POLITECNICO DI TORINO

Corso di laurea magistrale in

Ingegneria Energetica e Nucleare

Tesi di Laurea Magistrale

Analysis of the impact of COVID-19 pandemics on the Italian industrial energy consumption

Relatore

Prof.ssa Laura SAVOLDI

Correlatore

Daniele LEREDE

Candidato

Antonio OLIVA

ANNO ACCADEMICO 2020 – 2021

Abstract

Energy system models for the analysis of future scenarios are mainly driven by the set of energy service demands which define the broad outlines of socio-economic development throughout the selected time horizon. Energy system optimization models serve as a valuable tool to of inquiry for relevant decision-making insights about the evolution of a Reference Energy System (RES) – a simplified representation of the complex and dynamic real-world interactions related to energy production and consumption – over medium-to-long-term time scales. Among such models, encompassing the four dimensions of Energy, Economy, Engineering and Environment – the so called "4Es" – those belonging to the TIMES framework represent a widespread choice for the exploration of contrasted future scenarios. In TIMES models, the proper modelling of energy service demands in all the final consumption sectors is one of the fundamental pillars to build credible scenarios, needed to generate a set of coherent set of demand growth rates. Indeed, once that drivers and elasticities are chosen and associated to the different energy service demands, TIMES can endogenously build demand curves for each energy service accounted for in the model.

This thesis addresses the long-term effects of the Covid-19 pandemics on industrial production in Italy. Forecasts in 6 energy-intensive subsectors (Iron and Steel, Non-ferrous metals, Non-metallic minerals, Chemicals, Pulp and Paper, Other industries) are obtained through the application of Vector AutoRegressive models, to perform projections partly based on historical trends, without the need for external regressors. Results of the application of the method are computed in two cases, either considering or not the effects of the pandemics, showing a long-term reduction ranging from 3.5 ÷ 19.9 % in 2040, according to the subsector. A validation against the prescribed trends from the Italian Integrated National Energy and Climate Plan is also performed. As each industrial production trend acts itself as a driver in TIMES-Italia, the application to that model is presented to assess the impact on energy consumption forecasts. The results show how the long-term effects of the shock caused by the pandemics could lead, in the analyzed scenario, to a 10 % lower industrial energy consumption by 2040.

Contents

List of figures

List of tables

Acronyms

AAC	Alkali-activated cement-based
AIC	Akaike Information Criteria
BAU	Business As Usual
BIC	Bayesian Information Criterion
BF	Blast furnace
BOF	Basic oxygen furnace
BU	Bottom-up
CCS	Carbon capture and storage
DRI-EAF	Direct reduced iron-electric arc furnace
DMSFE	Discount Mean Squared Forecast Error
ENEA	Italian National Agency for new technologies, energy and sustainable economic development
ETP	Energy Technology Perspectives
ES	Exponential Smoothing
GDP	Gross Domestic Product
GEM-E3	General Equilibrium Model for Economy – Energy – Environment
GHG	Greenhouse Gases
HQC	Hannan-Quinn Criteria
ICIS	Independent Intelligence Services
IEA	International Energy Agency
IMF	International Monetary Found
IP	Industrial Production
ISIC	International Standard Industrial Classification
ISTAT	Italian National Institute of Statistics
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
MIDAS	Mixed-Data Sampling
MISE	Ministero dello Sviluppo Economico
MSE	Mean Squared Error
MSPE	Mean Squared Prediction Error
NACE	Statistical classification of economic activities in the European Community
OECD	Organization for Economic Co-operation and Development

OLS	Ordinary Least Squares
PLS	Penalized Least Squares
PNIEC	Italian Integrated National Energy and Climate Plan
RES	Reference Energy System
RMSE	Root Mean Squared Error
SEN	Italian National Energy Strategy
SSEM	Structural Simultaneous Econometric Models
sVAR	Sparse Vector AutoRegression
TIMES	The Integrated MARKAL-EFOM System
UNSTAT	UN Statistical Commission
VAR	Vector AutoRegression
VARMA	Vector-valued Autoregressive Moving Average
WEO	World Energy Outlook

Chapter 1

Introduction

The Covid-19 pandemic represents the biggest global crisis in generations, affecting almost all countries and more than 180 million people around the world[1]. Governments have been facing a real challenge where the taken measures consisted of trade-offs with respect to health, social, and economic issues that arose from such crisis. Spring 2020 has been the period where more than half of the world's population ended up with hard lockdown measures [1], a drastic decision adopted by most of the governments worldwide to foster social distancing, recognized as the only way to slow down the circulation of the disease, in absence of suitable medications and vaccines [2]. It is now plain how coronavirus and the choices made to contain it led to short-term shocks, mostly related to health and economy, and to long-term uncertain effects that still have to be deeply analyzed [3].

1.1 Covid-19 effects on the energy system

Among the various socio-economic sectors affected by the crisis, the energy sector has experienced different shocks. In [4], the short-term consequences of the pandemic, and in particular its effects on the energy situation in the hardest-hit countries are analyzed. The main outcomes show how there has been an average 25 % decrease in energy demand per week in countries in full lockdown, while in countries in partial lockdown the average decrease has been of 18 %.

Concerning the electrical sector, in Figure 1, included in a report published by the Independent Commodity Intelligence Services (ICIS) [5], the percentage deviations from the expected electricity demand for five different European countries in the period of the first lockdown measures, from March to July 2020, are shown. Italy has been one of the first countries to have reported coronavirus cases in Europe, and the full lockdown measures taken have led to a percentage decrease of about 25 % from the expected electricity demand in the last week of March. Moreover, it is also the country registering the most intermittent recovery from April to the mid of May.

France was the only EU economy to register a growth in June, while the other countries have faced partial lockdowns and other stringent measures to contain the virus spread. Whilst France has been the country with the highest speed of recovery,

[1] Google statistics: https://support.google.com/websearch/answer/9814707?p=cvd19_statistics&hl=en-IT&visit_id=637609018476960817-3123340833&rd=1 (accessed Jul. 3, 2021).

Germany has registered the lowest deviations in terms of electricity demand during the period highlighted by Figure 1.

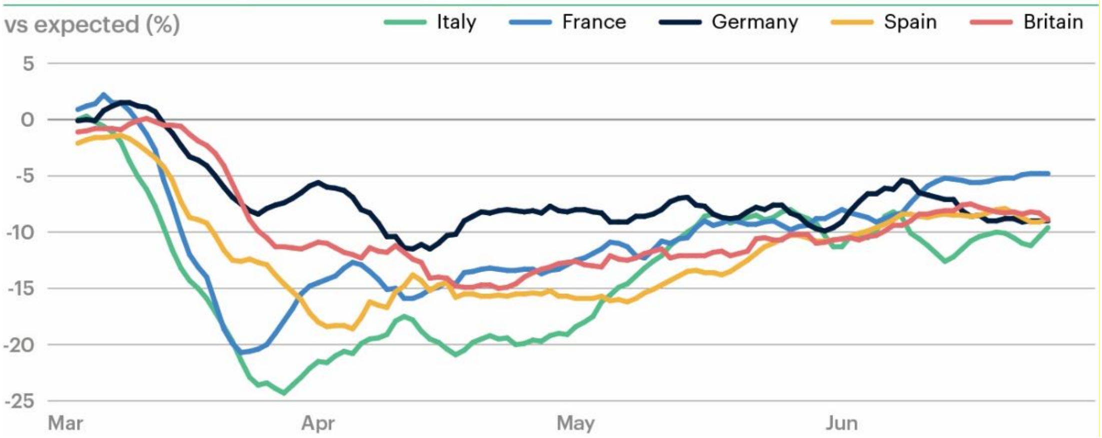

Figure 1: Electricity demand percentage deviation from expectations for different European countries from March to July 2020 [5].

This is just an example of how such crisis has raised major new uncertainties for the future of the energy sector. Other major key questions include the duration of the pandemic, the shape of the recovery, and how energy and sustainability can be included into the strategies that the governments will build for the recovery of their economies [6].

Many reports and publications have tried answering such questions so far, and performing reliable energy projections has been a challenge due to the many uncertainties derived from the crisis, that is still not completely concluded. IEA in its World Energy Outlook 2020 (WEO 2020) highlights the need of rethinking their approach to the problem, refocusing the work in order to deal with such uncertainties [6]. Indeed, given the current level of near-term uncertainty, the WEO 2020 is focused mainly in projections up to the next ten years, analyzing in particular different pathways and opportunities for a sustainable recovery from the current crisis. While, in fact, the Delayed Recovery Scenario (DRS) defined in the IEA's report considers a prolonged pandemic crisis with very high impacts on economy, health and energy sectors, the Sustainable Development Scenario (SDS) works backwards from the long-term climate goals, examining the necessary actions to introduce to achieve them.

Differently from such projections, where the results are based on assumption and constraints which do not analyze the response of the socio-economic drivers of the energy demands, this work of thesis focuses on this issue, suggesting a different approach to the problem.

The aim of this thesis is, then, to assess the long-term impact of the crisis on the industrial production in Italy, and how this affects the energy consumption.

1.1.1 Effects on the industrial energy consumption in Italy

In Italy, the final energy demand in the whole 2020 has decreased by about 10 % with respect to 2019 levels, with a total estimation of 115 Mtep [7]. This is mostly due to the decrease of the transport sector, with a decrease of 18 % compared to 2019 levels. Such contraction is primarily related to the reduction of the oil products demand for both terrestrial and air transport.

The analysis published by the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) for the year 2020 [7] shows how also the industrial sector has highly contributed to the energy demand decrease in 2020. In fact, such crisis has accentuated the already declining trend in 2019, with a decrease in the total energy consumption in such sector of 15 Mtep, about 7 % lower than the 2019 levels. During the year, the lowest values of industrial energy consumption has been registered in the second trimester, with a drop of 15 % compared to the same period in the previous year. Although the energy consumption in the industrial sector has returned to grow in the Sumer period, with a conjunctural increase of about 15 %, the balance at the end of 2020 remains negative with respect to 2019. Concerning the fuels, natural gas and electricity have registered the highest decrease in demand, with a decline of respectively 6 % and 7 % compared to 2019 levels.

The decline of the industrial energy consumption in 2020 follows the industrial production trend, where a new minimum is registered with respect to, at least, the last 15 years, as Figure 2 shows. The effects of the pandemic crisis accentuated an already present decreasing trend, with a change in the direction of the decoupling between industrial production and energy consumption started in 2013.

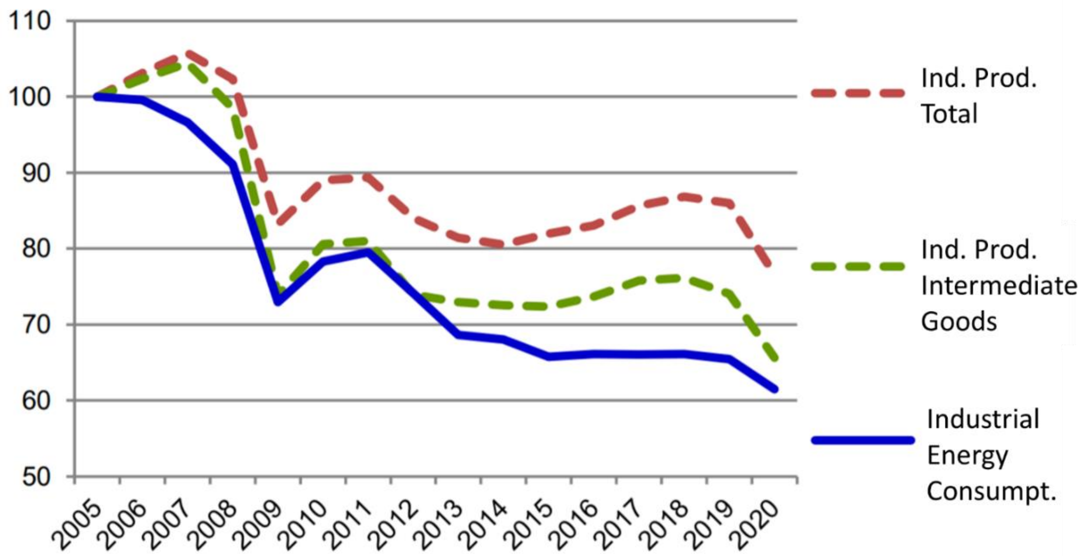

Figure 2: Italian industrial energy consumption (normalized at year 2005) compared to the total industrial production and of intermediate goods. Source: ENEA [7]

The economic shock registered in this year has raised new doubts on how the energy sector will respond on the long-term, as stated in the previous section. This work tries to partially answer such question by analyzing the Italian industry sector and performing projections on its final energy consumption, by means of the TIMES-Italia model. Indeed, energy system optimization models are probably the natural tool to generate possible future scenarios of the energy system or a sub-system at national, regional or global level over a time horizon of usually multiple decades [8].

1.2 Macro-scale energy models

Macro-scale models are generally used as supportive tools for policy making, in order to produce future projections for the evolution of the energy system and the analysis of the effects of targeted energy policies and climate change issues. Such models utilize techniques that range from mathematical programming to statistical and network analysis. Their main aim is to analyze the various components of an energy system and how they develop and interact together, in order to build cost-effective strategies for energy planning [8].

Macro-scale energy models can be generally classified on the basis of their approach.

Top-down models are general equilibrium econometric models, and they endogenously evaluate the responses of the economic system to different policies and scenarios. The energy sector is represented in a limited way, due to their market-oriented approach, and no detail on current and future technologies is present. Such technologies, in fact, are generally modeled through aggregated production functions defined for each economic sector [8]. For the mentioned reasons, top-down models are generally used to analyze the evolution of energy prices and macro-economic variables, but the effects of energy policies on the technological mix of the energy system are not considered.

On the other hand, bottom-up models are generally used to analyze the dynamics and interactions of the four dimensions of economy, engineering, energy and environment. Differently from top-down models, macroeconomic assumptions are here evaluated exogenously, and they represent input data to evaluate the energy demand required by the different services envisaged in the reference energy system.

In bottom-up models production, transformation and end-use technologies are described by means of technical (engineering approach) and economic (economic approach) parameters. Technical parameters are …; on the other hand, investment cost, fixed O&M cost and variable O&M costs represent the economic features necessary to put into force the minimum cost optimization algorithm.

In this case, the mix of technologies and commodities is determined by means of an optimization procedure in which the service demands are met, and the total discounted system cost is minimized. Such optimization is performed over the whole time horizon and considering user-defined constraints, considering the application of particular tax policies or emission constraints, and scenarios considering different assumptions for the evolution of certain energy or socio-economic features on the basis of different policy assumptions.

The Integrated MARKAL-EFOM System (TIMES) is one of the most popular bottom-up model generators. TIMES can be defined as “an economic model generator for local, national or multi-regional energy systems, which provides a technology-rich basis for estimating energy dynamics over a long-term, multi-period time horizon” [9]. Such framework allows to define partial equilibrium models where the optimization is based on the maximization of the total surplus, that is the sum of consumers surplus and producers surplus.

As already specified above, the projections of the main macro-economic drivers are determined exogenously in the top-down models. Such projections generally serve as an input in bottom-up energy models to determine the demand for energy services. An example of this is represented by the top-down General Equilibrium Model for Economy – Energy – Environment (GEM-E3) [10], where the determined projections of precise socio-economic indicators are utilized as an input in the TIMES models.

In TIMES models, the proper modelling of energy service demands in all the final consumption sectors is one of the fundamental pillars to build credible scenarios, needed to generate a set of coherent set of demand growth rates. Indeed, once that drivers and elasticities are chosen and associated to the different energy service demands, TIMES can endogenously build demand curves for each energy service accounted for in the model [11], generally according to Equation 1,

$$D_t = D_{t-1} \cdot [1 + \left(\frac{\delta_t}{\delta_{t-1}} - 1\right) \cdot e_{t-1}] \tag{1}$$

where D is the demand, t the time step, δ the demand driver and e the elasticity of the demand to its driver. While driver projections are usually established starting from external sources, elasticities are generally computed to project the entire economy according to a general trend, based on econometric observations.

Once all demands are defined for all the time periods envisaged in the TIMES model under exam, they will represent one of the most important and necessary constraint, the so-called “demand constraint”. The demand constraint drives the model ensuring that supply at least meets the demand specified in all periods and time slices, by ensuring that the sum of all the demand output commodity generated must meet the modeler-specified demand.

This work focuses on the assessment of the effects of the recent Covid-19 pandemics on socio-economic drivers and energy use forecasts for the industrial sector of Italy, one of the hardest-hit countries in Europe, both from the economic and health point of views [4]. In fact, a key question about the effect of the pandemic is whether it has affected the dynamics of the demand of energy services, which is the first driver of energy demand (together with the dynamics of energy technologies) [12]. Those dynamics depend, on one hand, on the evolution of the key socio-economic drivers, on the other hand on the elasticity of each energy service demand to these socio-economic drivers. Elasticities quantitatively describe the influence of each driver on the associated energy service demand [13]. They tend to change over time, and they have a decisive impact on the long-term evolution of energy demand [14]. In this sense, the aim of this work is to focus on these critical points and to suggest an approach based on traditional econometric models able to accurately project the socio-economic drivers of our interest.

1.3 Aim of the work

One of the main advantages of using general equilibrium models to derive demand trends for partial equilibrium model is that they ensure internal coherence among ocio-economic drivers' projections, guaranteed by the fact that they analyze the economic system as a whole, highlighting the importance of linkages between sectors.. Nevertheless, the limitations of such models have to be found in the high amount of data and human capital investment required [15].

As stated above, an exceptional situation such as the Covid-19 pandemic crisis leads to critical uncertainties about the future of social and economic issues, and this is also reflected on the analyses of macro-economic projections. In such cases, obtaining accurate projections, not biased from ad-hoc assumptions based on some aprioristic expectations, of socio-economic indicators for a bottom-up energy modeling becomes more difficult [16]. Moreover, as bottom-up models require projections of very specific data, such as sectorial outputs for various economic subsectors, which can be very difficult to find in literature, given the novelty of the case analyzed, and this adds complexity to the problem.

This work comes from the need of facing such difficulties, driven by the need to derive more compelling drivers for the future evolution of highly technologically detailed energy systems in macroeconomic models, and the analysis is focused on obtaining reliable projections of the production indicators for different industrial subsectors in Italy. Such indicators represent the drivers for the service demands of the entire industry sector in the TIMES model, and they well summarize all the discussed difficulties. In fact, at the time of the initial literature review for this work, no publications about projections of such data that also consider the Covid-19 outbreak effects have been found, while instead early analyses considered the very short-term effects on energy and electricity consumption, without analyzing the biggest picture of the impacts of the pandemic on the long-term future and on the drivers for energy demand. The change in energy intensity and electricity demand during lockdowns, assessed on different spatial scales in [17], leads to different considerations for energy recovery in areas of the world with the more diverse behaviors, while [18] investigates the electricity demand fluctuation between two subsequent years (2019 and 2020) in a limited spatial scale and [19] the most recent Chinese oil and electricity demand trends in order to drive future policy making in a sustainable recovery framework. As a common point, those and several other works all focus on final energy use providing future insights via a simulation approach. Nonetheless, they neglect the underlying socio-economic driver perturbations that would inevitably influence long-term energy service demand evolution.

During the last year, several works assessed the effects of the pandemics on energy and electricity use patterns, aiming at providing data and analysis solutions to overcome energy security and sustainability issues in a post-pandemic scenario. The change in energy intensity and electricity demand during lockdowns, assessed on different spatial scales in [17], leads to different considerations for energy recovery in areas of the world with the more diverse behaviors, while [18] investigates the electricity demand fluctuation between two subsequent years (2019 and 2020) in a limited spatial scale and [19] the most recent Chinese oil and electricity demand trends in order to drive future policy making in a sustainable recovery framework. As a common point, those and several other works all focus on final energy use providing future insights via a simulation approach. Nonetheless, they neglect the underlying socio-economic driver perturbations that would inevitably influence long-term energy service demand evolution.

The need of having a simple model that is capable of perform accurate econometric forecasts arises from the fact that, in case of sudden shocks or necessary adjustments to be made to the evolution of demand drivers in partial equilibrium models, only small datasets could be available, without any ad-hoc economic assumptions. Time series regression models represent the most reasonable choice, given their popularity for econometric projections due to their implementation simplicity [20].

In the case of such analyses, Vector AutoRegressive (VAR) models represent the most natural tool when dealing with time-series regression models. In fact, this types of models are able to determine mathematical correlations among the different time-series to forecast, and for this reason they are capable to maintain, at least in part, such internal coherence discussed for the general equilibrium models [15].

The difficulty in the definition of explanatory and exogenous variables and their projections in the future adds more difficulties to the ones already discussed, and could dangerously increase the overall inaccuracy of the results [21]. For the abovementioned reasons, Vector AutoRegressive (VAR) models come in hand, considering that the benefits of modelling all the variables to forecast jointly rather than one equation at a time are well established [22], without the need of trying to necessarily define exogenous variables. In the same context of the COVID-19 pandemics, similar tools have already been used for forecasting infection cases and deaths from the virus [23]: as in that work, while parameters for the models are estimated using real data, the accuracy of future projections may depend on many different, unpredictable variables. Nevertheless, the implementation of updated projections for socio-economic drivers is expected to radically impact on future energy consumption patterns, established by energy models, and cannot be neglected when working with a tool devoted to provide reliable and policy relevant insights as a TIMES model.

Attempts at modelling the pandemic crisis response of the evolution of economic indicators like GDP have been made [24], but the results are generally short-term projections which do not fit the needs of long-term analyses performed by energy system optimization models. Furthermore, even considering the possibility of extrapolating long-term forecasts starting from such results, the lack of literature that analyzes more specific socioeconomic drivers remains a problem. Thus, this work focuses on that need for a radical change of hypotheses and assumptions needed to drive the evolution of an energy system in an energy economy optimization model after such an unprecedented holistic crisis.

More in detail, here it is addressed the quantitative assessment of long-term effects of the economic shock caused by the pandemics on production trends in the Italian industrial sector using the TIMES-Italia. TIMES-Italia is a model instance of the TIMES family for the development of perspective energy scenarios for Italy. It was developed and maintained at ENEA, the Italian National Agency for new technologies, energy and sustainable economic development, primarily for the analysis of supporting strategies for public and transparent decision-making [25]. Indeed, TIMES-Italia was also used to produce the Italian National Energy Strategy (SEN) [26] on a time horizon extended until 2030. However, the tool is capable to perform long-term energy projections until 2050, and it describes the Italian energy system in its totality, starting from the upstream sector for import and extraction of primary supply resources, passing through energy transformation and distribution, to end-use devices for the satisfaction of final energy service demands in the transportation, buildings, industrial and agricultural sectors. Starting from 2006, the base year in which the energy system is represented matching actual energy consumption statistics provided by the national energy balances [27], energy consumption, and the related costs and emissions of greenhouse gases, among other aspects, are computed by the model on the basis of the available dataset of energy technologies which can be chosen to compose the future supply and production mix with the minimum total cost. The model can also be customized to account for particular user constraints on the maximum exploitation of some resources or putting a cap on the maximum amount of e.g., CO2 emissions over a certain time period.

This thesis is structured as follows. A first discussion on the TIMES model framework rationale is carried out, followed to an in-depth description of the Reference Energy System (RES) of the industrial sector in the TIMES-Italia. The methodology for the econometric projections is then discussed, and VAR models are described together with all the steps needed to build such models and determine valid and reliable projections. Therefore, it follows a description of the cross-validation approach utilized to choose the best model specifications related to the abovementioned needs. This is then discussed the validation of the model starting from results of projections performed from 2018, i.e., without considering the effect of the pandemic crisis. Such results have been compared with the projections considered in the Italian Integrated National Energy and Climate Plan (PNIEC) [28], which in turn are projections taken from the EU Reference Scenario 2016 [29], in order to further validate such results. The final projections that consider the 2020 data are then presented and discussed in Chapter 6. The TIMES-Italia forecasts of the energy consumption of the various industrial sectors are presented and discussed in Chapter 5, together with a comparison between the backcasting results of such model in the period 2006-2019 with the real data of final industrial energy consumption of Eurostat [30]. Eventually, the conclusions of the work are drawn in the final chapter.

Chapter 2

Fundamentals of the TIMES Model generator

The TIMES (The Integrated MARKAL-EFOM System) model generator was developed by IEA-ETSAP (Energy Technology Systems Analysis Program) context, to build long term energy scenarios and analyze energy and environmental topics [31]. Such model generator combines a technical engineering approach together with an economic approach. In fact, TIMES is a technology-rich, bottom-up model generator (engineering aspect), which uses linear-programming to produce a least-cost energy system (economic aspect), where the optimization is subject to a set of user constraints over a multi-period time horizon [11].

Bottom-up (BU) models present a wide description of technologies, for such reason are also defined as “technology rich”. They are very useful for modelling technological details for the energy systems to analyze, but they rely on a precise characterization of the various processes on which the model is built.

2.1 TIMES models rationale

TIMES model generator can be described as:

- **Technologically explicit** and **integrated**. TIMES is a technology-rich model in which each technology is described by the definition of a series of technical and economic parameters. This means that an explicit identification and distinction of each technology is performed in such model. Furthermore, the user can change at will the number of technologies and their relative topology, without modifying the model’s equation. For this reason, the model can be also described as data driven.
- **Multi-regional**. TIMES models like the global ETP-TIMES model can comprise several regional modules. The only limit on the number of regions in a model depends on the difficulty of solving the linear programming optimization. Energy and material trading variables link together the various modules. Given the existence of such linking variables the model can be seen as multi-regional, where actions taken in one region may as well affect the others.

- A model that computes a **partial equilibrium** on the energy market with **perfect foresight**.

2.1.1 Partial equilibrium

Once all the inputs, constraints and scenarios have been defined, the model tries to determine the energy system that meets the energy service demands over the entire time horizon through a least cost optimization. TIMES assumes perfect foresight, which means that there is a complete knowledge of future events concerning the investment decisions that are made. So, the optimization is made horizontally across all sectors, and vertically across all time periods.

An optimal mix of technologies and fuels at each period is then determined, along with the production and consumption of the different commodities and their prices. The optimization process seeks the equilibrium between supply and demand, because in this state the producer and consumer surplus is maximized.

