POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Ingegneria Edile

Tesi di Laurea Magistrale

Resorcefulness quantification for resilient communities and infrastructures

Relatore

Gian Paolo Cimellaro

Candidato

Alessandro Zona

A.A. 2018/2019

Summary

1 Introduction

1.1 Problem statement

Many frameworks to assess community resilience have been proposed in the state of the art; when weaknesses are revealed, frameworks propose to act at one of the resilience's dimensions level.

One of those is the resourcefulness, expression of the human factor, in terms of trust, preparedness, creativity, with regards to emergency management following natural disasters. It seems to be very tricky to quantify that and, in fact, there are no attempts in this respect in the state of the art. This thesis aims to be the first one, quantifying resourcefulness by means of a composite indicator.

1.2 Resilience - state of the art

Resilience of communities is a concept on which international research is focusing in the last years, involving many fields, such as civil engineering, statistics, sociology and economics.

In material sciences, resilience is defined as the ability to absorb energy and return to the original state; when it comes to social communities, resilience has been defined as the "ability of social units (e.g. organizations, communities) to mitigate hazards, contain the effects of disasters when they occur, and carry out recovery activities in ways that minimize social disruption and mitigate the effects of future earthquakes" (Bruneau, Chang et al. 2003); previous definition is limited to earthquakes, but we can extend it to any kind of natural disaster, *e.g.* hurricanes, floods, avalanches.

High community resilience implies minimal losses and damage in term of human lives, infrastructures, material and economic assets, so, at the same size of the disaster, a resilient community will suffer less damage and will take less time to return to the pre-disaster state. A mathematical definition of resilience can be expressed as:

$$R = \int_{t_0}^{t_1} [100 - Q(t)]\, dt \tag{1.1}$$

Where t_0 is the instant of the disastrous event, t_1 the time considered and $Q(t)$ the function of the system functionality, *i.e.* a measure of the performance of the infrastructure of a community, which can obviously range between 0% and 100%.

Resilience can be represented by Figure 1-1.
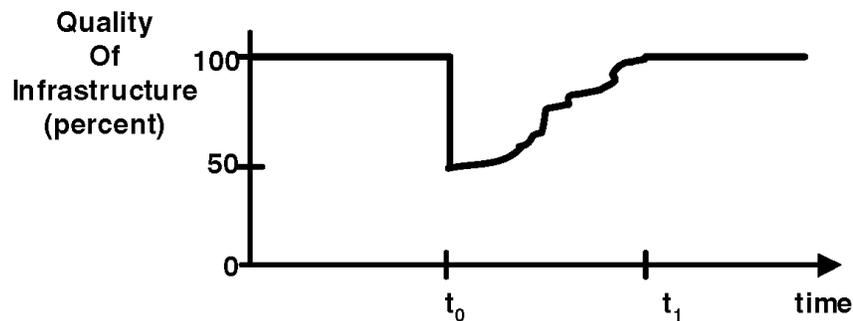
Figure 1-1: Measure of seismic resilience (Bruneau, Chang et al. 2003)

The question on how to enhance the resilience of a community can be dealt with introducing the "4-R" approach (Bruneau, Chang et al. 2003); it considers resilience as the result of contribution of four different dimensions:

- **Robustness**: the functionality that remains typical of the community following the occurrence of the event;
- **Rapidity**: the speed with which the community can reach at least the level of functionality pre-event;
- **Redundancy**: the ability of the community to offer more alternatives to manage the emergency;
- **Resourcefulness**: "the capacity to identify problems, establish priorities and mobilize resources when conditions exist that threaten to disrupt some element, system, or other unit of analysis" (Bruneau, Chang et al. 2003).

An improvement of any of these dimensions will increase the community resilience.

Before extending the research to the community-level, it's logical to evaluate resilience focusing on crucial infrastructures (*e.g.* water and power lifelines, hospitals, *etc.*), so that research has developed many frameworks to evaluate resilience for seismic resilience of a hospital (Cimellaro, Reinhorn et al. 2010), electric power system (Çağnan, Davidson et al.

2006) and so on; an exhaustive list and mutual comparison has been carried out by Cimellaro (Cimellaro 2016).

Nevertheless, also even some frameworks to evaluate resilience at community-scale have been proposed: *PEOPLES* framework, developed by Cimellaro (Cimellaro 2016), is one of them.

It aims to quantify resilience at the community level and for assessing the performance of infrastructures, taking into account seven dimensions, some of which referring to purely technological or economic issues, others to social and human ones.

