in terms of reduction of C02, global production of first generation liquid biofuels achieved positive results[10]. Rajagopal and Zilberman (2007) concluded that to effectively assess the contribution of biofuels in terms of reductions in C02 emissions should be measured during the whole lifecycle. (IEA has analysed the lifecycle of 60 biofuels productions and only eighteen exhibited a good performance in terms of emissions reduction). The assessment of biofuels environmental performance is complicated by the iLUC issue. ILUC (indirect Land Use Change) occurs when crops or land that would have otherwise been used for producing food or animal feed are used for growing biofuels, and existing agricultural production geographically shifts to new land areas created by converting natural areas (Croezen *et al.* 2010).

Nevertheless, today biofuels are the only alternative to fossil fuels for the transport sector.

Table 4 CO_2 net reduction by fuel

Biofuel	CO2 net reduction
Sugarcane ethanol	65%
Maize ethanol	15-20%
Wheat ethanol	15-20%
Soya biodiesel	95%

Source: AREA Science park 2007

[10] It is worth stressing that biofuel performances in terms of GHGs reduction significantly vary among countries and plants.

Resources competition, biodiversity loss and deforestation

In the last decade, biofuels feedstock production has been minimally satisfied trough reduction of food supply: subsequently, the most blatant effect on environment is represented by land use change. Land use change (LUC) occurs as lands are shifted from one use to another, for example, for urbanization or expansion of agriculture. Depending on the location and type of land converted, significant GHG emissions, biodiversity loss, competition for food production and property right violation may result. LUC caused by biofuels can occur both directly, as land use is shifted into biofuel crop production, or indirectly through market responses to supply and demand changes of biofuel crops and other related agricultural commodities. Resulting price fluctuations may incentivize these indirect land use changes (ILUC) elsewhere [11] (Broch, Hoekman and Unnasch, 2013, CIFOR, 2011)

In effect, biofuels feedstock has been mainly produced by cultivating previously uncultivated land. By doing so, biofuel production has been responsible for creating negative impacts on soils richness, water resources availability and, above all, on biodiversity. From 2004 to 2007, pro-biofuels cultivated area has grown from 1,3% to 25% of total cultivated land and, despite this, the land cultivated for food use has increased (FAO, 2008c).

The on-going growth of land used for biofuel production is causing deforestation of large areas in several countries, threatening the survival of many

[11] For the sake of clarity, we provide here some example of direct and indirect land use change. Item 1 illustrates an example of direct land use change, while the remainders represent a case of indirect land use change.
1. to produce rapeseed, and then rapeseed oil for biodiesel;
2. to produce rapeseed, and then rapeseed oil for non–biofuel, used to replace a unit displaced by biofuel consumption;
3. to produce rapeseed oil to replace soybean (for instance) that has been diverted to biofuel production directly or indirectly;
4. to produce rapeseed, and then rapeseed oil and meals, driven by an increase in demand by the livestock sector or the food processing industry related to general equilibrium effects (IFPRI, 2011)

animal and plant species. A working paper by CIFOR (2011) tried to shed light on why it is so difficult to evaluate biofuels role in deforestation. As the author explain, iLUC also works through the pricing mechanism, as the increased biofuels demand drives up commodity prices, hence increasing the pressure on global land and global ecosystems (Liska & Perrin 2009; Kim *et al.* 2009; Croezen *et al.* 2010). ILUC could also modify agricultural practices (e.g. farmers reacting to higher prices by using more fertilizers). Logically speaking, iLUC impact on deforestation and GHG emissions appears clear, but its scope and magnitude do not, because of the complexities of the economic and social systems that connect biofuel with global land use conversion. Despite methodological difficulties, numerous contributions elaborated on case study and specific regions dispelling false myth (such as Colombian biofuels massive forest clearing (Etter *et al.*, 2006)) and reporting evidence of huge environmental damages (as Malaysian palm oil production (Milieudefensie *et al.* (2008))

A well know and very sad example was reported by Nellemann, who in 2007 (UNEP, 2007) carried out a study on the survival of the Indonesian Orango tangos in the same regions where palm oil is cultivated. Beyond the Orango clear example, extensive monocultures for biofuels production may lead to a significant reduction of agro-biodiversity[12].

Converting natural ecosystems (that do not include marginal areas and abandoned areas), generates a carbon debt: the amount of the carbon debt depends on two factors, the type of land that is converted and the type of raw material that is cultivated (Fargione *et al.* 2008). Hence, for a comprehensive estimation of carbon debts it is necessary to implement a separate assessment for each type of crop that replaces each type of existing ecosystem. Up to now, only a

[12] The loss of variety of animals, plants and microorganisms, which are used directly or indirectly as food, fuel, medicines and fodder (FAO, 2004).

few biofuel crops, such as ethanol from sugar cane produced in the Brazilian Cerrado, compensate carbon debt in a few years, while most of the productions require decades or even hundreds years.

Table 5 Comparison of corn ethanol and gasoline GHGs with and without land-use change by stage of production and use

Net land use effect (grams of GHGs CO_2 equivalents per MJ of energy in fuel)							
Source of fuel	Making feedstock	Refining fuel	Vehicle operation (burning fuel)	Feedstock carbon uptake from atmosphere (GREET)	Land-use change	Total GHGs	% Change in net GHGs versus gasoline
Gasoline	+4	+15	+72	0	-	+92	-
Corn Ethanol (GREET)	+24	+40	+71	-62	-	+74 +135 without feedstock credit	-20% +47% without feedstock credit
Corn Ethanol plus land use change	+24	+40	+71	-62	+104	+177	+93%
Biomass Ethanol (GREET)	+10	+9	+71	-62	-	+27	-70%
Biomass Ethanol plus land use change	+10	+9	+71	-62	+111	+138	+50%

Source: Searchinger et. al, 2008

FAO has raised concerns about the so-called marginal areas. The increase in the amount of land devoted to biofuel production could exceed the limits of arable land by trespassing into lands that are now considered of little economic interest but still having high environmental value, and therefore protected for their environmental benefits. When the Indian government, with the "National Mission on Biofuels", declared its intention to convert 400,000 hectares of marginal land in plantations for biofuels production, many studies have shown that, in reality, almost all of these lands should be covered by biodiversity protection agreement and also constitute a public goods that provides a means of livelihood for the poor (Gundimeda, in 2005, estimated that in India the common property represent from 12% to 25% of income of the poor in rural areas).

Furthermore, Rossi and Lambrou (FAO, 2008a) showed that changes in land use may also be unexpected and several problems may be caused even by converting an apparently unused land. They pointed out that a non-resounding but still vital problem is caused to nomad shepherd whom travel whit ruminant livestock along rural regions. Nomads' survival depends crucially on the availability of pastures on their ways.

All the issues above certainly apply to large scale biofuel production. IFAD conclusions on small scale, rural production, seem to go in the opposite direction: poor communities in developing countries commonly use wood as energy source. By burning exclusively wood (untreated wood is a very poor energy carrier), these communities are causing an uncontrolled and unsustainable chopping of the flora, as well as significant health problems produced by indoor smoke. Replacing solid fuels (such as untreated wood) with liquid fuels sensitively reduce these negative effects.

Depending on regions and raw material type, large scale biofuel production needs a big quantity of water. In some countries, such as China, bioethanol crops require on average 2400 litres of water for the irrigation of a bulk of maize that serves to produce one litre of ethanol. The efficiency in the cultivation process is crucial to avoid waste of resources: while in Brazil 1150 litres of water are necessary to produce one litre of ethanol, in India it takes 3500 (De Fraiture *et al.*, 2007).

Water scarcity affects only some region of the globe and it can results from both climate/geographical and economic factors. In California, for example, none suffers from a lack of water and one can find extensive cultivation and green parks despite being basically a desert region. On the contrary, in tropical areas, where water is not scarce at all, water access may be prevented by the lack of infrastructure.

Hence, biofuels production represents both a threat and an opportunity: a threat when plantations subtract available water to rival uses, and an opportunity when is combined with infrastructures building, as long as they are also used for other purposes.

To summarize, first generation biofuels are produced by transforming agricultural raw material. Agricultural raw material production do requires land and water. Excluding services, land is required for basically all human activities (paper industry, building and furniture industry, pharmaceutics, cosmetics, cloth production etc. etc.). How much of the effects described above are attributable to biofuels is hard to say, nevertheless in many countries, biofuels policies have come to stringent environmental regulation.

The net energy ratio

In 2006, a study by the U.S. National Resources Defense Council highlighted that first generation biofuels generate, at the end of the process, only a small proportion of energy surplus, while advanced biofuels, such as lignocellulosic, have an energy efficiency of four times. Energy efficiency is calculated starting from lifecycle analysis: this allows defining the final energy content of biofuel and comparing it with the energy required from seeds production to end product. The net energy ratio in computed as follow:

$$\text{Net Energy Ratio(NER)} = \frac{e_b}{e_f}$$

When NER is >1, biofuels produce less energy than the amount employed for its production.

The estimation on the equation above presents several difficulties and diverse authorities have been committed to its calculation. According to estimation techniques, the NER can vary significantly. Table 7 displays some of these calculation attempts, reporting author names, and the year to which it refers to.

Table 6 NER by biofuels type

Biofuel	Energy balance	
	No by/co products	Including by/co products
Sunflaver biodiesel	2,0	3,1
Rapeseed biodiesel	1,7	3,0
Sunfluver SVO	2,8	4,3
Rapeseed SVO	1,8	3,4
Sugarbeet ethanol	1,1	2,2
Sugarcane ethanol	1,2-1,4	2,4
Maize ethanol	1,0-1,1	2,2-2,5

Source: AREA Science Park 2007

According to IFAD, although the biofuel sector is strongly characterized by significant economies of scale, when it comes to energy balance, labour intense small scale productions occurring in poor farms seem to be more environmentally friendly. Farmers produce small quantity of biofuels without using fertilizers, pesticides, or fuel powered apparatus (such as harvesting machines). In small scale productions, biofuels lifecycle is really short and may also be totally fossil fuels free, as in the jatropha oil production.

Table 7 NER calculated by different authorities, by raw material and country

Author and date	Raw material	Country	NER
Ho 1988	Maize	Usa	0,95
Marland, Turhollow 1990	Maize	Usa	1,28
Pimentel 1991	Maize	Usa	0,69
European Commission1994	Wheat/Beet	Ue	0,96-1,04
Keeney, De Luca 1992	Maize	Usa	0,9
Levy 1994	Wheat/Beet	France	1,91-1,31
Shapouri et al 1995	Maize	Usa	1,24
Gover 1996	Wheat	UK	0,93
Agri and agri food, Canada 1999	Maize	Usa	1,55
Wang 1999	Maize	Usa	1,42
Richards 2000	Wheat	UK	1,1
Pimentel 2001	Maize	Usa	0,69
Shapouri et al 2003	Maize	Usa	1,34
Ademe 2002	Wheat/Beet	France	2,05
Macedo et al 2003	Sugarcane	Brazil	8,3-10,2
Kim, Dale 2005	Maize	Usa	1,62

Source: Inea 2008

Social aspects

Gender inequalities

A less debated but still very important issue is represented by gender inequality linked to biofuels. The existence of large economies of scale makes biofuel industry difficult to access for vulnerable categories that cannot acquire fundamental input such as land ownership or access, credit and so on. In many poor and developing countries, women represent the majority of agricultural workers, but only a very small percentage owns the land. In countries such as Burkina Faso, women are excluded by law from property right or ownership of land, regardless of their education, social or economic position. In other countries, even when the law allows women to own land, as is the case in Zambia, social habits tend to prevail (for example, land is traditionally inherited by the male sons).

Nevertheless, landowners cultivating extensive monocultures usually prefer female workers, as they are commonly paid less and are more submissive.

Furthermore, recalling environmental aspect discussed in the previous section, it is worth noting that land use change and biodiversity loss affect women in a more arduous way with respect to men. In fact, marginal lands, and land with poor economic value are usually allocated to women, and there, women collect water and other assets necessary for survival.

Production standards and social dumping

In the last decade, international certifications have been more and more used to promote biofuels with low environmental impact and high social and economic benefits for export poor countries. On one hand, certifications practically represent the only way to access a market[13]: in fact, these tools can help seeds/crops/oils exporters and fuel producers to provide documentary evidence of their good practice, and allow biofuels distributors to access subsidies and other biofuels incentives. On the other hand, certifications represent a barrier that small entrepreneurs find very difficult to overcome. Certification, even when internally provided, has high costs both to adapt products and production to the required standards and to undertake compliance verifications. In addition, human capabilities such as literacy are fundamental in order to acquire a certification.

Under EU regulation, 20-20-20 goal must be achieved by importing biofuels and raw material in strict compliance with environmental and social sustainability criteria. In 2009, European Commission recognized a list of recommendations

[13] Even though there is no specific ban for “bad biofuels”, importing countries are quite attentive to social and environmental impacts. For example, palm oil biofuels have been deeply disapproved by international community because of the great damages caused to Indonesian flora, fauna and local communities. As a matter of fact, Pacific palm oil exports have been surpassed by Latin American ones.

from the Federal Polytechnic School of Lausanne. The Ecole Polytechnique Fédérale de Lausanne has promoted a laudable initiative to ensure that biofuel production does not take place in violation of internationally recognized values. The Institute has embarked on a long journey through which they managed to identify key issues to be addressed to ensure sustainability in biofuels production. They formulate a set of specific recommendations. Among others we recall the following.

Human and labour rights: biofuel production shall not violate human rights or labour rights, and shall ensure decent work and the well-being of workers.

Accordingly, the roundtable on sustainable biofuels stressed that producers shall adopt the "ILO's Decent Work Agenda," which developed clear strategic objectives to ensure consistency in worldwide standards of social protection.[14]

As frequently happen in extensive monocultures, child labour is a serious issue. According to ILO, child labour is a violation of fundamental human rights and several studies have shown how it can cause lifelong physical and psychological damages. Furthermore, child labour is closely linked to the so called poverty trap: it allow poverty to be perpetuated through generations by taking

[14] Some remarkable section of the ILO Agenda:

- Workers in biofuels projects will enjoy freedom of association, the right to organize, and the right to collective bargaining.
- Biofuels projects will ensure that no slave labour or forced labour is employed.
- Biofuels projects will ensure that no child labour is employed or allowed.
- Biofuels projects will be free of discrimination of any kind, whether in employment or opportunity, including in respect of wages and other working conditions, and in respect of social benefits.
- Workers' wages and working conditions will, in addition to respecting all applicable laws and international conventions to which the country in question is a party, as well as all relevant collective agreements, be determined by reference to, at a minimum, the conditions established for work of the same character or offered by comparable employers in the country concerned.
- Conditions of occupational safety and health for workers and communities involved in the operation of all biofuels programmes, including processing activities, will follow internationally recognized standards.

children away from education and limiting their social mobility. Child labour is a plague for poor people in most of the biofuels (mainly raw material) exporting countries and, additionally, it represent a clear example of social dumping, because it creates a reprehensible competitive advantage in raw materials production.

Some issues hampering the diffusion of second generation biofuels

As we have seen, second generation biofuels do not face many of the difficulties characterizing first generation. In fact, they are produced primarily from waste materials, their production do not compete with food, and do not subtract land and water from other uses. Of course, his is not true with regard to purposely grown raw material, such as poplars or Miscanthus.

As in the case of first generation, second generation biofuels confirm a close correlation with fossil fuel price, while they perform much better in terms of energy balance. According to IEA, producing maize bioethanol using second generation transformation processes makes a great difference: while each litre of fuel used in first generation production process produces between 1.3 and 1.65 litres of bioethanol, second generation techniques range between 4.4 and 6.6 litres. Better performances are shown also in terms of GHGs emissions and other harmful components reduction.

Despite all the advantages, second-generation biofuels are still far from large scale commercialization. In fact, two issues are hampering the development and the deployment of this source of energy, in particular lignocellulosic biofuels: high production costs of production and the need of technical capabilities.

Producing second generation biofuels requires huge initial costs to build and start up bio/thermo-chemical processing plants. So far, only few companies hold such a complex capability and the production is carried out by cooperating big firms, which in turn provides only a piece of the required equipment and expertise[15].

Operating costs are also very high and they vary according to the raw material and the processing process. First of all, raw material price varies depending on the type, which in turn depends on the weight of the biomass and its volume, the distance from the point of collection to the plant, the road network available and logistic infrastructures and so on. Then, raw material pre-treatment[16] makes a great difference in term of production cost. In 2007, Wyman *et al.* calculated capital requirement for the construction of a plant for the production of bio-chemical ethanol in the United States: estimations differ by many millions of dollars depending on the biomass pre-treatment system (Table 8).

Table 8 Production costs by pretreatments

Pretreatement	Initial investment US$billion	Annual production capacity (billion litre)	Fixed costs/litre (US$)
No pretreat.	200,3	34	5,88
Dilute acid	208,6	212	0,98
Hot water	200,9	166	1,21
Steam- explosion	190,4	200	0,95
AFEX	211,5	215	0,98
ARP	210,9	175	1,20
Lime	163,6	185	0,88
Pretreat. ideal	162,5	245	0,66

Source: Wyman et al. 2007

[15] Poet, one of the largest starch-ethanol manufacturers in US, added to its twenty-two operative plants an extra plant to produce second generation biofuel. To do so, the company had to work in partnership with Novozymes, which provides enzymes for biochemical processes, and Du Pont de Nemours, who provided fermentation facilities (Iea. 2008).

[16] Pretreatment is the first of three steps in the conversion of lignocellulosic biomass to ethanol. The next two sequential steps are saccharification and fermentation. The main objective of pretreatment is to "deconstruct" the plant cell walls, by destroying the tight lignin wrapping surrounding the biomass, resulting in the exposure of the carbohydrate polymers (mainly cellulose). The exposed carbohydrate polymers can then be easily converted to simple sugars in the second-stage saccharification step, and these sugars can then be fermented to ethanol (ISAAA website)

As commonly happens in high tech sector, second generation biofuels are characterized by a continuous evolution in scientific discoveries, so investors are constantly facing high uncertainty concerning technical obsolescence with the consequent risk of being excluded from the market.

Chapter 2. Environmental goals and technological innovation

Chapter overview

"Technological innovation can allow for the realization of environmental objectives in a manner which is less costly than would otherwise be the case" (Johnstone et al., 2008). This chapter provides a literature review of economics theory that elaborates on environmental policies and technological changes. The aim of the chapter is to contribute to shed light on the following question: if biofuels exist, improve energy security and reduce carbon and other pollutants emission, why they don't diffuse faster ?

Background and aim of the chapter

The concept of innovation and technological progress has changed its shape several times during the evolution of economic thought. It's possible to find some remarks about this subject in the work of the principal classical authors. Adam Smith took into account the relationships between technological change, labour sharing and structural economic change. In *The Wealth of Nations (1776)*, Smith's interpretation of the factors behind the increase of workers' productivity suggests that an impressive role has been played by "… the invention of a great number of machines which facilitate and abridge labour, and enable one man to do the work of many" (Smith,1981 p. 17). Also David Ricardo, in his *Principles of Political Economy (1817)* spoke about embodied technical progress and its effect on employment (Malerba, 2000).

Although the idea of "innovation" was already in place, Joseph Schumpeter is commonly considered the father of the economics of innovation. He clearly illustrated the differences between invention and innovation describing the former as "entirely immaterial whether an innovation implies scientific novelty or not…Innovation is possible without anything we should identify as invention and invention does not necessarily induce innovation, but produces of itself no economically relevant effect at all" (Schumpeter, 1939 p. 80).

In his thought, invention and, scientific and technological development must be considered as exogenous in the economic system. This conclusion influenced innovation economics approaches for several years. Recently this position evolved in a more comprehensive view: not only radical innovation is able to affect the technological change, but also minor and incremental innovation (Malerba, 2000).

The post-war theories exploited technological progress in order to explain the differences in economic growth. One of the main attempts was made in the Robert Solow growth model in the 1950s, in which technology was considered a truly public good. A more careful assessment pointed out that, considering technology as "manna from heaven", as an exogenous input in the production equation, does not allow for an effective explanation of the gaps in countries' growth, recognizing also that, if knowledge and information embody some features of public goods, technology must be considered as an appropriable good. If technology is not available for everyone in the same amount at the same time, differences in knowledge endowments can contribute to explain differences in economic growth across countries.

In the light of this conclusion, technology production becomes a means of access to increasing earnings and innovation efforts become an essential investment to achieve it. Appropriablity enables firms to play a fundamental role in the innovation process. It is worth mentioning though, that innovation effort is not a guaranteed path towards economically useful knowledge: several factors can influence the economic success of an innovation process.