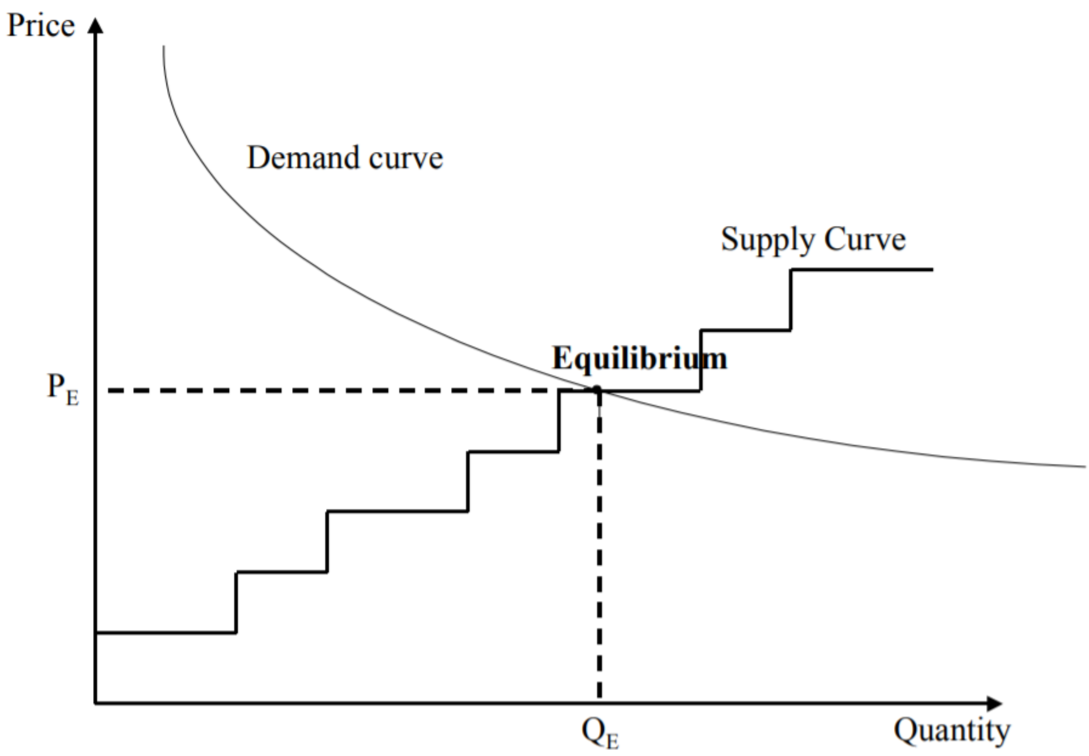

Figure 3: Market equilibrium in the case of an energy service in TIMES [11]

Figure 3**Errore. L'origine riferimento non è stata trovata.**shows the demand curve of an energy service, defined by the user via the specification of the own price elasticity of that demand. The supply-demand equilibrium point is represented by the intersection of the supply function and the demand function, and it corresponds to an equilibrium quantity Q_E and an equilibrium price P_E. Considering that TIMES equilibrium deals with many commodities simultaneously, the equilibrium is multi-dimensional, and Q_E and P_E are vectors and not scalars.

2.1.2 TIMES model architecture

The main output TIMES are energy system configurations, in which the end-use energy service demands are met through a least cost optimization, while the various defined constraints are observed. Figure 4 shows the TIMES model schematic and its outputs [32], and it can be summarized in the following main parts [33]:

- Resources, represented by the blue box: it includes current and potential availability of traditional and renewable energy sources, as also imports and trade.
- Conversion, red boxes: it includes all the technologies and processes related to the conversion of resources into usable energy.
- Consumption, represented by yellow boxes: the various uses, i.e., energy services, for the converted and distributed energy are described, along with the conversion technologies to usable forms of energy, such as light, heat, or refrigeration.
- Demands, represented in green: it is defined the amount of energy-services required, determined by modeling their relation to socio-economic exogenous drivers. Such drivers are generally determined by means of applied general equilibrium models like GEM-E3 [10].
- Outputs, represented by white exiting arrows: they are energy commodity prices, energy flows, GHG emissions, capacities of technologies, energy costs and marginal emissions abatement costs.

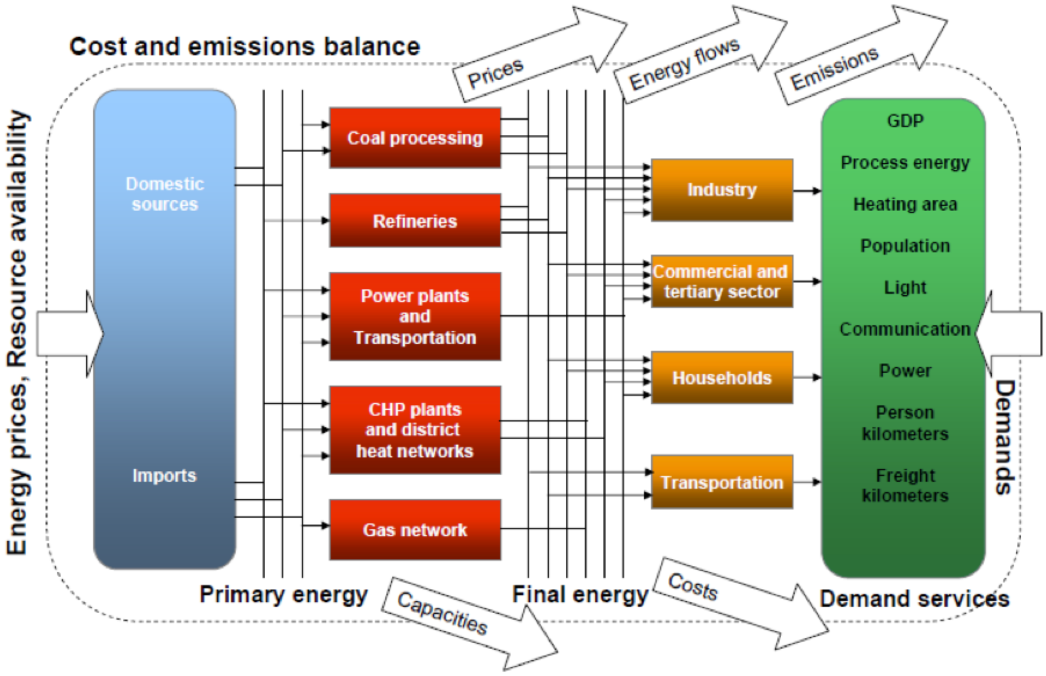

Figure 4: Schematic of TIMES inputs and outputs [32]

2.2 Reference Energy System

In TIMES, the technologically detailed feature is given by the description of a very large number of processes and technologies that produce and transform energy. All the steps between primary resources to end-use consumption are modeled. The energy supply-side (upstream sector) includes fuel mining, primary and secondary production, along with exogenous import and export. Energy is delivered to the end-use demand side through energy carriers, and those demands envisage all the activities included in the industrial, transportation, residential and commercial sectors.

Three main elements constitute every TIMES model [11]:

- **Technologies** (also called **processes**) transform commodities into other commodities. They can be primary sources of commodities (i.e., mining or import), conversion plants, energy-processing plants, or end-use devices such as cars or heating systems.
- **Commodities** are energy carriers, energy services, materials, monetary flows, and emissions. It is either produced or consumed by a technology.
- **Commodity flows** are what links processes with commodities. A flow represents, then, an input or an output of a particular process, and it is of the same nature as a commodity.

Through such units a Reference Energy System (RES) characterizing the country or region to analyze can be defined. Figure 5 shows a partial view of a RES, in which the blue boxes represent the processes, the vertical lines the commodities and the horizontal lines the commodity flows.

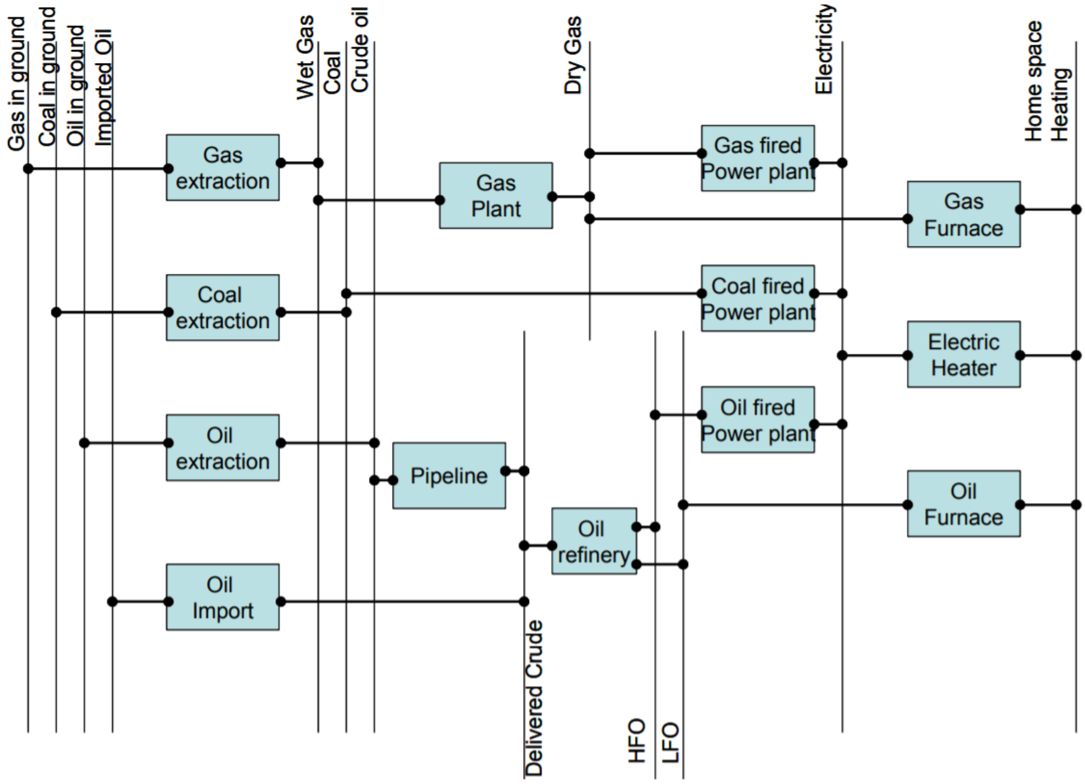

Figure 5: Example of a (partial) Reference Energy System [11]

2.3 Scenario analysis

TIMES models are generally used to analyze possible energy futures by defining contrasted scenarios. The scenario approach is quite useful given the long-time horizons simulated by such type of models, whereas econometric methods come in hand when shorter time scales are to be analyzed. In the definition of scenarios, the main drivers for the evolution of the energy system are not ad-hoc presupposed, but their projections are built starting from coherent assumptions. Four types of inputs need to be given to the model in order to well define scenarios: energy service demands, primary resources availability, sets of constraints, and the already mentioned set of technologies.

2.3.1 Energy service demands

Demand drivers (population, GDP, number of households, etc.) in the TIMES framework are obtained externally and given as inputs. For example, applied general equilibrium models like GEM-E3 can be used to derive such drivers trajectories [10].

The main drivers considered in the TIMES framework are: population, GDP, GDP per capita, number of households, and sectoral outputs.

The reference demand scenario can be constructed by computing a set of energy service demands trajectories. This is done by correlating such service demands to specific drivers, through the definition of elasticities of demands:

$$D_t = D_{t-1} \cdot \left[1 + \left(\frac{\delta_t}{\delta_{t-1}} - 1\right) \cdot e_{t-1}\right] \tag{1}$$

where D is the demand, t the time step, δ the demand driver and e the elasticity of the demand to its driver [11]. Different scenarios are then defined starting from different assumptions affecting the demand, such as emission constraints or different technology sets. The demands are endogenously adjusted in TIMES through the definition of another set of inputs, that are the elasticities of the demands to their own prices.

Equation 1 shows that energy services demands and drivers' projections are linked together through elasticities. Such elasticities are meant to describe any change in the patterns of energy service demands due to economic, technological, political, social or cultural factors, such as saturation effects like the decoupling of GDP and demand growth in developed countries. Elasticities usually take values close to 1, and the lower the elasticity the less influence of the driver on the energy demand service [13].

The analysis presented in this work deals with understanding how the industrial energy demand trajectories in Italy will be affected by the pandemic crisis, developing subsequent insights with long-term energy system optimization models like TIMES-Italy. In fact, a key question about the effect of the pandemic is whether it has affected the dynamics of the demand of energy services, which is the first driver of energy demand (together with the dynamics of energy technologies) [12]. Such dynamics depends on one hand on the evolution of the key socio-economic drivers, on the other hand on the elasticity of each energy service demand to these socio-economic drivers. Elasticities change over time, and they have a decisive impact on the long-term evolution of energy demand [14]. In this sense, the aim of this work is to focus on these critical points and to suggest an approach based on traditional econometric models able to accurately project the socio-economic drivers of our interest.

2.3.2 Supply component

Multi-stepped supply curves for primary energy and material resources, with each step representing a potential of the resource available at a particular cost, constitute another important input in the scenarios definition. Such potential can be expressed in different ways:

- Cumulative over the model horizon, like the amount of reserves of gas, crude oil or coal;
- Cumulative over the resource base, like the available areas for wind or solar energy conversion, or the available farmland for biocrops;
- Annual potential such as the maximum extraction rates, or the available potential of renewable resources.

The supply component may also include trading possibilities of commodities determined endogenously.

2.3.3 Policy component

One of the major impacts on the energy system is the introduction of policies, and for this reason their definition is a fundamental part for modelling a scenario. One example can be related to policies on GHG emissions, that could be ignored in a reference scenario, while they may be an important constraint in term of emission restrictions or taxes in different policy scenarios.

In the TIMES framework both micro measures like technology portfolios and wider policies like carbon taxes or nuclear policies can be simulated, given the technology richness of such model generator.

2.3.4 Techno-economic component

In TIMES, modelling the set of technical and economic parameters needed for converting primary resources into energy services may differ according to different analyzed scenarios. Such techno-economic parameters are presented as a set of technologies that transform some commodities into others (fuels, materials, energy services, emissions). Some of these technologies can be user imposed (i.e., starting from the energy mix at the base year) and others are simply available for the model to choose from, starting from the linear optimization problem.

As already said above, one of the strengths of a TIMES model lies on its technology richness, where both current and future technologies are defined, and the model can choose from such sets following the cost minimization paradigm.

Two types of technologies can be defined:

- Base year technologies, representing the existing technology mix, and defined starting from the model calibration of the RES at the starting year, considering as a reference historical data of the energy balance in such year.
- New technologies, representing future technologies from which the model can choose from, available from a precise user-defined year.

2.4 TIMES-Italia

In the Italian context, an energy model based on the TIMES framework has been developed by ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, primarily for the analysis of supporting strategies for public and transparent decision-making [25]. Indeed, TIMES-Italia was also used to produce the Italian National Energy Strategy (SEN) [26] on a time horizon extended until 2030. However, the tool is capable to perform long-term energy projections until 2050, and it describes the Italian energy system in its totality, starting from the upstream sector for import and extraction of primary supply resources, passing through energy transformation and distribution, to end-use devices for the satisfaction of final energy service demands in the transportation, buildings, industrial and agricultural sectors. Starting from 2006, the base year in which the energy system is represented matching actual energy consumption statistics provided by the national energy balances [27], energy consumption, and the related costs and emissions of greenhouse gases, among other aspects, are computed by the model on the basis of the available dataset of energy technologies which can be chosen to compose the future supply and production mix with the minimum total cost. The model can also be customized to account for particular user constraints on the maximum exploitation of some resources or putting a cap on the maximum amount of e.g., CO_2 emissions over a certain time period.

An aggregated scheme of such described RES is presented in Figure 6.

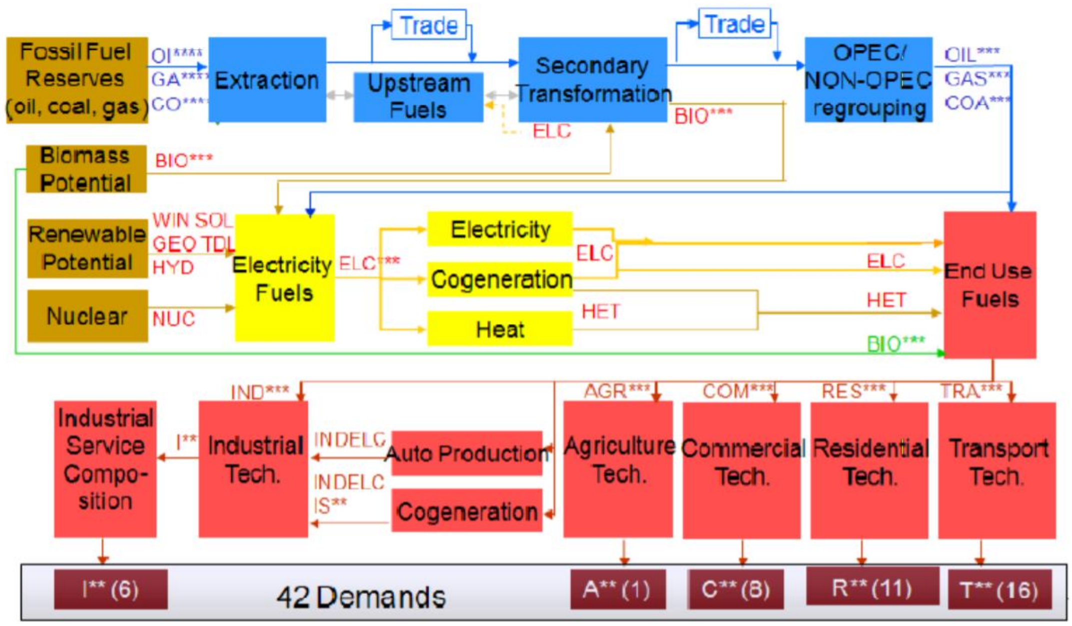

Figure 6: TIMES-Italia Reference Energy System [25]

2.4.1 Time horizon

TIMES-Italia allows to analyze time horizons until 2050, as discussed in the previous paragraph. In Table 1 the periods of the entire time horizon are shown. As it can be seen, the duration of such periods increases as we go towards 2050: one year for the first three periods, two years up to 2022, five years between 2025 and 2030, and ten years up to 2050. Having a lower duration of the first periods allows a deeper description of the already implemented policies, and as the uncertainty increases going in the future, the need of detail with smaller periods' length goes decreasing.

Table 1: Periods of TIMES-Italia time horizon.

2006	2007	2008	2010	2012	2014	2016
2018	2020	2022	2025	2030	2040	**2050**

Furthermore, a base year defined in 2006 allows to validate the model by comparing the results in the first years with the historical data at disposal.

2.4.2 Time slices

In TIMES-Italia each period of the year is divided in Times Slices, comprising the four seasons, and day and night hours, as presented in Table 2.

Table 2: Time slice in TIMES-Italia

	Summer	Fall	Winter	Spring
Day	0.125	0.115	0.105	0.115
Night	0.115	0.125	0.135	0.125
Peak	0.010	0.010	0.010	0.010

Such level of detail is fundamental to have a specific characterization of energy supply and demand. For example, some energy services concerning the demand are only active in some periods of the year, such as the residential heating during winter, or in some hours of the day, like the lightning during night. As for energy production, the same goes for renewable sources, only active in some periods of the year (hydroelectric) or in some hours of the day (photovoltaic).

Such points cannot be neglected, and for this reason some parameters characterizing various processes have different values depending on the time slices. At each time slice, the balance between supply and demand is ensured by the model.

2.4.3 Discount and depreciation rates

A constant real discount rate of 5 % is considered starting from the base year. Different discount rates can also be employed to differentiate innovative technologies from more mature ones.

A reasonable definition of depreciation rates are important in such energy models to well describe possible evolutions of the energy system. In TIMES-Italia the depreciation rate differs from sector to sector. In Table 3 it can be seen how in the industrial, commercial and agricultural sector the dicount rate is 30 %, in the residential sector is 60 %, and for transports it varies from 10 % to 45 %. Each investment in the energy sector is discounted with a 6 % rate.

*Table 3: Discount rates in TIMES-Italia. * Discount rate increasing for more advanced vehicles* [25].

Sector	**Discount rate**
Energy	6 %
Industry	30 %
Commercial	40 %
Residential	60 %
Transport*	
Bus and trucks	10 %
Cars and motorcycles	15 %

2.4.4 Greenhouse gases emissions

TIMES-Italia also allows to define different scenarios to perform environmental analyses examining the consequences of a certain environmental policy on the greenhouse gases (GHG) emissions. Each fuel in such model is associated an emission factor related to CO_2, N_2O and CH_4. At traditional fuels are associated static emission factors derived from literature, while at aggregated fuels are associated dynamic factors depending on the composition of such fuels. The emission of the different GHG defined above are then an output of the model, and different scenarios can be compared in such terms.

Chapter 3

Industry sector Reference Energy System

As discussed in Chapter 2, one fundamental step for energy system modeling requires the definition of the Reference Energy System (RES), which can be described as the set of “all the components related to the production, conversion, delivery and end-use of energy” [34]. A detailed RES for each analyzed sector allows the analysis of the mix of traditional and innovative technologies given by the optimization techniques in different scenarios.

This chapter presents the RES for the Italian industry sector as modeled in TIMES-Italia. The industrial module used in the current version of TIMES-Italia, used in this work for energy projections, is based on the open-source industrial demand technology database (Open-IDTD) by Lerede et al. [35], in which, among other things, the importance of having data transparency and result reproducibility when dealing with macro-scale energy modelling is stressed.

Globally, the industrial sector accounted for 37 % of total global final energy use in 2018 [36], while it accounted for about 20.5 % in Italy [37], being one of the most energy-intensive sectors means having a big impact on the results of energy modeling. For this reason, a detailed characterization of all the processes constituting such sector is necessary for analyzing the energy sector as a whole. In TIMES-Italia, the industrial sector is subdivided in six energy-intensive subsectors for each end-use energy service demand:

- Iron and steel, in which steel is produced.
- Non-ferrous metals, aluminum, copper and zinc production technologies are taken into account.
- Non-metallic minerals, including cement, ceramics, glass, and lime production technologies.
- Chemicals, and namely ammonia, chlorine, high value chemicals (HVC) and methanol.
- Pulp, paper and printing, contributing to the final paper demand.
- Other Industries, representing textile, food, beverage, tobacco, plant construction, machinery, and transport equipment sectors.

The technologies included in the TIMES-Italia database are based on the technological portfolio presented in the International Energy Agency (IEA) Energy

Technology Perspectives (ETP) Model [38], in which several low-carbon options are presented for the different industrial subsector included in TIMES-Italia.

In the following paragraphs a general description of how the energy intensive technologies are modeled in the TIMES framework precedes an in-depth analysis of the technologies characterizing the various industrial subsectors in TIMES-Italia.

For each sector, they are also presented the drivers related to their respective service demand. Such drivers follow the statistical classification of economic activities in the European Community (NACE) [39], a derived classification of the International Standard Industrial Classification (ISIC). NACE is the subject of legislation at the European Union level[2], which imposes the use of a uniform classification within all the Member States. It belongs to the international integrated system of economic classifications, which is based on the classifications made by the UN Statistical Commission (UNSTAT), Eurostat, and other national classifications. All of them are strongly related each to the others, and for this reason a comparability among other economic statistics produced worldwide by different institutions is made possible [39].

3.1 Energy intensive technologies

A set of technologies is modelled for each of the subsectors mentioned above, as shown in Figure 7.

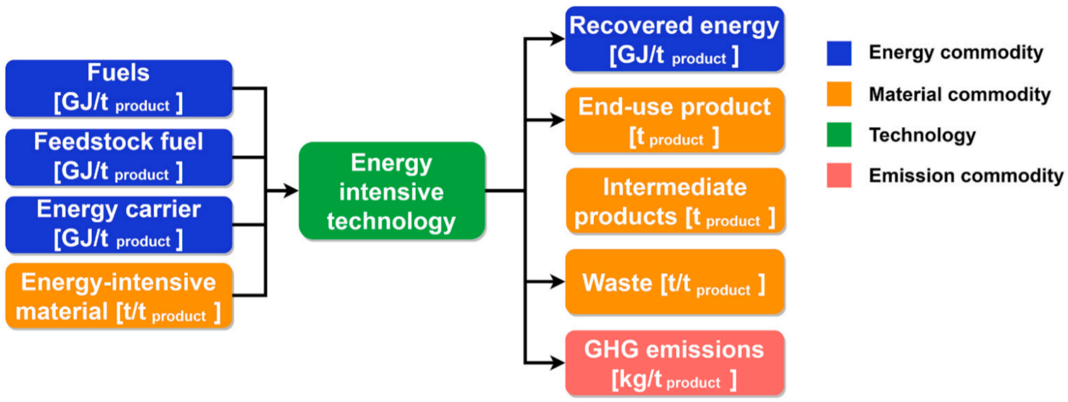

Figure 7: Structure of each energy-intensive industrial technology [35].

Entering in detail of Figure 7, the inputs of such technologies involved in the transformation process generally are:

---

[2] Council Regulation (EEC) No 3037/90 of 9 October 1990 on the statistical classification of economic activities in the European Community (OJ No L 293, 24.10.1990, p. 1) as amended by Commission Regulation (EEC) No 761/93 of 24 March 1993 (OJ No L 83, 3.4.1993, p. 1, and corrigendum, OJ No L 159, 11.7.1995, p. 31).

- **Fuels** expressed in $\mathrm{GJ/t_{product}}$, where $\mathrm{t_{product}}$ is the mass of the output product. Some examples of fuels are coal, natural gas, crude oil, etc.
- **Energy carriers** expressed in $\mathrm{GJ/t_{product}}$. Steam, electricity and machine drive are examples of energy carriers.
- **Feedstock fuels**, used in the chemical sector and expressed in $\mathrm{GJ/t_{product}}$. The difference with common fuels is related to the fact that feedstock fuels are not involved in combustion processes.
- **Energy-intensive materials**, expressed in $\mathrm{t/t_{product}}$.

The outputs of such technologies are, on the other hand:

- **End-use products**, expressed in $\mathrm{t_{product}}$;
- **Byproducts** successively used for producing other end-use products in other processes, expressed in $\mathrm{t_{product}}$.
- **Recovered energy carriers**, expressed in $\mathrm{GJ/t_{product}}$.
- **Greenhouse gas (GHG) emissions**, expressed in $\mathrm{kg/t_{product}}$.
- **Waste materials**, like the slag from steel production, expressed in $\mathrm{t/t_{product}}$.

Factors characterizing each technology are of three main types:

- **Technical factors**, like its energy input requirements, the starting date for the availability of the technology in the production system, the plant lifetime and the availability factor.
- **Economic factors**, such as investment cost or annual fixed O&M costs, both given in $\$/t_{final\ product}$.
- An **environmental performance factor**, which is actually automatically calculated by TIMES on the basis of the provided fuel emission factors and input commodity efficiencies.

3.1.1 Iron and steel

According to the industrial demand technology database in [35], steel production can be satisfied according to a set of thirteen technologies which comprehend the whole manufacturing process starting from raw materials to crude steel. The Open-IDTD also includes generic process defined for ferroalloys production, which considers a weighted average energy consumption of global production processes, but ferroalloys are not included as a demand commodity in TIMES-Italia. Concerning technologies that include carbon capture and storage (CCS), processes for smelting reduction are considered, blast and basic oxygen furnaces (BFBOF) and electric arc furnaces (DRI-EAF) [40]. Table 4 summarizes such technologies together with the starting date in the TIMES-Italia model, related to technology readiness, expressed in terms of its deployment state.

Table 4: Iron and steel technologies, comprising their starting date in TIMES-Italia model and their deployment state.