Theoretical algorithm of PEOPLES includes, as final step, recognizing resilience actions to enhance community's resilience; mitigation actions may refer to one of 4 Rs, as shown in Figure 1-2.
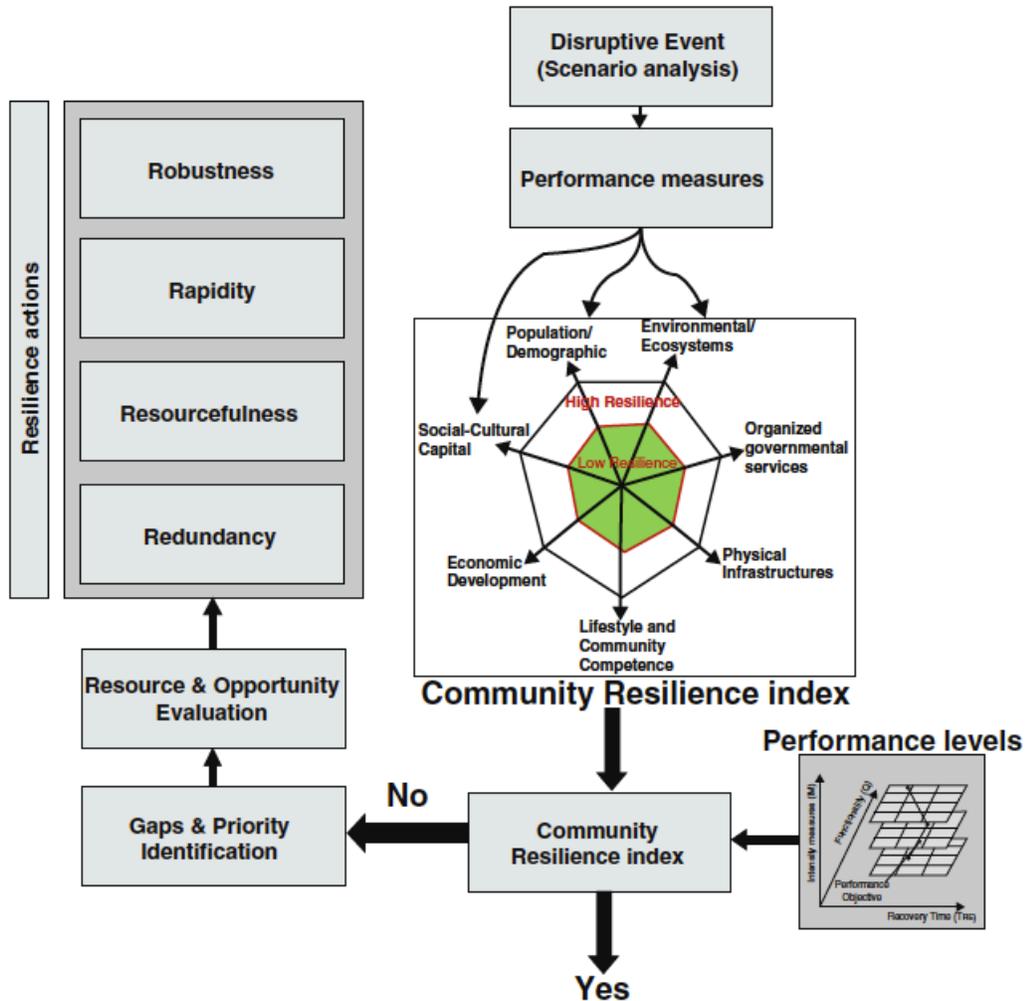
**Figure 1-2:** Methodology for resilience-based design (RBD) based on control (feedback loop) approach (Cimellaro 2016)

1.3 Resourcefulness – state of the art

In the field of emergency management, the concept of resourcefulness has been introduced to express the human factor of communities coping with natural disasters; many case studies (which will be described in section 2.4) of emergency management following natural disasters underline the importance of resourcefulness, and often the good results with which it replaces other aspects, frequently considered more important and which research traditionally has focussed on.

According to some, resourcefulness can be seen as parallel to resilience (Allen 2013), others have suggested that resourcefulness replace resilience when it comes to the behaviour of communities facing natural disasters (MacKinnon and Derickson 2013). More broadly, resourcefulness is considered as one of the dimensions in which resilience is divided (Bruneau, Chang et al. 2003), as has been discussed in section 1.2; evaluation of resourcefulness (and of the other three dimensions) allows, therefore, to look for weaknesses if a community shows an insufficient resilience.

Since discussion about resourcefulness is the main purpose of the present thesis, a more detailed description of the state of the art and some case studies to better frame the subject are needed, in order to be able to develop a reliable definition of resourcefulness and to set its arithmetical boundaries and constraints.

1.4 Shortcomings

State of the art about Resourcefulness points out more than one shortcomings.

In fact, the algorithm to compute community Resilience displayed in Figure 1-2 proposes to investigate on possible weaknesses at the 4-Rs levels: for example, if a lack of Redundancy is pointed out, then it will be clear to the policymakers which the 4 Rs they might invest on to enhance Resilience.

PEOPLES framework does not provide guidelines to know on which of the four Rs would be better to invest to enhance Resilience; indeed, this would require an in-depth analysis on each one of the 4 Rs, their connection among them and with Resilience and a specific framework to quantify them; without those, policymaking would be guided only by opinion and impressions and not by objective confirmations.