A very basic representation of the innovative effort is the so called "linear model" where new products or new processes come as the result of a simple sequence involving four different stages, which are: basic research, applied research, development and production, and diffusion. In such an unsophisticated perception of the innovation process, the relationships between stages appear plain: investment in basic R&D fosters applied research and applied research leads to inventions suitable for the market. R&D investment plays the key role in the linear model: in this perspective the greater the R&D expenditure, the greater the economically useful innovation. Despite the model remaining popular with decision makers, it is not easy to find regularities between the amount of local and foreign resources invested in R&D expenditure and the innovation production, and between both of these and the rate of economic growth, which suggests that R&D expenditure is not able to provide a complete explanation of this phenomenon. First of all, not all R&D expenditures are created equal: "long-term and general R&D is likely to yield fewer economic returns in the short run than short-term applied R&D.

The same could be said of public vs. private investment in R&D" (Malecki,1991 in Rodriguez-Pose, 1999). Even considering R&D from a more global point of view, it is clear that these are not sufficient to reveal why some areas are more successful in transforming R&D into economically useful activities.

Recently, the idea of innovation process has evolved from the linear model towards a more sophisticated notion: empirical studies have shown that economic systems are characterized by a large variety of ways in which innovation activities take place and spread their effects on economic growth and employment. A comprehensive list of all factors able to affect the production of economically useful knowledge is not easy to fill. The complexity of these relationships brought to define the concept of "innovation system". In this context, innovation appears as an iterative and non-linear process (that could be linear by chance) that seems to be affected by factors related to peculiar historical, geographical and social features. The idea came up in the early 1990s from Lundvall. It is possible to find the core of his idea in the introduction of his book *National System of Innovation*. He wrote, “It is assumed that learning is predominantly an interactive and, therefore, a socially embedded process which cannot be understood without taking into consideration institutional and cultural context. Specifically, it is assumed that the historical establishment and development of the modern nation state was a necessary prerequisite for the acceleration of the process of learning which propelled the process of industrialization in the last centuries. Finally, it is recognized that the traditional role of the nation state in supporting learning processes is now challenged by the process of internationalization and globalization”(Lundvall, 1992 p.1).

Lundvall’s approach was followed by several attempts to complete the innovation framework investigating on the linkages between local institutional networks and innovative capacity: in particular, evolutionary economics elaborated on technology like a subject naturally included in the “evolutionary process”. As Bernard de Mandeville pointed out ‘What a Noble as well as Beautiful, what a glorious Machine is a First-Rate Man of War. …We often ascribe to the Excellency of Man’s Genius, and the Depth of his Penetration, what is in reality owing to the length of Time, and the Experience of any Generations, all of them very little differing from one another in natural Parts of Sagacity’. (Mandeville 1714 vol. II, pp. 141-1423 in Dosi, 2009)

Analysis characterised by the use of a combination of these approaches allowed for the definition of some fundamental features concerning essential conditions for the innovation process: “proximity, local synergies, interaction and the importance of ‘inter-organizational networks, financial and legal institutions, technical agencies and research infrastructures, education and training systems, governance structures and innovation policies” (Crescenzi, Rodríguez-Pose, 2006

p. 4). Nevertheless, those are only a few of the factors that can affect the innovation production: the firms ownership and the competition level, the object of the production (Dosi *et al.*1988).

An efficient approximation of the local and social condition involved in the process of innovation production is the concept of social filter which stands for a set of "conditions that render some courses of action easier than others" (Morgan 2004 in Crescenzi, Rodríguez-Pose, 2006) , determining the adoption rate of innovation and transformation into economically useful knowledge. In other words, the social filter represents a proxy of all those conditions able to point out the reason why the innovation effort not always successfully occurs.

From the social filter perspective, spatial dimension is a crucial factor that can deeply affect the innovation production: proximity can make the knowledge and information flows easier. Considering information as a not always codifiable good, "face to face" contacts remain the fastest and cheapest way to acquire new knowledge. Moreover, proximity may influence the behaviour of firms: economic agents situated in the same region are more likely to experience the same performance, sharing the same economic and social background.

Furthermore, the idea of innovation system applied to different spatial scales defining different system degrees: the analysis of the spatial boundaries of technological spillovers and knowledge externalities suggest that innovation is more likely to be a regional process than a national one. Several authors demonstrated that, despite the boundaries of the State Territory, innovative activities tend to cluster in specific regions and industries where knowledge spillovers reach a critical mass (Jaffe, 1989).

Moreover, the concept of National System of Innovation, which is define as "…the elements and relationships which interact in the production, diffusion and use of new, and economically useful, knowledge ... and are either located within or rooted inside the borders of a nation state." (Lundvall, 1992), has to be reinterpreted in the light of the globalization process: most of the features of the National System of Innovation have been changing in different ways. For example, when it comes to consider new lifestyles, the role of institutions such as the education system changes its importance when it faces the (widely exploited) possibility of migration. The latter is only one of the examples that can demonstrate how dynamic is the innovation process, which continually change shape in space and time, suggesting that academic conclusions and policy maker decisions have to be frequently re-examined in order to keep up with behavioural modification, with adjustments market and industrial structure and with changing in "social filter".

Starting from Schumpeter' seminal work *The theory of economic development* (Schumpeter, 1934), the debate on the relationship between technological innovation and economic development has not been able to provide clear evidence on which are the drivers of the first one and the effects on the second one.

In particular, there is no robust evidence on which is the relationship between environmental innovation and economic performance while there is a big consensus on environmental policy ability to stimulate technological change. In this chapter we will try to recap the determinants of innovation in general and then focus on the determinants of environmental innovation. The third section of the chapter is dedicated to the evolutionary and co-evolutionary innovation economics and presents the concept of technological lock-in, path dependency models. The last section is entirely devoted to carbon lock-in.

Innovation and eco-innovation

Innovations in environmentally-friendly technology, from now on EIs (Eco-Innovations) show most of the typical characteristic of Innovation *tout court,* but present also a specific feature: their application possibly will decrease harmful consequences of many productive activities.

The common form of Innovation can be summarized as follow. Innovation can be:

- Incremental: implying slight modifications to the state of the art, which is usually perceived as a "we do the same thing in a better way";
- Radical: which is a breakthrough in products (a new product or a substantially different one) or processes;
- Small-scale: representing a technological change only for a precise firm, field or consumers' category;
- Large-scale: involving the transition of the whole economy towards a new paradigm;
- Organizational: referring to smart habits or changes in practice which may lead to improvements in goods production of use.

When it comes to EIs, technologies show some specific features related to final products, production processes or the so-called end-of-pipe solutions. Environmentally friendly products are thought to be less polluting by having a longer life cycle, being less harmful or energy efficient or by triggering a smaller quantity of waste material. The same can be said about green "process-integrated" technologies. End-of-pipe technologies try to remedy to existing pollution by capturing or neutralizing it.

Following the approach adopted by Johnston *et al.* (2008), the determinants of innovation can be divided into those related to market and firm-level factors and those induced by policies.

To the first category belong Schumpeterian and neo Schumpeterian view of innovation and their opponents. According to Schumpeter (1942), there is a positive relation between innovation and markets where imitation prevention is easier and resources availability is higher. Thus, monopoly (or similar) would be the perfect environment for R&D expenditure proliferation. Later, neo Schumpeterian argued that innovation originated from firms acknowledgment of technological opportunities.

Twenty year after, Arrow (1962) argued perfect competition can create stronger incentives to innovation.

Both Arrow and Schumpeter theories have been repetitively confirmed by empirical studies shifting the attention toward firm size as a possible determinant of innovation (Syrneonidis, 1996).

A more univocal conclusion has been drawn by Criscuolo et al. (2005) demonstrating a positive correlation between special scope of the market and innovative attitude of the firms operating in that market.

The second category refers to the role played by policies which are particularly significant in the case of EIs.

EIs are designated to internalize several environmental externalities: distinct policies have to be implemented by distinct authorities. In some cases, the disjunction of policy competencies can lead to implement a bunch of policy conflicting one another (Goel and Hsieh, 2006).

This policy incoherence may be cause by different underlying policy objectives and, as we discuss in the next pages, by objectives of interconnected sectors and institutions.

Evolutionary economics: innovation in perspective

In 1982, Nelson and Winter pointed out that technological change may occur as a uncertain, groping, disorderly, and error-ridden processes. Such a complex phenomenon results particularly difficulty to incorporate within the orthodox theoretical scheme[17]. Furthermore, by looking at firms R&D expenditure, they appreciated that "much of firm behaviour could be more readily understood as a reflection of general habits and strategic orientations coming from the firm's past than as the result of a detailed survey of the remote twigs of the decision tree extending into the future." (Nelson & Winter, 1982, p.vii-viii). Starting from their dissention from orthodox economic theory, the authors depict a model where instead of optimizing agents, perfect information, total and unbounded rationality, and static equilibrium, there are patterns of survival in the population of firms based on routines who play the same role of genes in biological evolutionary theory. These routines include decisions on investments, R&D expenditures, diversification strategies, etc. [18]

Since then, evolutionary economics elaborated on different dimension of innovation and technological change.

To the aim of this study it is particularly interesting to recall evolutionary theory contribution to innovation, industrial dynamics and structural evolution,

[17] "Creative intelligence, in the realm of technology as elsewhere, is autonomous and erratic, compulsive and whimsical. It does not lie placidly within the prescriptive and descriptive constraints imposed by outsiders to the creative process . . . " (Nelson and Winter, 1982, p.viii)

[18] The authors categorized firms depending on how they behave accordingly to 3 routines: operating, investments and high level review/revision.

intended as a change in products, production processes, firms changes and growth and changes in competences boundaries (Malerba, 2000). In this field of research, evolutionary economists came to define a macro level of industry evolution taking into consideration the emergence of new industries, creation of and modification of products and techniques, the development of firms and technological networks and so on. In this scientific area, the challenge of empirical contribution is crucial (Nelson, 1999).

Malerba (2000) listed some of the evidences observed in empirical analyses:

- Structural evolution is linked to continuous development of products, productions techniques, and to the hunt of incremental and radical innovation,
- Technological discontinuities are followed by incremental innovations and then again by discontinuities,
- New access occur in particular periods, but do not necessarily occur after technological discontinuities,
- Initial conditions could vary depending on specific characteristics of new technologies and their proximity to existing one; structure and competencies of the firm producing similar or substitute goods,
- The nature of the firm itself could continuously change over industrial evolution,
- Policies and public institutions strongly influence structural evolution,
- There are massive institutional and organizational differences among countries.

From a theoretical point of view, evolutionary economics provided several interesting contribution in the industry evolution macro level. In particular we refer to co-evolutionary processes.

Co-evolutionary processes : the carbon lock in

At the end of the 20th century, many researchers begun to investigate co-evolutionary approaches to shed light on technological development where innovation in turn influences and is influenced by social, economic and cultural environment where they occur (Rip & Kemp, 1998, Kemp, 2000). Co-evolution refers to joint and interconnected changes between technology, competencies, strategies and organization of a firm, market, demand and institutions occurring during the evolution of a certain industry. In 1994, Nelson elaborated on an evolutionary theory of economic growth that brought together appreciative theorizing regarding growth and formal theorizing. The first suggests that a new technology develops along a rather common path from its birth to its maturity, and that firm and industry structure 'coevolve' with it. The other is concerned with the development of institutions in response to changing economic conditions, incentives, and pressures (Nelson, 1994).

Co-evolutionary theory explains that new alternative technologies have to face huge barriers generated by initial advantages of their competitors. Such advantage could lock the economic system in the use of a pre-existing technology (the so-called technological lock-in). A specific aspect of co-evolution refers to processes affected by path dependency (David 1985) in which the success of new technologies depends on the capacity of the whole market to overcome the advantages of the existing technology. Arthur and others have argued that increasing returns to adoption (positive feedback) lead to lock-in of current technologies, preventing the take up of potentially superior alternatives. Arthur (1994) identified four major classes of increasing returns:

- Scale economies: existing technology often has significant 'sunk costs' from earlier investments, meaning that firms will be reluctant to invest in more sustainable alternatives.

- Learning effects: learning act to improve products or reduce their cost as specialised skills and knowledge accumulate through production and market experience, the so-called 'learning-by-doing' (Arrow, 1962).
- Adaptive expectations: increasing adoption improves familiarity and reduces uncertainty and both users and producers become progressively confident about quality, performance and durability of the current technology leading to a lack of 'market pull' for more sustainable alternatives.
- Network or co-ordination effects: network effects occur when advantages increase for agents using all the same technology (E.g phone companies). Infrastructures develop based on the attributes of existing technologies (as clearly happened in fossil fuel based sectors) create a barrier to the adoption of a more sustainable technology with different attributes leading to lock-in, especially in large technological systems, such as electricity generation or transportation. (Foxton 2003, Foxton & Kemp, 2005).

Increasing returns play an important role also in institutional changes. In order to regulate human interactions, institutions imposes constraints such as legislation, economic rules and contracts, and informal constraints, such as social conventions and codes of behaviour, sometimes creating drivers and barriers for social and economic change, Pierson (2000) argues that, as well as markets, political institutions are particularly prone to increasing returns having high start-up costs and being subject to time and resource consuming learning processes and adaptive expectations. In the light of this, particular political institutions, such as regulatory frameworks, create a path dependency. As modern technological systems are deeply embedded in institutional structures, institutional lock-in can interact with and reinforce the drivers of technological lock-in.

The carbon lock-in: a review of its theorization by Unruh[19]

Technological and institutional lock-in has fundamental effects on of innovation for environmentally friendly development, and the policy framework needed to promote it.

In his paper, Unruh (2000) presented an investigation on interlocking technological, institutional and social forces which could generate policy inertia towards climate change mitigation. Industrialized countries have become locked into fossil fuel-based energy systems through a path-dependent process driven by technological and institutional increasing returns to scale. This phenomenon, called carbon lock-in, arises through a combination of systematic forces that continue to spread fossil fuel-based infrastructures in spite of their well-known environmental adverse impact and the apparent existence of cost-neutral, or even cost-effective, alternatives. According to Unruh's perspective, carbon lock-in arises from systemic complex interactions among technologies and institutions. He termed these interactions Techno-Institutional Complex (TIC), which have to be seen as complex systems of technologies embedded in a powerful conditioning social context of public and private institutions. Once locked-in, TIC are difficult to displace and can lock-out alternative technologies for extended periods, even when the alternatives demonstrate improvements upon the established TIC.

The author provided a clear response to the question proposed at the beginning of this chapter: if biofuels exist, improve energy security and reduce carbon and other pollutants emission, why they don't diffuse faster?

---

[19] This section is entirely based on "Gregory C. Unruh. 2000. Understanding carbon lock-in. *Energy policy* 28. Pp. 817-830" and "Gregory C. Unruh. 2002. Escaping carbon lock-in. *Energy policy* 30 pp. 317-325". In his papers he author included a great number of historical evidence of the carbon lock-in thesis. For the sake of conciseness we omit them and we do not provide further references.

The first answer is given by increasing returns: if the economy is functioning efficiently, new technology cannot be cost-neutral when the total costs of their adoption are considered. Thus, while low carbon technologies may reduce energy price, they must have other overlooked costs that make their adoption economically untenable. Given the evidence that economically viable opportunities do exist, alternative explanations for the slow diffusion have focused on the idea of complex barriers to adoption.

The transport sector provide a self-evident example of this complexity: it is composed of numerous interconnected technological systems including vehicles, roads, fuel pumps stations, etc., managed by a series of public and private social institutions. Moreover, a vehicle itself can be seen as a complex technological system composed of several subsystems such as the engine, the fuel tank, etc., which in turn represent a further subsystem. The structure of this multilevel system represent the so-called dominant design, where investments tend to ignore pre-commercial research and development stages, focusing rather on competition among commercially viable technologies. Repeating investments in dominant design commit in turn a firm, an entire industry and then an entire market to the dominant design trajectory and create lock-in. In highly connected technological systems, these interdependencies are amplified between the many subsystems that co-evolve together with the primary network. Once again, transport sector provide a clear example of the phenomenon because the deployment of a fully operative vehicles system depended upon the co-development of manifold subsidiary technologies and industries. Such technological inter-relatedness can create lasting barriers to competing technologies.

The interdependent growth of industry networks and subsystems involves substantial coordination efforts that can be provided through standards and conventions (such as voltage, programming languages, octane standard etc.)

reducing or eliminating investment and institutional uncertainties. Given the complexity of this industrial and institutional network, the establishment of a new technology can require harmonized industry efforts and elimination of existing standards. Risk adverse firms may not engage this process when there are uncertainties about the potential responses of and to the new technology[20]. When innovative firm try to introduce a new technology, other firms may encumber the creation of a new standard strengthening the dominant design. The constant re-investment in the commercially available technology creates a self-reinforcing feedback that can lock-in existing technological. Likewise, financial institutions usually make loans to dominant design producer, with have definite ability to payback debt further enhancing lock-in. Additionally, formal and informal societal institutions may generate non-market forces of lock-in through coalition building, voluntary association, lobby, media and the emergence of societal norms and customs.. Finally, the involvement of government is fundamental because of the ability of legislation to override market forces. Policies can create additional incentives and modify the rules of the game. In the evolution of a technological system, government intervention can remove market uncertainty about the direction of technological development through policy, and thus favour a specific design

In this perspective, Governments have the power to push new technology. Nevertheless, once a dominant design is established, institutions tend to foster only incremental change for long periods and it takes time to come to escaping institutional path dependency and remove unproductive policies.

As illustrated in this section, technological systems and institutions can become closely connected in a self-referential system that Unruh named the

[20] “Unless innovators are confident that a technology will become the new dominant design and bring along supporting networks, it may be judged too risky to make the required irreversible investments to market the new technology.” (Unruh, 2000, pp.)

Techno-Institutional Complex (TIC). Historically, TIC facilitated the adoption of carbon intense technologies becoming nowadays the setting for techno-institutional lock-in, slowing the emergence of alternative technological solutions.

Once acknowledged that TIC can create carbon lock-in, further, higher level reasons for the slow uptake of carbon-saving technology can be found. Techno-institutional lock-in indicates that there are structural forces that make it difficult to deviate from the path of existing techno-institutional systems. Even with the growing of evidence of substantial environmental risk, government policy makers may not recognize the importance of lock-in phenomenon.

Chapter 3. Building BioPat

Chapter overview

This chapter describes the methodology, characteristics and potential use of BioPat, a dataset containing patents in the field of biofuels. The innovative methodology we use aims to solve drawbacks related to how patent data are allocated and organized in international databases. In order to create a database which includes patents strictly related to the investigated field, we propose an original method based on keywords, rather than on International Patent Classification (IPC) codes. Starting with a systematic mapping of biofuel production processes, we built a simplified but comprehensive description of the technological domain related to the production of biofuels by applying so-called process analysis. The keyword selection relies on an iterative approach, based on an analysis of recent scientific literature. The database was finalized with a series of interviews with experts in the biofuels sector, and compared with IPC-based biofuel codes, revealing improved accuracy when selecting data using our methodology.

Background and aim of the chapter

The measurement of innovative activities is a rather challenging task and a great number of different science and technology indicators have been identified in the litreature (Sirilli, 1997). The main input indicator relies on research and development (R&D) expenditure, while the most used innovation output indicators are based on patent data. Both types of indicators have strong limitations since not all research efforts translate into the introduction of innovations and not all innovations are patented. For our purposes, specific and systematic information on private R&D expenditures in the biofuels sector are not available while access to patent data makes it possible to collect information on the evolution of the innovative performance of economic systems by looking at the volume of patents registered and granted (Johnstone *et al.*, 2010).

As already mentioned, the use of patents has its pros and cons. The advantages of using patents as a proxy of innovation are manifold. A single patent provides information on relevant aspects of the innovative process such as the geographical origin of the innovation, its relevance in terms of technological progress, the previous stock of knowledge that allowed the development of new technological knowledge, the inventors and the owners of the patent and the usefulness of patented knowledge for subsequent innovations. On the other hand, using patents as a proxy for innovation presents several relevant issues (Griliches, 1990). In particular, even though there are very few cases of non-significant inventions wich have not been patented (Dernis & Gaullec, 2001; Dernis & Kahn, 2004), only a limited part of the whole produced innovations are applied for intellectual protection (Archibugi and Pianta, 1996). Moreover, there is an intrinsic variability of patents' value (Jaffe and Trajtenberg, 2002) and a different propesity to patent across countries and sectors (Johnstone *et al.*, 2010).