Product	Technology	Starting date	Deployment state
Steel	Blast furnace-basic oxygen furnace (BF-BOF)	Base year	Traditional
	Direct reduced iron-electric arc furnace (DRI-EAF)	Base year	Traditional
	Steel from scrap – EAF	Base year	Traditional
	Smelting reduction – BOF	2006	Innovative/Commercial
	BF-BOF with CCS	2020	Demonstration phase
	BF with top gas recovery – BOF	2020	Demonstration phase
	DRI-EAF with CCS	2020	Demonstration phase
	HIsarna – BOF	2025	Demonstration phase
	HIsarna – BOF with CCS	2025	Demonstration phase
	Hydrogen direct reduction – EAF	2030	Demonstration phase
	Ulcored with CCS	2030	Demonstration phase
	Ulcolysis	2030	R&D phase
	Ulcowin	2030	R&D phase

The driver for the Iron and steel service demand is related to the statistical classification of economic activities in the European Community (NACE), and it corresponds to the industrial production index of the manufacture of basic metals and fabricated metal products, except machinery and equipment (code CH, divisions 24-25). The reference data is derived from ISTAT Database [41].

3.1.2 Non-ferrous metals

The eleven non-ferrous metals manufacturing technologies characterized in Open-IDTD are shown in Table 5. Aluminum is modeled by considering six different technologies, while for copper, zinc, tin, titanium and niobium a single technology each is considered. Two steps can be determined for aluminum, that are the alumina production represented by the Bayer process [42], and the output of such process is given as input for aluminum technologies except the secondary route [43] and the kaolinite reduction [44].

Table 5: Non-ferrous metals technologies, comprising their starting date in TIMES-Italia model and their deployment state.

Product	Technology	Starting date	Deployment state
Alumina	Bayer process	Base year	Traditional
Aluminum	Hall-Héroult	Base year	Traditional
	Secondary aluminum	Base year	Traditional
	Hall-Héroult with inert anodes	2030	Innovative/Commercial
	Carbothermic reduction	2050	Demonstration phase
	Kaolinite reduction	2050	Demonstration phase
Copper	Primary copper production	Base year	Traditional
Zinc	Zinc production	Base year	Traditional

Non-ferrous metals service demand is derived from the driver related to the industrial production index of the manufacture of basic precious and other non-ferrous metals. In the ISTAT Database, this corresponds to the division 244, a sub-category of the manufacture of basic metals.

3.1.3 Non-metallic minerals

Concerning Non-metallic mineral manufacture, in TIMES-Italia eleven technologies are considered, as shown in Table 6. A clinker production step before cement blending is the traditional process modeled for cement production [45]. Innovative processes like Alkali-activated cement-based (AAC) blinders [46] and belite cement [47] are also modeled. Lime and ceramics production are modeled according to a single technology, while glass production considers two processes relying either on fossil fuels or electricity.

Table 6: Non-metallic minerals technologies, comprising their starting date in TIMES-Italia model and their deployment state.

Product	Technology	Starting date	Deployment state
Clinker	Dry process	Base year	Traditional

	Wet process	Base year	Traditional
	Dry process with post-combustion CCS	2025	Demonstration phase
	Dry process with oxy-fuel combustion CCS	2025	Demonstration phase
Cement	Cement blending	Base year	Traditional
	Alkali-activated cement-based blinders	2010	Innovative/Commercial
	Belite cement	2010	Innovative/Commercial
Lime	Long rotary kiln	Base year	Traditional
Glass	Fossil fuel-fired furnace	Base year	Traditional
	All-electric furnace	Base year	Traditional
Ceramics	Ceramics production	Base year	Traditional

The driver corresponding to such subsector is the production index of the manufacture of rubber and plastics products, and other non-metallic mineral products (code CG, divisions 22-23) [39].

3.1.4 Chemicals

The chemical sector is characterized by twenty-three technologies, eight for high value chemicals, which include olefins and aromatics, five for ammonia production, six for methanol production and three for chlorine production, as in Table 7.

Table 7: Chemicals technologies, comprising their starting date in TIMES-Italia model and their deployment state.

Product	Technology	Starting date	Deployment state
HVC	Naphta steam cracking	Base year	Traditional
	Ethane steam cracking	Base year	Traditional

	Gas oil steam cracking	Base year	Traditional
	LPG steam cracking	Base year	Traditional
	Propane dehydrogenation	2010	Innovative/Commercial
	Naphta catalytic cracking	2011	Innovative/Commercial
	Methanol-to-olephins	2015	Innovative/Commercial
	Bioethanol dehydration	2020	Demonstration phase
Ammonia	Natural gas steam reforming (NG SR)	Base year	Traditional
	Naphta partial oxidation	Base year	Traditional
	Coal gasification	Base year	Traditional
	Synthesis via electrolysis	2015	Innovative/Commercial
	Biomass gasification	2025	Demonstration phase
	NG SR with CCS	2025	Demonstration phase
Methanol	NG SR	Base year	Traditional
	Coke oven gas steam reforming	Base year	Traditional
	LPG partial oxidation	Base year	Traditional
	Coal gasification	Base year	Traditional
	Synthesis via electrolysis	2015	Innovative/Commercial
	Biomass gasification	2025	Demonstration phase
Chlorine	Mercury cell	Base year	Traditional
	Diaphragm cell	Base year	Traditional
	Membrane cell	Base year	Traditional

The driver associated to the chemicals sector in TIMES-Italia is the average of two industrial production indexes:

- Manufacture of chemicals and chemical products, code CE division 20.
- Manufacture of pharmaceuticals, medicinal chemical and botanical products, code CF, division 21.

3.1.5 Pulp and paper

Table 8 shows the six traditional pulp-processing technologies included in TIMES-Italia, and one generic process representing the paper production. CCS technologies are not considered for the pulp and paper sector.

Table 8: Pulp and paper technologies, comprising their starting date in TIMES-Italia model and their deployment state.

Product	Technology	Starting date	Deployment state
Pulp	Mechanical pulping	Base year	Traditional
	Semi-chemical pulping	Base year	Traditional
	Kraft process	Base year	Traditional
	Sulfite process	Base year	Traditional
	Recycled fiber pulping	Base year	Traditional
Paper	Paper production	Base year	Traditional

Pulp and paper service demand is determined from the driver related to the industrial production index of the manufacture of paper and paper products, code CC, division 17 [41].

3.1.6 Other Industries

The remaining industrial technologies are aggregated in a final subsector, in which a generic process is modeled. The associated driver for the energy service demand related to such sector is the industrial production index of the total manufacturing industry, which is labeled with the code C in the NACE classification [39].

Chapter 4

Analysis of the industrial historical data

The following paragraphs give a deeper description of the historical data utilized to train the VAR model. The general trend of each dataset is firstly showed considering the yearly aggregated data from 1990 to 2020. More in detail, an assessment on how each sector has responded to the 2009 crisis can help to understand what to expect for this current crisis caused by the Covid-19 pandemic. This assumes that a similarity in terms of response to an economic crisis exists within each sector, but such practice can be seen in different econometric analyses in literature. For example, Foroni et al. [24] perform a GDP nowcast in different European countries based on autoregressive models, and an intercept correction adjustment [48] is performed in order to force a behavior in the post-pandemic projection similar to the one had after the Great Recession in 2009. A final analysis is made by comparing the industrial production values of the different sectors in the first quarter of 2021 with the values had in the same period in 2019 and 2020. Being the most recent data at disposal for the projections, they hold the information related to the expected pace of the crisis recovery, and, therefore, they help to build some expectations on the results, especially for the short-term.

4.1 Iron and steel

Figure 8 shows the historical data of the industrial production index of iron and aggregated yearly from 1990 to 2020. It corresponds to the manufacturing volume of the industrial subsector normalized in base 100 at year 2015. A slow but steady increase prior to the 2009 crisis can be seen, and a constant and lower post-crisis trend is present, underlying how there has not been a full recovery from the Great Recession. Furthermore, 2020 values suggest how the pandemic crisis has had a big impact on this sector.

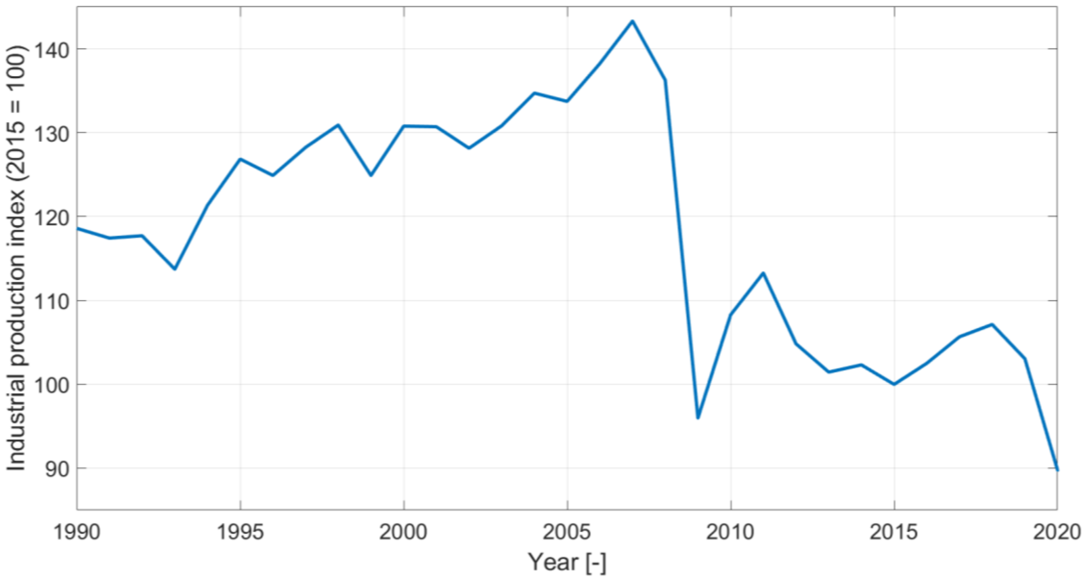

Figure 8: Historical data of the Iron and steel industrial production index

Such index is presented in the ISTAT Database as monthly values, and the first quarter of 2021 shows how there has been a full recovery with respect to 2019 values. In particular, in Figure 9 it can be seen that in March and April 2021 the industrial production of iron and steel has been even higher to the 2019 values. The percentage increase with respect to 2019 has been of 6.11 % and 9.00 %, respectively for March and April. Such months in 2020 represent the first lockdown measures taken by Italy, and with respect to that the increase in 2021 has been respectively of 63.6 % and 119 %. Such results can suggest that a fast recovery from the crisis is possible, considering that in the last two months the industrial production is even higher than the one registered in the same period of 2019.

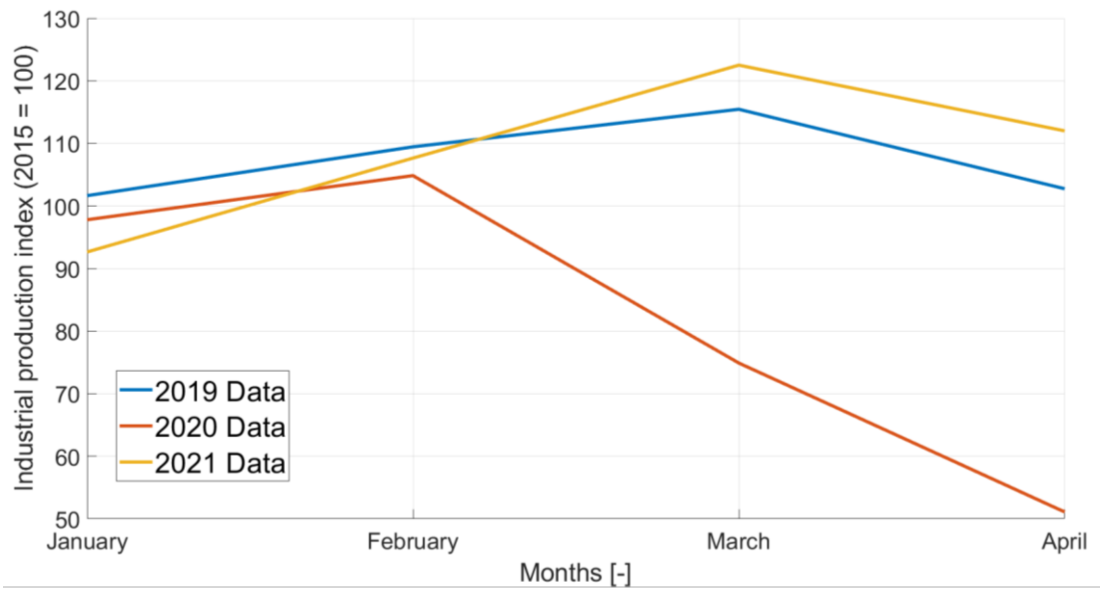

Figure 9: Comparison of the Iron and steel industrial production for the first quarter of 2019, 2020 and 2021.

4.2 Non-ferrous metals

The yearly industrial production index (base 100 at year 2015) of the Non-ferrous metals sector is shown in Figure 10. In this case, the pre-2009 trend is quite constant on average, with an initial increase up to year 2000 and a subsequent decrease. Similarly to the iron and steel sector, after the 2009 crisis the industrial production has been steady at a lower value with respect to the pre-crisis data. In this case, Covid-19 crisis has accentuated the decreasing trend started in 2019, but the values are not the lowest of the total period analyzed. The lowest values are, in fact, registered in 2013.

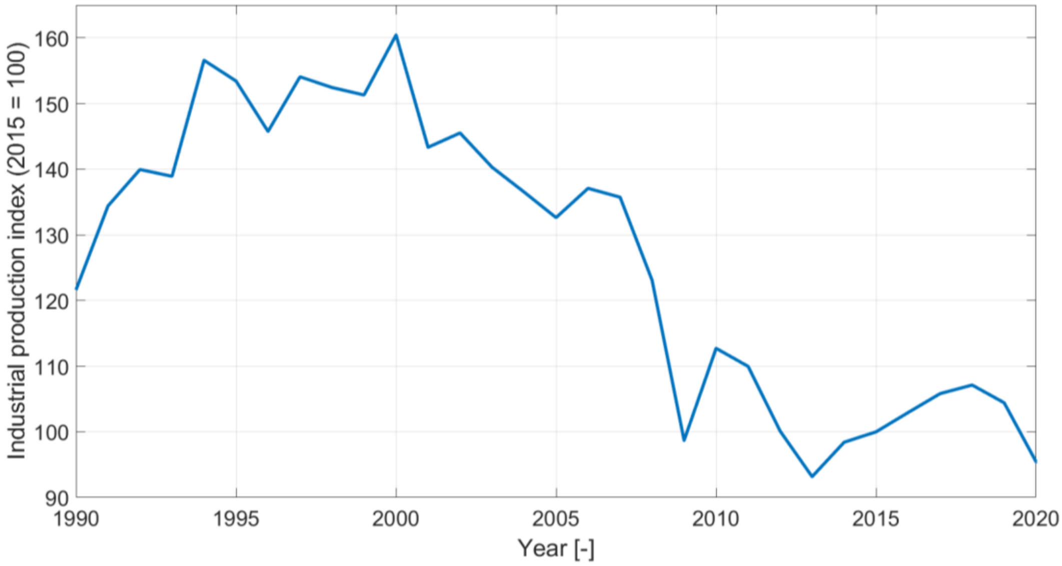

Figure 10: Historical data of the Non-ferrous metals industrial production index

The first quarter of 2021 shows for such data the same recovery pattern registered for the iron and steel sector. In fact, Figure 11 shows how the values in March and April of 2021 have been higher than the ones had in 2019. The percentage increase with respect to 2019 has been of 11.8 % and 15.9 %, respectively. Concerning the 2020 values, the increase in 2021 has been of 50.1 % and 59.4 %. Such results, along with the ones discussed for Figure 10, can suggest a fast recovery from the pandemic crisis, considering the low total impact in 2020 and the fast improvement had in the first months of 2021.

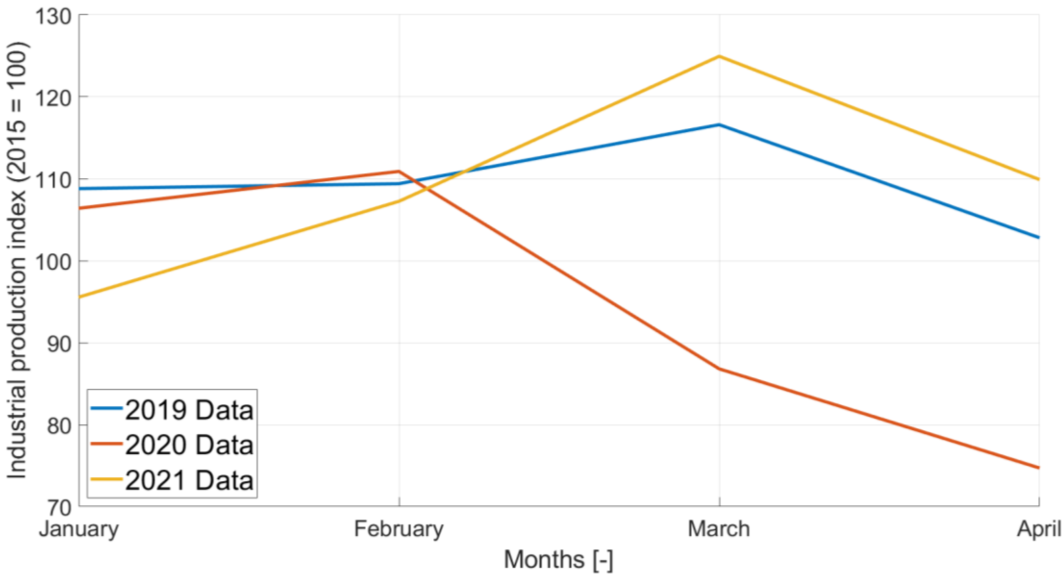

Figure 11: Comparison of the Non-ferrous metals industrial production index for the first quarter of 2019, 2020 and 2021.

4.3 Non-metallic minerals

Concerning the non-metallic minerals sector, the yearly industrial production index (base 100 at year 2015) is shown in Figure 12. A similar pattern with iron and steel and non-ferrous metals sectors is present, with respect to the clear difference in the trend between pre- and post-2009 crisis. In fact, the values up to 2007 are increasing, with a maximum in the year 2001, while after the Great Recession the trend presents a slow decrease. The Covid-19 pandemic has amplified such situation, and the 2020 has become the worse year in the past 30 years, at least. It is interesting to notice, however, how such decrease in 2020 is comparable in amplitude with some other years in the past, especially the 2011.

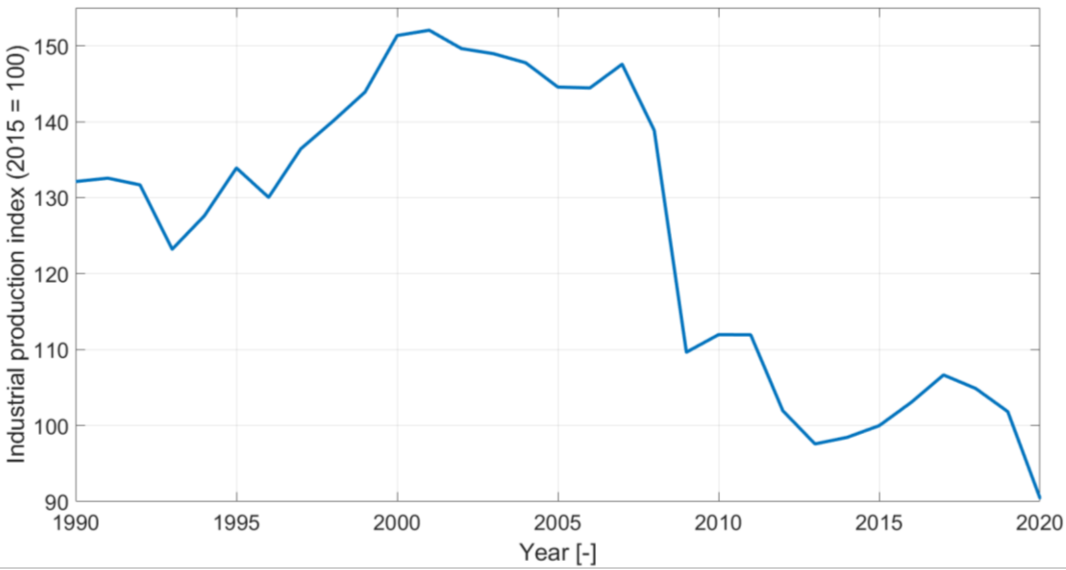

Figure 12: Historical data of the Non-metallic minerals industrial production index

In Figure 13 is presented the first quarter of 2021 compared with the same quarter in 2020 and 2019 for the industrial production index of the Non-metallic minerals sector. Differently from the iron and steel and non-ferrous metals sectors, the recovery with respect to 2019 has already started in February, with a 6.02 % percentage increase. As for March and April, the recovery has been of respectively 8.60 % and 8.79 % compared to 2019. Furthermore, the percentage increases of 50.6 % and 161.4 of the industrial production index of March and April 2021 with respect to 2020 show both the high impact of such crisis and the high resilience of such sector, which has returned to pre-pandemic values.

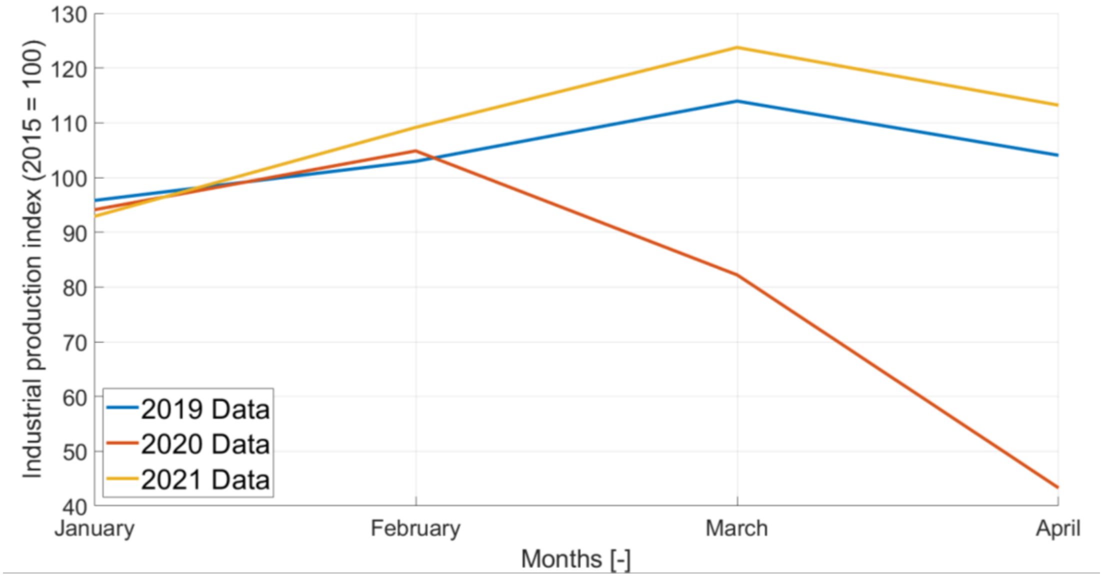

Figure 13: Comparison of the Non-metallic minerals industrial production index for the first quarter of 2019, 2020 and 2021.

4.4 Chemicals

The chemicals sector presents many differences both in terms of historical trend and crisis response. Figure 14 shows the yearly industrial production index (base 100 at year 2015) of such sector, and the first main dissimilarity with respect to the previously discussed sectors is its increasing trend in the whole period analyzed. A first fast increase is present up to the year 2000, while a full recovery of the 2009 crisis has been registered, with a returning of the industrial production index to pre-crisis values in 2017. Furthermore, 2018 registers the highest value in the past 30 years. Besides, the pandemic crisis has mildly hit such sector, which presents a steeper decrease in the years 2018-2019.

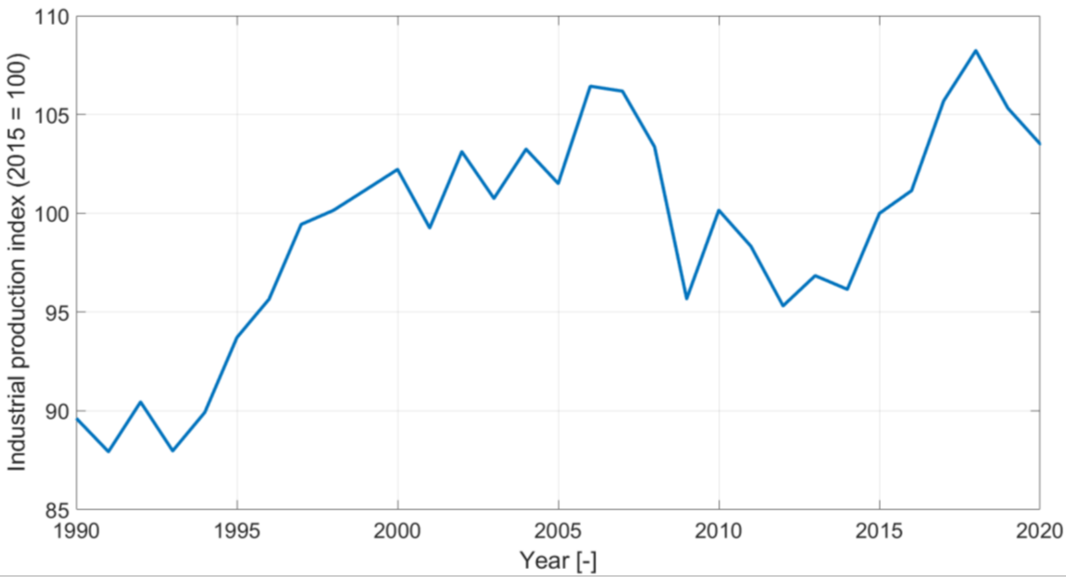

Figure 14: Historical data of the Chemicals industrial production index

If the 2020 crisis has not hit the chemicals sector as hard as the previously discussed sectors, Figure 15 shows how the recovery in the first months of 2021 has been more moderate, highlighting how a decreasing trend is already present from 2018. In fact, only in March 2021 it can be noted a higher industrial production index with respect to 2019, with a percentage increase of 4.47 %. The period from February to April 2021 presents an average percentage decrease of 0.36 % with respect to the same period in 2019, while if only the last two months are considered, the situation improves, with an average increase of 1.02 %. The increase from the 2020 data in March and April is also moderate, with respectively 9.13 % and 9.63 %, showing both the low impact of the crisis on such sector and its slow recovery from it.