For our purposes, another important problem has to be taken into account. A patent usually has a very standard object: a chemical formula, a variation or an improvement in a natural process or a mechanical, artistic or even immaterial device. Once registered, the patent receives a code that classifies its content. Classification is fundamentally a technical problem referring to how patent data are allocated and organized in national and international databases. Every patent office provides each patent with an internal code that includes a reference to the object of the invention. An international code named IPC (International Patent Classification) is associated with the internal code which allows the classification of patents by following a hierarchical criterion (from 8 main fields to almost 70.000 subgroups) based on chemical and technological principles, only occasionally related to manufacturing sectors. In particular, the resulting classification is only of limited usefulness when it identifies a specific sector which does not fit the criteria used in the classification, as in the biofuels sector.

The aim of this chapter is therefore to illustrate a possible methodology for building a sector-specific patent database and showing how it can be potentially used for economic analysis. Despite the well-known limitations related to the use of patent data in innovation studies, in order to draw a picture of sectoral technological patterns, a valuable option is to build a database that tries to identify precisely the entire universe of patents strictly related to the biofuels sector. To do this, we must first adopt an early approach suggested by Hekkert *et al.* (2007) in order to map the actors which participate in the Biofuels Innovation System systematically by means of a process analysis. In the following, we first describe the IPC system and the Green Inventory database. We then provide details of the adopted keyword methodology and after that we give first descriptive results drawn from the collected database. The conclusions provide a synthetic discussion of the reached objectives and future research developments.

The IPC system and the Green Inventory database

During the last century, the increasing amount of patents registered daily worldwide and the great number of interactions among patents offices made the adoption of a uniform system of patent classification necessary.

The first attempt to create a global market for patents came with the founding of the World Intellectual Property Organization (WIPO), as a United Nations agency. WIPO was established by the WIPO Convention in 1967 with a mandate from its Member States to promote the protection of intellectual property (IP) throughout the world through cooperation among states in collaboration with other international organizations.

The will to foster closer international cooperation in the industrial property field, and to contribute to the harmonization of national legislation in that field led in 1971, after 15 years of international cooperation, to the Strasbourg Agreement concerning International Patent Classification (which entered into force on October 7th 1975). The huge number of patents (and related documents) created two main problems the treaty had to deal with: the administrative processing of the patent applications and the maintenance of the search files containing the published patent documents.

According to the 2011 version of the IPC guide, "the Classification, being a means for obtaining an internationally uniform classification of patent documents, has, as its primary purpose, the establishment of an effective search tool for the retrieval of patent documents by intellectual property offices and other users, in order to establish the novelty and evaluate the inventive step or non-obviousness (including the assessment of technical advance and useful results or utility) of technical disclosures in patent applications" (IPC Guide, 2011, p. 1).

The International Classification divided the universe of patents into 8 sections, 20 subsections, 118 classes, 624 subclasses and over 67,000 groups (of which approximately 10% are main groups and the remainder are subgroups). Each of the sections, classes, subclasses, groups and subgroups has a title and a symbol, and each of the subsections has a title. Each classification term consists of a sequence of symbols: the first one is a capital letter which represents the section. The letter is followed by a two digit number which represents the class and then by another capital letter that stands for the subclass. The subclass is then followed by a 1 to 3 digit "group" number, an oblique stroke and a number of at least two digits representing a "main group" or "subgroup". Hence, the IPC is a hierarchical system, with layers of increasing detail. The following represents an example of the classification: A01B1/00 symbolizes Human Necessities (Section A); Agriculture (Subsection title); Agriculture, Forestry, Animal Husbandry, Hunting, Trapping, Fishing (Class A01); Soil working in agriculture or forestry, Parts, details or accessories of agricultural machines or implements in general (Subclass A01B); Hand tools (Group A01B1); subgroup not specified (A01B1/00).

These different sections allow distinctions to be made between patents belonging to categories which sporadically present an economic importance (such as the case presented above, hand tools used in agriculture). On the contrary, the IP classification is not suitable when the focus of the research does not match an existing section.[21] (e.g., harvest tools). Several attempts have been made to provide a crosscutting interpretation of the standard classification.

The first category of attempts consists in a top down approach that relies on the IPC class and aims to define its content. Accordingly, one approach consists in the exploitation of the linkages between classes assigned to the same patent by considering those appearing together as a "class family" where IPC codes are seen as technologies. So, different IPC codes can refer to sub-technologies that a patent covers and they can be interpreted as a "class family" when they appear frequently together. An empirical application of this approach is represented by the IPC co-occurrence method (Antonelli et al., 2010; Kraft et al., 2011). Co-occurrence could be efficiently used when the investigated sector is already defined per se (as in the weapon case) or when it is relatively easy to be identified.[22] In other cases, such as

---

21 As an example, IPC contains a whole section for weapons. When interested in the investigation of innovation in the weapons industry, researchers should refer to the IPC code F41. On the contrary IPC do not contains an ad hoc code for the cosmetic industry (that would probably fall under several categories in the chemistry section).

[22] This is the case of sectors that are sufficiently wide to cover an entire section of the IPC, such as information and communication technologies, biotechnologies or nanotechnologies, where ad hoc

biofuels or other crosscutting sectors, the use of this method may end up in rough and unpredictable results.

A second category gathers attempts which try to identify the classes which are suitable for containing a patent related to the investigated object (OECD, 2011).

The "IPC Green Inventory" database (GI) falls into the latter category and was developed by the IPC Committee of Experts in order to facilitate searches for patent information relating to Environmentally Sound Technologies (ESTs), as listed by the United Nations Framework Convention on Climate Change (UNFCCC).

ESTs are currently scattered widely across IPC in numerous technical fields. The GI allows all ESTs to be collected in one place. Following the IPC system, the ESTs are presented in a hierarchical structure. According to the WIPO web site, two steps were required to create the GI. First, a list of technologies was completed by the UNFCCC as a basis for the work of the IPC Committee of Experts who identify the related IPC places. In order to identify the IPC places correctly, the experts can use the IPC Catchword Index, the IPC term search and their expertise in the relevant technical areas in order to collect all the green-related IPC places under the specific category. Hence, the inventory consists of a list of IPC classes characterized by the fact that they are suitable for containing patents related to a green technology.

Among the ESTs, for our purpose, we considered 44 IPCs (40 subgroups and 4 subclasses) that identify the biofuels sector.

---

selection tools are also available in search engines such as OECD Patstat. Patents related to these industries are scattered in more than one section but could be easily identifiable. Another example is given by fertilizers industry. Fertilizers belong to class C05, an *ad hoc* category in the chemistry section, but fertilizers distributors or fertilizing practices are included in the A01 class, which contains patents related to agricultural activities. Thus, a patents referring to a new fertilizer and its distribution practice would fall under both classes, ideally making a "class family" that identifies the fertilizers industry.

In Annex II we list the IPC subgroups and subclasses, the number of patents included in them (accordingly to Thomson Routers as of February 2011) and the technology associated with the different IPC codes.

As already mentioned, the classes above are suitable for containing patents related to the object specified in the GI (last column). It is worth remembering that these objects, which refer to the related IPC class, are not the IPC class object. For example, the first class (first row) A01H, which, according to GI, is suitable for containing patents related to liquid biofuels obtained by genetically engineered organisms, can actually contain, according to the IPC, all the patents that fall into the category (subclass title) "new plants or processes for obtaining them, plant reproduction by tissue culture techniques".

At present, the GI website does not display any statistics on the effective number of patents in each class that are also coherent with the object assigned (as a sort of validation). Hence, in order to shed light on the accuracy of the GI databases, we validated a sample of patents included in the IPC classes indicated above by asking a team of experts from the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) to check their coherence. Additionally, we asked the group of experts to distinguish between patents with a direct application in the biofuel production process and an indirect one. We downloaded the description field of the whole universe of patents belonging to these classes for USPTO, WIPO and EPO and eliminated the duplicates (each patent can fit in more than one class) ending up with 107,161 elements from which we selected a 1% sample.

The results of the expert validation showed that on average, only 25% of the patents included in the sample have a direct application in the biofuels sector. This percentage significantly varies among the patent offices. Such a result confirmed our intuition regarding the limits associated with the identification of patents through the IPC system in the biofuels sector.

The BioPat methodology

Setting a proper methodology to select patents in a rather specific sector is not an easy task. As shown by the experts' validation on the GI, the IPC class selection fails to extrapolate the classes that are supposed to identify a single economic sector, maintaining a high risk of considering external elements. Moreover, considering the huge variety of raw material and processes available for biofuel production that often overlap with other manufacturing sectors, it is highly probable that the GI classification does not catch all the patents that have a direct or an indirect application in the investigated field. Moreover, the method usually adopted by several international organizations, which considers all patents directly or indirectly linked with each other in a single family, is not appropriate when it comes to working on a small sector (or on a limited number of patents) because the smaller the sector, the higher the likelihood of catching external elements.

In order to tackle the lack of specificity from an economic point of view, several researchers have developed different methodologies essentially based on the exploitation of catchword tools and litreature scrutiny. The last decade's litreature on keyword analysis basically consists in selections of words from already existing keyword lists or the extraction of keywords from titles and, at least, abstracts of patents and scientific publications.

The litreature followed three main approaches:

- co-word study based on the keywords proposed by experts (Looze and Lemarie, 1997);

- use of descriptors chosen by professional indexers employed in patent offices and search engines (Coulter *et al.*, 1998);
- extraction of keywords from titles and abstracts of patents (Corrocher *et al.*, 2007).

These three approaches are characterized by strong differences. The first two are based on an attempt to describe the sector using words that are commonly considered sector specific, whereas the last one seeks to eliminate the arbitrarity of the selection process. In fact, Corrocher *et al.* (2007) pointed out that the ex-ante selection of the keyword procedure might reflect preconceptions, different backgrounds and points of view of the words' selectors and differences in the trainings and backgrounds of professional indexers. As a result, the authors decided to identify the most frequent sequential triples of words without imposing any priority constraint on the selection of keywords. The authors argue that triples of words within patent abstracts can identify technological domains that can be compared with the existing IPC technological classes.

Unfortunately, the method which looks ex-post for the triples of words is more appropriate when it comes to investigating a sector that is sufficiently wide to cover an entire section of the IPC (which is not the case for biofuels). Moreover, it is also more appropriate when the novelty of patents is based on engineering contents, which are more likely to fit into ad hoc classes.

On the contrary, the patents related to biofuels are spread across several IPC classes because the technology that characterizes the sector basically consists of thermo/bio-chemical processes and very common raw materials that can find applications in several fields.

Since we realized that the subjectivity of the selection process could represent a big challenge for the research outcome, we tried to make the process as objective as possible. We then decided to consult technical experts in the field of biofuels. We interviewed exponents of ENEA who helped us describe the process of biofuel production. This team of technical experts completed and validated the list of keywords derived from the scrutiny of a large number of scientific publications and the keyword list extracted by Scopus, a powerful search tool which provides access to a large number of scientific publications and patents office databases.

The choice and classification of keywords derives from recent scientific litreature which gives us the empirical basis of the process analysis. The search for keywords was divided into 2 different steps: the first one was dedicated to a search for "raw material" keywords, where a relevant number of technical and scientific papers was analyzed in order to pick out the terms describing the biomass used (or potentially used) to produce biofuels. The second step consisted in an accurate description of the "transformation process" currently known in biofuel production, including pre-treatment processes, chemical agents involved in the process and technical instrumentation used in it. Keywords were then tested on Scopus (www.scopus.com). At the same time, Scopus allows you to check if patents exist containing the selected keywords. Hence, the final selection of the keywords comes from an iterative procedure which allows results from scientific articles to be compared with patent results. This first step led to selecting several keywords which showed positive results both in patents and articles via Scopus. These keywords were submitted to the ENEA experts (see Annex II).

Finally, we improved the traditional keyword methods that look for keyword matches only in the patent's titles and abstracts. According to the IPC terms of reference, patent novelty is usually classifiable following two main principles: a patent can be characterized by engineering content or by bio-chemical content. The latter is true for the biofuels sector and represents the explanation of the crosscutting shape that it assumes in the IPC classification. In light of this, we decided to expand the use of keywords to the "patent descriptions" and "patent claims" fields in order to exploit the possibility of catching all patents that have a hypothetical, and not necessarily direct, function in the biofuel production process.

The patents were downloaded using Thomson Innovation, a single, integrated solution that combines intellectual property, scientific literature, business data and news with analytic, collaboration and alerting tools in a robust platform. With Thomson Innovation, we were able to export up to 30,000 records into .csv formats in one single operation. Thomson Innovation has the world's most comprehensive collection of patent data from major patent authorities, specific nations and proprietary sources exclusive to Thomson Reuters.

All process-specific and raw material keywords were used in the Thomson innovation jointly with a more general keyword (such as bio-diesel, bio-ethanol, bio-gas, bio-fuels) in order to exclude patents that share the same raw materials or transformation processes (in particular pharmaceutics and cosmetics are strongly related to the biofuels sector). Afterwards, some testing searches were implemented with a few selected keywords in order to verify the response of the Thomson database to the inputs. The Thomson search engine also allows symbols to be used as a means of catching variations of the same word, as well as plurals. For instance "fermented sugar" was entered as "ferment* sugar*", catching in this way a combination of different words such as "fermenting sugars" or "ferment sugarcane" and so on.[23]

---

[23] The list of keywords was built on different levels. First, we tested searches with selected keywords in order to verify the response of the TI database to the inputs and exclude the possibility that keywords list was inadequate. Second we created sub-groups of terms standing for the same keyword. For example, raw material terms derived from literature scrutiny were usually in English, while some inventors tend to use scientific Latin or Greek names. Scientific names were attributed to the same original (English) keyword.

Furthermore, we carried out a special search using general keywords in the "applicant" field, hypothesizing that a firm called "The Biofuel Company" deals with patent inventions related to biofuels.

Using Thomson Innovation, patents can be downloaded from national and international patent data offices. We focused our research on the European Patent Office (EPO), World Intellectual Property Organization (WIPO) and United States Patent and Trademark Office (USPTO) as described in Table 2.2.

With regard to raw material keywords, the search on Thomson was carried out as follows: by using Boolean operators "OR" and "AND" we selected all the patents (kind code A1 and B1 from 1/01/1990 to 31/12/2010) containing the keywords among a fixed set of general keywords introduced with the Boolean operator OR (at least one of the term must appear) and a more specific one (added one by one to the fixed set), with the Boolean operator AND. Multiple words were added in quotation marks[24].

Table 9. Data available on Thomson Innovation

WIPO APPLICATIONS	
Published international patent applications, fully searchable, language: 70% English, 15% German, 5% French, 1% Spanish	1978 – present
UNITED STATES	
US Granted, fully searchable, language: English	1836 – present
US Applications, fully searchable, language: English	2001 – present
EUROPE	
European Granted, potentially 31 countries, fully searchable, language: 60% English, 30% German, 10% French	1980 – present
European Applications, potentially 31 countries, fully searchable, language: 60% English, 30% German, 10% French	

24 For example: Nannochloropsis (an alga) AND "renewable *ethanol" OR "green *diesel" OR *methanol OR *buthanol OR biomethane OR biomethiletere OR "Synthet* fuel*" OR biodiesel OR "renewable fuel*" OR biofuel* OR etc.

With regard to the transformation process, keywords were used with the same sequence of fixed terms representing the general name of biofuel products (with Boolean OR, kind code A1 and B1 from 1/01/1990 to 31/12/2010) and a second level containing all general terms (added one by one with the Boolean AND) for production process such as transesterification, Fischer-Tropsch, anaerobic digestion and so on[25].

An important advantage of the adopted methodology is that by selecting patents related to ex-ante classified keywords, economically significant categories can be assigned to patents derived from each keyword, allowing us to distinguish between separate ages of the sector.

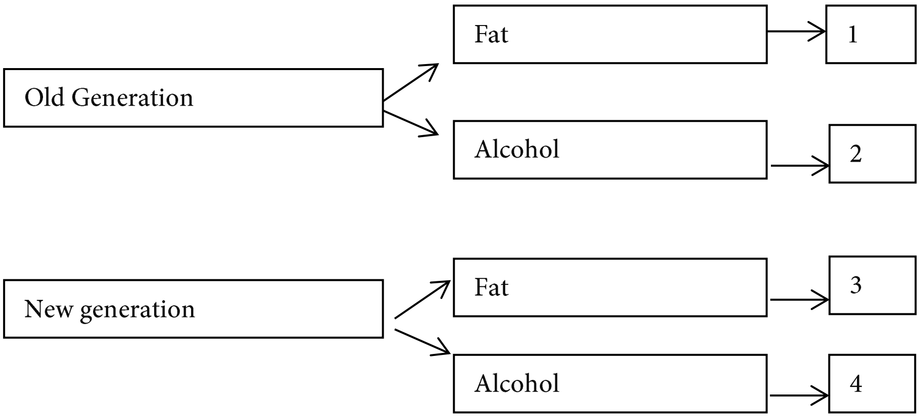
According to an evolutionary perspective, radically new knowledge looms up as a result of research efforts in unexplored regions of knowledge space. Once the new technology is discovered, subsequent innovations are mainly incremental, and knowledge recombination is more likely to occur inside a strongly delimited technological field. In other words, knowledge behaves as a complex network and a transition towards new orders occurs by means of a mix of slight intra-sectoral and extra-sectoral adjustments, depicting a long periods of incremental variations (Prigogine, 1987, Arthur, 1989). At a certain point, radically new technologies dissipate their lucrative power and incremental knowledge redirects innovative effort towards radically-new opportunities. This stage is characterized by the recombination of more and more distant technologies (Saviotti, 1996; Antonelli *et al.*, 2010).

25 After that, we verified if the downloads could represent a significant part of the whole universe achieved using only the general keywords. The huge specific outcome obtained by using the general keywords strongly reinforces the choice of working with selected specific keywords rather than working on a broader definition of biofuels (e.g. Karmarkar-Deshmukh and Pray, 2009) or on IPC codes (e.g. OECD documents).

In order to map technology evolution in the biofuels sector, beside the distinction between production input and stages (e.g., raw materials and transformation process) and final product (fat, alcohol and gas), different "generations" were used to classify patents. According to IEA (2008) first generation biofuels, which are mainly produced from agricultural crops and traditional oleaginous plants (such as palm and colza), are characterized by mature commercial markets and well known technologies. On the contrary, second and third generation technologies are currently developing. Second generation biofuels are obtained from non-food crops, especially from forestry, agricultural and industrial residues (in particular lingocellulosic material) or dedicated energy crops (such as miscanthus). Third generation biofuels are mainly related to algae and GMO. Hence, for the sake of clarity, in our work we distinguished between old and new generations. This classification method adopted is based on the following assumption: the actual technology used to produce biofuels, which includes raw materials, techniques knowledge, tools and machineries, is considered the current technological knowledge stock. Within this knowledge stock, two main technological categories can be discerned: "old generation" and "new generation", both for raw material and process keywords, which are related and include the entire supply of technologies for biofuel production. Making use of the exclusion principle, it is easy to define everything that is not in the old category as belonging to the new category.

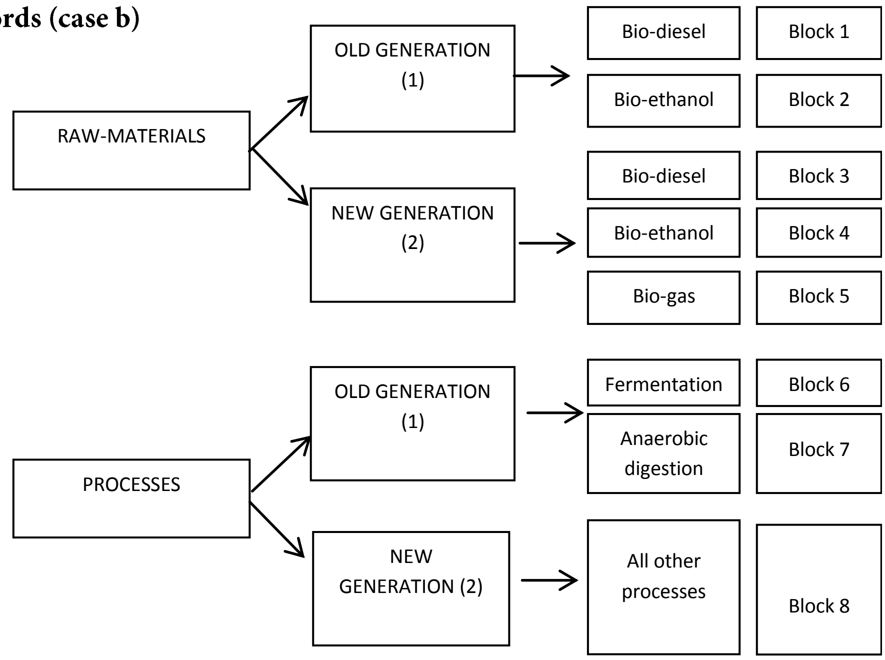
The raw material keywords can be divided into several categories which help to identify the patent's content: chemical agents, agricultural waste/crop, agricultural waste/ligno, algae, crops, GMO, ligno, livestock, oleaginous, sugar, urban waste and non-urban waste. Some keywords can overlap with more than one category. Obviously, different combinations are possible and numerous categories can be created. As an example, in Figure 12 and 13 we provide a combination of keywords and categories.

Figure 12Exemplificative alternative structures of database and classifications using keywords (case a)



TYPE	KEYWORD	BLOC	GEN.	FAT	ALCOOL
algae	Chlorella vulgaris	3-4	2	1	1
food	Corn	2	1	0	1
algae	Dunaliella tertiolecta	3-4	2	1	1
Food	Maize	2	1	0	1
sugar	Sorghum	2	1	0	1
ligno	Miscanthus	4	2	0	1
oleaginous	Jatropha	3	2	1	0
sugar	Bagasse	2	1	0	1

Figure 13. Exemplificative alternative structures of database and classifications using keywords (case b)



TYPE	KEYWORD	BLOC	GEN	DIES	ETHA	GAS
algae	Chlorella vulgaris	3-4	2	1	1	0
algae	Dunaliella tertiolecta	3-4	2	1	1	0
livestock	Anaerobic digestion	8	1	0	0	1
crop	Corn	2	1	0	1	0
crop	Maize	2	1	0	1	0
crop	Colza	1	1	1	0	0
crop	Soybean	2	1	0	1	0
ligno	Switchgrass	4	2	0	1	0
ligno	Miscanthus	4	2	0	1	0
ligno	Poplars	4	2	0	1	0
livestock	edible tallow	3-5	2	1	0	1
livestock	animal manure	3-5	2	1	0	1
oleaginous	palm oil	1	1	1	0	0
oleaginous	vegetable oil	1	1	1	0	0
oleaginous	coconut oil	1	1	1	0	0
oleaginous	Jatropha	3	2	1	0	0
sugar	Sugarcane	2	1	0	1	0
sugar	Sorghum	2	1	0	1	0
sugar	Bagasse	2	1	0	1	0

Database structure and preliminary descriptive statistics

The database was obtained using Thomson Innovation, which provides access to all the available information on patents. The collected information consisted of the 72 different fields listed in Table 10 that can be classified as follows:

1. Patent identification (international, national and office codes, patents' class)
2. Patent object (title, description, claims, abstract)
3. Patent owners (applicants, inventors, assignee, buyers)
4. Patentability process stages and dates (from the application to granted patent)
5. Patent opposition (other claims on the invention)
6. Patent quality (citation)

The information provided by the database can be used to study the impact of technological change on biofuel production, which is supposed to be large considering the weight of innovation effort on biotechnological sectors. It will also be possible to study the evolution of the sectoral innovation system using indicators that capture the dynamics of innovations, their concentration in terms of geographical location, holding companies and inventors.

Table 10. Information available in the BioPat database

Publication Number, Title (Original), Title (English), Abstract, Abstract (English), Claims, Claims Count, Claims (English), Description, Assignee/Applicant, Assignee/Applicant First, Assignee – Standardized, Assignee – Original, Assignee - Original w/address, Assignee Count, Inventor, Inventor First, Inventor – Original, Inventor - w/address, Inventor Count, Publication Country Code, Publication Kind Code, Publication Date, Publication Month, Publication Year, Application Number, Application Country, Application Date, Application Year, Priority Number, Priority Country, Priority Date, Priority Year(s), Related Applications, Related Application Number, Related Application Date, Related Publication Number, Related Publication Date, PCT App Number, PCT App Date, PCT Pub Number, PCT Pub Date, IPC – Current, IPC Class, IPC Class Group, IPC Section, IPC Subclass, IPC Subgroup, IPC Class First, IPC Class Group First, IPC Section First, IPC Subclass First, IPC Subgroup First, ECLA, US Class, US Class – Main, US Class – Original, Locarno Class, Cited Refs – Patent, Count of Cited Refs – Patent, Cited Refs - Non-patent, Count of Cited Refs Non-patent, Citing Patents, Count of Citing Patents, Citing Pat 1st Assignee, Litigation (US), Opposition (EP), Opposition (EP) – Opponent, Opposition (EP) - Date Filed, Opposition (EP) – Attorney, Language of Publication

The information collected can help to solve the problem of defining and measuring the magnitude of inventions and the problematic distinction between the cost of producing invention and the value it creates, containing many items of information such as the identity and the location of applicants and inventors, the technological area of the invention and citation of previous patents. The latter is a fundamental part of the total amount of information contained in the database. It follows a cumulative view of the process of technological change (Weitzman, 1996 and 1998) so that each inventor benefits from the work of colleagues before, and in turn contributes to the base of knowledge upon which future inventors build.

All information provided by the "patent opponent" section can be qualitatively exploited, among numerous possibilities (Scellato *et al.*, 2011) to verify if, due to existing connections between biofuel production and plants and, moreover, due to inter-linkages between biofuel raw materials and pharmaceutical raw materials (Harhoff & Reitzig, 2004), limitations to the patentability of living materials affect the innovation process of the sector. Starting from the TRIPs' model (Art. 27)[26], two main trends can be distinguished: a moderately liberal pattern represented by the U.S. patent system, and a more restricted system as designated by the European directive and, to some extent, by the EPO practice. "Since the adoption of the Agreement, the differences in the treatment of biotechnological inventions among developed countries have been reduced, but not eliminated", noting "plant varieties and animal races are not patentable in Europe, while they are eligible for protection in the USA" (UNCTAD-ICTSD, 2005, p. 388).

Differences in USA and EU patentability limitations and exclusions are just one of the aspects that can be studied. Patent applications can be viewed as a noisy indicator of the success of the innovation process, with the "propensity to grant a patent" possibly varying over institutions[27] (de Saint-Georges and van Pottelsberghe de la Potterie, 2011). Nevertheless, different regimes in patenting procedure are strongly reflected in the number of patents, the length of

[26] The Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement is Annex 1C of the Marrakesh Agreement Establishing the World Trade Organization, signed in Marrakesh, Morocco on 15 April 1994. The TRIPS agreement introduced intellectual property law into the international trading system. In 2001, the Doha declaration clarified the scope of TRIPS, stating, for example, that TRIPS can and should be interpreted in light of the goal "to promote access to medicines for all" and should respect the traditional knowledge of tribal communities. The declaration also mentioned the patentability of living materials. TRIPS also specifies that the protection and enforcement of all intellectual property rights shall meet the objectives of contributing to the promotion of technological innovation and the transfer and dissemination of technology, to the mutual advantage of producers and users of technological knowledge and in a manner conducive to social and economic welfare and a balance of rights and obligations.

[27] In fact, the USPTO is often criticized for its propensity to grant many low quality patents. See The Economist (March 17, 2011) and Lemley and Sampat (2008).

patentability iter and the scientific quality of the patents (that is mainly tested by using information on citation or through network analysis). Finally, comparing patents from different institutions can reveal which organization manages the possessed information better, making this information clear and available to everyone.

Patents citations represent a useful tool to skip over the variability problem in terms of patent value by quantifying the impact of knowledge contained in a specific patent on subsequent innovation through the analysis of citation data (Narin *et al.*, 1997; Jaffe and Trajtenberg, 2002). A patent can be weighted with the number of received citations. The number of patent citations can be used to characterize the technological and economic impact of a given invention providing a more meaningful measure of inventive output than a simple patent count. Moreover, patent citations can also represent an important instrument for studying some aspects of knowledge diffusion and technological spillovers (Jaffe *et al.*, 1993) such as the geographical distribution of citations, inventors and patentees (the issue of patent quality will be extensively discussed in chapter 3).

All the patents downloaded using our methodology amount to 1,293,197 patents (21% EPO, 59% USPTO, 20% WIPO). Then, using this initial information, we tried to make the database suitable for our purposes. First of all, in order to link each patent with the nationality of a specific applicant, we looked for country codes in the variable "assignee address" obtaining information on numerous countries. This allowed us to create a panel database that raises the number of studied countries, listed in Table 2.4, to 37[28].

At the present stage, given the difficulty of managing data deriving from different patent offices at the same time, we decided to start with an analysis of

[28] 37 represents the highest number of countries considered so far in a environmental technology field. For instance, Johnston *et al.* (2009) considered 25 countries.

data collected from the EPO source since it significantly reduces data management problems compared with other sources.

With regard to EPO patents, we subsequently asked the team of experts from ENEA to validate our database. We started validating the same classes indicated in the GI filtered with our keywords. The sample was built as follows: we took the EPO patents in our database; we selected the patents that shown at least one IPC class indicated by the GI, eliminated the duplicates and delivered 1% of the selected patents to the experts from ENEA.

Table 12 displays the number of patents divided by patent office for the main countries considered here[29].

The results of the validation are summarized in Table 13 which shows that our methodology allowed the percentage of patents actually related to the sector to be doubled. Additionally, the share of patents directly related to the investigated sector also increased.

In order to provide some preliminary descriptive evidence deriving from the collected information, Figures 2.2 show the evolution of patenting activity registered at the EPO since 1990 for US, Japan and EU countries. We display here only patents application, as common practice in the literature (Johnstone *et al.*, 2010; Picci, 2010), even though the typical caveats may apply. In particular we acknowledge that EPO is showing a very high rate of application refusals and patents application would not properly reflect the production of new knowledge. Nevertheless, patents applications offer a clear picture of productive innovative effort. Moreover, working with granted patents obliges us to cut the series considering the years required for the patents acceptance process (the lag between

[29] Our methodology results particularly effective for EPO because the address contained in the variable is consistent in all records. As shown by Table 2.5, the variable "assignee address" is not exploitable for USPTO.

application date and grant date), wasting 4 years of information, as suggested by EPO online forum.

Table 11. Selected countries in BioPat for descriptive statistics

US (United States of America), TH (Thailand), SG (Singapore), SE (Sweden), RU (Russia), PT (Portugal), NZ (New Zeeland), NO (Norway), NL (Holland), MY (Malaysia), MX (Mexico), LU (Luxemburg), KR (South Korea), KP (North Korea), JP (Japan), IT (Italy), IN (India), ID (Indonesia), HK (Hong Kong), GR (Greek), GB (Great Britain), FR (France), FI (Finland), ES (Spain), DK (Denmark), DE (Germany), CN (China), CH (Switzerland), CA (Canada), BR (Brazil), BE (Belgium), AU (Australia), AT (Austria), AR (Argentina), AE (Arab Emirates).

Table 12- Count of records and share of patents by main country and patent office

Country	Count	Share	EPO	WIPO	USPTO	EPO %	WIPO %	USPTO %
US	272,234	21.1	81,038	103,124	88,072	30.5	39.6	11.5
JP	129,683	10.0	79,158	5,465	45,060	29.8	2.1	5.9
DE	84,675	6.6	20,693	6,882	47,100	7.8	6.5	6.1
CA	55,348	4.3	3,100	7,528	44,720	1.2	2.9	5.8
GB	40,288	3.1	15,481	17,717	7,090	5.8	6.8	0.9
CH	28,633	2.2	11,153	10,787	6,693	4.2	4.1	0.9
FR	26,715	2.1	8,405	5,827	12,483	3.2	2.2	1.6
NL	18,433	1.4	8,937	5,802	3,694	3.4	2.2	0.5
Others	535,224	41.4	7,150	49,761	478,313	2.7	19.1	62.4

Table 13. Validation of BioPat for EPO patents: percentage of patents related to the biofuels sector

	GI	Share of biofuels related patents between direct and indirect application	GI filtered by keywords	Share of biofuels related patents between direct and indirect application
Direct application	5%	28	15%	40
Indirect application	14%	72	23%	60
Total	19%		38%	

It is worth reminding that patenting procedures are usually quite slow and 2010 application do not precisely reflect innovation activities of the very antecedent years, especially because applicants tend to apply for a patents in the national office before applying to an international one (De Rassenfosse *et al.* 2012). Fig. 14 shows the evolution of patenting activity for EU[30] , Japan and US from late '80s to 2010 as captured by the BioPat database and filtered by using GI classes.

The first graph displays a simple patent count whereas the second one weights patents by citations received. As the following chapter will extensively discuss, patents citations suffer from a severe time bias due to the evidence that older patents have more time to be cited. According to the literature, time bias may invalidate what clearly shown by data because we do not know whether the trend in citations attitude reflect a "real" phenomenon (more self-reliant patents or more original ones) or because of time lag (Hall *et al.*, 2001).

Since the computation of a sectoral time lag is beyond the scope of this study, we consider on average an estimation made by Bacchiocchi and Montobbio (2004) for EPO patents. According to their work, time bias should be depleted in 5-6 years, because EPO patents have little chance to receive a citation beyond the first 5 years after the publication. Even taking into account a 5 years time lag, some Figures show a fall in the citation attitude.

In addition, as shown by Table 14, in BioPat highly cited patens are very few and the oldest one has been applied in 1994. It is worth stressing that hose very uncommonly cited patents (note that on average the higher number of citations

[30] EU stands for European counties and include Germany, France, Italy, Spain, Portugal, United Kingdom, Sweden, Denmark, Norway, Switzerland, Belgium, The Netherlands, Luxembourg, and Greece.

received in BioPat/GI is 9) may bias the interpretation of citations trend in the last decades of the Figures.

Table 14. Highly cited patens

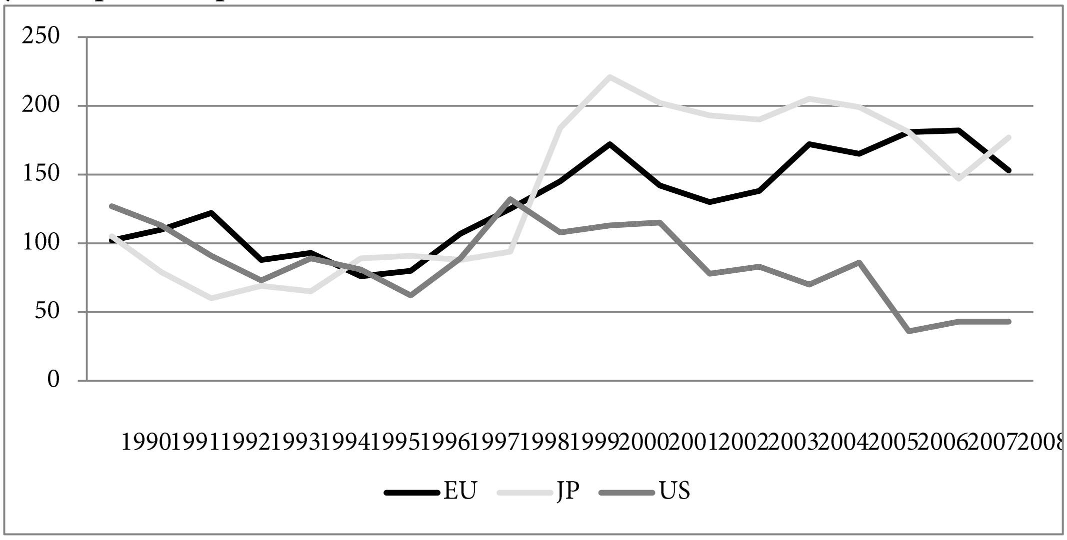
Patent	Application Date	Number of forward citations	Country
EP323753A1	29-Dec-1988	91	Canada
EP60057A1	24-Feb-1982	92	Usa
EP308341A1	16-Sep-1988	92	France
EP640619A1	16-Aug-1994	94	Usa
EP375091A1	18-Dec-1989	94	Germany
EP344029A1	26-Apr-1989	95	Belgium
EP531372A1	08-May-1991	95	Denmark
EP385962A1	20-Feb-1990	97	Usa
EP219756A1	06-Oct-1986	97	Germany
EP609079A1	27-Jan-1994	99	South Africa
EP155476A1	31-Jan-1985	99	Usa

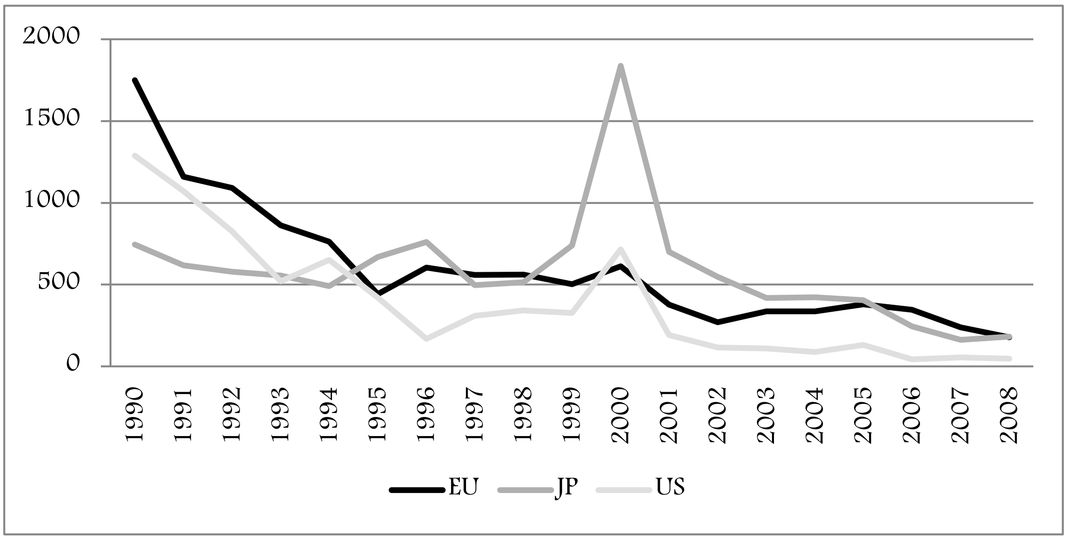
The Figures below show the general trend of biofuels patenting activity from 1990 to 2008. Figure 14 displays a clear common peak occurring at the end of 20th century, perfectly in line with the increasing attention given to climate change issues as well as to growing concerns for oil price volatility and peaks. Nevertheless, there are divergent trends in patent applications between US on the one side and Japan and EU on the other side. From 1999 EU and Japan lead the patenting activity in this sector showing a stable increasing trend, while the US reveals a persistent decline in patenting activity.

The Figure showing sum of forward citation gives us interesting information about Japanese patent activity. Japanese patents are usually more cited than the others. We looked into BioPat to verify whether there was a bunch of highly cited Japanese patents capturing citations peak occurring in year 2000. We found that the citations (usually 9 citations per patent) are quite balanced among a great number of patents revealing an average good quality of Japanese inventions.

The categorization of patents developed with our methodology using specific groups of keywords allows us to shed light on the specific composition of these trends. In particular, we present here the dynamics of patents for conventional and advanced generation biofuels, here represented by food-related biofuels and algae-related biofuels, respectively.

Figure 14. Number of patents and number of patents weighted by forward citations by year, Japan, European countries and United States of America



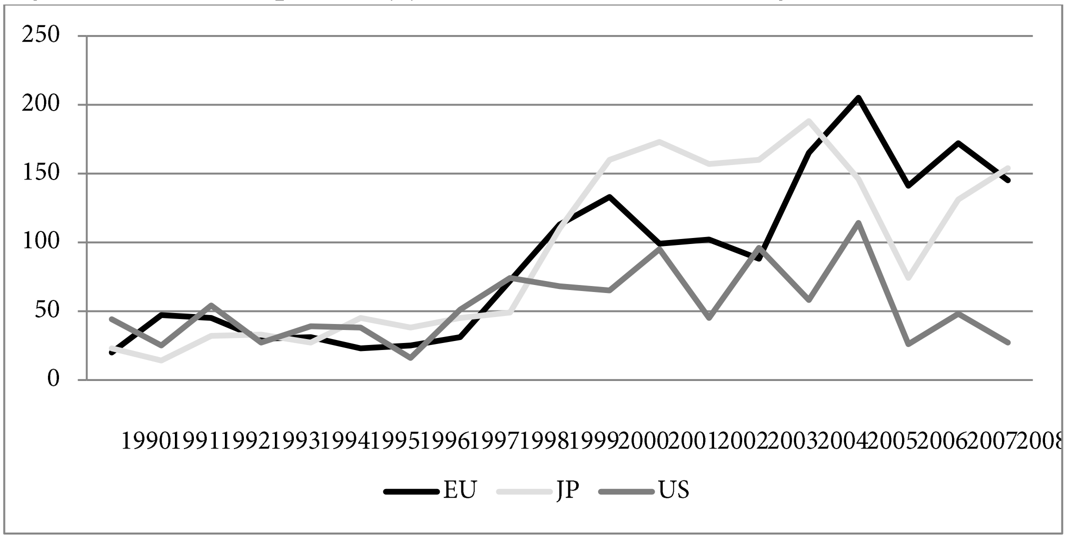


Source: own elaborations on BioPat

Figure 15, referring to food-crop related biofuels, seems to capture the variability of the second decade of the global trend. A global positive trend can be observed for Japan and EU, whereas US shown a more flat development.