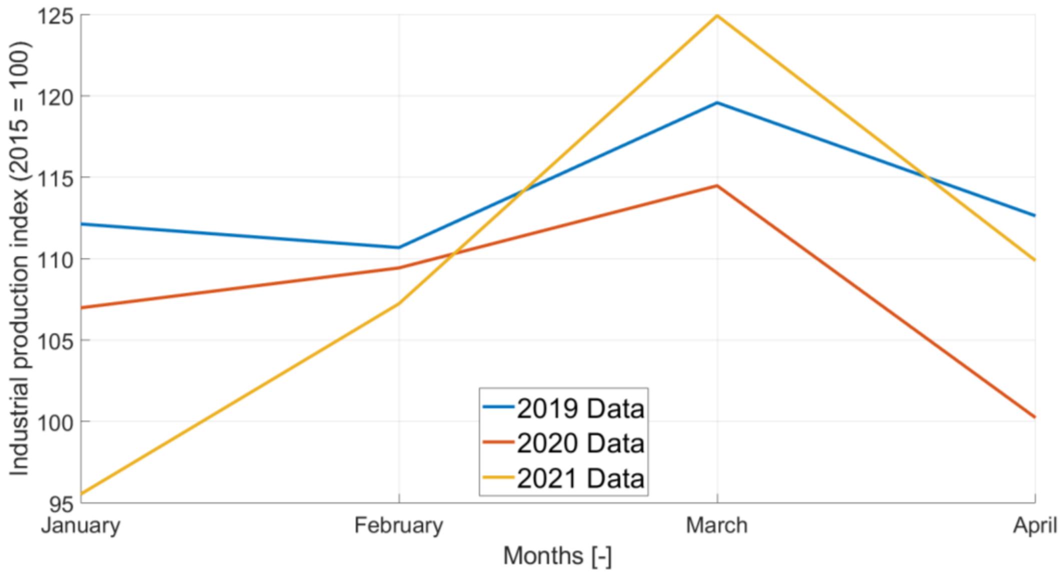

Figure 15: Comparison of the Chemicals industrial production index for the first quarter of 2019, 2020 and 2021.

4.5 Pulp and paper

The pulp and paper sector presents various similarities with all the sectors discussed above. In fact, like the iron and steel sector, Figure 16 shows a constant increasing trend up to 2007, also being the year with the highest value in the period analyzed. The trend is more moderately increasing on average after the crisis, with no recovery from it registered up to 2020. Covid-19 crisis has hit such sector even more modestly than the chemicals industry, with industrial production index values in 2020 comparable with the ones in 2019.

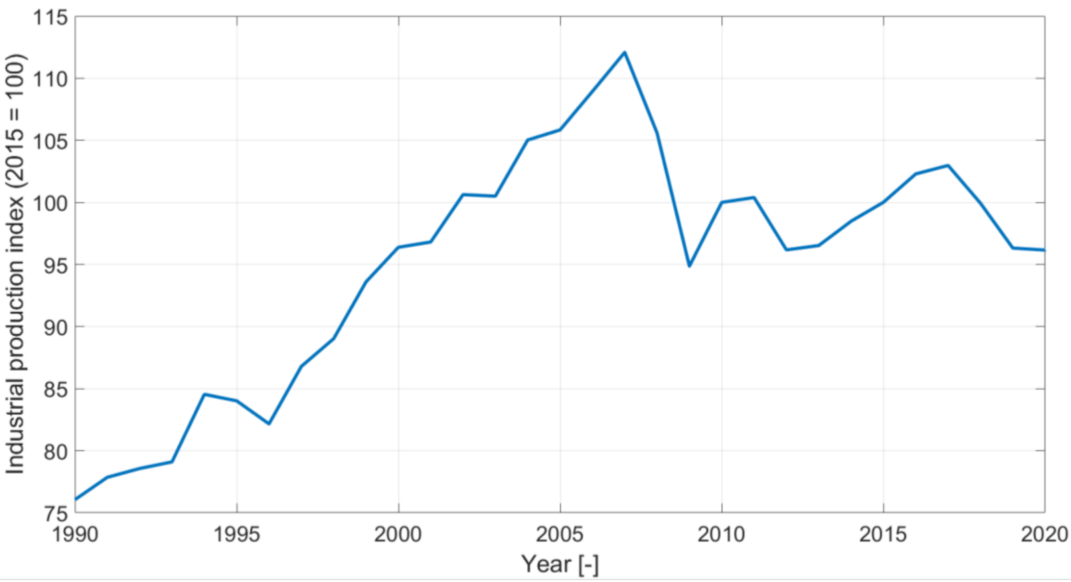

Figure 16: Historical data of the Pulp, paper and printing industrial production index

Such low impact of the 2020 crisis can be seen by comparing the industrial production indexes in the period from January to April in 2019, 2020 and 2021, as presented in Figure 17. In fact, 2021 data shows a high recovery in March and April with respect to 2019, with a percentage increase of respectively 2.45 % and 8.25 %. Furthermore, the increase from 2020 in the same period is of 8.47 % and 13.0 %, respectively, which is much lower with respect to sectors such as iron and steel and non-ferrous metals. The reason is related especially to the small decrease in 2020 compared to 2019, as it is shown in a monthly basis in Figure 17 for the first lockdown period, and yearly in Figure 16.

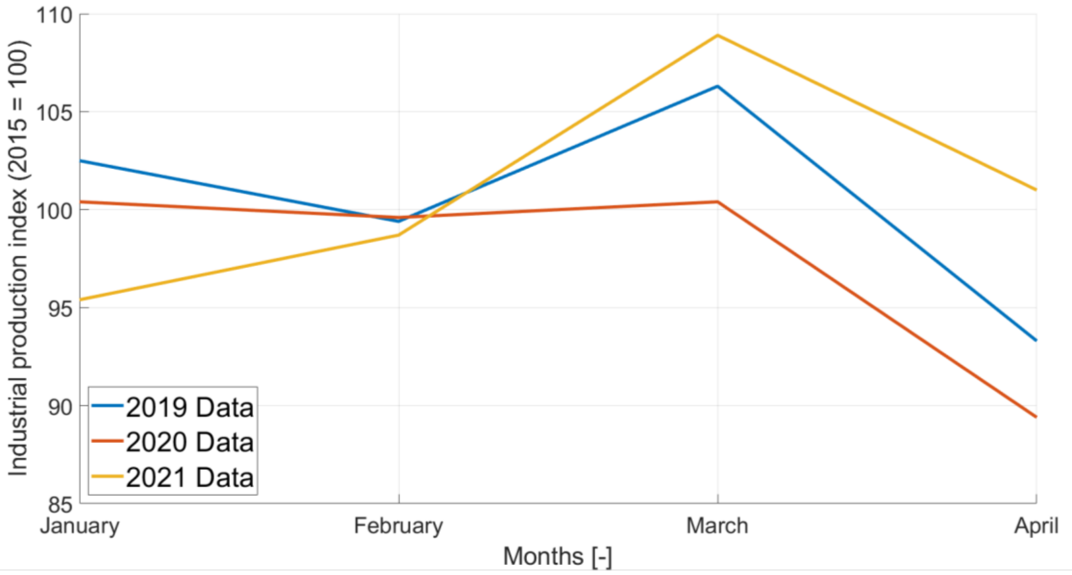

Figure 17: Comparison of the Pulp, paper and printing industrial production index for the first quarter of 2019, 2020 and 2021.

4.6 Other industries

The historical trend of the industrial production index (base 100 at year 2015) aggregated yearly for the other industries is shown in Figure 18. The behavior is very similar to the one seen in the iron and steel, non-ferrous metals, and non-metallic minerals sectors, with the 2009 crisis splitting in two the graph. A small increase in the pre-crisis period is, in fact, translated to an average stagnation after 2009. The pandemic crisis has hit very hard such sector, registering in 2020 the lowest industrial production index in the past 30 years, and a steep decrease with respect to 2019 which is second only to the 2008-2009 drop-down.

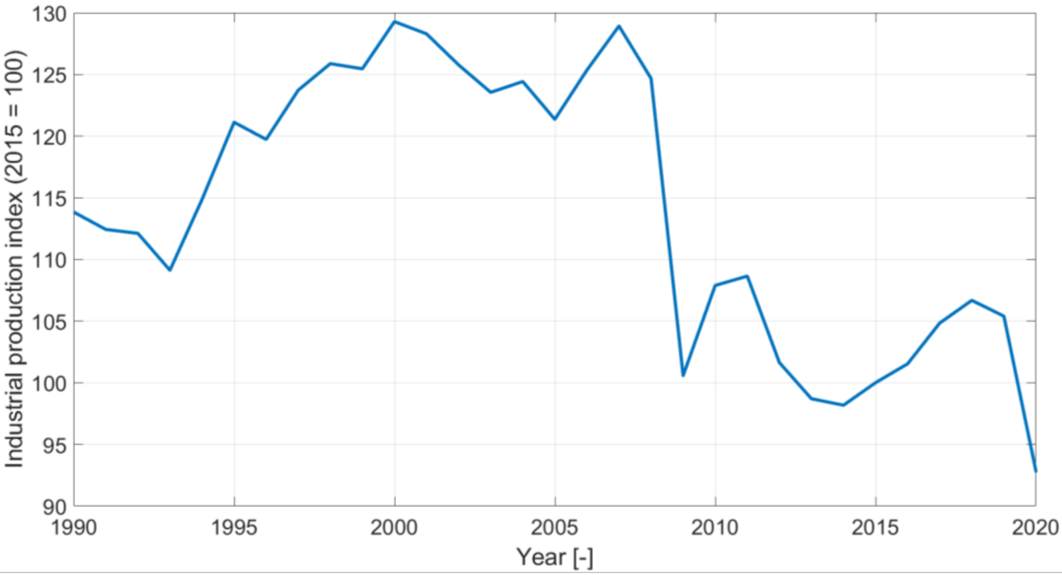

Figure 18: Historical data of the Other industries industrial production index

Such response to the crisis period can be also seen in the 2020 data for March and April with respect to 2019, as shown in Figure 19. However, an hint of recovery has been registered in March and April 2021, which present a percentage increase with respect to 2019 data of 3.22 % and 5.26 %, respectively. The increase of 2021 data in the same period compared to 2020 is very high, with 45.9 % in March and 92.2 % in April. This suggests, as for the non-metallic minerals sector, both the high impact of such crisis and the high resilience of such sector, which has returned to pre-pandemic values.

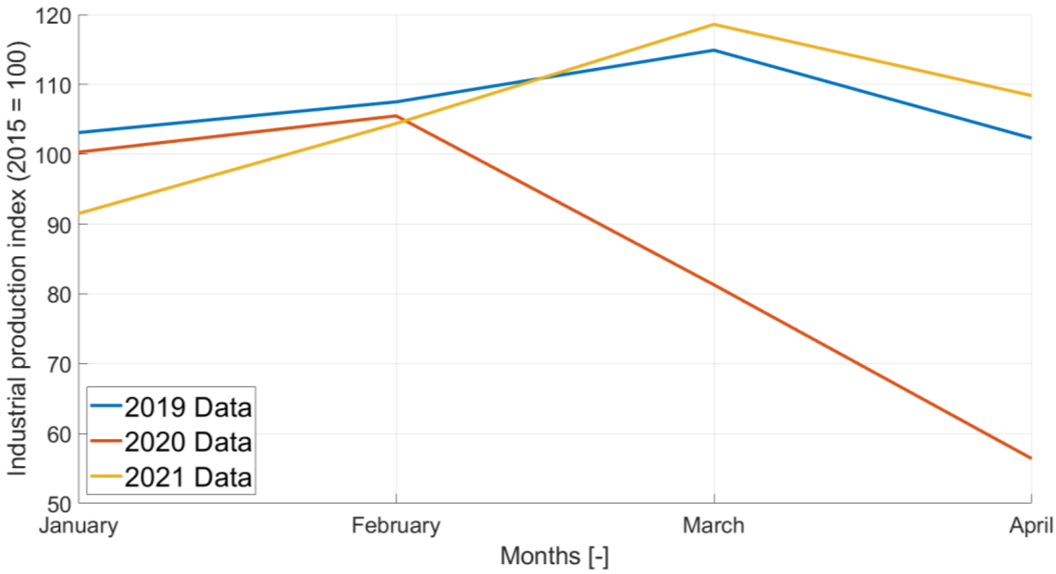

Figure 19: Comparison of the Other industries industrial production index for the first quarter of 2019, 2020 and 2021.

Chapter 5

Vector AutoRegressive Models

Vector AutoRegressive models are usually used to forecast multiple time series variables in a single system of equations, and to analyze the dynamic impact of the disturbance factors contained in the system variable. They were firstly developed in control theory, where linear dynamic systems were identified by vector-valued autoregressive moving average (VARMA) and state-space representations [49]. VAR models were successively introduced in economic analyses by Granger and Sims. The Granger causality concept was proposed by Granger in order to determine whether a time series x_t brings relevant information to forecast a variable y_t [50]. Sims proposed in [51] an alternative to large scale structural simultaneous econometric models (SSEM) [52] that treat some variables as exogenous by ad- hoc assumptions, not supported by solid theories. VAR models have been widely used since then, especially for testing Granger-causal relationships among macroeconomic variables, like government spending and taxes on economic output [53].

The variables in VAR models are treated as a-priori endogenous, and their success depends mostly on their simplicity and their forecast accuracy. They are also used in economic analysis to understand the interrelationship between variables, by means of tools like impulse response and variance decomposition [54].

A VAR model can be seen as a generalization of univariate autoregressive models. Its main assumption is that all the variables to forecast affect each other, and they are endogenous.

It is generally presented with the notation $VAR(d)$, where d represents the number of lagged variables considered for the regression. For simplicity let us consider a two variable VAR(1), i.e. with one lag:

$$\begin{aligned} y_{1,t} &= c_1 + \phi_{11,1} y_{1,t-1} + \phi_{12,1} y_{2,t-1} + \varepsilon_{1,t} \\ y_{2,t} &= c_2 + \phi_{21,1} y_{1,t-1} + \phi_{22,1} y_{2,t-1} + \varepsilon_{2,t} \end{aligned} \tag{2}$$

where $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise error terms. The lagged variables of $y_{1,t}$ and $y_{2,t}$ are represented respectively by $y_{1,t-1}$ and $y_{2,t-1}$. In general, the coefficients $\phi_{ii,m}$ and $\phi_{ij,m}$ describe respectively the relation between the *m-th* lag of variable y_i on itself, and the relation between the *m-th* lag of variable y_j on y_i [20]. Given its simplicity, the coefficients are estimated by a simple Ordinary Least Squares (OLS) regression.

Forecasts for each variable in the system are generated in a recursive manner. Assuming VAR(1) model described in Equation (2), the one-step-ahead forecasts can be written as:

$$\begin{aligned} \hat{y}_{1,\mathrm{T+1|T}} &= \hat{c}_1 + \hat{\phi}_{11,1}\hat{y}_{1,T} + \hat{\phi}_{12,1}\hat{y}_{2,T} \\ \hat{y}_{2,\mathrm{T+1|T}} &= \hat{c}_2 + \hat{\phi}_{21,1}\hat{y}_{1,T} + \hat{\phi}_{22,1}\hat{y}_{2,T} \end{aligned} \tag{3}$$

This is the same form as Equation (2) except that the errors have been set to zero and parameters have been replaced with their estimates. The process can be iterated for all future time periods by replacing the unknown values of y_1 and y_2 with their forecasts [20].

VAR models can be defined "as a directed network of interactions among the individual time series" [55], as showed in Equation 2. Generalizing such equation, a VAR model of order d, with the notation VAR(d), can be written for a p-dimensional process $Y_t = (Y_{t1}, \dots, Y_{tp})$ as:

$$\begin{gathered} Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_d Y_{t-d} + \varepsilon_t \\ Var(\varepsilon_t) = \Sigma_\varepsilon \end{gathered} \tag{4}$$

where $\Phi_1, \dots, \Phi_d$ are $p \times p$ matrices called transition matrices, which describe the temporal relationships among the variables. Σ_ε is the error covariance matrix, in which additional dependences among the various processes are present, while ε_t represents the white noise. As stated above, in forecasting problems the transition matrices have to be determined, and this is generally done with simple tools like an Ordinary Least Squares (OLS) regression.

5.1 Sparse Vector AutoRegressive (sVAR) models

The elements of the transition matrices $\Phi_1, \dots, \Phi_d$ in Equation 4 are often described as being part of a Granger-causal network, while the transition matrices are also called adjacency matrices. In macroeconomic analyses the Granger-causal networks are determined from a dataset consisting of a stationary time series process in the form of a vector $\{Y_1, \dots, Y_T\}$, with T generally very large.

One of the most important issues in VAR modeling is related to the identification of relevant variables in the regression. In fact, neglecting relevant variables brings to spurious correlations among the time series, meaning that a non-causal relationship has been determined, due to the absence of such relevant factor. This brings to incorrect estimations of the elements of the Granger-causal network, with consequent inaccuracies in the forecasts. Such problem has been highly discussed in literature. Christiano et al. [56], for example, states that a counterintuitive increase in inflation from the unexpected monetary tightening in the post-war US economy is caused by having neglected forward looking variables in the VAR model.

A high-dimensional VAR framework can help to include all the relevant parameters and relationship among them. Nonetheless, such approach may result in high standard errors of OLS estimates or even in the infeasibility of the regression [57], and such problem is strictly related to the dimensionality of the VAR. Vector autoregressive models are intrinsically high-dimensional, considering that the number of parameters grows quadratically. In fact, considering to fit a VAR of order d for p time series, the number of parameters is given by $d \cdot p^2$. Furthermore, the need of hundreds of variables is increasing in recent applications in different research areas, from macroeconomics to genetics. This brings to the need of making structural assumption to aim to low-dimensionality [55].

Another common approach to manage such dimensionality issue is to force some shrinkage on the coefficients of the transition matrices [58], like the Bayesian shrinkage, which is the most popular approach [59]. In the recent years the implementation of regularization methods in VAR models for time series forecasting has been analyzed, especially for methods like the lasso [60] and its variants. Such methods are defined as penalized least squares (PLS) optimization problems, and they can be efficiently solved with iterative optimization algorithms, such as coordinate descent [61] and generalized gradient descent [62]. The shrinkage of coefficients towards a zero value makes the matrix of regression coefficients in the VAR model sparse, for this reason we refer to sparse VAR (sVAR) models when dealing with such methods. In the paragraphs below, a detailed description of the lasso method and its variants is discussed, based on [57].

5.1.1 Traditional lasso method

The lasso (least absolute shrinkage and selection operator) method was firstly proposed in [60], while different structural variants can be seen in [63] and [64]. It corresponds to a penalized least squares (PLS) method which brings to zero some coefficients of the VAR model, and for this reason both estimation and variable selection are performed simultaneously.

Considering a $k \times 1$ vector $a = (a_1, \dots, a_k)'$, a = (a1, . . . , ak) 0 , the L_1 and L_2 norms of a are defined respectively as $\|a\|_1 = \sum_{j=1}^{k} |a_j|$ and $\|a\|_2 = \sqrt{\sum_{j=1}^{k} a_j^2}$. The lasso of a h-step-ahead direct forecast model of a generic time series variable $y_{i,t+h}$ is obtained by the problem:

$$\min_{\mu_i,\phi_{1,i},\dots,\phi_{d,i}} \sum_{t=d}^{T-h} \left\| y_{i,t+h} - \mu_i - \sum_{j=1}^{d} \phi'_{j,i} y_{t+1-j} \right\|_2^2 + \lambda \sum_{j=1}^{d} \left\| \phi_{j,i} \right\|_1 \qquad (5)$$

where μ_i is the intercept of the regression equation and λ the so called regularization parameter. The second term of Equation 5 is the penalty term, which shrinks some coefficients towards zero. The sparsity of the transition matrix depends on the value of the regularization parameter, the larger is λ, the sparser is the model. In high-dimensional macroeconomic analyses lasso has some limitations. In fact, in such analyses the number of predictors is generally higher than the number of observations, while lasso uniquely determines a number of predictors at most equal to the observations. Furthermore, macroeconomic variables are often highly correlated, and in this case lasso does not perform well in terms of forecasting [60].

5.1.2 Elastic net and group lasso

The elastic net proposed in [63] tries to solve the drawbacks of lasso through a combined implementation of the L_1 and L_2 penalties. The related minimization problem is:

$$\sum_{t=d}^{T-h} \left\| y_{i,t+h} - \mu_i - \sum_{j=1}^{d} \phi'_{j,i} y_{t+1-j} \right\|_2^2 + \lambda \sum_{j=1}^{d} \left((1-\alpha) \left\| \phi_{j,i} \right\|_1 + \alpha \left\| \phi_{j,i} \right\|_2^2 \right) \qquad (6)$$

where α is a tuning parameter that weights the two penalty functions, and can assume the value 0 or 1. Lasso corresponds to $\alpha = 0$, while $\alpha = 1$ leads to the so-called ridge regression. In [63] it is described how the elastic net is able to perform well in case of high correlations among the variables, and how such method can select more predictors than observations. This makes such method suited when dealing with macroeconomic data.

The last approach described here is the group lasso, firstly proposed in [64]. This method needs that groups of variables have to be defined in advance, in order to automatically include or exclude some sets of variables. Having to firstly specify group of variables makes group lasso more restrictive than the elastic net, but when dealing with macroeconomic variables this is rarely a problem, considering that such parameters have generally specific attributes that help to group them properly.

Let us have the time series variable y_t divided into L groups. Such variables can be then written as $y_t = \left(y_t^{1\prime}, \dots, y_t^{L\prime}\right)'$. The minimization problem for group lasso is then:

$$\sum_{t=d}^{T-h}\left\|y_{i,t+h}-\mu_i-\sum_{l=1}^{L}\sum_{j=1}^{d}\phi_{j,i}^{l}{}'y_{t+1-j}^{l}\right\|_2^2+\lambda\sum_{l=1}^{L}\sum_{j=1}^{d}\sqrt{d_l}\left\|\phi_{j,i}^{l}\right\|_2 \tag{7}$$

where d_l is the dimension of y_t^l and $\phi_{j,i}^l$ is the vector of regression coefficient of dimension $d_l \times 1$. In order to avoid favoring high dimension groups $\sqrt{d_l}$ is present in the penalty term. Group lasso has also the characteristic that all the coefficient are penalized by the L_2 norm, and this means that the whole group of variables is either dropped or not. It is also interesting to notice that grouping each variable in single groups means having $d_l = 1$, leading to the traditional lasso case described in Equation 5.

5.1.3 Selection of regularization and tuning parameters

An important step in lasso methods is to choose an appropriate value for λ and α. In case of series that are not time-dependent, this can be done by cross-validation techniques, such as K-fold cross-validation [65]. When dealing with time-dependent series, a similar procedure can be applied. The regularization parameters can be, in fact, selected by minimizing the Mean Squared Prediction Error (MSPE) of rolling window forecasts. The data sample is splitted in two periods, respectively called training ($t = 1, \dots, T_1$) and test period ($t = T_1 + 1, \dots, T$). Then, an iterative procedure is implemented, where h step-ahead forecasts of y_{i,T_1+j} based on the observation $(y_j, \dots, y_{T_1+j-h})$. The MSPE is finally calculated for several values of λ, and the value of regularization parameter that minimize such error is selected for the lasso method:

$$MSFE_{i,h}^{\lambda}=\frac{1}{T-T_1}\sum_{j=1}^{T-T_1}\left(y_{i,T_1+j}-\hat{y}_{i,T_1+j|T_1+j-h,\dots,j}\right)^2 \tag{8}$$

where $\hat{y}_{i,T_1+j|T_1+j-h,\dots,j}$ is the h step-ahead forecast. In a very similar way the values of α and λ are determined for elastic net and group lasso.

Chapter 6

Application and Validation of VAR Models

In this thesis, all the analyses and forecasts have been performed in the R computing environment, by means of the `HDeconometrics` package [66], capable of estimating the transition matrices of a VAR model through the lasso penalized least squares method.

To perform accurate projections of such datasets it is important to address not only the historical trends, but also exploiting the structural interdependencies among such sectors. This is the main reason to choose a multivariate regression model like VAR over univariate models like ARIMA or Exponential Smoothing, and such statement will be discussed by comparing the VAR results with projections performed by means of such univariate regression models.

The main problem of VAR models is the overparameterization derived from the number of variables to analyze and coefficients needed to describe all the relationships among the lagged values. In such cases, a sparse VAR modelling approach comes in hand [67]. In fact, it is reasonable that not every coefficient describing the relations among different lagged values introduces significant information to improve the regression accuracy, and for this reason a sparse VAR model (sVAR) can be implemented, where most of the autoregressive coefficients are set equal to zero. The methods to select which parameters can be set to zero are automatically implemented by the fitting functions in the R `HDeconometrics` package.

6.1 Industrial historical dataset

The analyzed data consists of the monthly industrial production index of six different Italian industrial sectors. As stated by the Board of Governors of the Federal Reserve System (US) [68], the industrial production index (IPI) is a monthly economic indicator measuring real output in the manufacturing, mining, electric, and gas industries, relative to a base year. It also measures capacity, estimating the production levels that could be sustainably maintained, and capacity utilization, the ratio between actual output and capacity. It is generally expressed as an index level relative to a base year, and this means that they do not express absolute production volumes, but the percentage change in production relative to the base year. Within the general IPI, many other sub-indices provide a detailed look at the output of specific industries' subsectors.

Figure 20 shows the historical data of the industrial production index of all the six Italian industrial sectors, as categorized in TIMES-Italia, during the period 1990-2020. It shows a clear annual seasonal pattern, and this cannot be neglected when constructing a VAR model.

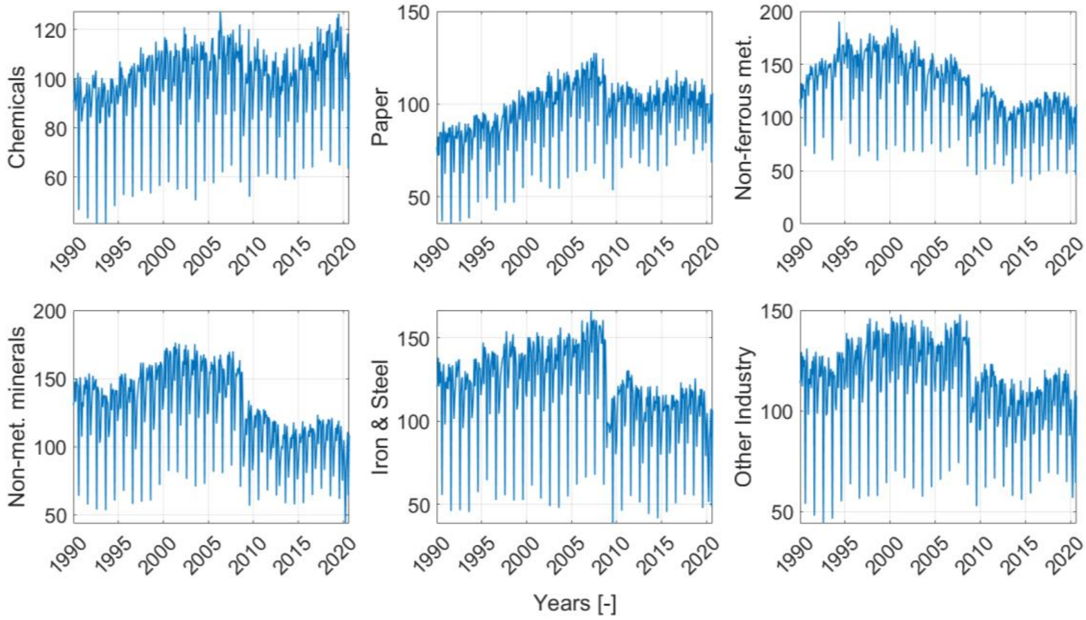

Figure 20: Historical series of the Italian monthly industrial production index for all the industrial subsectors.