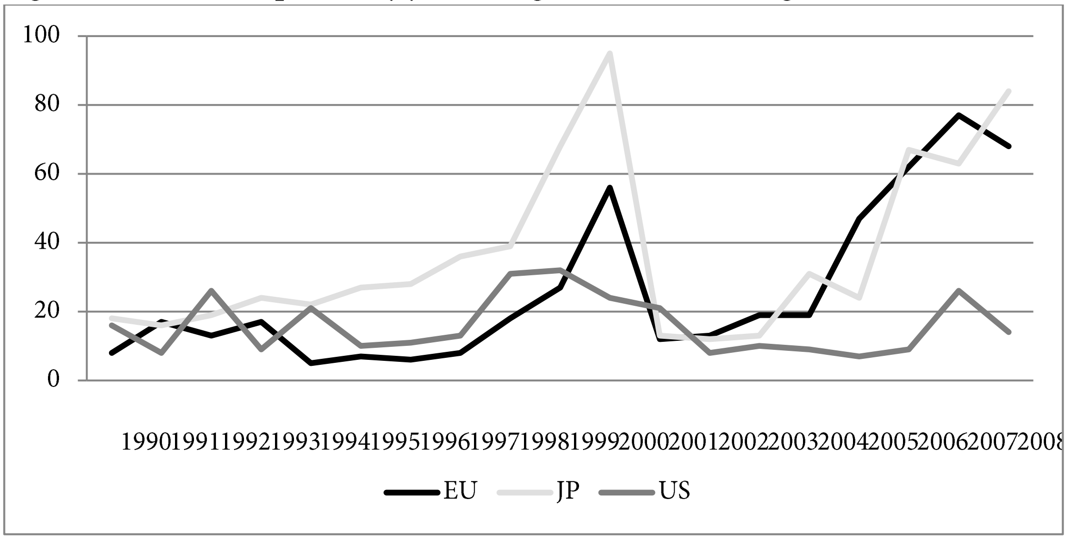
Algae-related biofuels described in Figure 16 show a very interesting trend with a peak in the year 2000, in line with other technologies, and an increasing trend for EU and Japan after 2001.

Figure 15. Number of patents by year for food-related technologies



Source: own elaborations on BioPat

Figure 16. Number of patents by year for algae-related technologies



Source: own elaborations on BioPat

Further characterizations of the biofuels sector through BioPat

As already mentioned, the time frame between radically-new knowledge and a new radically-new knowledge is characterized by a mix of slight intra-sectoral and extra-sectoral adjustments, depicting the so-called incremental innovation. When these modifications of the "state of the art" lose their lucrative appealing, innovative effort is steered towards the hunt of new opportunities leading to the recombination of more and more distant technologies.

In order to verify whether this hypothesis is confirmed in the biofuels sector, we can give a first look at the emergence of new technological fields (as represented by IPC classes) where biofuels patents trespass in.

Starting from BioPat EPO applications included in the GI list, we observed that 6,835 patents have been placed into more than 6,000 IPC subgroups across 30 years (1980-2010), that is to say that only few biofuels patents have been classified in the same subgroup. This impressive finding reveals two distinct aspects: on the one hand, IPC subgroups are increasing exponentially losing any appeal when it comes to identify an economic sector and, on the other hand, once again biofuels established themselves as a very crosscutting sector.

We performed the same research using a 4 digit categorization and we found that from 1980, biofuels patents have been placed in 233 unexplored IPC subclasses, both new classes in BioPat/GI (classes that do not appear in the BioPat/GI database before year 2000) and radically new classes in the IPC system (classes that were created by WIPO after year 2000).[31]

Starting from January 2000, 46 new IPC subclasses (referring to 42 different patents) appear (Table 3), denoting a behaviour that we cannot univocally identify

[31] The emergence of new classes in the IPC system does not allow us to easily distinguish between this particular phenomenon and the peculiar expansion of the biofuels sector.

as a propensity to a higher dynamism.[32] Given the crosscutting applicability of biofuels-related inventions, we do not know whether the sector is expanding itself or it is simply interacting with different technological fields. In other words, an invention may physiologically find multiple uses in several areas: reasonably we cannot conclude that, even if biofuels and the printing industry are interacting, printers production can be included in the biofuels sector (while we may suppose that cellulose pre-treatment has manifold applications).

In addition to considerations related to the amount of new IPC subclasses, we can perform a more qualitative analysis by looking at the IPC subclasses objects. Table 15 shows 29 of the 46 new IPC subclasses selected in order to reduce repetitions among similar classes (similar class objects). From a first look we can acknowledge that during the last decade, the biofuels sector started facing some barriers due to peculiar characteristics of its products: shipping difficulties linked to different organic nature of biofuels, metals as fuels additives to improve engines adaptability, pipes production, issues related to fuel conversion into electricity, cooling for engines and so on. Moreover, we can observe that BioPat/GI does not catch any patents included in C13J (extraction of sugar from molasses) and C13B (production of sucrose) before 2005. A reasonable explanation of this surprisingly late appearance may be that co-production of sugar and bioethanol is a recent issue and that the biofuels sector has just recently crossed the frontiers of very traditional and not innovative production processes.

As shown by Table 15, the biofuels sector has also interacted with fertilizers and feed industries, as the full exploitation of by-products, co-products and waste material is one of the key elements of the international biofuels technology roadmap (IEA, 2011).

[32] It is worth noting that we started from BioPat/GI database. Thus, new classes are not necessarily included in the GI list, but they always appear with at least one GI class.

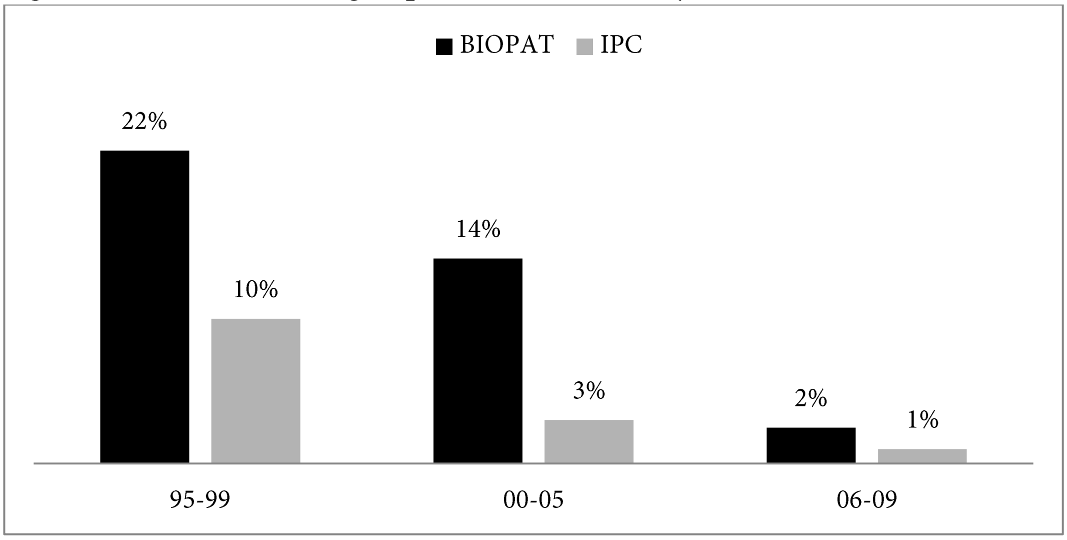
Table 15. Selection of new IPC classes included in BioPat/GI from 2000

Year	Patent	IPC	Class object
2000	EP1167861A1	F17C	vessels for containing or storing compressed, liquefied, or solidified gases; fixed-capacity gas-holders; filling vessels with, or discharging from vessels, compressed, liquefied, or solidified gases
2000	EP1048614A1	C01G	compounds containing metals
2000	EP1548454A1	G01R	measuring electric variables; measuring magnetic variables
2000	EP1215172A1	H05B	electric heating; electric lighting not otherwise provided for
2000	EP1145755A1	F28D	heat-exchange apparatus, not provided for in another subclass, in which the heat-exchange media do not come into direct contact
2000	EP1683780A1	F25J	liquefaction, solidification, or separation of gases or gaseous mixtures by pressure and cold treatment
2001	EP1123896A1	C12F	refrigeration or cooling; liquefaction or solidification of gases
2001	EP1142963A1	B41N	printing plates or foils
2001	EP1320388A1	C05B	phosphatic fertilisers
2001	EP1340800A1	B60K	arrangement or mounting of propulsion units or of transmissions in vehicles; arrangements in connection with cooling, air intake, gas exhaust, or fuel supply, of propulsion units
2002	EP1482087A1	C01D	compounds of alkali metals, i.e. lithium, sodium, potassium, rubidium, caesium, or francium
2002	EP1293128A1	A22C	processing meat, poultry, or fish
2003	EP1431374A1 EP2292722A1	F02D	controlling combustion engines
2003	EP1515916A1	B63B	ships or other waterborne vessels; equipment for shipping
2004	EP1616968A1	C21B	manufacture of iron or steel
2005	EP1712561A1	D01B	mechanical treatment of natural fibrous or filamentary material to obtain fibres or filaments
2005	EP1778851A1	C13B	production of sucrose; apparatus specially adapted therefor
2005	EP1778851A1	C13J	extraction of sugar from molasses
2005	EP1806401A1 EP2295566A1	B62D	motor vehicles; trailers
2006	EP1967248A1	F04F	pumping of fluid by direct contact of another fluid or by using inertia of fluid to be pumped
2006	EP2078746A1	A01M	apparatus for the destruction of noxious animals or noxious plants
2007	EP2037202A1	F28F	details of heat-exchange or heat-transfer apparatus, of general application
2007	EP2044845A1	C13D	production and purification of sugar juice
2007	EP2077311A1	F23B	methods or apparatus for combustion using only solid fuel
2008	EP2149625A1	C25B	electrolytic or electrophoretic processes for the production of compounds or non- metals
2008	EP2212426A1	E04H	buildings or like structures for particular purposes; swimming or splash baths or pools; masts; fencing; tents or canopies, in general
2008	EP2227324A1	F24J	production or use of heat
2009	EP2098365A1	F16L	pipes; joints or fittings for pipes; supports for pipes, cables or protective tubing; means for thermal insulation in general
2009	EP2103914A1	G01F	measuring volume, volume flow, mass flow, or liquid level; metering by volume

Even though it is possible to extrapolate some interesting information from this simple analysis, IPC classes do not seem to be the best data to investigate on the dynamism of a crosscutting sector, because it is not possible to discriminate from sector dynamic and technological fields interactions (such as vessel for biodiesel transportation) and an invention suitable for a variety of applications in clearly separated fields (such as cellulose production, which can be related to both ethanol production or fabrication of tents).

In order to conclude that the biofuels sector has experienced an expansive dynamic we had to compare the percent variation of BioPat (EPO, at least one GI class per patent) IPC classes with the percent variation of IPC classes in general. This allows us to discriminate between a real sectoral expansion and a multiplication of the patent classes due to updated versions of the classification system. The Figure 17 shows the percent variation for 2 lustrum, corresponding to IPC version 7, version 8 and the period 2006-2009.

Figure 17. Variation in classgroup number in the IPC system and in BioPat



Source: our elaboration on BioPat and Wipo IPC statistics

We considered variation in the number of IPC classgroup the higher detailed class specification available for the whole period (subgroup did not exist in previous versions of the IPC system). The Figure clarify that there is a considerable increase in BioPat IPC groups which is not explained by the growth in IPC groups *per se*.

Through the appearance of new IPC classes in the BioPat/GI database it is possible to give a broad picture of the expansion over time of the biofuels sector. In addition, we have explored the possibility that an expansion could be observed outside the GI classes. As a preliminary check we went back to TI and we performed a search using the keywords list (the general list including broad terms such as biofuel or biodiesel) specifically excluding GI classes. This allowed us to isolate the non-GI patents included in BioPat. These patents are likely to be directly linked to the biofuels sector thanks to the keywords selection tool. In order to make sure we were excluding non-biofuel patents as much as possible, we searched only in title and abstract excluding the description field. The outcome of this procedure is extremely interesting: we found an amount of patents numerically equivalent to the BioPat/GI database.

Chapter 4. An adjusted text mining approach to network analysis of biofuels patent data

Chapter overview

In recent years several attempts have been made to depict technological trends and innovation's features in several fields and a particular attention has been paid to the hunt of the right proxy for innovation. Among others instruments, patents seem to offer great advantages and a wide range of uses: patents represent a deep source of information, fairly often harmonized and extensively available.

Through patents exploitation it is possible to show basic statistic description and more sophisticated elaboration such as complex indicators and visualization methods. In this work we will discuss some traditional patents analysis and illustrate how one of the most recent techniques, the network analysis, is dealing with some outstanding issues of the same.

The proliferation of studies in the field of social network analysis has brought to the creation of a number of tools that consent visualization, elaboration and interpretation of patents data, allowing researchers to understand and explain the relationship between technology fields (IPC classes, keywords), patent applicants, inventors, patent documents, etc. (Sternitzke et al., 2008).

In this work we applied network analysis of patent data to the biofuel sector in order to understand the transition of transports towards sustainable fuels.

Background and aim of the chapter

According to this study, the transport sector conversion to sustainable fuels represents a drastic shift in industrial and consumption pattern which cannot be achieved by doing things slightly different. Energy is a fundamental input in almost all economic activities and newly invented products have to compete with well understood and fully implanted technologies which are very difficult to modify (Hughes, 1987). This phenomenon, also known as carbon lock-in (see Chapter 2), has been studied by Unruh (2000) and represents an extension of a more general technological lock-in. The idea is that existing fossil fuel based technologies are not fully adaptable to new forms of energy, resulting in a deceleration of the transition towards green technologies.

According to evolutionary approach, such a huge transition brings along issues of uncertainty and complexity that have to be dealt with specific 'transition policies': "Policy support is needed and warranted because new technologies are often only promises" (Nill & Kemp, 2009 , p. 672). As Rosenberg found in the Black Box, (1982) an innovation commonly requires the environment to be adapted to its use because it rarely meets user needs, producer know how and it may also require accessory technology and purposely-built infrastructure.

Once again, biofuels match the 'perfect' case study of a self-consistent sector[33] because transition of the transport sector towards sustainability consists in several important changes that are taking place along the whole biofuels' lifecycle. In the light of this, it is crucial to map properly in-sector technological evolution.

The aim of this research is twofold: the first one is to develop an original research method based on network analysis to be applied to narrowly defined technological domain; the second one is to apply such a new methodology to the biofuels sector as a case study, in order to map connections inside this technological domain.

From the methodological side, we provide the following contributions:

---

[33] Biofuels sector cover the entire biofuel production value chain

- we develop a method to look at patent connections in a specific technological domain by replacing patent citations with keywords;
- we adopt an appropriate measure for text similarity when comparing patent documents;
- we propose a method for the identification of the cut-off value in the definition of the network;
- we adopt a method for the identification of sub-network topics.

The paper is organized as follow:

Biofuels as a complex technological domain

Even though the hunt for alternative solutions to fuel scarcity in the transport sector began the day after the first petroleum crisis almost four decades ago, the technology and the knowhow needed to produce biofuel in a non-commercial scale was acquired at the beginning of the 20th century. The reasons for wanting alternatives are increasing over time: high and volatile prices of oil, uncertainty of supplies, uncertainties related to reliance on supply from politically unstable regions, raised awareness of pollution damages, climate change concerns etc. (Sterner and Coria, 2011).

In this intricate status quo, biofuels has been acknowledge as one of the key answers for the transport sector explaining the proliferation of domestic and international policies fostering biofuels and global declarations of intent addressing this issue. Biofuels promotion has raised concerns about the effective environmental-social-economic benefits originated by the explosion of biofuel production and use especially due to: great uncertainty related to net savings of greenhouse gas emissions compared to traditional fossil fuels and to the scope of the impact of biofuels production on prices of raw material diverted from food and feed purposes (Holmgren, 2012).

Both concerns presented above are strictly interlinked to technological state of the art in the sector: biofuels have great potential in terms of emissions saving and are environmentally friendly from many side, but innovative efforts are required in order to fully exploit this endowment. The issue of direct and indirect land use and land use change is closely associated to the use of by-products and production residues as biofuels feedstock (as an example it is possible to mention the use of sugar as direct input in ethanol production versus ethanol produced by means of lignin content of sugar bagasse) and inefficient use of energy units during biofuels processing (Fargione et al., 2008; Searchinger et al., 2008), can be dealt through proper innovative solutions.

Although the debate biofuels vs food has recently cooled down recognizing that it was much exaggerated (Sterner and Coria, 2011), a strong interest remains in 2nd generation biofuels, which are define as biofuels manufactured from agricultural and forest residues and from non-food crop feedstock, possibly cultivated in poor-quality land (IEA/OECD, 2008).

In this context, significant investment in R&D is occurring and biofuels innovation is booming and innovative activities are very intense.

Patent analysis and patent quality

As mentioned in Chapter 3, the measurement of innovative activities is mainly based on two different inputis a rather challenging task and a great number of different science and technology indicators have been identified in the literature (Sirilli, 1997). The main input indicator relies on research and development (R&D) expenditure, while the most used innovation output indicators are based on patent data. Both types of indicators have strong limitations as not all research efforts translate into the introduction of innovations and not all innovations are patented (Archibugi & Pianta, 1996).

In a desert of data, patents have demonstrated to be a quite manageable proxy of innovation. Nevertheless, patents have pros and cons. The advantages of using patents as a proxy of innovation are manifold. On the other hand, using patents as a proxy for innovation presents several relevant issues especially related to patents heterogeneity in terms of patent quality (Griliches, 1990).

The hunt of a comprehensive and univocal definition of patent quality is not an easy one. To simplify what stated by Burke & Reitzig (2007), patent quality may refers to the techno-economic quality created by the patent‘s underlying invention; and the legal quality created by the patent‘s reliability as an enforceable property right. For the purpose of this study we discuss only the patent quality related to the techno-economic aspects[34].

As the techno-economic quality heterogeneity decreases the meaning of patent counts as measure of innovation output, empirical research habitually weights patent counts by indicators of the quality of patents, usually observable quantitative characteristics (such as the number of citations a patent receives by posterior patents, the number of claims, numbers of inventors, size of the patent family, the number of patents cited as prior art) or more complex index (such as the patent’s generality index and originality index) (Griliches, 1990; Baron & Delcamp, 2010).

Among all, since the seminal work of Garfield et al. (1964), the relevance of the number of citations a patent receives as valuable tool to quantify the impact of the ‘publication channel’ on technology development (Narin et al., 1997) has been

[34] For a vast discussion of the procedural quality of patents see Scellato et al. (2011)

widely assessed and confirmed. Starting from Trajtenberg (1990), it has been pointed out that patents that cite a high number of patents play a greater role in consumers and producer welfare, Harhoff et al. (1999) demonstrate that assignee value more those of their patents that are more frequently cited, Lanjouw & Schankerman, (1999) prove that the more patents are cited, the more expected to be opposed or to be included into technological standards (Rysman & Simcoe, 2008). Afterwards, Harhoff & Reitzig (2004) demonstrated that high frequently cited patents are more expected to be litigated. In 2003, Giummo shows that there is a high positive correlation between the number of citation a patent receive and its probability to be licensed. In a different approach, Lanjouw & Schankerman (2004) analyze the causes of changes in measured research productivity (the patent/R&D ratio) focusing on three aspects: the level of demand, the quality of patents, and technological exhaustion. Using a factor model, the authors build a minimum variance index based on the numbers of claims, forward citations, backward citations, and patent family size, arguing that multiple indicators largely decreases the measured variance in patent quality. Moreover, through the factor analysis they demonstrate that using a common determinant of many indicators rather than a single one permits the reduction of the noise and enhances the capabilities of indicators to estimate patent quality.