This is done by implementing centered seasonal dummy variables [69], that are capable to address pre-specified seasonality patterns. Considering the case of this work, where the aim is to forecast monthly data, Table 9 shows the dummy variables needed, represented by $d_{i,t}$.

It is important to notice that only eleven dummy variables are needed to code twelve categories. That is because the eleventh category, December in this case, is captured by the intercept, and is specified when the dummy variables are all set to zero. Each coefficient related to the dummy variable can be interpreted as a measure of the effect of that category relative to the omitted category [20]. In such case, the coefficient of $d_{1,t}$ associated with January will measure the effect of January on the forecast variable compared to the effect of December.

Table 9: Seasonal dummy variables for monthly data forecasting.

	$d_{1,t}$	$d_{2,t}$	$d_{3,t}$	$d_{4,t}$	$d_{5,t}$	$d_{6,t}$	$d_{7,t}$	$d_{8,t}$	$d_{9,t}$	$d_{10,t}$	$d_{11,t}$
01/1990	1	0	0	0	0	0	0	0	0	0	0
02/1990	0	1	0	0	0	0	0	0	0	0	0
03/1990	0	0	1	0	0	0	0	0	0	0	0

04/1990	0	0	0	1	0	0	0	0	0	0	0
05/1990	0	0	0	0	1	0	0	0	0	0	0
06/1990	0	0	0	0	0	1	0	0	0	0	0
07/1990	0	0	0	0	0	0	1	0	0	0	0
08/1990	0	0	0	0	0	0	0	1	0	0	0
09/1990	0	0	0	0	0	0	0	0	1	0	0
10/1990	0	0	0	0	0	0	0	0	0	1	0
11/1990	0	0	0	0	0	0	0	0	0	0	1
12/1990	0	0	0	0	0	0	0	0	0	0	0
01/1991	1	0	0	0	0	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…	…

It has to be noticed that, without a seasonal adjustment, it would be difficult to process any kind of trend that the series can have, considering the high fluctuations of the data.

6.2 Cross-validation procedures for VAR model construction

An important step for the VAR model construction is selecting the best number of lagged variables in the system. Information criteria such as Akaike, Bayesian and Hannan-Quinn criteria (AIC, BIC and HQC respectively) are generally used to select the number of lags to be included [20], but since we are only interested in forecasting, a cross-validation approach has been utilized. Cross-validation is widely employed in regression and prediction of time series to assess the accuracy of a forecast model by averaging predictive errors across mutually exclusive data subsamples. Estimates can then be used to select the most accurate model among multiple candidates [70].

More in detail, a time-series cross-validation procedure based on a rolling forecast origin has been performed [20]. This is an iterative process, where the first iteration has consisted in considering a two-years training set in the period 1990-1991. The test set comprises a one-step ahead forecast (i.e., the projection at January 1992), and from this a forecast error can be estimated comparing the result with the historical value. In the second iteration, the training set is increased by one month ahead, and the test set consists of the forecast at February 1992. This procedure is iterated until the training set corresponds to the last month in the historical dataset. Figure 21 shows the first six steps of the procedure, considering the historical values of the chemicals industrial production index as an example.

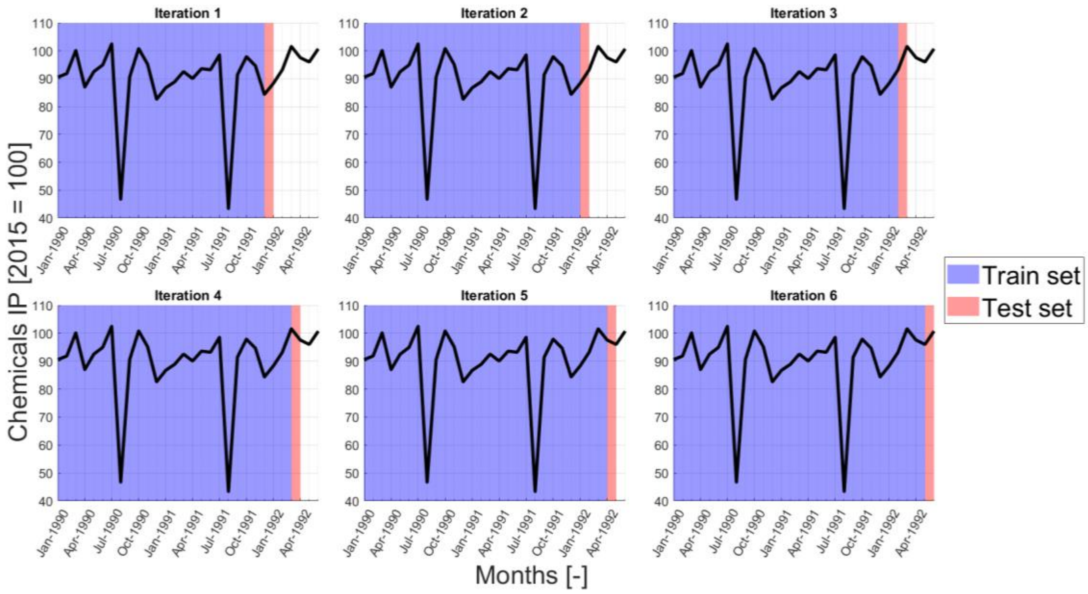

Figure 21: Cross-Validation procedure for model selection.

The final output of this procedure is a vector containing the various forecast errors computed, and a Root Mean Square Error (RMSE) can be determined to have a quantitative evaluation of the overall forecast accuracy of the selected model. In fact, such cross-validation has been performed for various VAR models with different number of lags, and it can be seen that the most accurate model with the lowest RMSE is the VAR(12) model, that is with 12 monthly lagged values, see Figure 22. This can be expected given the strong annual seasonality of the time series analyzed.

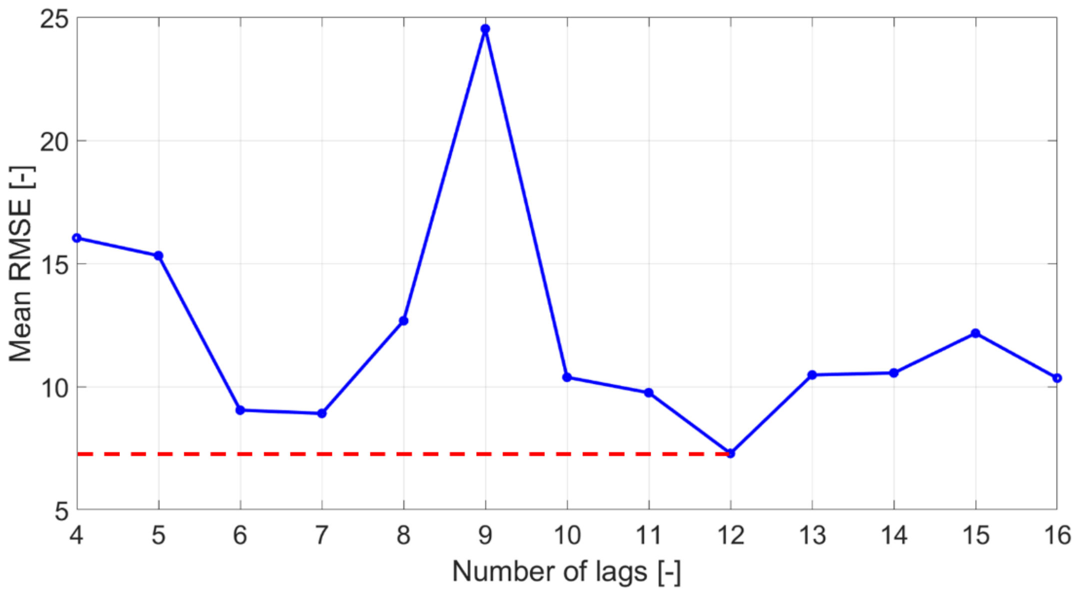

Figure 22: Selection of the number of lags from RMSE minimization.

6.3 VAR model projections results

The validation of the forecast model described above has been made by considering the historical series up to 2017 and performing forecasts starting from 2018 up to 2040, comparing the results with the projections considered in the Italian Integrated National Energy And Climate Plan (PNIEC) [28], which in turn are projections taken from the EU Reference Scenario 2016 [29]. The results of the validation are reported in Figure 23 - Figure 28, showing the post-2020 projections of the VAR model, along with their 95% confidence bounds.

Figure 23 shows the projections of Iron and steel production, and as it can be seen how the pre-pandemic projections seem to well follow the PNIEC projections, especially in the mid- and long-term. Even though major differences arise for the first years of the projections, VAR projections are more in line with the decrease of 2019 registered in the historical data, compared to PNIEC.

Post-pandemic VAR projections present a fast recovery from the crisis, as expected from the analysis on the historical data performed in Paragraph 4. Furthermore, the long-term trend presents many similarities with the PNIEC projections, with only a slighter lower steepness. Note that, despite the growing rate, the production remains significantly smaller than that it was in 2007 (peak level), but Post-pandemic projections are able to reach pre-pandemic (2019) historical levels already starting from 2022. Eventually, the Iron and steel sector shows in 2040 just a - 3.5 % deviation with respect to PNIEC and Pre-pandemic projections.

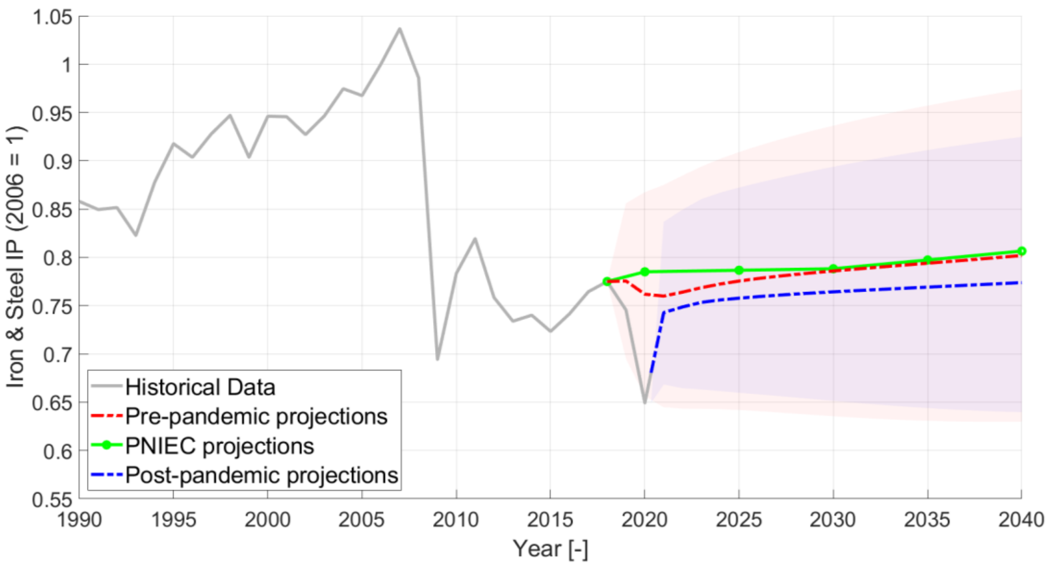

Figure 23: Iron and steel industrial production projections. The light red area encloses the 95 % confidence bounds related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related to Post-pandemic projections

Figure 24 shows the projections of Non-ferrous metals production. In this case the similarities with PNIEC projections can be noticed mostly in the long-term. Nonetheless, similarly to iron and steel, the pre-pandemic VAR projections better follow the decreasing trend of the latest years, further dejected by the 2020 economic crisis.

Post-pandemic projections present a strong resilience, presenting again a shock response similar to the one had after the Great Recession. Even in this case, Post-pandemic projections are able to reach pre-pandemic historical levels starting from 2022, as expected in Paragraph 4.2, but 2000 peak production levels are never reached again. In the long-term there is still an increasing trend, even if slower than that of Pre-pandemic projections. Non-ferrous metals production in 2040 is computed to be just 6.5 % lower than in PNIEC and Pre-pandemic projections.

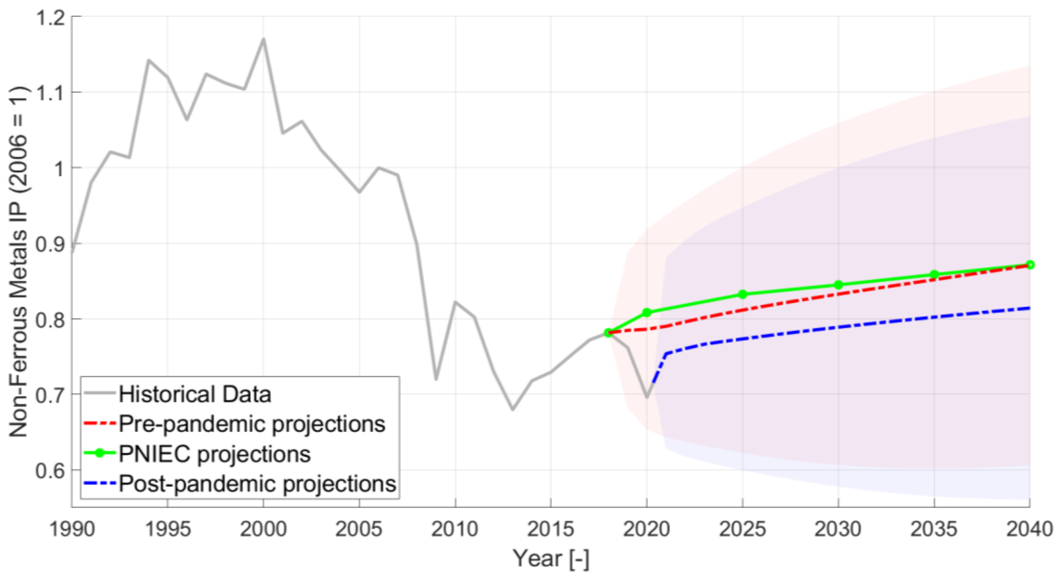

Figure 24: Non-ferrous metals industrial production projections. The light red area encloses the 95 % confidence bound related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related to Post-pandemic projections

Figure 25 shows the projections for the chemical sector. Such sector presents a stronger increasing trend with respect to the previously shown Iron and steel and Non-ferrous metals, but the pre-pandemic VAR projections present again a similar behavior to the one discussed above. Pre-pandemic projections present a first stagnation phase up to 2020 which better follows the historical data with respect to PNIEC, while in the long term the two projections tend to overly.

Post-pandemic projections show a small bouncing effect on the short-term, in line with the post-2009 crisis recovery behavior, and no stagnation in the medium term. However, the pandemics has a very time-limited effect on this set of forecasts, and the growth highlighted by the historical series is confirmed in projections using both Pre-and Post-pandemic drivers. The Chemicals sector is the one showing the highest growth rate and the highest absolute growth with respect to the 2006 value, despite the pandemics, but also a wide range between PNIEC/Pre-pandemic and Post-pandemic projections, with a 6.25 % difference.

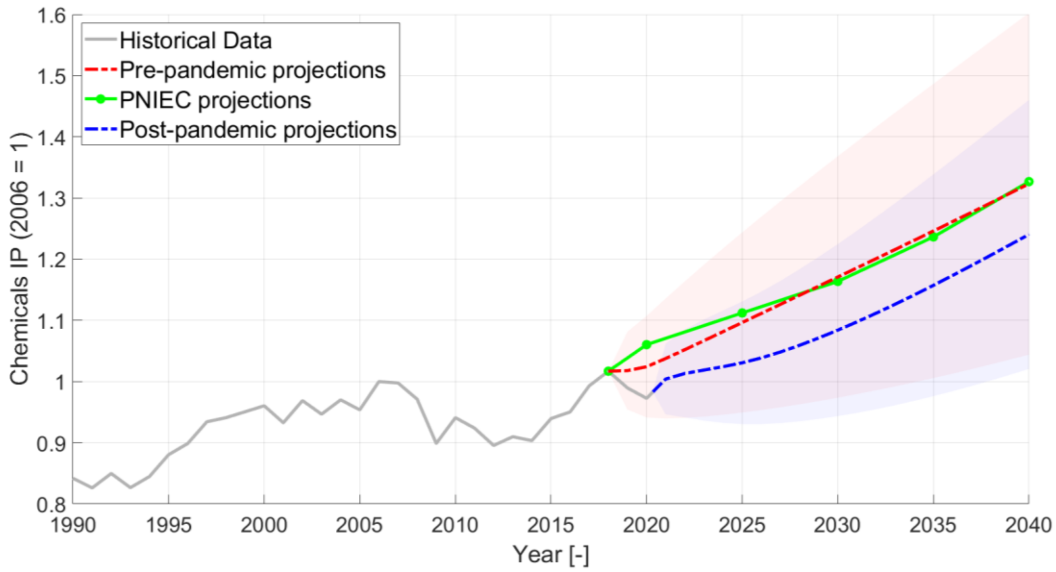

Figure 25: Chemicals industrial production projections. The light red area encloses the 95 % confidence bound related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related to Post-pandemic projections.

The projections of the non-metallic minerals sector production are reported in Figure 26. Similarly to the other sectors discussed above, the short-term bounce is more in line with the decreasing trend of the historical data in 2019, compared to PNIEC results. Such bounce is then followed by a strong increasing trend, being able to reach in 2040 comparable levels with respect to 2006. Both in the mid- and long-term, Pre-pandemic projections are mostly overlying the PNIEC curve.

Post-pandemic projections also present a strong increasing trend, with a full crisis recovery reported in 2023, as expected from the historical analysis described in Paragraph 4.3. In 2040, such trend in the Post-pandemic projections leads to a 4.55 % difference in 2040 compared to PNIEC/Pre-pandemic.

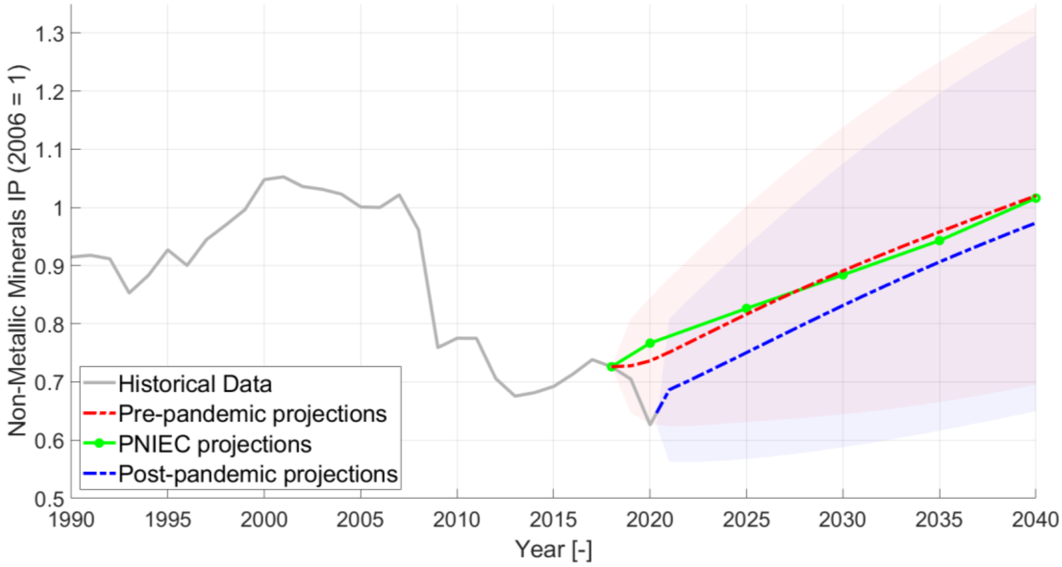

Figure 26: Non-metallic minerals industrial production projections. The light red area encloses the 95 % confidence bound related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related to Post-pandemic projections.

The results for the Pulp, paper and printing sector are presented in Figure 8, and in this case of the Pre-pandemic trend, VAR projections are in line with PNIEC even in the short-term. The strong increasing trend leads to 2007 values, which present the peak in the analyzed period, already in 2026.

Post-pandemic projections tend to follow in 2021 the increasing bounce started in the first months of the year, as described in Paragraph 4.5, reaching 2018 values. In the long-term the trend is increasing, but with a slower pace with respect to Pre-Pandemic/PNIEC projections, and the 2007 values are reached in 2038. The reason for the lower increase in the projection of Pulp and paper production can be investigated by giving a look at its historical series, and considering that the recovery from the 2009 crisis was very slow. The results of the projections lead to a 10.5 % difference between PNIEC/Pre-pandemic and Post-pandemic projections in 2040.

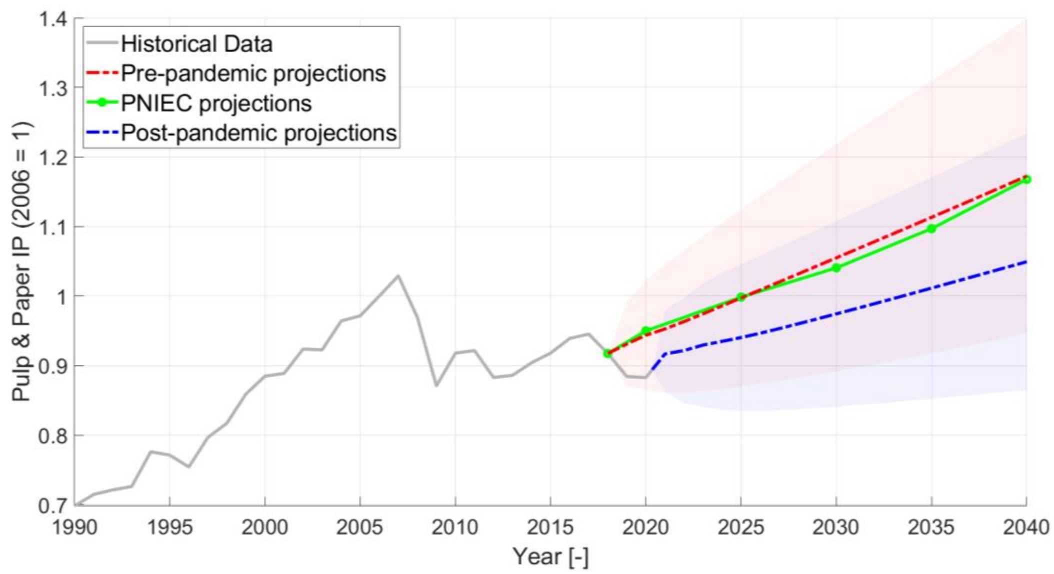

Figure 27: Pulp, paper and printing industrial production projections. The light red area encloses the 95 % confidence bound related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related toPost-pandemic projections.

The remainder of the industrial production is grouped under "Other industries", even though it represents an important share of the total Italian industrial production. Results for Other industries production trend is shown in Figure 28. The Pre-pandemic VAR future projections do not differ much from the results of Non-metallic minerals in Figure 26, with an initial bounce that follows the historical values of 2019. The main difference from the Non-metallic minerals can be seen in the long-term, where a flattening trend starts to happen in 2026, contrary to PNIEC. Nevertheless, the Pre-pandemic projections acceptably follow the PNIEC trend in the period of interest for the analysis.

The Post-pandemic projections, on the other hand, present a good response from the pandemic crisis in 2021, in line with what it has been showed in Paragraph 4.6. It follows a stagnation up to 2026, which is also the year where a full recovery from the crisis is reported. In the long-term, the steepness of the projections' curve presents higher values than that of Pre-pandemic projections, more in line with PNIEC results. The difference between PNIEC/Pre-pandemic and Post-pandemic projections in 2040 for such sector is of 3.65 %, and, differently from Pulp and paper production, 2007 peak production levels are never reached again.

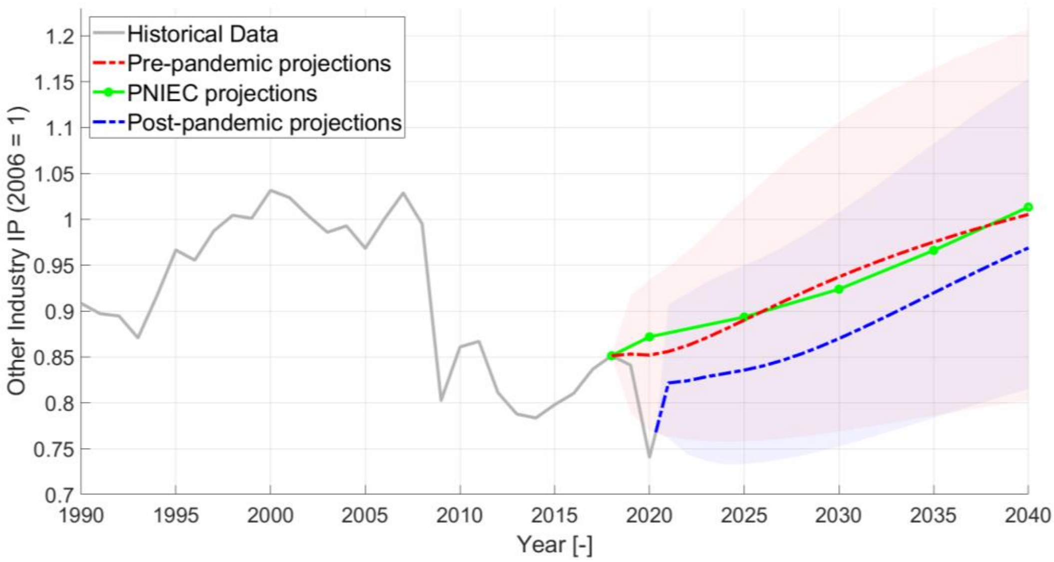

Figure 28: Other Industries industrial production final projections. The light red area encloses the 95 % confidence bound related to Pre-pandemic projections, while the light blue area represents the range enclosed within the 95 % confidence bounds related to Post-pandemic projections

In Table 10, the plotted behavior of the pre-pandemic VAR projections is translated in terms of average annual growth rates and compared to the annual growth rates of the added values of the industrial sectors in the baseline and PNIEC scenarios, taken as reference for the model validation.

Table 10: Average annual growth rates of VAR projections and of the added value of the industrial sectors in the PNIEC scenarios (in brackets).