As mentioned above, patents citations may be helpful in a different way representing an important instrument to study the features of innovations' spillover, such as its geography, surveying inventors whose patents are cited and patentees whose patents made citation (Jaffe et al., 1993). However, patent analysis suffer from some weaknesses, especially when it comes to use the so called backward citation (made by a patent) and forward citation (received by a patent) to build a patent network. First, citations are time biased: older patents may have more citations due to mere time lag (Hall et al., 2002). Second, citations, especially in the European Patent System[35], may reflect individual links among few patents, making difficult to grasp the overall picture. Third, citations have no capability to capture self-citations, making impossible the discrimination between internal and external relations and forth the density of the patent web in a complex industry automatically affects the average number of citations because, independently of its quality, a patent will be more cited if its content is included in a technological area where the propensity to patent is high and, for the same reason, a patent in such a crowded web will have more previous art to cite than a patent in less dense field (Baron & Delcamp, 2010).

While not denying the usefulness of patent citations for the construction of patent quality indicators, in this study we deviate toward an alternative approach, that result more effective when it comes to elaborate on a very narrow selection of patents. In fact, we are not interested in studying the inter-sectoral connection among biofuels patents and those that they cite, ending up analyzing patents that are not in the biofuels domain, but we are more concerned about measuring the relevance of certain production paradigms and understanding the intra-sector closeness among different technologies.

Network analysis

A Network can be identified in different type of analyses: several actors interacting for a single purpose, knowledge spillovers across countries or institutions, inventions building incremental capabilities. Many authors exploit the same source of information, a patent, discriminating networks from the

[35] This is not particularly true for USPTO patents. According to USA legislation, a patent has to cite mandatory each single patent that may have influenced the applied one in order to skip unpredictable oppositions related to not clearly acknowledged references. The logic result of this practice consists in a proliferation of non-discriminated citation that may eventually help in the construction of a wide patent network.

analytical unit that is taken into account (Huang et al. 2004; Narin, 2000; Consoli & Mina, 2009). To the aim of this work we will consider an innovation network, based on the innovation content, as separate from an innovators network[36], based on innovative actors.

A patent object (an invention) shows interactive relations with other patents as it were a living actor. This interaction can be represented by mean of a network: social network analysis traditionally explores the connections (also called "ties" or "edges" or "links") among 2 or more actors ("nodes" or "vertices"). When taking patent as a source of information, nodes can represent inventors, applicants, countries but also invention itself. In this context, a node can be represented by many variables.

Historically, citations and co-occurrence of classification codes have been extensively used to observe patents relationship [37](see Chapter 3). Citations are usually defined bibliometric data. They are strongly harmonized and easy to understand. Together with innovators' information and other information, bibliometric data are known as "structural data". Many researchers have developed patents analyses based on them but, on the other hand, structural data suffer from some limitation in terms of explanatory and creative capacity (Lee et al. 2009).

In the present work we take into consideration the patent content as analytical unit and we propose an alternative approach to structural data analysis.

Data mining, in particular text mining is used to extract information from text document. Here we propose a slightly modified version of traditional text mining methods to depict patents linkages, in order to skip over well recognized

[36] The importance of network of innovators has been analyzed in Powell and Grodal (2006).

weaknesses of traditional approaches (see section above) and exploit the advantages of unstructured data.

According to text mining logic, a text document can be characterized by means of keywords extracted using the Text Mining Algorithm (Weiss e al. 2005). The more two texts share the same keywords, the more similar they are. The application of this algorithm can easily results in a collection of non-discriminating terms poorly useful in the definition of text similarity. In the light of this, we used text mining as an initial screening instrument and then we combined it with expert knowledge and statistical relevance[38.]

Methodology

Data Source and sample selection

Because of the inaccuracy of the traditional methods (IPC, patents families, citations) in detecting Biofuel-related patents, we adopted a database built using a keyword approach.

As described in Chapter 3, BioPat is a database containing patents (directly and indirectly) related to the biofuel sector.

Considering that the visual tool is not able to show sectoral relations when the data set is not perfectly coherent with the investigated field, and when is too large to clearly observe connections, we started from BioPat patents to select a representative sample of the biofuels sector. We proceeded as follow:

1. We sorted the database by publication number and we selected those rows containing a code beginning with E in order to pick only EPO patents (for further information on classifications numbers and code see Chapter 3)

[38] A very recent work of Cecere et al. (2012) uses co-occurrence of IPC classes to depict patents interactions Different approaches are also possible (see Chapter 3)

Explanation: we limited our analysis to EPO for two main reasons. First of all, EPO patents are less noisy and easier to handle (see chapter 2, section XX) than WIPO and USPTO data. Second, working on more than one patent office increase the risk of considering redundant information and insert a bias due to differences in institutional practices such as a different propensity to grant, different procedural time length, different training in patents reviewers, different interpretations of Kind Codes and so on.

Number of patents: 250.638

2. We selected A1 and A2 Kind Code.

Explanation: we decide to work with patent application because it has been acknowledged that application better represent the innovative effort of firms and institutions. There are several reason why patents do not successfully reach the grant stage and only few of those having reference to a lack of awareness on the existing knowledge (that we can interpret as a “false” invention).

Number of patents: 112.925

3. We eliminated duplicates sorting the database by publication number.

Explanation: Publication number univocally identifies a patent. BioPat has been built using a methodology that allow for more than a few repetitions in the patents download. For example, it is likely that a patent that contains terms referring to biodiesel produced by algae (such as chlorella) would be captured also in the research that uses FAME as keyword. In the same way, an invention related to a method to increase starch level in several cereals, it has been downloaded once for each different cereal name (such as oat, wheat, corn and so on).

Number of patents: 55.747

4. We eliminated redundant patents using Levenshtein minimum distance on Claims text and we selected those patent belonging to at least 1 International Patent Class included in the Green Inventory database (see annex I).

Explanation: there are several reasons why the same patent comes out with two or more different publication number. These reasons are mainly related to procedural issues and changes in patents attributes, but rarely involving changes in patents inventive content. In the light of this we consider patents sharing the same title, abstract and claims, as a unique patent. We performed the elimination by using an algorithm which allows for slight in-text modifications in order to avoid spelling error and minor text changes. The decision of working on patents belonging to Green Inventory IPC classes is merely related to database validation. As shown in section 2.X, so far only this part of the database has been validated by experts and up to now we do not have any precise information on the level of sectoral coherence of the whole database.

Number of patents: 6.835

5. We selected those patents showing in title and abstract a positive result for very general main keywords: *thanol (such us ethanol, bioethanol, methanol, biobuthanol etc.) biodiesel (including variations such as bio diesel and bio-diesel) and biofuel (including variations such as biofuels, bio

fuel, bio-fuel, etc.). Patents whose main purpose was not related to biofuel production were not totally excluded by the procedure, but the alien patents included in the dataset are estimated to be less than 5%.

Explanation: To make sure the sample matched the biofuel sector as closely as possible; patents were filtered searching for main keywords only in fields that usually contain words strictly referring to patent's purpose, such as title ad abstract. The procedure described above responds to a very diffuse and acknowledged issue related to the manifold usages of biofuels raw materials and production processes. In particular with the main keyword "ethanol" it is reasonable to assume that not all patents showing a positive result in the keywords search would have production of ethanol as their final target. This method could represent a valid and replicable alternative to a human performed screening and revision (Charlita de Freitas & Kaneko, 2012).

Number of patents: 273

6. We eliminated non-English text by applying a R package called 'textcat' categorize texts by computing their n-gram profiles, and finding the closest category n-gram profile.

Explanation: by searching for main keywords, we captured few non-English document (both because those keywords are usually the same also in other languages and because there is a text field called "Title/Abstract/Claims English" where original document are often, but not always, translated into English). Because some patents are registered in international offices with original documents and English translations are not compulsory for all the text fields, it not possible to apply a text mining approach in text written in other languages.

Number of patents: 246

7. We validated the procedure screening 50 entries and we found 1 no "strictly biofuel" patent.

Keywords vectors

Using the sample of 246[39] patents we computed the frequency of keywords occurrence in patent documents to create keyword vectors.

We built keywords vectors as follow:

- We took all the keywords proposed in BioPat
- We searched for positive results in 4 different text fields of the 246 patents included in the sample (title, abstract, claims, description) obtaining 246 keywords vectors.

Es.

	K1	K2	K3	K4	K5	K6
P1	2	0	0	6	1	0
P2	3	0	0	10	0	3
P3	1	0	2	2	0	0
P4	2	0	0	1	1	7

[39] This number is sufficiently small to allow visual representation of the network but is big enough to be a significant sample of the BioPat EPO GI database (~6800). According to expert validation (see Chapter 3) only a percentage of the patents included in the BioPat EPO GI database (~40%) are actually related to the biofuel sector. Hence, the network represents ~10% of the database.

- We kept those keywords appearing more than once at least in two different patents, ending up with 109 terms

Es.

	K1	K2	K3	K4	K5	K6
P1	2	0	0	6	1	0
P2	3	0	0	10	0	3
P3	1	0	2	2	0	0
P4	2	0	0	1	1	7

- We divided the frequency of keyword occurrence by the number of words present in the text fields in order to account for different length of patents documents.

As shown by matrix$_{14x14}$ (Table 16), keywords occurrence can vary significantly. Obviously, very general and highly sector-related keywords (such as yeast or ethanol) show an extremely frequent occurrence with respect to the average of the used terms.

Table 16. Keywords occurrence

Patent	K6	K9	K14	K16	K17	K20	K22	K23	K30	K35	K39	K45	K59	K60
EP0424117	0	0	0	0	0	4	0	0	0	0	2	4	0	0
EP0470312	0	7	0	0	0	47	0	49	3	1	55	1	0	1
EP0511238	0	0	1	4	0	48	0	0	5	3	6	66	5	21
EP0512768	0	0	0	0	0	28	0	0	0	0	0	0	0	0
EP0527758	0	6	0	0	0	25	0	1	1	0	39	25	1	1
EP0565647	0	3	0	0	0	8	1	0	0	0	1	72	0	0
EP0576621	1	1	0	2	1	410	16	0	89	13	70	298	14	42
EP0645094	4	0	0	0	0	34	0	0	0	1	4	37	0	3
EP0645456	0	0	0	0	0	113	2	0	0	2	2	45	0	0
EP0718392	0	0	2	0	0	41	22	9	1	0	0	0	0	0
EP0741794	0	1	0	0	1	87	4	2	62	1	0	98	0	47
EP0745576	0	0	0	0	0	20	0	29	0	0	0	0	0	0
EP0770685	0	0	0	0	0	23	0	0	0	0	16	0	0	1
EP0792932	0	7	0	0	0	18	0	1	0	12	9	0	0	0

Performance of patent network analysis

We use the final list of keyword obtained from the above procedure, using both absolute and relative frequency, to compute incidence matrix to calculate the similarities among vectors.

Yoon and Park (2004) and literatures originated from their work (Chang et al., 2011 and 2012,) proposed a measure of similarity based on the Euclidean distance: the less the distance the more the similarity. The use of this measure presents some crucial deficiencies. In particular, when comparing zeros Euclidean distance treats missing information as identical information. For the sake of argument, consider the following keywords vectors:

	K1	K2	K3	K4	K5	K6	
Pa	0	0	0	0	0	0	
Pb	0	0	0	1	0	0	TOT
E.D.	0	0	0	1	0	0	1

	K1	K2	K3	K4	K5	K6	
Pc	4	0	0	6	1	3	
Pd	5	0	1	5	1	3	TOT
E.D.	1	0	1	1	0	0	3

According to the Euclidean distance, patent a and patent b are more similar than patent c and patent d, because the sum of the distances is 1 vs 3. We acknowledge that a non-co-occurrence of keywords (zeros) has to be considered as missing information and not as a clue for similarity.

Furthermore, consider K1 and K3 of the second two vectors: the Euclidean distance measure the similarity as 1 in both case but we tend to believe that K1 denotes a more stringent relation between the two patents with respect to K3.

Additionally, the Euclidean distance does not take into account the importance of higher numbers with respect with small ones. In fact, it is reasonable to assume that the higher the number, the heavier the similarity:

	K1	K2	K3	K4	K5	K6
Pa	1	1	0	0	0	2
Pb	1	1	0	0	1	1
E.D.	0	0	0	0	1	1

	K1	K2	K3	K4	K5	K6
Pa	5	0	11	6	17	3
Pb	6	0	11	6	18	3
E.D.	1	0	0	1	0	0

According to the Euclidean distance the two couple of patents are similar with the same magnitude, while it is logic to deduce that the second two are much more connected.

In order to skip over the weaknesses described above, we propose an alternative measure of similarity given by the cosine function. Given two keywords vectors (V) and (t) the number of time n terms occur in patents documents Vj =(t1j, t2j, t3j,…., tnj), Vk =(t1k, t2k, t3k,…., tnk):

$$(1)\ \mathrm{sim}(V_j, V_k) = \cos V_j, V_k = \frac{V_j V_k}{\|V_j\|\|V_k\|} = \sum_{i=1}^{n} t_{ij} t_{ik} \Big/ \sqrt{\sum_{i=1}^{n} t_{ij}^2 \sum_{i=1}^{n} t_{ik}^2}$$

The cosine function allows discriminating between similarities among zeros or among zero and a positive number and similarities among numbers. Additionally, the cosine function weights more those similarities occurring between two large numbers. Table 17 shows a 4x4 sample of the 246x246 original matrix containing the association value (1).

Table 17. Matrix of association value

	EP0527758	EP0818535	EP0884391	EP0898616
EP0527758	**0**			
EP0818535	**0,109163588**	**0**		
EP0884391	**0,777092708**	**0,814174709**	**0**	
EP0898616	**0,084507930**	**0,010843517**	**0,778228497**	**0**

The nxn matrix resulting after computing the similarities among vectors (association value) is called adjacency matrix(Table 18).

In order to define whether an edge exists or not, we need to define a cut-off value that determines the threshold of a "connection" or a "non-connection" between two nodes. That is, the connectivity between Patent A and patent B is considered strong and set to 1 in the incidence matrix if the association value between the patents is larger than cut-off value. Otherwise, the two patents are considered non-connected and the element of the incidence matrix is set to 0. The network becomes denser as the cut-off value becomes lower, whereas it becomes sparser as the cut-off value becomes higher. The cut-off value is usually determined in a subjective way by applying a trial-and error procedure in order to reach, at some point, a reasonable value so that the structure of the network acceptably appears.

We tried to skip over this subjective evaluation by using a measure able to give the simpler depiction of the network while keeping its nodes as much connected as possible. To do so, we use the Krackhardt efficiency:

$$\text{Efficiency}= 1 - \left[\frac{\text{V}}{\text{Max V}}\right]$$

Where the numerator is the number of redundant edges and the denominator is the maximum possible number of redundant edges. Hence, the efficiency is one minus the weighted average component inefficiency weighted by component size. Efficiency vary from 0 (min) to 1(max).

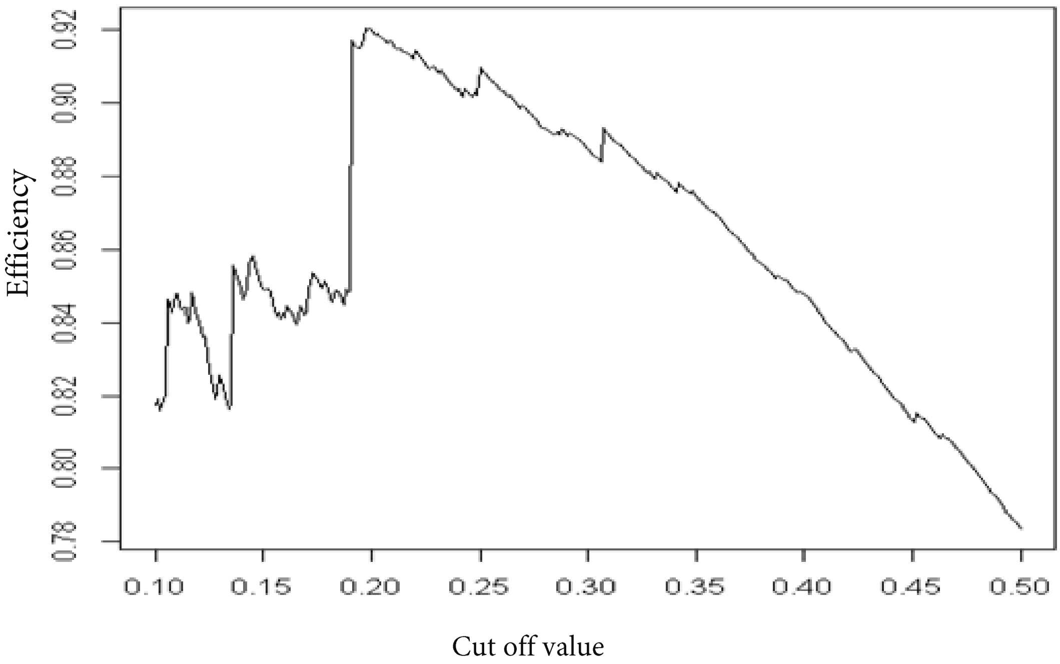
The highest efficiency is reached when the cut-off value is set to 0.2. Figure 3. X show efficiency evolution for values from 0.1 to 0.5 containing connected and not connected patents according to the most efficient cut-off. This matrix is commonly known as adjacency matrix, a nxn matrix where the non-diagonal entry represents a connection (1) or the absence of connection (0) of patent i with patent j.

In Table 18 we display a 4x4 example of the 246x246 original matrix

Table 18. Matrix of adjacency

	EP0527758	EP0818535	EP0884391	EP0898616
EP0527758	0	1	0	1
EP0818535	1	0	0	1
EP0884391	0	0	0	0
EP0898616	1	1	0	0

Figure 3.x Krackhardt efficiency Max = 0.197



Network identification

The issue of collecting patents related to biofuels is basically linked to the cross-cutting usefulness of inventions: a patent can have a very standard object with no distinctive application, making impossible the identification of a specific category depicting the sector. In the light of this, the only way to analyze biofuels' patents is to take into consideration generic but pertinent patents and useful inventions to a greater or lesser extent intended for the production of non-biofuels products. Hence, under patent classification systems, biofuels sector results randomly scattered around undetectable patent classes. It is possible to conclude that biofuels categorization is latent within a non-sector specific classification and pro-Biofuels patents have to be identified among disparate inventions, as for instance, by making use of keywords.

Several attempts have been made in order to provide a strong tool to evaluate the real effectiveness of text mining results. Tseng et al (2007) described a series of text mining techniques suggesting different approaches to avoid the unpleasant and common exercise of collecting useless words. Lee et al. (2009) confronted two different approaches in keyword usage: "all-keyword approach" and "major-keyword approach". The second one is once more divided into two categories according to selection method: more frequently repeated keywords in patents documents or high variance of appearance of keywords among documents.

We propose here an alternative statistics selection method that uses text mining output to define which topics are included in our network. Starting from the acknowledgment that the hunt for hidden categories (the issue described above) is statistically equivalent to a cluster analysis of latent topics, justifying the use of an estimation model for topics selection as instrument to define network's subtopics, we do so by mean of likelihood based model and a goodness-of-fit analysis.

A topic model depicts multivariate count data as multinomial observations parameterized by a weighted sum of latent topics. With each observation $x_i \in (x_1 \dots x_n)$ a vector of counts in p categories, given total count $m_i = \sum_{j=1}^{p} x_{ij}$, the K-topic model is

$$x_1 \sim MN\ (\omega_{i1}\theta_{1+\dots+}\omega_{iK}\theta_K, m_i)$$

where topics $\theta_k = [\theta_{k1} \dots \theta_{kp}]'$ and weights ω_i are probability vectors. The topic label is due to application of the model in (3) to the field of text analysis. In this context, each x_1 is a vector of counts for terms in a document with total term-count m_i, and each topic θ_K is a vector of probabilities over words. Documents are thus characterized through a mixed-membership weighting of topic factors and, with K far smaller than p, each ω_i is a reduced dimension summary for x_1 (Taddy, 2011).

The application of this selection method brought us to discriminate 6 topics (Figure 18). Each patent has been assigned to one single topic with a certain probability (Figure 19).

Table 19 illustrates the first 30 keywords characterizing each topic. The identification of the topic’s object brought us to interpret the 6 topics as follow:

- Fat-based biofuel production through transesterification
- Alcohol-based biofuel production through lingo-cellulosic material
- Biochemical compound for alcohol production
- Fat-based biofuel additives
- Biochemical processes
- Alchol-based biofuel production by starch fermentation

Figure 18. Selection of number of topics

N=6 => max Bayes factor

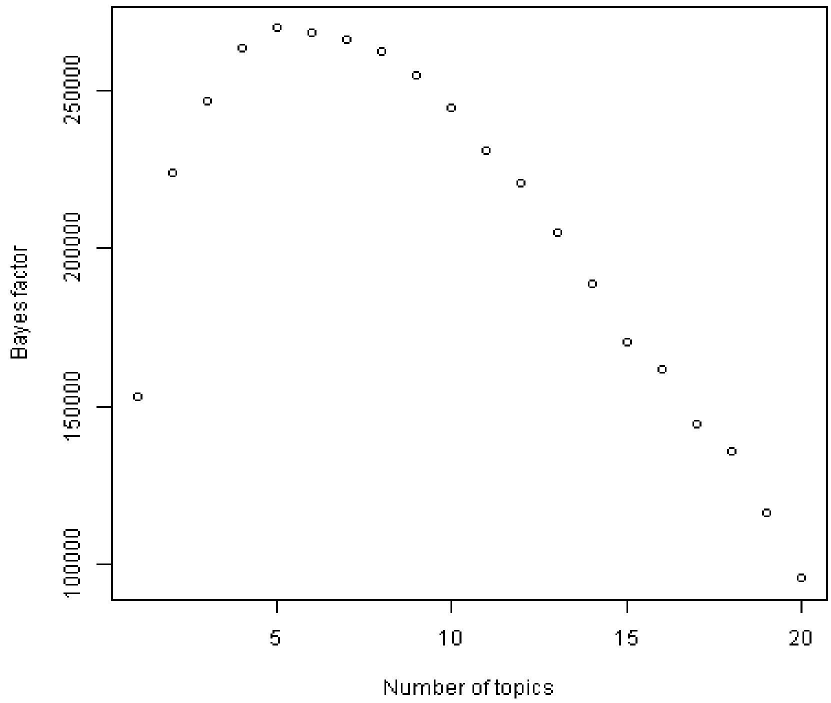


Figure 19. Probability document-topic

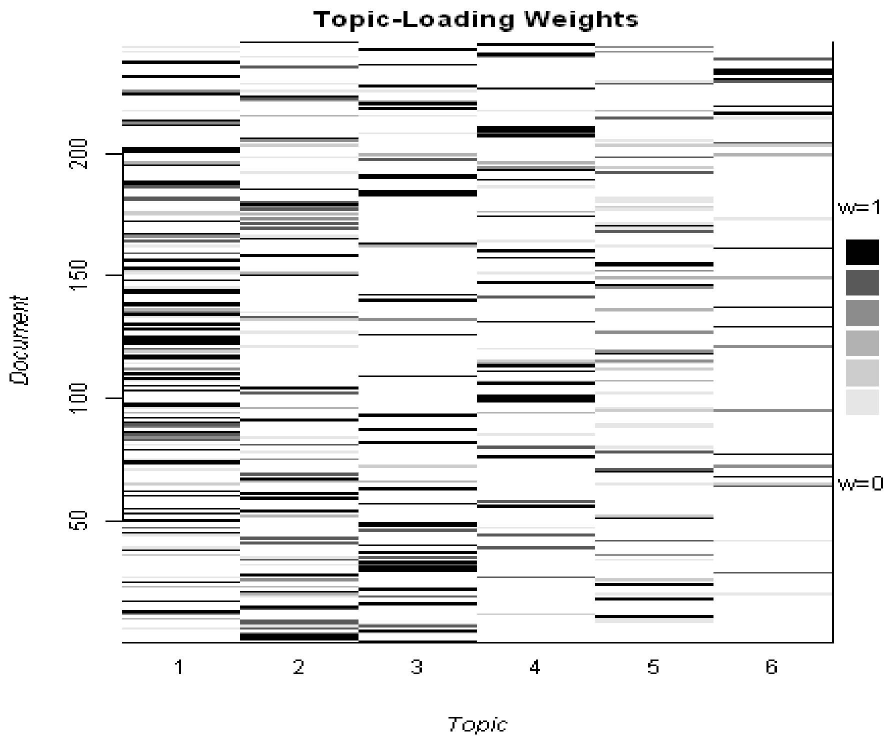


Table 19. Stemmed keywords-topic

1	2	3	4	5	6
acid	ferment	sequenc	fuel	process	ferment
oil	process	gene	acid	product	ethanol
fatti	ethanol	cell	oil	materi	embodi
reaction	acid	acid	ester	system	wast
process	product	dna	biodiesel	water	enzym
biodiesel	materi	enzym	composit	embodi	method
ester	yeast	express	addit	ethanol	starch
catalyst	produc	yeast	fatti	plant	step
alcohol	cellulos	xylos	methyl	heat	mixtur
product	method	ethanol	mixtur	gas	process
prefer	lignocellulos	strain	compris	biomass	compris
method	feedstock	plant	diesel	stream	prefer
produc	step	product	blend	includ	separ
mixtur	compris	ferment	prefer	produc	includ
step	concentr	encod	engin	ferment	vapor
methanol	biomass	transform	atom	tank	water
temperatur	cell	activ	ethanol	control	mean
obtain	strain	seq	alcohol	flow	vessel
separ	sugar	compris	compound	separ	produc
phase	use	host	contain	biofuel	materi
glycerol	glucos	contain	alkyl	solid	sugar
water	pretreat	produc	weight	contain	plant
transesterif	hydrolys	promot	carbon	concentr	hydrolyz
glycerin	hydrolysi	prefer	select	pressur	temperatur
triglycerid	contain	recombin	gasolin	temperatur	citrus
contain	enzym	plasmid	consist	method	product
materi	prefer	protein	embodi	solut	dehydr
compris	solut	method	method	vessel	acid
fat	temperatur	select	hydrocarbon	acid	glucoamylas
reactor	separ	use	combust	andor	finedivid

Visual and quantitative analysis

In this section we present a selection of map resulting from the cut off 0.2 (the one that maximize the efficiency test) and from the topic selection. The first map (Figure 3.x) represents the whole network: highly connected patents are squared-shaped while colors represent the topic. By looking at this first map, we can stress some remarkable evidence: alcohol-based and fat based biofuels creates two separate agglomerate of patents. In the fat-based biofuels agglomerate it is possible to find alcohol-based patents, while the same cannot be said of the opposite. This is easily explained by the fact that ethanol is a input of fat-based biofuels, while the reverse does not occur. The two agglomerates are kept together by a pink patent which stands for a alcohol-based biofuel production through lingo-cellulosic material. The shape of the nods is given by the degree of centrality of the patent. The centrality is computed by mean of the Technology Centrality Index (TCI). TCI is defined as

$$C_D(n_i) = \frac{d(n_i)}{g-1}$$

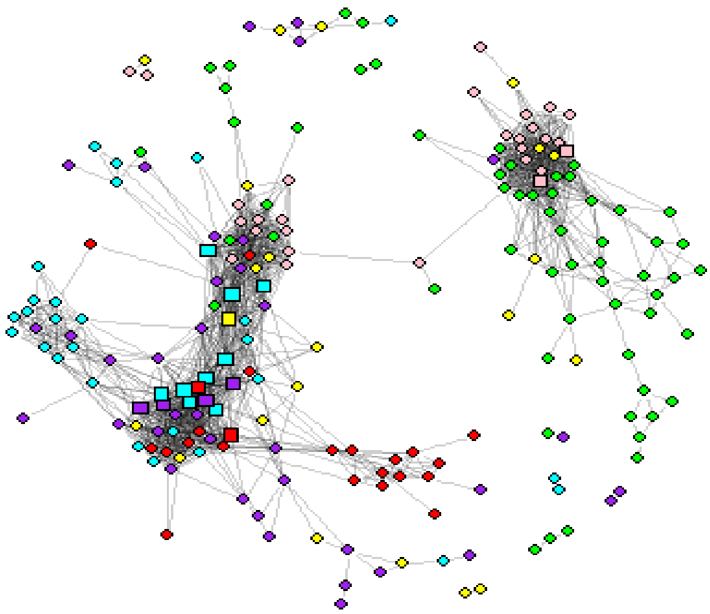
where $d(n_i)$ is the number of ties which are incident with patent i and g in the total number of patents. The centrality index in a patent network indicate the ratio of the number of tied links to all g-1 other patents. Thus, the higher the TCI the greater the relevance of the patent with respect to other patents (Yoon & Park, 2004).

The advantage of the TCI compared to his conceptually equivalent in citation analysis (current impact index) is that it is much more informative and reliable. In citation analysis, the importance of a patent is given by the frequency of citation received. However, this measure suffers from several drawbacks as illustrated in previous paragraphs and the current impact index may be artificial and even misleading. On the contrary, in network analysis the connection between two patents is represented by the similarity between their keywords vectors. Therefore, links among patents indicate a more reliable value of the frequency, given by the number of ties a patent originates and, a qualitative measure of association, given by the number of keywords that two patents have in common.

TCI is displayed in Table 20 for the first 10 highly connected patens. The Table informs also about the nationality and nature of the assignee (private or public) and the topic. The highest TCI is shown by a patent which fall under a very crosscutting category (biochemical compound for alcohol production). This patent has been granted to a Japanese public firm, perfectly in line with information on R&D public expenditure in this country. Nevertheless, as the Table displays (and R&D analysis confirms), private sector is the central actor in patent activity. The second highest TCI patent presents an interesting characteristic: it consist of alcohol-based biofuel production through lingo-cellulosic material, a second generation technology.

Figure 20. Biofuel network, by topic and number of connections

Topic: red=1; yellow=2; pink=3; blue=4; purple=5; green=6 Square= highly connected



Patent	TCI	Country	Topic	Assignee
EP1291428	0.15573770	JP	3= Raw material production for alcohol-based biofuel	Public
EP1766028	0.15163934	SE	2=Alcohol-based biofuel production through lingo-cellulosic material	Private
EP1468093	0.15163934	NL	3	Private
EP1682650	0.14754098	SE	3	Private
EP1330513	0.14344262	US	3	Public
EP2226387	0.13934426	CH	6=Alcohol-based biofuel production by starch fermentation	Private
EP2235146	0.13524590	DE	4=Fat-based biofuel additives	Private
EP1888759	0.13524590	FI	6	Private
EP1813590	0.13524590	US	2	Private
EP1972679	0.12704918	DE	4	Private

Figure 21. Network by number of ties

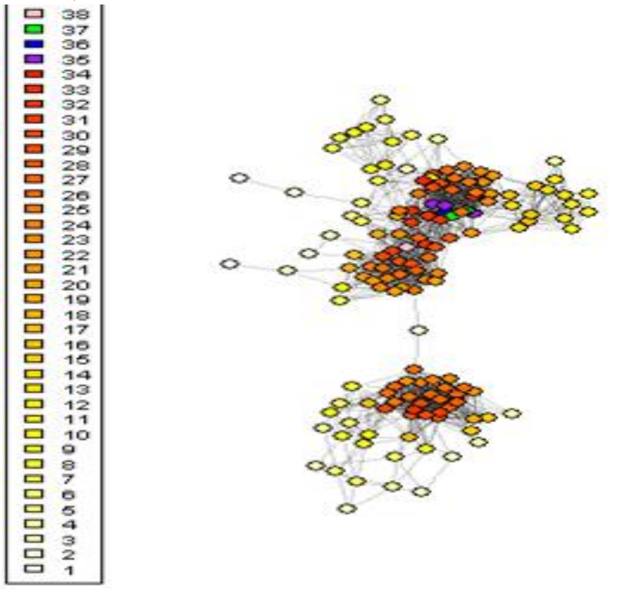


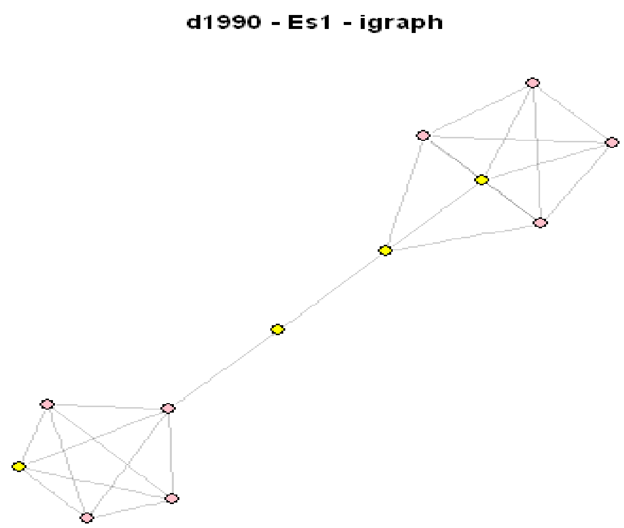
Table 20. Number of ties and keywords

Patent	Ties	N°K	Keywords
EP1291428	38	10	ammonium - bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*mass - enzyme - ferment - purif - starch - sugar - synthesis - yeast
EP1766028	37	16	bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*mass - cellulos - enzyme - ferment - forest[-]*residue - genetically[-]*engineered - hemi[-]*cellulos - hydrolysate - hydrolysis - lignin - ligno[-]*cellulos - purif - sugar - treatment - yeast
EP1468093	37	14	bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*mass - cellulos - corn - enzyme - ferment - genetically[-]*engineered - hydrolysate - ligno[-]*cellulos - mesophilic - starch - sugar - xylanase - yeast
EP1682650	36	14	additive - bio[-]*ethanol|(?!bio[-]*)ethanol - cellulos - corn - enzyme - ferment - genetically[-]*engineered - hemi[-]*cellulos - hydrolysate - lignin - ligno[-]*cellulos - purif - sugar - yeast
EP1330513	35	18	ammonium - bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*fuel - bio[-]*mass - cellulos - corn - enzyme - ferment - genetically[-]*engineered - hemi[-]*cellulos - hydrolysis - lignin - purif - saccharification - starch - sugar - treatment - yeast
EP2226387	34	29	SSF - additive - animal[-]*waste - barley - beet - bio[-]*diesel - bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*fuel - bio[-]*mass - bio[-]*methanol|(?!bio[-]*)methanol - cassava - cello[-]*bio[-]*hydrol - cellulos - corn - enzyme - ferment - grain - hemi[-]*cellulos - hydrolysis - maize - pre[-]*treatment - saccharification - sorghum - starch - sugar - transgenic[-]*plant - treatment - xylanase - yeast
EP2235146	33	21	additive - ammonium - bio[-]*diesel - bio[-]*ethanol|(?!bio[-]*)ethanol - bio[-]*fuel - bio[-]*methanol|(?!bio[-]*)methanol - castor - coconut - cooking[-]*oil - fatty[-]*acid[-]*methyl[-]*ester - maize - palm[-]*oil - peanut[-]*oil - purif - rape[-]*seed - siloxane - soy[-]*bean - synthesis - tallow - trans[-]*esterification - vegetable[-]*oil

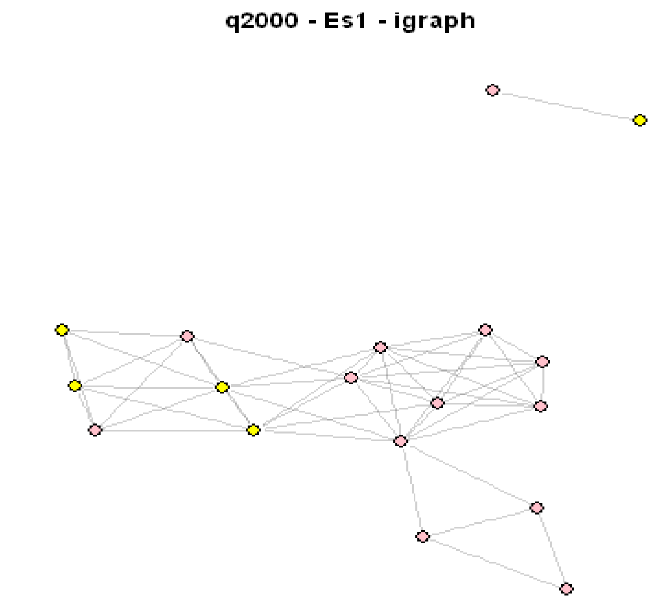
The map above (Figure 21) illustrates the whole network. Colours stand for number of ties (links) incident with a node. The pink node represents the Japanese patent (highest TCI). This map allows visualizing central patents in the network. Table 20 displays number of ties, number of keywords and keywords for the more central patens. Keywords are ordered by frequency of occurrence.

Finally, we displays three maps which divide the network in three different time span: 1990-1999, 2000-2004, 2005-2010. This is particularly interesting because it shows that, even though the boom in patenting activity occurred in the "post Kyoto years" the selected sample, which include patents unequivocally and directly linked to biofuels production, contains mainly patents applied in the last lustrum of the 21th century. This is likely because the lucrative appealing of biofuels is gaining obviousness (because of public policy fostering their expansion) and inventors are recognizing a profit opportunity in stressing the suitability of their invention for biofuels production. For these maps we combined topics in old (pink) and advance (yellow) generation patents.

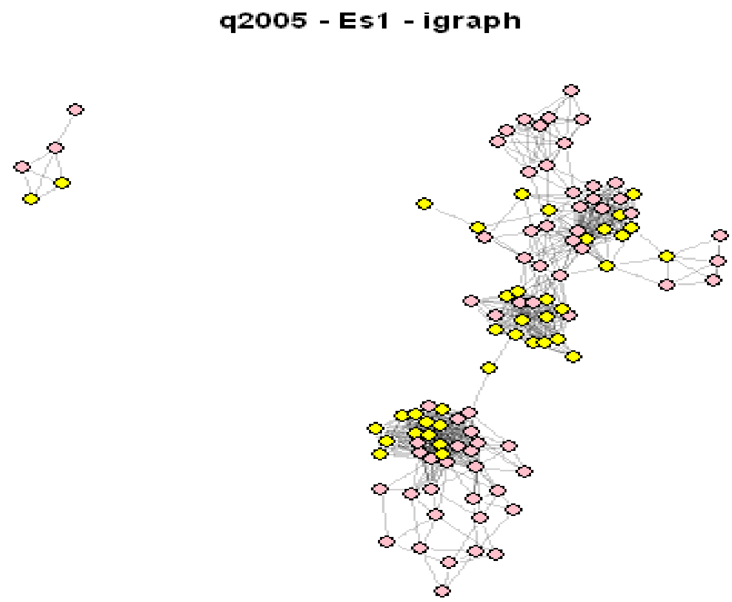
Figure 22. Network by years
1990-2000



2000-2004



2005-2009



Conclusion

In the present work we tried to shed light on whether and why there was a slow uptake of the biofuels for transports, and which is the role of biofuels innovation in the long term sustainability of the sector. Some remarkable conclusion can be summarized as follow.

Up until recent times, biofuels were not competitive in most of the cases, both from an economic and an environment point of view. Excluding ethanol production in Brazil and few other examples, none of the biofuels firms would have entered the business in the absence of subsidies, commercial protection and blending mandates. This is now changing in some countries, especially those not facing particularly stringent resources constraints, notably the US.

An expansion brake can be observed even in those countries, remarkably in Brazil, where biofuels have been competitive with or without public intervention. The reasons that limit biofuels growth can be divided into economic and technological.

In order to understand economic reasons, it is worth reminding that the main part of commercially available biofuels can be classified as "traditional". First generation biofuels rely on feedstock availability and price and the energy sector, which is essentially larger than the agricultural one, would end up absorbing too many natural resources.

The 2011 IEA roadmap for liquid biofuels (IEA, 2011) estimates that to meet biofuel demand, agriculture would have to supply (at current yields, current conversion rates/technologies/etc.) around 65 Exajoules of biofuel feedstock, occupying 100 million hectares of land, 2/3 of the world's cropland or the entire cropland of developing countries (2 per cent of the total agricultural land), in 2050. This raises significant concerns considering the competition for land and other resources (notably water) from demands for food and for additional biomass.

In other words this would suggest that the energy sector should absorb supply from agriculture, as in fact is doing, up until prices for agricultural feedstocks meet their parity prices levels of alternative (to biofuels) energy sources. Considered the perfectly elastic demand and assumed an inelastic supply, prices move quickly up their parity price levels, and thus limiting uptake of biofuels. Given the resources constrain and the rigidity of first generation biofuels supply, an increasing biofuels demand will be sustainably satisfied only with a significant contribution coming from advanced biofuels. Patents network analysis confirms that in the last decade, advanced biofuels have been gaining importance in the innovation panorama of the whole biofuels sector. In particular, lingocellulosic biofuels patents represent a great part of the new inventions and results highly connected (and thus central) according to our methodology.