	2018-2020 [%]	**2020-2025 [%]**	**2025-2030 [%]**	**2030-2035 [%]**	**2035-2040 [%]**
Iron and steel	-0.57 (0.43)	0.36 (0.04)	0.27 (0.04)	0.20 (0.23)	0.20 (0.23)
Non-ferrous metals	0.20 (1.1)	0.64 (0.59)	0.52 (0.30)	0.46 (0.32)	0.44 (0.30)
Chemicals	0.24 (1.4)	1.4 (0.96)	1.3 (0.91)	1.3 (1.2)	1.2 (1.4)
Non-metallic minerals	0.46 (1.8)	2.1 (1.5)	1.8 (1.4)	1.5 (1.3)	1.3 (1.5)
Pulp, paper and printing	0.94 (1.2)	1.1 (1.0)	1.1 (0.83)	1.1 (1.1)	1.0 (1.3)
Other industries	0.61 (0.80)	0.88 (0.49)	1.0 (0.67)	0.80 (0.90)	0.61 (0.96)

As it can be seen, such growth rate values present various differences for the VAR model results and the PNIEC projections, especially in the short-term. Note, however, that the 2019 projections of the VAR model seem to better follow the historical decreasing trend, and this good short-term accuracy is important to have reliable forecasts of the post-pandemic recovery. Table 11 shows the absolute values of such projections, in terms of average annual industrial production index, expressed as industrial output normalized as a base value of 100 at the year 2005. The percentage deviations between the values are also showed, and it can be seen how in the short-term, every VAR projection confirms slightly lower results than the PNIEC values, but better following the decreasing historical values, as seen in the previous figures. In the long-term, all the computed results tend to get closer to the reference projections, by maintaining a slight underestimation for the whole period as for iron and steel and non-ferrous metals, or by also having overestimations in some years. Nonetheless, the difference in the two results is low, as already qualitatively visible from the figures above.

Table 11: Average annual industrial production projections according to the VAR model adopted in this paper and PNIEC projections (base 2015 = 100), along with the percentage deviation of the VAR value with respect to PNIEC estimates.

		2020	**2025**	**2030**	**2035**	**2040**
Iron and steel	VAR	105.3	107.2	108.6	109.8	110.9
	PNIEC	108.5	108.7	109.0	110.2	111.5
	Percentage deviation [%]	3.04	1.38	0.368 %	0.364	0.541
Non-ferrous metals	VAR	107.8	111.3	114.1	116.8	119.4
	PNIEC	110.8	114.1	115.8	117.7	119.5
	Percentage deviation [%]	2.81	2.56	1.47	0.790	0.0922
Chemicals	VAR	109.0	116.7	124.6	132.6	140.8
	PNIEC	112.8	118.4	123.9	131.6	141.2
	Percentage deviation [%]	3.52	1.43	- 0.613	- 0.774	0.268
Non-metallic minerals	VAR	106.4	117.9	128.7	138.4	147.3
	PNIEC	110.8	119.4	127.7	136.2	146.8
	Percentage deviation [%]	4.14	1.22	-0.782	- 1.55	- 0.399
Pulp, paper and printing	VAR	102.8	108.6	114.9	121.3	127.7
	PNIEC	103.5	108.8	113.3	119.5	127.2
	Percentage deviation [%]	0.673	0.115	- 1.37	- 1.46	- 0.398
Other industries	VAR	106.8	111.6	117.5	122.2	126.0
	PNIEC	109.3	112.0	115.8	121.1	127.0
	Percentage deviation [%]	2.34	0.374	-1.44	-0.939	0.798

Table 12 represents the industrial production VAR projections with and without the 2020 pandemic effects. Overall, all the industrial sectors seem to be affected by the crisis, but with different amplitudes, both in the short- and long-term. In fact, sectors like Chemicals and Pulp and paper present a better response at year 2020 to the pandemic crisis, differently from the other sectors. The Iron and steel and

Other Industries industrial sectors seem to get closer to the pre-pandemic results in the long-term, while the rest tend to follow the same trend of the pre-pandemic projections, maintaining constant the shift in the values. An exception is the Pulp, paper and printing sector, that presents in the long-run a larger difference in the two projections.

Table 12: Industrial production pre- and post-pandemic VAR projections. Average annual values of the monthly industrial production index (base 2015 = 100).

	2020		**2025**		**2030**		**2035**		**2040**	
	Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
Iron and steel	105.3	89.76	107.2	104.8	108.6	105.7	109.8	106.3	110.9	107.0
Non-ferrous metals	107.8	95.41	111.3	106.0	114.1	108.1	116.8	110.0	119.4	111.6
Chemicals	109.0	103.5	116.7	109.7	124.6	115.4	132.6	123.2	140.8	132.0
Non-metallic minerals	106.4	90.46	117.9	108.4	128.7	120.1	138.4	130.9	147.3	140.6
Pulp, paper and printing	102.8	96.16	108.6	102.4	114.9	106.1	121.3	110.2	127.7	114.3
Other industries	106.8	92.83	111.6	104.7	117.5	109.0	122.2	115.3	126.0	121.4

6.4 Univariate regression results comparisons

It has been said how demand drivers (population, GDP, number of households, etc.) in the TIMES framework are obtained externally and given as inputs. For example, applied general equilibrium models like GEM-E3 can be used to derive such drivers trajectories [10]. One of the main advantages of using such general equilibrium models is having an internal coherence of the socio-economic drivers' projections given by analyzing the economy as a whole, without neglecting any type of interdependence among its various actors. The limitations of such models have to be found in the high amount of data requirements and human capital investment required.

Projections like the ones discussed in this work can be non-trivial to perform when dealing with very specific data such as the various industrial subsectors, and at this it is added the issue of a scarce literature. From such difficulties it has arisen the need of having a simple model that is capable of perform accurate forecasts, starting from small datasets and not biased from ad-hoc assumptions based on some aprioristic expectations.

Time series regression models represent the most reasonable choice, given their popularity for econometric projections due to their implementation simplicity. Exponential smoothing (ES) and ARIMA models represent the two most widely used models to time series forecasting, providing complementary approaches to the problem. In fact, while ES models analyze the trend and seasonality in the data to forecast, ARIMA models focus on the autocorrelations that can characterized the data [71].

Nevertheless, when trying to perform projections of multiple time series, such models cannot address any structural interdependencies that can arise among the data. Multivariate regression models like VAR are the most natural tool to consider such correlations among the different time series to forecast, as already stated, and for this reason they are capable to maintain, at least in part, such internal coherence discussed for the general equilibrium models.

In the following paragraph, a small description of the ARIMA and ES models it is carried on. Successively, the Pre- and Post-pandemic projections performed by means of such models are showed and compared to the PNIEC projections. Highlight the criticalities of such models results is useful to understand the need of having chosen for this analysis a more complex model like VAR[3] over the most common time series forecasting models.

6.4.1 ARIMA and Exponential Smoothing general description

A quick description of ARIMA and Exponential Smoothing models is presented in this section, based on [20].

ARIMA models are used to describe the pattern of time series data, generally for forecasting purposes. They may include autoregressive terms (AR), moving average terms (MA), and differencing operations (I). They differ from the exponential smoothing models because they look for the autocorrelations in the data.

An extension of this family of models, is the so-called SARIMA, used to model time series with seasonal components, i.e., a regular pattern of changes in the data that repeats over a precise time period.

Their general form is written as follows:

$$ARIMA\ (p, d, q)(P, D, Q)_S$$

[3] The complexity of VAR models is intrinsically given by the fact that VAR models can be seen as a generalization of univariate regression models [71], as it will also be clear in the following paragraphs.

where the seasonal part of the model is written using the upper notation. The S represents the seasonality.

The non-seasonal and seasonal autoregressive terms are respectively represented by p and P, while q and Q represent the moving average terms. The terms d and D indicate respectively the presence of a non-seasonal and seasonal differencing operation, in order to eliminate any kind of nonstationary behavior.

Exponential Smoothing is used to forecast univariate time series data, where the prediction is a weighted sum of past observation, with the weights decaying exponentially as the observations get older. This is an alternative of the ARIMA methods, and they are often used together for benchmarking and comparisons of the results.

This is a big family of methods, and like the ARIMA models they can be classified depending on the trend and seasonal components.

6.4.2 ARIMA and ES models construction

All the analysis and forecasts have been performed in R. A first step has been defining the best ARIMA and ES models that minimize the AIC through the R built-in functions *auto,arima()* and *ets()*. Because the time series is relatively long, a training and a test set have been used to compare the models, rather than time series cross-validation [6]. This brings to much faster results, without losing too much accuracy for not having considered all the data at our disposal. A training set for the model selection from January 1990 to December 2013, and a test set from January 2014 to December 2019 to compare the forecasting accuracy of the models have been defined. The year 2020 has been excluded in this part of the analysis on purpose, because the outliers from the pandemic crisis could bias this selection phase. In fact, as much as we are interested in modeling the crisis recovery, the work also aims at a long-run forecast, where the global trend is strongly affected by outliers. As the results will show, the presence of the 2009 recession (similar in some ways to the present crisis) in the training dataset seems to bring to reasonable results in terms of post-pandemic recovery, having similarities with the post-2009 trend.

The nowcasts from 2014 to 2019 performed by the various models have been compared to the real data. The final results consist of a weighted average forecast combination [7], where the weights are higher for the models that have shown lower MAPE in the nowcasts. It is, in fact, well known how forecast combinations often lead to increased forecast accuracy [8][9].

As last step, all the historical and forecasted data have been annualized through averaging the monthly values.

Here follows a practical example describing all the various steps of the analysis for one of the industrial subsectors (pulp, paper and printing), while all the results are shown below.

The annualized historical data well shows the overall trend of the series, as well as the subsector's behavior to the 2009 crisis. This is an important information for critically assess the forecast results. For example, in Figure 2 it can be seen how for the pulp and paper sector the general increasing trend has been stopped by the 2009 recession, and the further recovery has been very small. From this, it can be expected a very small recovery from the today crisis.

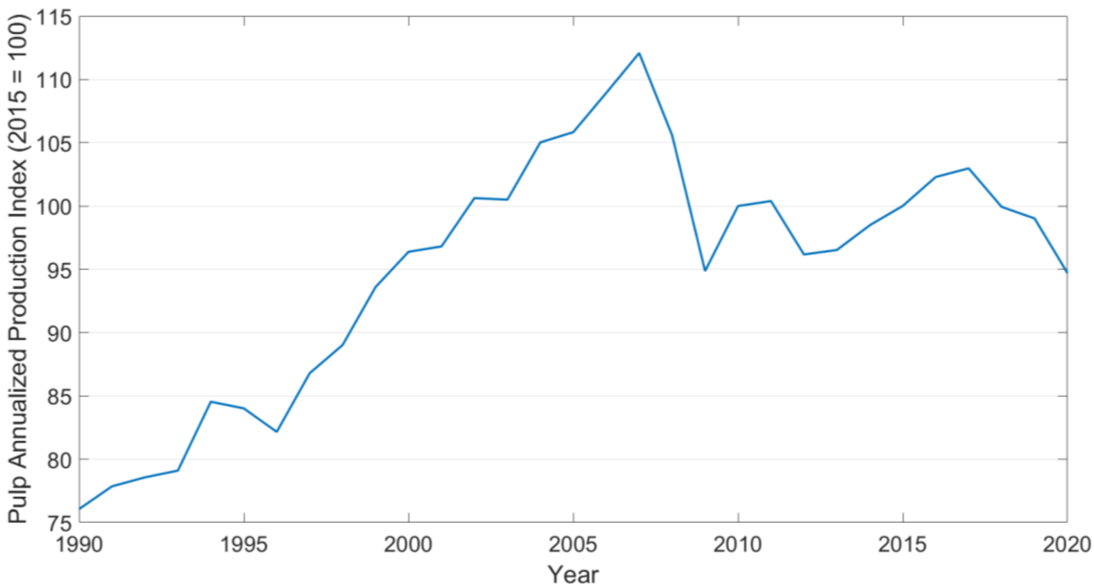

Figure 29: Pulp, Paper and Printing annualized production index (2015 = 100)

The R function *auto.arima()* auto selects the ARIMA model with the lowest AIC. From the Table 13 it can be seen that the $\mathbf{ARIMA(1,1,3)(0,1,1)_{12}}$ has the lowest AIC.

Table 13: Performances of different ARIMA models considering non-seasonal differencing of the time-series.

Model	AIC	Model	AIC
$\mathbf{ARIMA(0,1,2)(0,1,1)_{12}}$	1571.22	$\mathbf{ARIMA(2,1,0)(0,1,1)_{12}}$	1569.66
$\mathbf{ARIMA(0,1,2)(2,1,0)_{12}}$	1577.73	$\mathbf{ARIMA(2,1,0)(2,1,0)_{12}}$	1566.34
$\mathbf{ARIMA(0,1,3)(0,1,1)_{12}}$	1563.50	$\mathbf{ARIMA(2,1,1)(0,1,1)_{12}}$	1564.91
$\mathbf{ARIMA(0,1,3)(2,1,0)_{12}}$	1569.06	$\mathbf{ARIMA(2,1,1)(2,1,0)_{12}}$	1566.48
$\mathbf{ARIMA(0,1,4)(0,1,1)_{12}}$	1563.29	$\mathbf{ARIMA(2,1,2)(0,1,1)_{12}}$	1566.74
$\mathbf{ARIMA(1,1,2)(0,1,1)_{12}}$	1567.39	$\mathbf{ARIMA(3,1,0)(0,1,1)_{12}}$	1567.24
$\mathbf{ARIMA(1,1,2)(2,1,0)_{12}}$	1572.12	$\mathbf{ARIMA(3,1,0)(2,1,0)_{12}}$	1566.93
$\mathbf{ARIMA(1,1,3)(0,1,1)_{12}}$	1562.23	$\mathbf{ARIMA(3,1,1)(0,1,1)_{12}}$	1566.95

Despite this first analysis, forecasting with this model brings to completely inaccurate results. Figure 30 well shows such inaccuracy, considering the diverging confidence bounds and the strong decrease in the projections, contrarily from what it is expected given the historical trend.

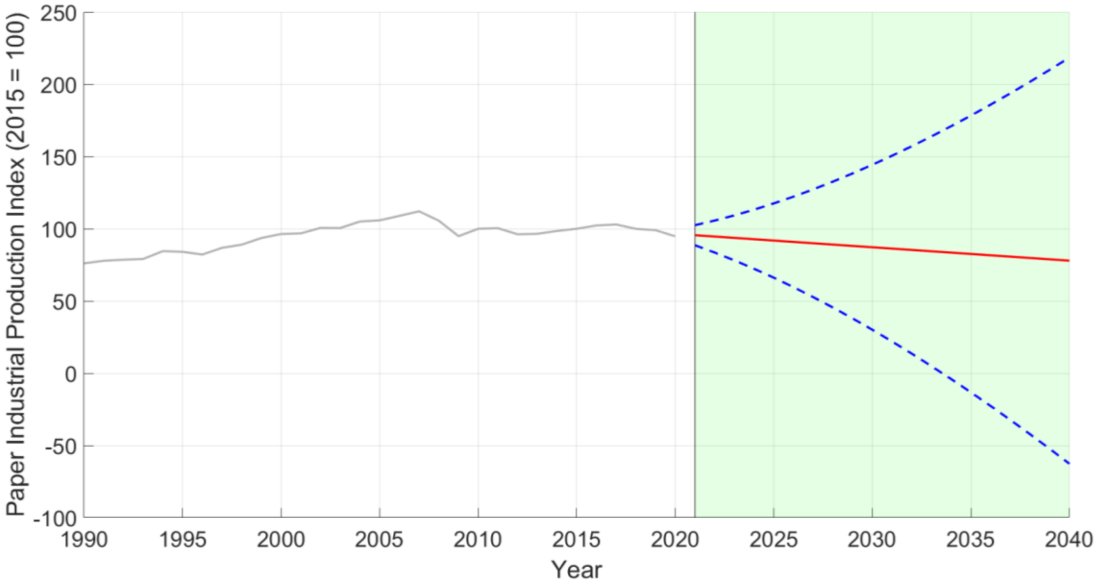

Figure 30: Forecast results for ARIMA(1,1,3) (0,1,1)$_{12}$

Even choosing other models in the table do not improve the accuracy. For this reason, the model selection has been performed again by forcing the function *auto.arima()* to work without non-seasonal differencing. The final models have been chosen for comparison (2 ARIMA and 1 ES):

- $\mathbf{ARIMA(3,0,0)(0,1,2)_{12}}$: AIC = 1535.86
- $\mathbf{ARIMA(3,0,0)(2,1,0)_{12}}$ **with drift**: AIC = 1564.82
- $\mathbf{ETS(A, A_d, A)}$: AIC = 2461.5

Table 14 summarizes the MAPE determined from the nowcasts of each model and the real data.

Table 14: Final selection of the univariate regression models for performing the forecast combination

Model	MAPE [%]
$\mathbf{ARIMA(3,0,0)(0,1,2)_{12}}$	4.282
$\mathbf{ARIMA(3,0,0)(2,1,0)_{12}}$ **with drift**	4.976
$\mathbf{ETS(A, A_d, A)}$	4.292

At this point, the three models can be combined to obtain the final results. As stated above, the forecast combination consists of a weighted average of the forecasts with respect to their MAPE value in the training. In this case, the ARIMA model without drift has the highest weight, while the ARIMA model with drift has the lowest.

The results of the projections for the Pulp and paper and Non-Metallic minerals subsectors are presented in Figure 31 – Figure 32. It is showed how the Pre-Pandemic and Post-pandemic projections present very similar long-term linear trends, as if they were shifted horizontally. The two Pre-pandemic projections highly differ from the PNIEC projections, which present a higher general increase up to 2040. Compared with the VAR projections in Figure 26 and Figure 27, a higher inaccuracy is also present in terms of amplitude of the confidence bounds, ranging in 2040 in an interval much higher than the one of the entire historical dataset.

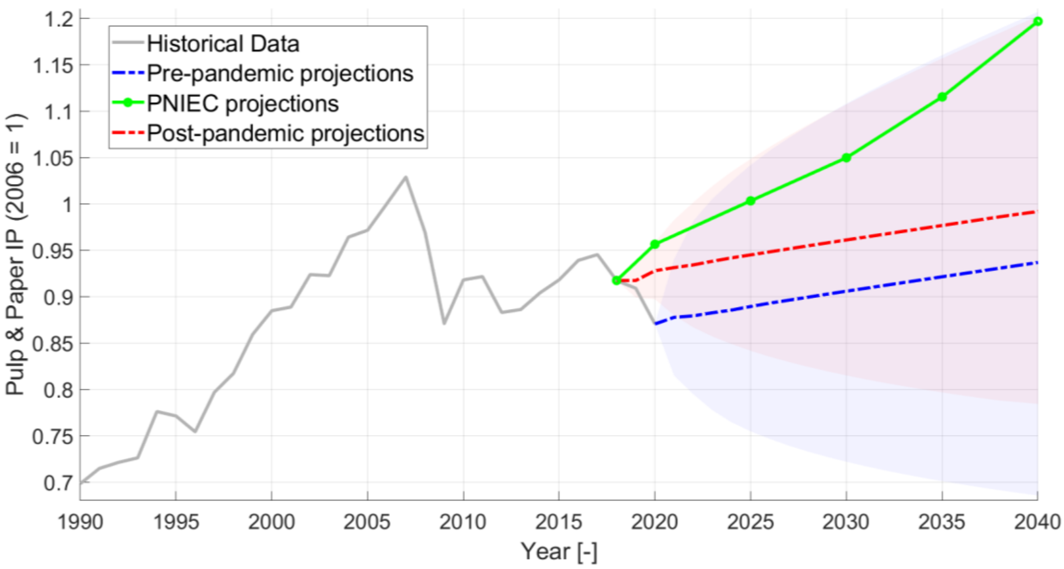

Figure 31: ARIMA and ES forecast combination of the Pulp and paper sector. The red and blue areas represent the 95 % confidence bounds of respectively the Pre-pandemic and Post-pandemic projections.

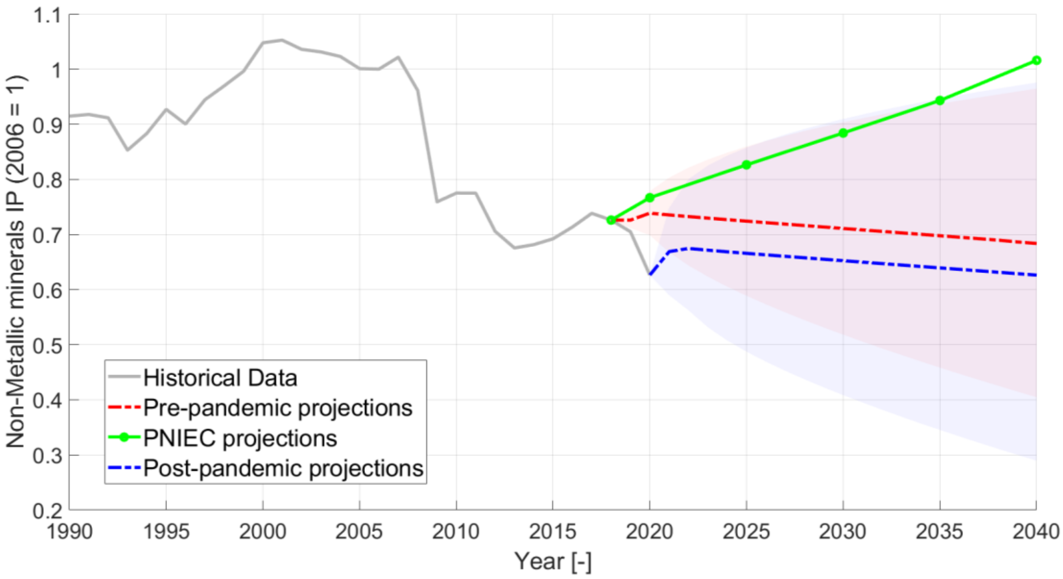

Figure 32: ARIMA and ES forecast combination of the Non-metallic minerals sector. The red and blue areas represent the 95 % confidence bounds of respectively the Pre-pandemic and Post-pandemic projections.

Furthermore, the decreasing trend in the projections of Non-metallic minerals is a perfect example of the limitations that univariate regression models like ARIMA and ES present in this type of analyses. In fact, if on one hand the projections for such sector seem to perform well in following the general historical trend, the final results go in a completely opposite direction with respect to PNIEC. This is mostly due to the fact that such models do not work with any information on the linkages that can be present among the various industrial sectors, differently from multivariate regression models like VAR. In fact, in the VAR results showed in Figure 26, the increase of the Non-metallic minerals industrial production projections derives from having considered such relationships among the different sectors.

For this reasons, the advantages of having chosen a multivariate regression model to perform such analysis are clear. In fact, VAR models have proven to be able to partially recreate one of the main strengths of general equilibrium models, that is the internal coherence among the various sectors that have to be analyzed[4]. The only drawback can be related to the overall accuracy in terms of the confidence bounds of the projections, which remain relatively high. However, the amplitude of such intervals can be exploited when performing the energy consumption projections by means of a sensitivity analysis. Such analysis have the advantage of defining different future pathways without the need of defining any constraint, differently from what it is generally done in a scenario analysis approach.

[4] Such internal coherence can be defined as partial being that it is based only on mathematical relationships (i.e., the correlations among the time-series), and not on macroeconomic assumptions.

Chapter 7

Energy consumption projections

Once industrial production projections are computed via VAR models and validated against PNIEC projections, their application to an energy system modeling framework is presented here, using the TIMES-Italia model deeply discussed in Chapter 1 and focusing on the outcomes for the Italian industrial sector.

Both Pre-pandemic and Post-pandemic industrial production projections are used to act as drivers in Equation 1 for the calculation of energy service demand throughout a time scale from 2006 to 2040. The time horizon is divided into a user-chosen number of time-periods, each period containing a (possibly different) number of years, which are then identified according to a single milestone year per period [11]. The current version of TIMES-Italia is characterized by time intervals with increasing amplitude: the first two periods represent single years, then the duration of each time step increases to two years until 2022, to five years between 2025 and 2030, and to ten years between 2030 and 2040. Therefore, the identified milestone years are presented in Table 15.

Table 15: Milestone years considered in the TIMES-Italia

Milestone years →	**2006**	2007	2008	2010	2012	2014
2016	2018	2020	2022	2025	2030	**2040**

For this reason, even though industrial production drivers were calculated in this paper for each single year after 2018 (or 2020 in the case of Post-pandemic projections), the remaining drivers calculated through GEM-E3 follow the abovementioned time-step resolution. As TIMES interpolates the user-defined timeseries data (among which drivers) only for the milestone years, and then uses the value at the milestone year as a representative value for the whole period [9], the computed industrial production rates are averaged to representative values for each milestone years in order to be used in TIMES-Italia.

Since energy service demands for the industrial sector identified in TIMES-Italia exactly match physical industrial production, elasticities from Equation 1 are set to a value equal to 1, making the industrial production growth both driver and service demand growth rate.

Those demand drivers are then coupled to the highly technologically detailed Reference Energy System encompassed in TIMES-Italia and described in Chapter 2 to calculate the exact match of energy supply and demand needed to satisfy the given energy service demand.

Results from a business as usual (BAU) scenario are presented here, without considering the adoption of any particular policy measure neither to contrast the increase of CO_2 emission nor any limitation on the use of specific energy carriers e.g., coal and gas, just to show the effects of the modified drivers on the energy consumption pattern when applied to TIMES-Italia.

In Paragraph 7.1 it will be described a backcasting analysis concerning the TIMES-Italia results compared to the Eurostat balances [30] has been performed, in order to validate such results with the real data. In fact, as already stated in Chapter 2, one of the most important advantages of starting the TIMES model optimization in 2006 is the possibility of performing a validation procedure based on the comparison of the first year results with historical data.

The first, and most expected outcome from Figure 10, which presents the total energy consumption patterns in the 6 Italian industrial subsectors modeled in TIMES-Italia and the whole industry sector, is that lower demand corresponds to lower energy consumption, as in the case of projections obtained using Post-pandemic industrial productions.

7.1 Energy final consumption backcasting analysis

Concerning the benchmark against Eurostat historical series for the period 2006-2019, the Iron and steel is one of the most critical subsectors in terms of inconsistencies between the TIMES-Italia results and the historical data. Figure 33 shows how TIMES-Italia seems to overestimate the energy consumption starting levels for the iron and steel sector. The difference keeps happening throughout the whole backcasting time scale and is reflected for the total industrial energy consumption. The initial 45 PJ difference in 2006 is constantly flattened due to the strongly decreasing energy consumption trend, resulting from the uptake of the electric arc furnace for steel production. Another important point to highlight is how the drop in energy consumption in 2009 due to the crisis is not present in TIMES-Italia due to how the milestone years are defined, but this is a minor issue considering that the trends of the backcasting results and of the Eurostat data present the same steady decrease.