Technological reasons refer chiefly to the vastly discussed blendwalls. A blendwalls is the limit at which no more biofuel can be blended into fossil fuel due either to regulation and infrastructural reasons. The relevance of breaking infrastructural blendwalls has been analysed through patent data. The evidence suggests that biofuels are expanding their technological domain and much of their new applications cover those inventions which aim to solve the infrastructural barrier (biodiesel additives, new tanks, new engines, new shipping facilities, etc.). It is not possible to conclude that issues concerning "infrastructural adaptability" will be dealt in the very next years, but the patent analysis point out that the question has been taken into serious consideration.

Regulation blendwalls are strictly related to biofuels availability, and some countries will be able to break them in less than one year (E.g. USA). How many countries will break the wall, and how much overall is hard to say. The extent of the fall will mainly depend on how fast the agricultural feedstock supply will increase (that in turn depend on how much R&D will go into agriculture and how strong will be the socio-political opposition to growing food for energy purposes) and on how rapid new generation biofuels will start producing biofuels in an economically viable way.

Annex I. Green Inventory biofuels classes

A01H	20,189	Biofuels –Liquid fuels -From genetically engineered organisms
A62D 3/02	431	Harnessing energy from manmade waste - Anaerobic digestion of industrial waste
B01D 53/02	3120	Harnessing energy from manmade waste - Landfill gas- Separation of components
B01D 53/04	4423	Harnessing energy from manmade waste - Landfill gas- Separation of components
B01D 53/047	1491	Harnessing energy from manmade waste - Landfill gas- Separation of components
B01D 53/14	2948	Harnessing energy from manmade waste - Landfill gas- Separation of components
B01D 53/22	3498	Harnessing energy from manmade waste - Landfill gas- Separation of components
B01D 53/24	109	Harnessing energy from manmade waste - Landfill gas- Separation of components
B09B	6613	Harnessing energy from manmade waste - Landfill gas
C02F 11/04	576	Harnessing energy from manmade waste - Industrial waste - Anaerobic digestion of industrial waste
C02F 11/14	669	Harnessing energy from manmade waste - Industrial waste - Anaerobic digestion of industrial waste
C02F 3/28	1365	Biofuels – Biogas
C07C 67/00	9671	Biofuels – Liquid fuels - Biodiesel
C07C 69/00	15443	Biofuels – Liquid fuels - Biodiesel
C10B 53/00	1089	Pyrolysis or gasification of biomass
C10B 53/02	In previous	Biofuels - Solid fuels - Torrefaction of biomass
C10G	17625	Biofuels – Liquid fuels - Biodiesel
C10J	2795	Pyrolysis or gasification of biomass
C10L 9/00	412	Biofuels - Solid fuels - Torrefaction of biomass
C10L 1/00	2713	Biofuels – Liquid fuels
C10L 1/02	In previous	Biofuels – Liquid fuels – Vegetable oils /Biodiesel / Bioethanol
C10L 1/14	1958	Biofuels – Liquid fuels
C10L 1/182	503	Biofuels – Liquid fuels - Bioethanol
C10L 1/19	672	Biofuels – Liquid fuels – Vegetable oils /Biodiesel
C10L 3/00	1757	Integrated gasification combined cycle (IGCC)/ Biofuels - Biogas
C10L 5/00	759	Biofuels – Solid fuels / Harnessing energy from manmade waste – agricultural waste
C10L 5/40	In previous	Biofuels - Solid fuels - Torrefaction of biomass
C10L 5/42	In previous	Harnessing energy from manmade waste – agricultural waste - Fuel from animal waste and crop residues
C10L 5/44	In previous	Harnessing energy from manmade waste – agricultural waste - Fuel from animal waste and crop residues

C10L 5/46	In previous	Harnessing energy from manmade waste - Landfill gas – Municipal waste
C10L 5/48	In previous	Harnessing energy from manmade waste - Industrial waste / Biofuels - Solid fuels
C11C 3/10	925	Biofuels – Liquid fuels - Biodiesel
C12M 1/107	489	Biofuels – Biogas
C12N 1/13	243	Biofuels – From genetically engineered organisms
C12N 1/15	11575	Biofuels – From genetically engineered organisms
C12N 1/21	27080	Biofuels – From genetically engineered organisms
C12N 15/00	16555	Biofuels – From genetically engineered organisms
C12N 5/10	30000	Biofuels – From genetically engineered organisms
C12N 9/24	2754	Biofuels – Liquid fuels - Bioethanol
C12P 5/02	414	Biofuels – Biogas
C12P 7/06	1159	Biofuels – Liquid fuels - Bioethanol
C12P 7/14	104	Biofuels – Liquid fuels - Bioethanol
C12P 7/64	1931	Biofuels – Liquid fuels - Biodiesel
D21C 11/00	983	Harnessing energy from manmade waste - Industrial waste - Pulp liquors

Annex II. Examples of keywords

fame	eicosapentaenoic acid	peanut

	scenedesmus	
fatty acid methyl esters	corn	oil-bearing organisms
fatty acid ethyl esters	maize	jatropha curcas
free fatty acid	cassava	jatropha
lipids as feedstock	grain	babassu coconut
lipids microbial organisms	soybean	helianthus tuberosus
fatty acyl-acp thioesterase	genetically engineered microbes	oleaginous microorganisms
fatty acyl-coa/aldehyde reductase	genetically modified crops	rhodotorula glutinis
fatty aldehyde decarbonylase	ligno-cellulosic	medicago sativa l.
acyl carrier protein	perennial grasses	nut shells
volatile fatty acids	forest	sugarcane
microbial lipids	panicum virgatum l	beet
microbial hosts	perennial plant	sorghum
trichosporon	phalaris	sugar esters
agricultural feedstocks	alfafa	bagasse
starch	reed canarygrass	fermentable sugars
corn cobs	fibrous plant materials	cooking oil
corn stover	switchgrass	wet organic wastes
cereal straw	bark	monosodium glutamate wastewater
forest harvest residues	wood shavings	urban wood residues
husks	chip boards	ammonium
chlorella vulgaris	garden mulch	animal waste
spirulina maxima	vegetative grasses	anlage
nannochloropsis sp.	miscanthus	excreta
scenedesmus obliquus	prairie grass	feed mixture
dunaliella tertiolecta	short rotation forest species	fibrobacter succinogenes
scenedesmus dimorphus	eucalyptus	kalium
chlorella emersonii	poplars	lignocellulose
chlorella protothecoides	lignin	liquid manure
chlorella minutissima	cellulose	microorganisms
dunaliella bioculata	hemicellulose	ruminococcus albus
dunaliella salina	wood process residues	sewage
microalgae oil	wheat chaff	siloxane
phaeodactylum tricornutum	animal fat	sulphide
vegetable oil	edible tallow	digested sludge
soya oil	animal manure	fibrous material
untreated raw oils	granular sludge	hydrolysate
oilseed rape	porcine pancreatic lipase	liquid manure
coconut oil	rapeseed	mesophilic bacteria
jojoba (limited to biodiesel)	palm oil	microbial consortia
canola oil (limited to biodiesel)	organic material	sludge
methanogenic bacteria	animal slurries	treated wastewater

References

Archibugi, D., Pianta, M. (1996). Measuring technological through patents and innovation surveys. *Technovation 16 (9)*, 451-468.

Bacchiocchi, E., Montobbio, F. (2004). *EPO vs. USPTO Citation Lags.* CESPRI WP n. 161.

Baron, J., Delcamp, H. (2010). *Patent quality and value in discrete and cumulative innovation.* Cerna Working Paper 2010-07.

Biofuels cities. (2008). *Vehicle warranties and the use of biofuels. An overview of vehicles available within the EU which have manufacturers warranty allowing the use of blends of biofuels.* SenterNovem.

Biofuels Digest. (n.d.). *Biofuels Digest.* Retrieved 01 27, 2012, from http://www.biofuelsdigest.com/

Blume, S. S. (1992). *Insight and industry: on the dynamics of technological change in medicine.* Boston: MIT Press.

Borgatti S. P., E. M. (s.d.). UCINET 6.0 Version 1.0. Harvard: Analytic Technologies Publishers.

Burke, P., Reitzig, M. (2007). Measuring patent assessment quality - Analyzing the degree and kind of (in)consistency in patent offices‘ decision making. *Research Policy 36 (9)*, 1404-1430.

Chang P.L., W. C. (2009). Using patent analyses to monitor the technological trends in an emerging field of technology: a case of carbon nanotube field emission display. *Scientometrics 82*, 5-19.

Chang, P.L., Wu, C.C., Leu, H.J. (2012). Investigation of technological trends in flexible display fabrication through patent analysis. *Displays 33*, 68-73.

Charlita de Freitas, L., Kaneko, S. (2012). Is there a casual relation between ethanol innovation and the market characteristics of fuel in Brazil? *Ecological Economics 74*, 161-168.

Consoli D., M. A. (2009). An Evolutionary Perspective on Health Innovation Systems. *Journal of Evolutionary Economics 19*, 297-319.

Coombs R., M. J. (2000). Organizing for innovation: coordinating distributed innovation capabilities. In M. V. Nicolai J. F., *Competence, governance and entrepreneurship.* Oxford: Oxford University Press.

Corrocher N., M. F. (2007). Schumpeterian Patterns of Innovative activity in the ICT field. *Research Policy 36*, 418-432.

Costantini, V., & Crespi, F. C. (s.d.). In V. Costantini. Springer.

Costantini, V., Crespi, F. (2012). Public policies for a sustainable energy sector: regulation, diversity and fostering of innovation. *Journal of Evolutionary Economics*, DOI:10.1007/s00191-010-0211-3.

Coulter, N. M. (1998). Software engineering as seen through its research literature: a study in co-word analysis. *Journal of the American Society for Information Science 49 (13)*, 1206-1223.

Crawshaw, E. H. (2006). *Brevetto n. EP1674553.* Netherlands.

De Looze, M-A., Lemarie, J. (1997). Corpus relevance through co-word analysis: an application to plant proteins. *Scientometrics 39 (3)*, 267-280.

De Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., Van Pottelsberghe de la Potterie, B. (2012). The Worldwide Count of Priority Patents: A New Indicator of Inventive Activity . *Research Policy, Forthcoming.*

de Saint-Georges, M., van Pottelsberghe de la Potterie, B. (2011). *A quality index for patent systems.* Universite Libre de Bruxelles: Working Papers ECARES 2011-010, ULB.

Dernis, H., Guellec, D. (2001). *Using patent counts for cross-country comparisons of thecnology output.* Paris: OECD.

Dernis, H., Kahn, M. (2004). *Triadic patent families methodology.* . Paris: OECD.

Dosi, G., Freeman, C., Nelson, R.R., Silverberg, G., Soete, L. (1988). *Technical Change and Economic Theory.* London: Pinter Publishers.

Dosi, G., Nelson R. R. (2009). Technical Change and Industrial Dynamics as Evolutionary Processes. *LEM working paper*.

EC. (2005). *Biomass action plan. COM(2005) 628 final (7.12.2005)*. Brussels: European Commission.

EC. (2006). *European Union biofuel strategy. COM(2006) 34 final (8.2.2006)*. Brussels: European Commission.

EC. (2007). *Renewable energy roadmap. COM(2006) 848 final.* Brussels: European Commission.

Economist, T. (March 17, 2011). Patent reform: the spluttering invention machine – America’s patent system has problems; a new law would fix only a few.

EurObserv' ER. (s.d.). *EurObserv' ER indicators.* Tratto il giorno 05 17, 1012 da http://www.eurobserv-er.org/

EUROSTAT. (2007). *Energy, transport and environment indicators.* Eurostat Pocketbooks.

Garfield, E., Sher, I., Torpie, I. (1964). *The use of citation data in writing the history of science.* Philadelphia: Institute for scientific information.

Giummo, J. (2003). *Should All Patentable Inventions Receive Equal Protection? Identifying the Sources of Heterogeneity in Patent Value.* Berkeley: Discussion paper of the University of California.

Goto A., M. K. (2007). Construction of a Japanese Patent Database and a first look at Japanese patenting activities. *Research Policy 36*, 1431-1442.

Griliches, Z. (1990). Patent statistics as economic indicators: a survey. *Journal of economic literature 28 (4)*, 1661-1707.

Hall, B. H., Jaffe A. B, Trajtenderg, M. (2001). *The NBER patent citations data file: lessons, insights and methodological tools.* NBER WP 8498.

Hall, B.H., Jaffe, A.B., Trajtenberg, M. (2002). Market value and patent citations. *NBER Working Paper*.

Harhoff, D., Narin, F., Scherer, F.M., Vopel, K. (1999). Citation frequency and the value of patented inventions. *The Review of Economics and Statistics 81*, 511-515.

Harhoff, D., Reitzig, M. (2004). Determinants of opposition against EPO patent grants: the case of pharmaceuticals and biotechnology. *International Journal of Industrial Organization 22 (4)*, 443–480.

Hekkert, M. P., Suurs, R. A. A., Negro S.O., Kuhlmann S., Smits R.E.H.M. (2007). Functions of innovation systems: a new approach for analyzing technological change . *Technological Forecasting & Social Change 74*, 413–432.

Huang, Z et al. (2004). International nanotechnology development in 2003: Country, institution, and technology field analysis based on USPTO patent database. *Journal of Nanoparticle Research* 6, 325-354.

Hughes, T. P. (1987). The evolution of large technological systems. In T. P. Wiebe Bijker, *The Social Construction of Technological Systems: New Directions in the sociology and History of technology* (p. 51-83). MIT.

IEA. (2007). *World energy outlook 2007.* Paris: Iea.

IEA. (2008). *From 1st to 2nd generation biofuels technologies. An overview of current industry and R&D activities.* Iea.

IEA. (2010). *Medium Term Oil and Gas Markets 2010.* Paris: OECD/IEA.

IEA. (2011). *Technology roadmap. Biofuels for transport.* Paris: OECD/IEA.

IPC Guide. (2011). Retrieved from http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf

Jaffe A., Trajtenberg M. (2002). *Patents, Citations, and Innovations: A Window on the Knowledge Economy.* Cambridge, Mass: MIT Press.

Jaffe, A. B. (1989). Real Effects of Academic Research. *The American Economic Review, 79(5)*, 957–970.

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

Johnstone, N., Hascic, I., Ostertag, K. (2008). Environmental policy, Technological Innovation and patent activity. In AA.VV., *OECD Studies on Environmetal Innovation. Environmetal policy, technological innovation and patents* (p. 17-63). Paris: OECD.

Johnstone, N., Hascic, I., Popp, D. (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental & Resource Economics 45 (1)*, 133-155.

Karki M. (1997). Patent citation analysis: a policy analysis tool. *World Patent Information.*

Karmarkar-Deshmukh, R., Pray, C. E. (2009). Private sector innovation in biofuels in the United States: induced by prices or policies? *AgBioForum 12 (1)*, 141-148.

Lanjouw, J., Schankerman, M. . (1999). *The quality of ideas: measuring innovation with multiple indicators.* NBER Working Paper.

Lanjouw, J., Schankerman, M. (2004). Patent quality and research productivity:measuring innovation with multiple indicators. *Economic journal 114 (495)*, 441-465.

Lee, S., Yoon, B., Park, Y. (2009). An approach to discovering new technology opportunities: keyword-based patent map approach. *Technovation 29*, 481-497.

Lemley, M., Sampat, B. N. (2008). Examiner characteristics and the patent grant rate. *Stanford Law and Economics Olin Working Paper 369.*

Lundvall, B. A. . (1992). *National Innovation Systems: Towards a Towards a Theory of Innovation and Interactive Learning.* London and New York: Pinter Publishers.

Malerba F., O. L. (1996). Schumpeterian Patterns of Innovation are Technology Specific. *Research Policy 25*, 83-117.

Malerba, F. (2000). *Economica dell'innovazione.* Carocci.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research Policy 31*, 247-264.

Malerba, F. (2004). *Sectoral System of Innovation.* Cambridge University Press: Cambridge.

Metcalfe J.S., J. A. (2005). Emergent Innovation Systems and the delivery of clinical services: the case of intra-ocular lenses. *Research Policy 34*, 1283-1304.

Michel, J., Bettels, B. (2001). Patent citation analysis. A closer look at the basic input data from patent search. *Scientometrics, 51*, 185–201.

Narin, F. (2000). Tech-Line® background Paper. In J. Tidd, *Measuring Strategic Competence.* London: Imperial College.

Narin, F., Hamilton, K. S., Olivastro, D. (1997). The increasing linkage between U.S. technology and public science. *Research Policy 26 (3)*, 317-330.

Nomalera, O., Verspagenbc, B. (2008). Knowledge Flows, Patent Citations and the Impact of Science on Technology. *Economic Systems Research, 20: 4*, 339-366.

OECD. (1997). *National System of Innovation.* Paris: OECD.

OECD-FAO. (2010). *Agricultural Outlook 2010-2019.* Paris: OECD-FAO.

Pachys, F. (2011). Importance Factor - New Patent-Valuation Index. *Concord 2011. Conference on Corporate R&D.* European Commission.

Picci, L. (2010). The internationalization of inventive activity: a gravity model using patent data. *Research Policy 39*, 1070-1081.

Powell, W. W., Grodal, S. (2006). Network of Innovators. In J. M. Fagerberg, *The Oxford Handbook of Innovation* (p. 56-85). Oxford: Oxford University Press.

Prigogine, I. (1987). Exploring complexity. *European Journal of Operational Research 30*, 97-103.

Rodríguez-Pose, A. (1999). Innovation prone and innovation averse societies: economic performance in Europe. *Growth and change, 30 (1)*, pp. 75-105.

Rodriguez-Pose, A. C. (s.d.). R&D, Spillovers, Innovation Systems and the Genesis of Regional Growth in Europe. *ERSA conference papers* (p. 371). European Regional Science Association.

Rysman, M., Simcoe, T. (2008). Patents and the Performance of Voluntary Standard Setting Organizations. *Management Science 54, n° 11*, 1920-1934.

Saviotti, P.P. (2009). Knowledge Networks: Structure and Dynamics. In A. S. A. Pyka, *Innovation Networks, Understanding Complex Systems* (p. Chapter 2). Berlin Heidelberg : Springer-Verlag .

Scellato, G., Calderini, M., Caviggioli, F., Fronzoni, C., Ughetto, E., Kica, E., Rodriguez, V. (2011). *Patqual. Study on the quality of patent system in Europe* . Tender MARKT/2009/11/D.

Schumpeter, J. (1939). *Business Cycle: a Theoretical, Historical and Statistical Analysis of the Capitalistic Process.* New York: McGraw-Hill.

Schumpeter, J.A. (1942). *Capitalism, Socialism and Democracy.* New York: Harper.

Sirilli, G. (1997). Science and technology indicators: The state of the art and prospects for the future. In G. Antonelli, & N. De Liso, *Economics of Structural and Technological Change.* London: Routledge.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* Indianapolis: Liberty Fund.

Sterner, T., Coria, J. . (2012). *Policy Instruments for Environmental and Natural Resource Management. Second Edition.* RFF Press. Routledge.

Sternitzke, C., Bartkowski, A., Schramm, R. (2008). Visualizing patent statistics by means of social network analysis tools. *World Patent Information 30*, 115–131.

Trajtenberg, M. (1990). A Penny for Your Quotes: Patent Citations and the Value of Innovations. *The RAND Journal of Economics 21, n° 1*, 172-187.

Tsen, Y. H., Lin, C.J., Lin, Y.I. (2007). Text minin2007g techniques for patent analysis. *Information processing and Management 43*, 1216-1247.

UNCTAD-ICTSD. (2005). *Resource Book on TRIPS and Development.* New York: Cambridge university press.

Unep. (July 2010). *Global trends in green energy 2009: new power capacity from renewable sources tops fossil fuels again in US, Europe.* Press release.

Unruh, G. C. (2000). Understanding carbon lock-in. *Energy Policy 28*, 817-830.

van Raan A.F.J., T. R. (1993). The neural net of neural network research. *Scientometrics 26*, 169–192.

Verspagen B. (2007). Mapping technological trajectories as patent citation networks: a study of the history of fuell cell research . *Advances in Complex Systems 10*, 93-115.

Vinkler P. (1994). The origin and features of information referred in pharmaceutical patents. *Scientometrics 30*, 283-302.

Wasserman S., F. K. (1994). *Social network analysis: methods and application.* Cambridge: Cambridge University Press.

Weitzman, M. L. (1996). Hybridizing growth theory. *American Economic Review 86*, 207-212.

Weitzman, M. L. (1998). Recombinant growth. *Quarterly Journal of Economics 113*, 331-360.

Yoon B., P. Y. (2004). A text-mining-based patent network: analytical tool for high-technology trend. *Journal of high technology mangement research 15*, 37-50.