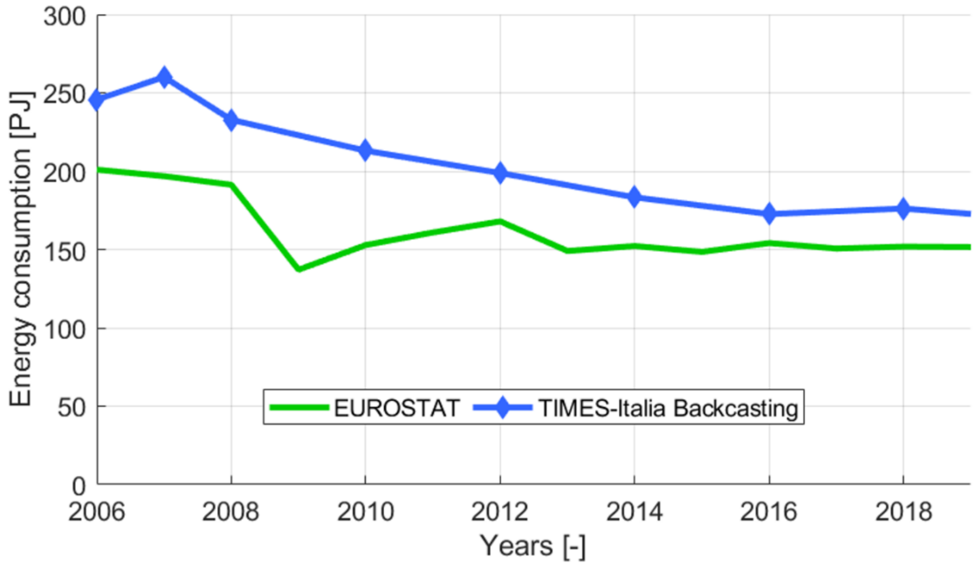

Figure 33: Energy consumption backcasting for the Iron and steel subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

Figure 34 shows the backcasting analysis for the Non-ferrous metals subsector. In this case, the starting point corresponds to the one of the Eurostat balances, and the two trends are comparable in the whole period of the analysis. The major difference can be found in the year 2008, where TIMES-Italia results show a higher decrease than the one in the Eurostat values. In fact, TIMES-Italia seems to anticipate the decrease happened up to 2013, but the differences between the two curves remain reasonably small.

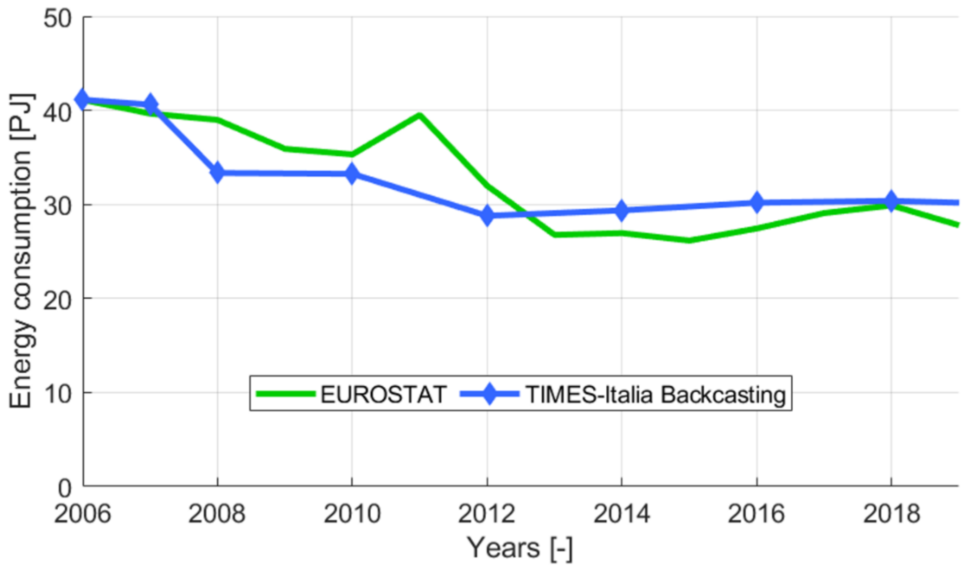

Figure 34: Energy consumption backcasting for the Non-ferrous metals subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

The energy consumption of the Non-metallic minerals sector presents a very similar starting point with respect to the Eurostat data in 2006, as showed in Figure 35.

Furthermore, the differences between the two curves remain at acceptable values in the whole period from 2006 to 2019, with the TIMES-Italia results mildly overestimating the real energy consumption. Similarly to the other sectors, the effect of the 2009 crisis is lowly anticipated, as expected due to the definition of the milestone years.

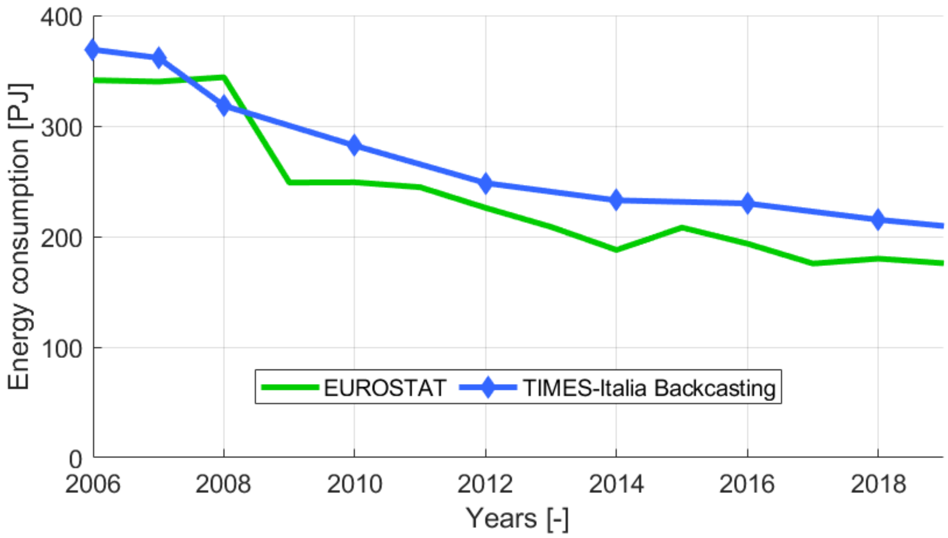

Figure 35: Energy consumption backcasting for the Non-metallic minerals subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

Along with Iron and steel, the Chemicals sector presents the highest deviation from the Eurostat values, as showed in Figure 36. In such sector, the initial 39 PJ difference between Eurostat and TIMES-Italia consumption in 2006 is slowly broadened due to the lower decreasing rate in energy consumption in TIMES-Italia. Furthermore, TIMES-Italia does not present any decrease in 2007 compared to the Eurostat, and the 2009 crisis effect is more moderate.

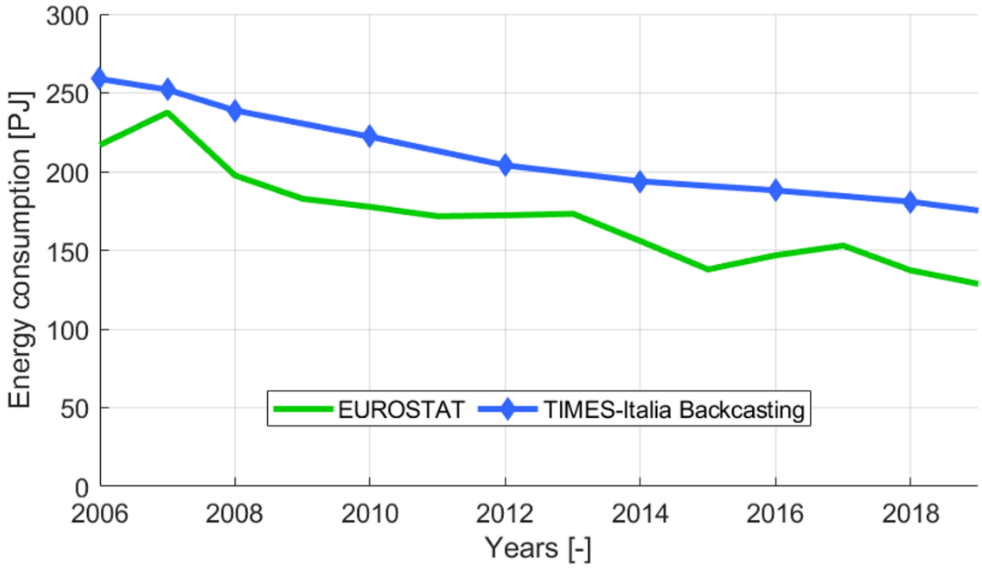

Figure 36: Energy consumption backcasting for the Chemicals subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

Pulp and paper is the sector presenting the most similarities in terms of energy consumption with respect to the Eurostat balance, as showed in Figure 37. The starting point in 2006 is practically the same as Eurostat, and the two curves are overlying in many points. Differently from the other sectors discussed above, the first three years of the backcasting are perfectly coherent with the historical data, resulting in a detailed representation of the 2009 crisis effects.

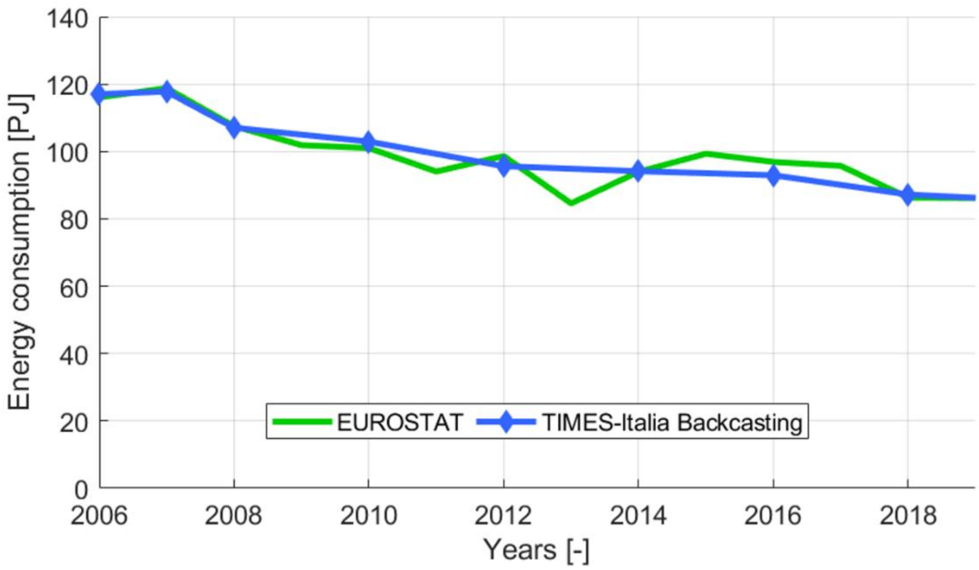

Figure 37: Energy consumption backcasting for the Pulp and paper subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

The Other industries subsector corresponds to the highest share of energy consumption with respect to the other sectors, as showed in Figure 38. For this reason, the high similarities both in the initial value in 2006 and in the whole trend between

TIMES-Italia results and Eurostat data present an important point in favor of the reliability of such model concerning the forecasting results. Similarly on how it has been seen in other industrial subsectors, TIMES-Italia results are overestimated compared to the Eurostat, but the differences remain acceptable.

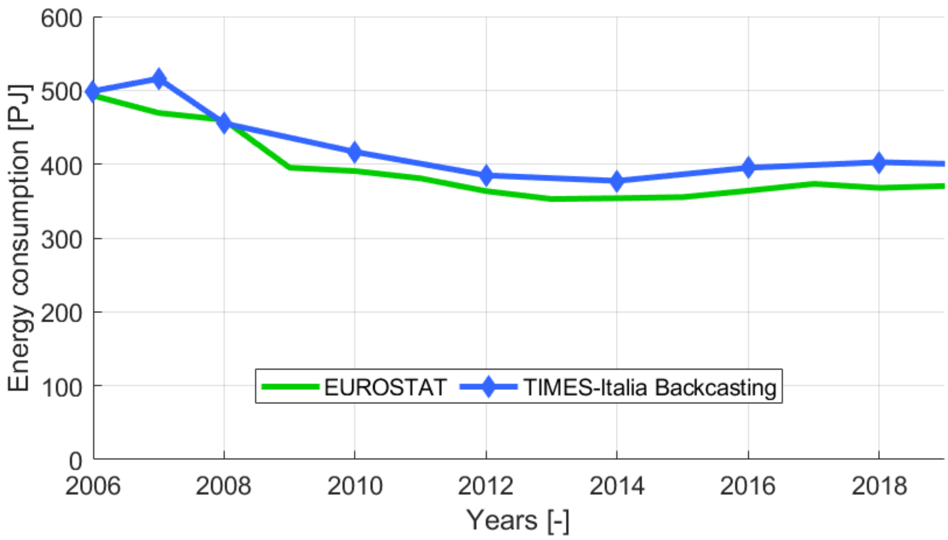

Figure 38: Energy consumption backcasting for the Other industries subsector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

The main reason for the different starting points of TIMES-Italia results and Eurostat data is to be found in the change of the accounting method for the compilation of Eurostat energy balances. Indeed, TIMES-Italia calibration was performed using the 2009 version of the IEA energy balances for OECD countries [27], before the update in the accounting methodology.

Figure 39 shows that such differences are not negligible when the energy consumption of the total industry sector is considered, with a constant overestimation of TIMES-Italia. Such issue can be solved by a careful recalibration of the model, starting from the energy balances in the base year. Nonetheless, such results are acceptable for the analysis performed in this work, considering that the general trend is of higher importance over the actual values for assessing the results response to a change in the input drivers. In fact, analyzing the change of the Industrial sector forecasts due to the pandemic crisis means, above all, comparing the general trends of Pre- and Post-pandemic projections, together with the historical trends. Figure 39 highlights such strong similarity between the lowering model results values and the Eurostat data.

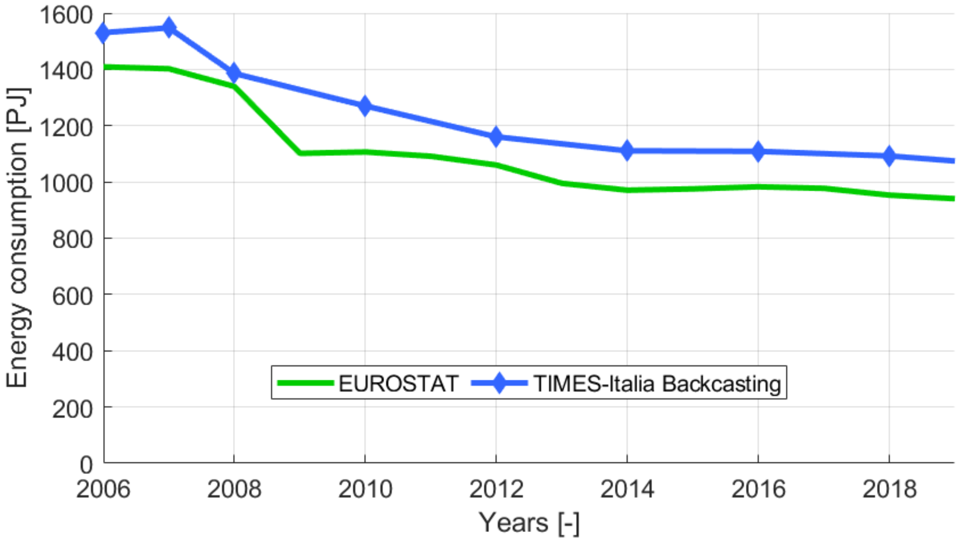

Figure 39: Energy consumption backcasting for the total industry sector using TIMES-Italia. The TIMES-Italia backcasting is benchmarked against Eurostat historical series 2006-2019.

7.2 Energy final consumption forecasting results

Even though the economic shock caused by the pandemics shows its effects in most subsectors, according to the strong bump highlighted for 2020 industrial production, the energy consumption reduction due to the use of Post-pandemic drivers is almost imperceptible in the Iron and steel subsector (Figure 40), where also relevant demand growth is not taken into account when looking at demand growth rates in Figure 23, after 2025. Pre-pandemic projections also present a small decrease in 2020, followed by an increasing trend that flattens in the long-term. Post-pandemic projections show the same behavior, with a more pronounced flattening trend.

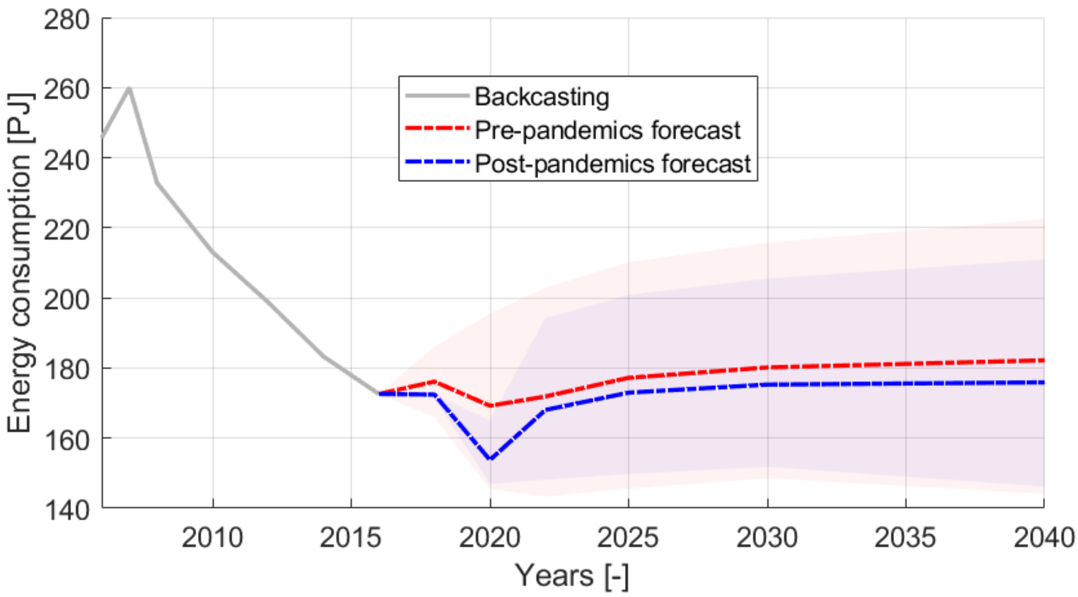

Figure 40: Energy consumption forecasting for the Iron and steel subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

Concerning the Non-ferrous metals subsector, the low energy consumption levels (always < 50 PJ) bring to almost negligible differences in the two alternative projection sets, as showed in Figure 41. In this case, the two trends are very similar, but shifted by a difference related more on the one between 2018 and 2019 values of the drivers (Figure 24), and not from the pandemic crisis. In fact, both in Figure 24 and Figure 41 it can be seen how a total recovery from the crisis happens in the first years of the projections.

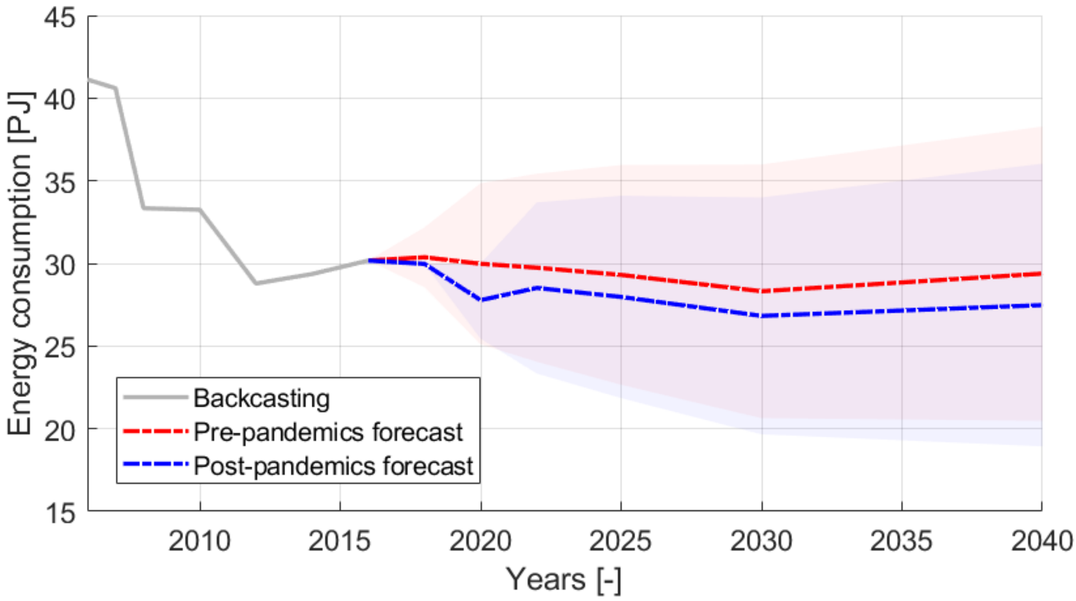

Figure 41: Energy consumption forecasting for the Non-ferrous metals subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

In Figure 42 the sector where one of the lowest demand gap between Pre-pandemics and Post-pandemics projections was highlighted in Figure 26 (Non-metallic minerals, with 4.55 % difference in 2040), presents a very similar difference of just 4.5 % when coming to energy consumption. In the long term, in fact the two projections tend to converge.

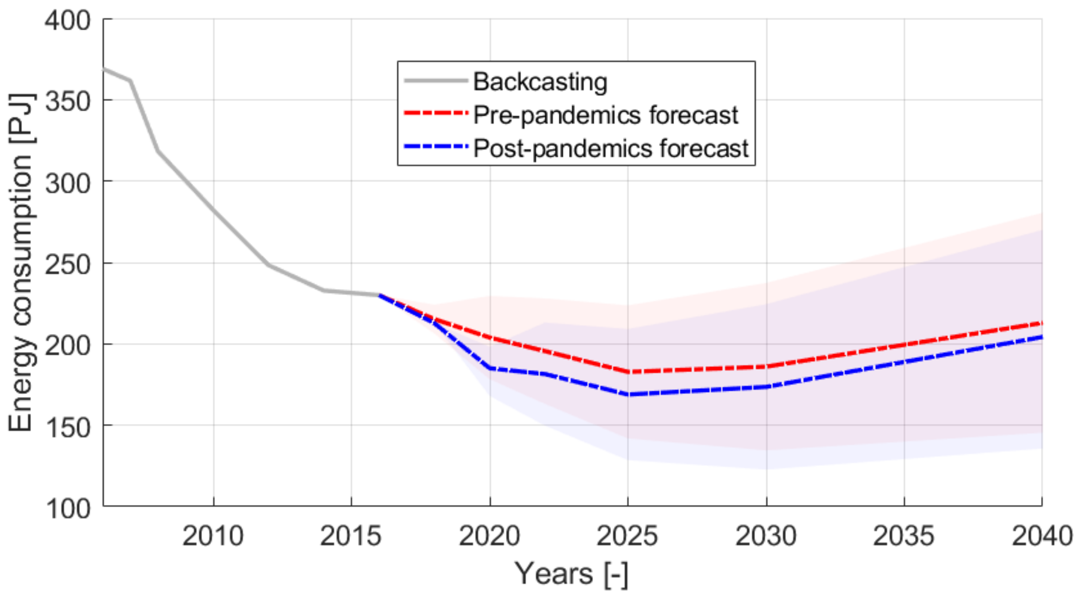

Figure 42: Energy consumption forecasting for the Non-metallic minerals subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

On the other hand, it is interesting to notice how the optimization process encompassed in TIMES can lead to a stronger energy consumption reduction in a sector where demand growth, even when considering Post-pandemic drivers, leads to a large growth with respect to 2006 levels. Such a strong demand growth, happening for instance in the Chemicals sector (Figure 43), is reflected on an energy consumption level which is the 40 % lower in 2040 with respect to 2006. Indeed, Figure 43 shows how Chemicals energy consumption is the 5 % and the 11 % lower in projections using Pre-pandemic and Post-pandemic drivers, respectively, in 2040 with respect to 2006.

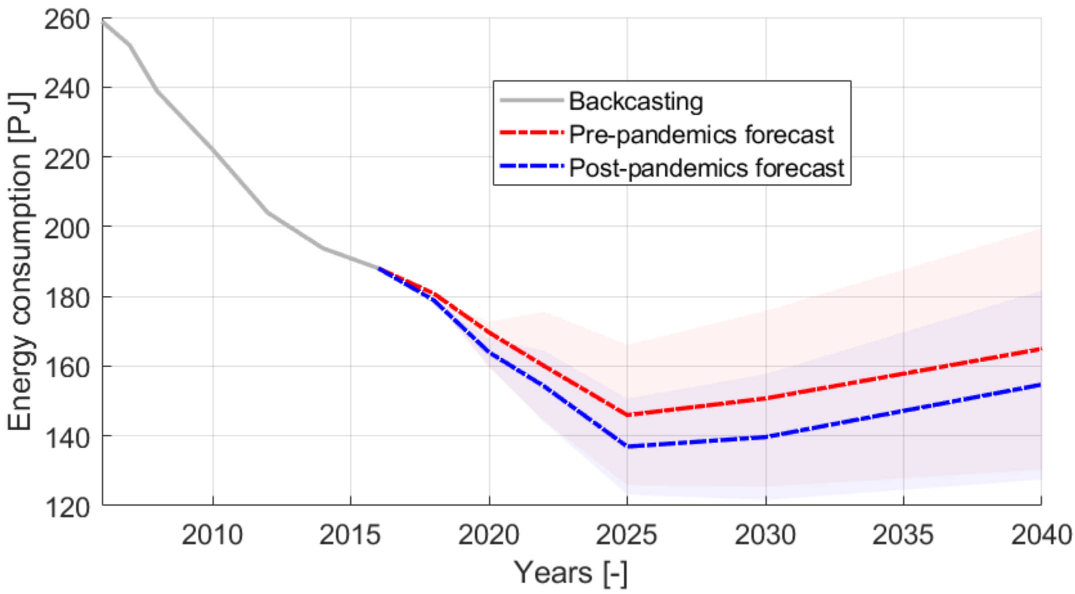

Figure 43: Energy consumption forecasting for the Chemicals subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

Figure Figure 44 shows how the Pulp and paper sector presents the same decreasing behavior of the Chemicals compared to its respective driver increase showed in Figure 27. Furthermore, Pre-pandemic and Post-pandemic forecasts tend to diverge on the long term, following the same behavior of the driver projections.

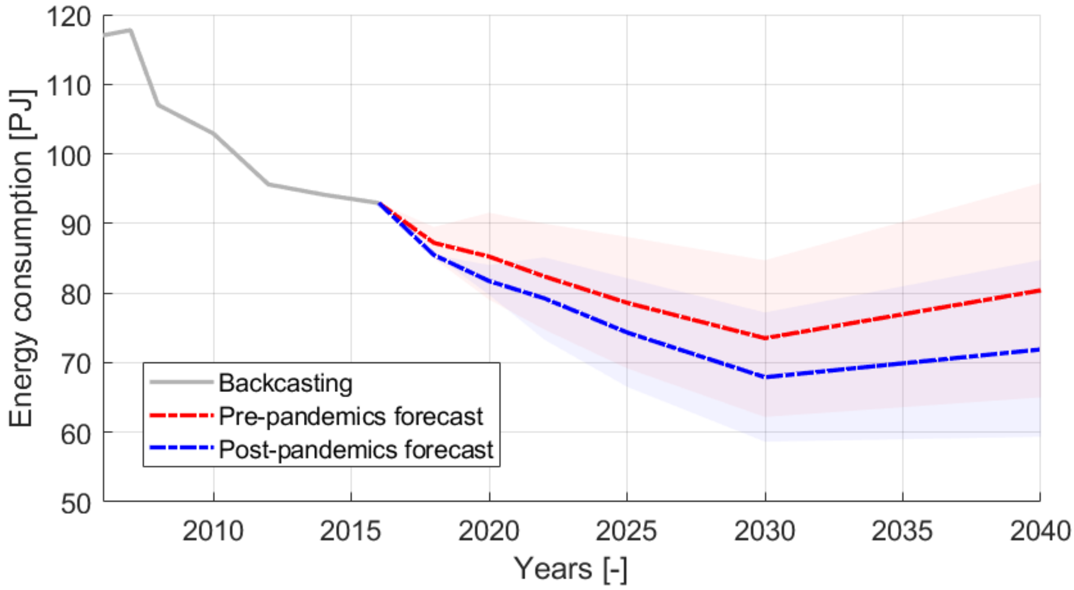

Figure 44: Energy consumption forecasting for the Pulp and paper subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

The Other industries subsector also presents an important difference between the Pre-pandemic and Post-pandemic energy forecasts, as showed in Figure 45. Such projections show a stagnation in the mid- and long-term, but a small increase in the Post-pandemic forecast brings the two curves to a smaller deviation in 2040.

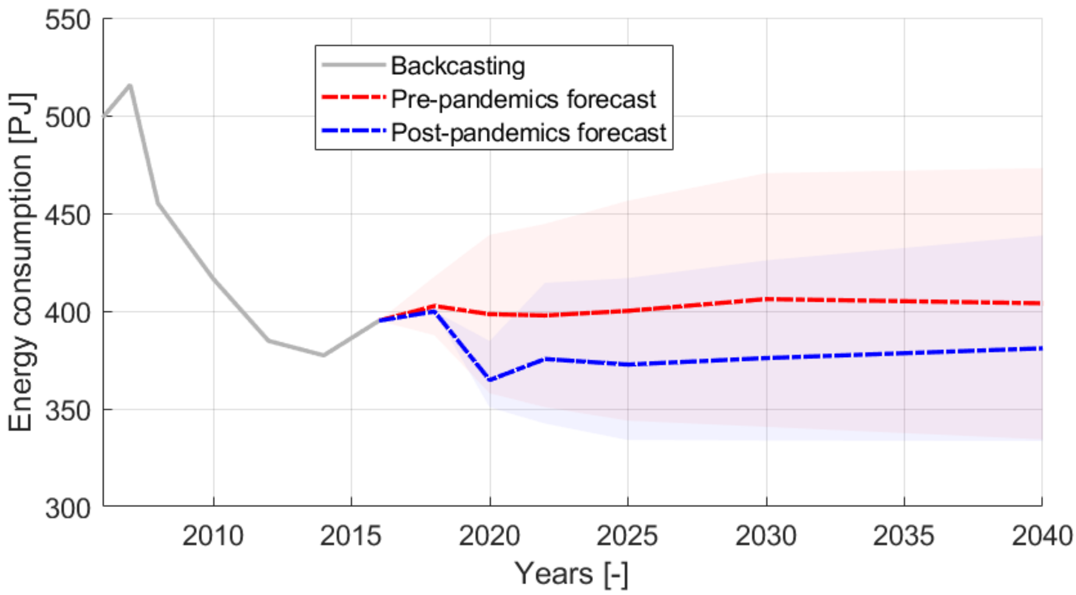

Figure 45: Energy consumption forecasting for the Other industries subsector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

Eventually, concerning total industrial energy demand, the 7 % energy consumption range between projections using Pre- and Post-pandemic drivers, visible in Figure 46 for the year 2020, shows even more long-term effects when highlighting a 60 PJ difference, corresponding to a 5 % over total consumption still in 2040.

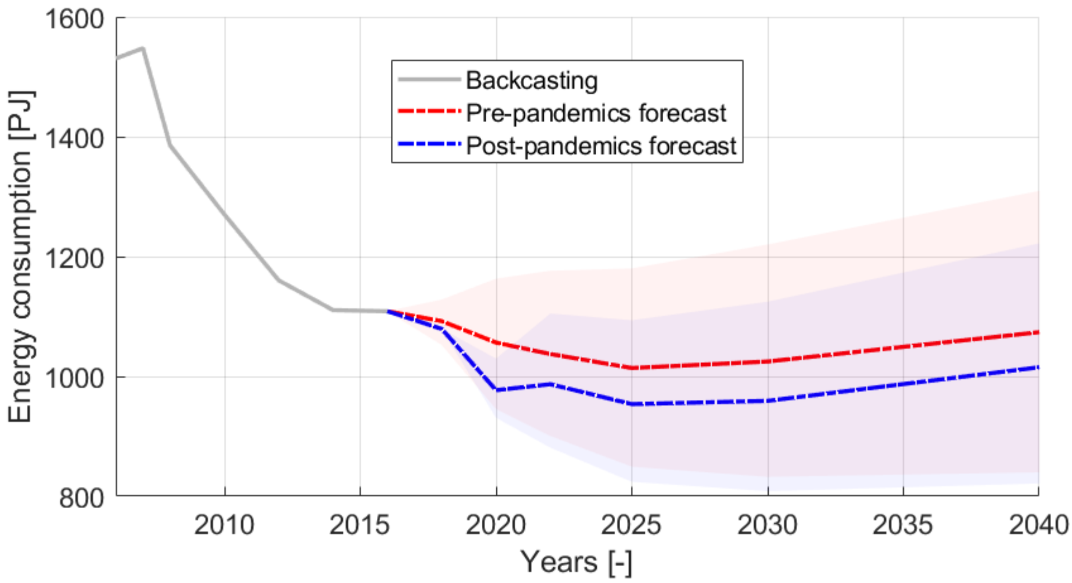

Figure 46: Energy consumption forecasting for the total industrial sector using TIMES-Italia. Projections obtained using the upper and lower bounds of the 95 % confidence interval as drivers are enclosed within the light red (Pre-pandemics) and light blue (Post-pandemics) areas.

Overall, the energy consumption projections results can be summarized in two main points. First of all, both the Pre-pandemic and Post-pandemic projections seem to show the same trend on the long-term, with a difference mostly related to the one registered from 2018 and 2019 in the historical data of the drivers. In fact, the second important point showed by such results is that in almost all the sectors a fast recovery from the pandemic crisis is registered, as expected from the historical data of the industrial production indexes analyzed in Chapter 4.

Chapter 8

Conclusions

8.1 A ‘methodological’ objective: VAR models projections

VAR models have proved themselves to be a valid approach to obtain reliable projections, that have a reasonable global trend and minimize the information loss of the post-pandemic period. The results discussed in this thesis show how a simple model like VAR can present very good similarities with respect to more complex models like the ones used for constructing the PNIEC scenarios, even without dealing with exogenous terms. This represents how such a simple model can be a valid alternative to obtain reliable forecasts in cases where unexpected periods like the pandemic crisis make the already rare past projections unusable.

In fact, one of the strengths of multivariate regression models such as VAR is its capacity to obtain such internal coherence that more complex general equilibrium models present. This is done in a VAR model by analyzing the interdependencies among the time-series to project. Although such coherence is given by mathematical correlations and not by economic assumptions, the Pre-pandemic results have showed to be in line with PNIEC results.

The fact that the Post-pandemic VAR projections in the short-term seem to follow the same behavior of the post-2009 recession represents a further qualitative validation of the model, considering that in econometric forecasting this similarity is often forced with ad hoc approaches to improve the short-term accuracy [24].

Overall, all the industrial sectors seem to be affected by the crisis, but with different amplitudes, both in the short- and long-term. In fact, sectors like Chemicals and Pulp and paper present a better response at year 2020 to the pandemic crisis, differently from the other sectors. The Iron and steel and Other Industries industrial sectors seem to get closer to the pre-pandemic results in the long-term, while the rest tend to follow the same trend of the pre-pandemic projections, maintaining constant the shift in the values. An exception is the Pulp, paper and printing sector, that presents in the long-run a larger difference in the two projections.

8.2 An objective of 'merit': final energy consumption forecasts

The industrial production trends computed using VAR models have then been applied to the energy system modeling framework TIMES-Italia, suited for the analysis of the Italian energy system on the long run, in a Business As Usual scenario in order to highlight the effects of the applications of the newly computed drivers.

One of the most important advantages of the TIMES-Italia model is the fact that its optimization starts in 2006, and this gives the possibility of performing a validation procedure based on the comparison of the first year results with historical data.

Such validation procedure has been performed comparing TIMES-Italia results with Eurostat data, highlighting different criticalities and strengths.

The main reason for the different starting points of TIMES-Italia results and Eurostat data is to be found in the change of the accounting method for the compilation of Eurostat energy balances. Indeed, TIMES-Italia calibration was performed using the 2009 version of the IEA energy balances for OECD countries [27], before the update in the accounting methodology.

Such differences are not negligible when the energy consumption of the total industry sector is considered, with a constant overestimation of TIMES-Italia. Such issue can be solved by a careful recalibration of the model, starting from the energy balances in the base year. Nonetheless, such results are acceptable for the analysis performed in this work, considering that the general trend is of higher importance over the actual values for assessing the results response to a change in the input drivers. In fact, analyzing the change of the Industrial sector forecasts due to the pandemic crisis means, above all, comparing the general trends of Pre- and Post-pandemic projections, together with the historical trends. The backcasting resultsFigure 39 highlight such strong similarity between the lowering model results values and the Eurostat data.

Concerning the energy projections, Both the Pre-pandemic and Post-pandemic seem to show the same trend on the long-term, with a difference mostly related to the one registered from 2018 and 2019 in the historical data of the drivers. In fact, the second important point showed by such results is that in almost all the sectors a fast recovery from the pandemic crisis is registered, as expected from the historical data of the industrial production indexes analyzed in Chapter 4.

Acknowledgements

Un ringraziamento speciale alla mia famiglia, in particolare a mia madre e mio padre: è grazie a loro sostegno e alla loro presenza se oggi sono riuscito a raggiungere questo traguardo.

Grazie alla mia metà, senza la quale questo ultimo anno sarebbe stato tre volte più difficile.

Grazie a mia nonna, che nonostante le distanze mi ha sempre mostrato infinito affetto e amore.

Ringrazio la Prof.ssa Laura Savoldi, per avermi guidato e supportato nella fase più importante del mio percorso accademico, offrendomi sempre grande disponibilità e ascolto.

Grazie a Daniele, per avere avuto la pazienza di seguirmi e darmi sempre i giusti consigli.

Bibliography

[1] A. Sandford, "Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement.," *Euronews*. https://www.euronews.com/2020/04/02/coronavirus-in-europe-spain-s-death-toll-hits-10-000-after-record-950-new-deaths-in-24-hou (accessed Jul. 05, 2021).

[2] A. Atalan, "Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective," *Ann. Med. Surg.*, vol. 56, pp. 38–42, 2020, doi: 10.1016/j.amsu.2020.07.001.

[3] D. Allain-Dupré, I. Chatry, V. Michalun, and A. Moisio, "The territorial impact of COVID-19 : managing the crisis across levels of government," 2020.

[4] International Energy Agency (IEA), "Global Energy Review 2020," 2020. doi: 10.1787/a60abbf2-en.

[5] W. Peck, "French power demand recovery outstripping other countries," *Independent Commodity Intelligence Services (ICIS)*, 2020. https://www.icis.com/explore/resources/news/2020/06/29/10524282/french-power-demand-recovery-outstripping-other-countries (accessed Jun. 24, 2021).

[6] IEA, *World Energy Outlook 2020*. 2020.

[7] F. Gracceva *et al.*, "Analisi trimestrale del Sistema Energetico Italiano: Anno 2020," 2021.

[8] S. Bhattacharyya and G. Timilsina, "A review of energy system models," *Int. J. Energy Sect. Manag.*, vol. 4, pp. 494–518, 2010.

[9] G. Goldstein, A. Kanudia, A. Lehtila, U. Remme, and E. Wright, "Documentation for the TIMES Model Part III," *IEA Energy Technol. Syst. Anal. Program.*, no. April, pp. 1–78, 2016.

[10] P. Capros *et al.*, *GEM-E3 Model Documentation*. 2013.

[11] R. Loulou, G. Goldstein, A. Kanudia, A. Lettila, and U. Remme, "Documentation for the TIMES model: PART I. Energy Technology Systems Analysis Programme," no. July, p. 151, 2016.

[12] L. C. Hunt and D. L. Ryan, "Economic modelling of energy services: Rectifying misspecified energy demand functions," *Energy Econ.*, vol. 50, pp. 273–285, Jul. 2015, doi: 10.1016/j.eneco.2015.05.006.

[13] S. Simoes *et al.*, *The JRC-EU-TIMES model. Assessing the long-term role of the SET Plan Energy technologies*, no. EUR 26292 EN. 2013.

[14] R. Fouquet, "Long-run demand for energy services: Income and price elasticities over two hundred years," *Rev. Environ. Econ. Policy*, vol. 8, no. 2, pp. 186–207, 2014, doi: 10.1093/reep/reu002.

[15] J. Gilbert and E. Tower, "'Standard' CGE Model from Chapter 27 of Gilbert and Tower (2013)." Nov. 09, 2012, [Online]. Available: https://econpapers.repec.org/RePEc:uth:gt2013:2012-43.

[16] C. Lewis and N. Pain, "Lessons from OECD forecasts during and after the financial crisis," *OECD J. Econ. Stud.*, vol. 2014, pp. 9–39, 2015.

[17] P. Jiang, Y. Van Fan, and J. J. Klemeš, "Impacts of COVID-19 on energy demand and consumption: Challenges, lessons and emerging opportunities," *Appl. Energy*, vol. 285, p. 116441, 2021, doi: https://doi.org/10.1016/j.apenergy.2021.116441.

[18] A. Abu-Rayash and I. Dincer, "Analysis of the electricity demand trends amidst the COVID-19 coronavirus pandemic," *Energy Res. Soc. Sci.*, vol. 68, p. 101682, 2020, doi: https://doi.org/10.1016/j.erss.2020.101682.

[19] N. Norouzi, G. Zarazua de Rubens, S. Choupanpiesheh, and P. Enevoldsen, "When pandemics impact economies and climate change: Exploring the impacts of COVID-19 on oil and electricity demand in China," *Energy Res. Soc. Sci.*, vol. 68, p. 101654, 2020, doi: https://doi.org/10.1016/j.erss.2020.101654.

[20] R. J. Hyndman and G. Athanasopoulos, "Forecasting: principles and practice, 3rd Edition," 2021. .

[21] R. Ashley, "On the relative worth of recent macroeconomic forecasts," *Int. J. Forecast.*, vol. 4, no. 3, pp. 363–376, Jan. 1988, doi: 10.1016/0169-2070(88)90104-5.

[22] L. Kilian, "Handbook of research methods and applications in experimental economics," in *Handbook of Research Methods and Applications in Experimental Economics*, N. Hashimzade and M. A. Thornton, Eds. Edward Elgar Publishing Limited, 2019, pp. 1–426.

[23] Q. Wang, Y. Zhou, and X. Chen, "A vector autoregression prediction model for COVID-19 outbreak," *arXiv*, no. Cdc, pp. 1–14, 2021.

[24] C. Foroni, M. Marcellino, and D. Stevanovic, "Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis," *Int. J. Forecast.*, no. 2468, 2020, doi: 10.1016/j.ijforecast.2020.12.005.

[25] M. Gaeta and B. Baldissara, *IL MODELLO ENERGETICO TIMES-Italia Struttura e dati*. ENEA, 2011.

[26] Ministero dello Sviluppo Economico; Ministero dell'Ambiente e della Tutela del Territorio e del Mare, "Strategia Energetica Nazionale (SEN), 2017," 2017.

[27] OECD-IEA, *Energy Balances of OECD Countries (2009 Edition)*. 2009.

[28] Ministry of Economic Development; Ministry of the Environment and Protection of Natural Resources and the Sea; Ministry of Infrastructure and Transport, "Integrated National Energy And Climate Plan," Italy, 2019.

[29] European Commission, "EU Reference Scenario 2016," *EU Ref. Scenar. 2016*, p. 27, 2016, [Online]. Available: https://ec.europa.eu/energy/sites/ener/files/documents/ref2016_report_final-web.pdf.

[30] Eurostat, "Energy balances." https://ec.europa.eu/eurostat/web/energy/data/energy-balances (accessed Jun. 24, 2021).

[31] R. Loulou, G. Goldstein, and K. Noble, "Documentation for the MARKAL

Family of Models," *Int. Energy Agency*, no. October, pp. 1–389, 2004.

[32] U. Remme, G. A. Goldstein, U. Schellmann, and C. Schlenzig, "MESAP / TIMES --- Advanced Decision Support for Energy and Environmental Planning," in *Operations Research Proceedings 2001*, 2002, pp. 59–66.

[33] Scottish Government, *DRAFT CLIMATE CHANGE PLAN: The draft third report on policies and proposals 2017-2032*. Edinburgh: Scottish Government, 2017.

[34] J. M. Allwood, V. Bosetti, N. K. Dubash, L. Gómez-Echeverri, and C. von Stechow, "Glossary," in *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler,* vol. 3, Cambridge, New York: Cambridge University Press, 2014, pp. 113–136.

[35] D. Lerede, C. Bustreo, F. Gracceva, M. Saccone, and L. Savoldi, "Techno-economic and environmental characterization of industrial technologies for transparent bottom-up energy modeling," *Renew. Sustain. Energy Rev.*, vol. 140, no. February, p. 110742, 2021, doi: 10.1016/j.rser.2021.110742.

[36] IEA, "Tracking Industry 2020," 2020. https://www.iea.org/reports/tracking-industry-2020.

[37] IEA, "Italy energy statistics." https://www.iea.org/countries/italy (accessed Jun. 24, 2021).

[38] IEA, "ETP Clean Energy Technology Guide," 2020. https://www.iea.org/articles/etp-clean-energy-technology-guide.

[39] Eurostat, *NACE Rev. 2 – Statistical classification of economic activities in the European Community*. Luxembourg, 2008.

[40] IEA, *Energy technology perspectives 2017*. Paris: IEA Publications, 2017.

[41] "Istat Statistics." http://dati.istat.it/Index.aspx?lang=en&SubSessionId=9001105c-878f-49de-9048-052c55785a6c (accessed Feb. 17, 2021).

[42] K. C. Popat and T. A. Desai, "Alumina," *Biomater. Sci. An Introd. to Mater. Third Ed.*, pp. 162–166, 2013, doi: 10.1016/B978-0-08-087780-8.00018-8.

[43] ETSAP, "Technology brief, Aluminium production," 2012.

[44] J. Green, *Aluminum recycling and processing for energy conservation and sustainability*. Russel Township, 2007.

[45] F. Schorcht, I. Kourti, B. M. Scalet, S. Roudier, and L. D. Sancho, *Best Available Techniques (BAT) Reference Document for the Production of Cement, Lime and Magnesium Oxide*. Publications Office of the European Union, 2013.

[46] F. Pacheco Torgal, J. Labrincha, C. Leonelli, A. Palomo, and P. Chindaprasirt, *Handbook of alkali-activated cements, mortars and concretes*. 2015.

[47] IEA-CSI, "Technology Roadmap, low-carbon transition in the cement industry," *Int. Energy Agency*, p. 66, 2018.

[48] M. P. Clements and D. F. Hendry, *Forecasting Non-Stationary Economic Time Series*. Cambridge, U.K.: The MIT Press, 1999.

[49] E. J. Hannan and M. Deistler, "The Statistical Theory of Linear Systems," vol. 70, SIAM, 2012.

[50] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.

[51] C. A. Sims, "Macroeconomics and Reality," *Econometrica*, vol. 48, no. 1, p. 1, 1980, doi: 10.2307/1912017.

[52] P. Chen and J. Frohn, *On the Specification and Estimation of Large Scale Simultaneous Structural Models*. Berlin, 2006.

[53] O. Blanchard and R. Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *he Q. J. Econ.*, vol. 117, no. 4, pp. 1329–1368, 2002.

[54] H. Lütkepohl, "Vector autoregressive models," in *Handbook of Research Methods and Applications in Empirical Macroeconomics*, N. Hashimzade and M. A. Thornton, Eds. Edward Elgar Publishing Limited, 2013, pp. 139–164.

[55] S. Basu, "Modeling and Estimation of High-dimensional Vector Autoregressions," 2014.

[56] L. J. Christiano, M. Eichenbaum, and C. L. Evans, "Chapter 2 Monetary policy shocks: What have we learned and to what end?," *Handbook of Macroeconomics*, vol. 1, no. PART A. Elsevier, pp. 65–148, Jan. 01, 1999, doi: 10.1016/S1574-0048(99)01005-8.

[57] Y. Nakajima and N. Sueishi, "Forecasting the Japanese macroeconomy using high-dimensional data," *Japanese Econ. Rev.*, 2020, doi: 10.1007/s42973-020-00041-z.

[58] W. Nicholson, D. Matteson, and J. Bien, "BigVAR: Tools for Modeling Sparse High-Dimensional Multivariate Time Series," pp. 1–39, 2017, [Online]. Available: http://arxiv.org/abs/1702.07094.

[59] M. Banbura, D. Giannone, and L. Reichlin, "Large bayesian vector auto regressions," *J. Appl. Econom.*, vol. 25, no. 1, pp. 71–92, 2009.

[60] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Stat. Soc. Ser. B*, pp. 267–288, 1996.

[61] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *J. Stat. Softw.*, vol. 33, no. 1, 2010.

[62] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[63] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. R. Stat. Soc. Ser. B*, vol. 67, no. 2, pp. 301–320, 2005.

[64] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. R. Stat. Soc. Ser. B*, vol. 68, no. 1, pp. 49–67, 2006.

[65] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Stat. Surv. 4*, pp. 40–79, 2010.

[66] G. Vasconcelos, "gabrielrvsc/HDeconometrics: Implementation of several econometric models in high-dimension.," 2020. https://rdrr.io/github/gabrielrvsc/HDeconometrics/ (accessed May 28, 2021).

[67] C. A. Sims and T. Zha, "Bayesian Methods for Dynamic Multivariate Models," *Int. Econ. Rev. (Philadelphia).*, vol. 39, no. 4, p. 949, 1998, doi: 10.2307/2527347.

[68] Board of Governors of the Federal Reserve System, "Industrial Production and Capacity Utilization," *Stat. release G.17*, 2013.

[69] N. R. Draper and H. Smith, "'Dummy' Variables," *Applied Regression Analysis*. pp. 299–325, Apr. 09, 1998, doi: https://doi.org/10.1002/9781118625590.ch14.

[70] M. Stone, "Cross-Validatory Choice and Assessment of Statistical Predictions," *J. R. Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 111–147, Apr. 1974, [Online]. Available: http://www.jstor.org/stable/2984809.

[71] G. Hyndman, R.J., & Athanasopoulos, "Forecasting: Principles and Practice," *3rd ed, OTexts: Melbourne, Australia*, 2021. https://otexts.com/fpp3/ (accessed Feb. 17, 2021).

Appendix: R script for VAR projections

```
library(ggplot2)
library(forecast)
library(lubridate)
library(dplyr)
library(FitAR)
library(HDeconometrics)

setwd("C:/Users/mayto/Desktop/Uni/tesi magistrale/Dati Italia")
getwd()

# Open the file containing the historical dataset

Industry <- read.csv(
  file="Industry_R.csv",
  stringsAsFactors = FALSE
)

# Adjust the database matrix

Industry$date = seq(from = as.Date("1990/01/01"), to =
as.Date("2021/04/01"), by = 'month')


sector = data.frame(Industry[,2:7])

date = Industry$date

ind_data <- ts(sector,frequency=12,start=c(1990,1))

colnames(ind_data) <-
c('chemicals','nonferrous','siderurgy','nonmetal','paper','other')

# Create the matrix of seasonal dummies

x=ts(ind_data[,1:6],freq=12,start=c(1990,1))
x = seasonaldummy(x,h = NULL)
x1 = x[1:376,]
x1 = ts(x1,frequency = 12, start=c(1990,1))
x2 = x[377:612,]
x2 = ts(x2,frequency = 12, start=c(2021,5))

e <- matrix(rep(NA, 2016), nrow = 6)

# Cross-validation for lag selection: change the p value, and choose
the one that minimizes e

for (i in 30:375)
{
  var1 <- lbvar(window(ind_data[,1:6],end=c(1990,i)), p=12, lambda =
0.05, xreg = window(x1,end=c(1990,i)))
  var_for <- predict(var1,h=1,newdata = window(x1,end=c(1990,i+1)))
  e[1,i-29] <- ind_data[i+1,1]-var_for[,1]
  e[2,i-29] <- ind_data[i+1,2]-var_for[,2]
  e[3,i-29] <- ind_data[i+1,3]-var_for[,3]
  e[4,i-29] <- ind_data[i+1,4]-var_for[,4]
  e[5,i-29] <- ind_data[i+1,5]-var_for[,5]
```

```
  e[6,i-29] <- ind_data[i+1,6]-var_for[,6]
}
sqrt(mean(e^2, na.rm=TRUE))

# Perform fit and forecast

var1 <- lbvar(ind_data, p = 12, lambda = 0.05, xreg = x1, ps =
FALSE)
var_for <- predict(var1,h=236,newdata = x2, interval = "confidence")
var_for_ts <- ts(var_for,frequency=12,start=c(2021,5))

# Adjust the results for the plots: change accordingly the sector to
show

df_forec <- data.frame("date" = seq(from = as.Date("2021/05/01"), to
= as.Date("2040/12/01"), by = 'month'),
                       "inds" = c(var_for_ts[,1]))
df_hist <- data.frame("date" = c(date),
                      "inds" = c(sector$chemicals))
df_lower <- data.frame("date" = seq(from = as.Date("2021/05/01"), to
= as.Date("2040/12/01"), by = 'month'),
                       "inds" = c(var_for_ts[,2]))
df_upper <- data.frame("date" = seq(from = as.Date("2021/05/01"), to
= as.Date("2040/12/01"), by = 'month'),
                       "inds" = c(var_for_ts[,3]))

# Pass from monthly to yearly data

industry_2020 <- rbind(df_hist,df_forec)

industry_2020_annual <- industry_2020 %>%
  group_by(year(date)) %>%
  summarize(inds_year = mean(inds))

industry_2020_annual <-
ts(industry_2020_annual$inds_year,frequency=1,start=c(1990,1))

data_lower <- df_lower %>%
  group_by(year(date)) %>%
  summarize(inds_year = mean(inds))

data_upper <- df_upper %>%
  group_by(year(date)) %>%
  summarize(inds_year = mean(inds))

data_lower <- ts(data_lower$inds_year,frequency=1,start=c(2021,1))
data_upper <- ts(data_upper$inds_year,frequency=1,start=c(2021,1))

ts.plot(industry_2020_annual,data_lower,data_upper,gpars = list())
```