Nevertheless, state of the art pointed out neither any attempt to quantify Resourcefulness nor consensus about its definition, so that a comparison among the 4 Rs to investigate on community weaknesses is, at the moment, totally impossible.

Moreover, none of the proposed frameworks tried to evaluate community resilience through a mathematical formulation of the 4 Rs, above all due to the difficulty to translate the meaning of Redundancy and Resourcefulness into mathematical quantities and to analyse the interconnections between these two concepts and the others.

1.5 Thesis contributions

Also considering what explained in section 1.4, this thesis has to be intended as the first attempt in the state of the art of quantification of Resourcefulness for communities Resilience framework.

Obviously, that cannot be independent on a careful and thorough description of the current state of the art, in order to identify the best definition of Resourcefulness, its theoretical principles and the most proper factors that can be used to quantify it.

The framework that will be described can be used, therefore, to guide policymakers to enhance communities resilience and it can provide food, from a methodological point of view, for similar works on the other 4 Rs.

1.6 Thesis organisation

This thesis is organised at the manner of a traditional research work.

Chapter 2 mainly deals with two aspects: it analyses the current state of the art about Resourcefulness in finer detail than what has been done in section 1.3 and it aims to provide a series of case studies to better understand the concept; moreover, it sets the theoretical principles that have to be respected by the algorithm from a mathematical point of view.

Chapter 3 will focus on the methodology used to quantify Resourcefulness, tapping into the current state of the art about the construction of composite indices; it will provide, as final result, an algorithm, in the form of a MATLAB spreadsheet, to compute Resourcefulness for a given community.

Chapter 4 will discuss the indicators selection, providing, for each indicator, the motivation for which it is appropriate to include it, with references to case studies or to the state of the art.

Chapter 5 will analyse a case study (United States) following the developed methodology, checking its behaviour and pointing out any weaknesses.

Chapter 6 will wrap up the debate and discuss the results, paving the way for future developments.

2 Resourcefulness

2.1 Definition

Resourcefulness has been defined by as "the capacity to identify problems, establish priorities and mobilize resources when conditions exist that threaten to disrupt some element, system, or other unit of analysis" (Bruneau, Chang et al. 2003).

It has been introduced to express the human component in the behaviour of a community coping with a disaster; indeed, it's obvious the influence that factors like the moral, sense of belonging and trust in the decisional system can have on the emergency management.

A resourceful community could, for example, be able to understand that a wooden table can work as a raft following a flood, or a school gym as a shelter for refugees, and so on.

The importance of this factor and the insufficiency of a concept of Resilience without Resourcefulness have been well described by MacKinnon and Derickson (MacKinnon and Derickson 2013): they have defined Resourcefulness as the "capacity to engage in genuinely deliberative democratic dialogue to develop contestable alternative agendas and work in ways that meaningfully challenge existing power relations" and see it as an alternative to Resilience, instead of an its dimension; Resilience is here seen as lacking social justice.

More often, researchers with a more technical point of view consider Resourcefulness as a component of Resilience: for example, Resourcefulness has been also defined as "the ability to skilfully prepare for, respond to, and manage a crisis or disruption as it unfolds" (Berkeley, Wallace et al. 2010). A logical consequence is that Resourcefulness is a feature of the community which manifests before and after the shock: a good preparation (e.g. planning, training) for the event, from a technological and human point of view, is a symptom of Resourcefulness, but also an excellent management of the available resources and information sharing after the disaster happened.

Main definitions of resourcefulness as dimension of resilience are summed up in Table 2-1.

It's clear that, based on the definitions given so far, a mathematical formulation and a consequent numerical quantification of Resourcefulness would be very hard to develop, since it particularly depends on human intuition and ability to collaborate; indeed, there are no attempts in this respect in the state of the art.

Table 2-1: Summary of definitions of Resourcefulness (intended as a dimension of Resilience) in the state of the art.

	Resilience dimensions	Definition of Resourcefulness
(Bruneau, Chang et al. 2003)	- Robustness - Rapidity - Redundancy - Resourcefulness	"Capacity to identify problems, establish priorities and mobilize resources when conditions exist that threaten to disrupt some element, system, or other unit of analysis."
NIAC (2009)	- Robustness - Resourcefulness - Rapid recovery	"Ability to skilfully prepare for, respond to, and manage a crisis or disruption as it unfolds."
(Berkeley, Wallace et al. 2010)	- Robustness - Resourcefulness - Rapid recovery - Adaptability	"Ability to skilfully manage a disaster as it unfolds. It includes identifying options, prioritising what should be done both to control damage and to begin mitigating it, and communicating decisions to the people who will implement them. Resourcefulness depends primarily on people, not technology."
(Brown 2015)	- Resistance - Rootedness - Resourcefulness	"Resourcefulness encompasses the resources that people can draw on, but also the capacity to use them at the right time, in the right way."

A new definition of Resourcefulness will be adopted and the rest of this work will be developed with reference to this one: Resourcefulness is defined as the "capacity to identify problems, establish priorities, allocate and mobilize resources before, during and after an event that threaten to disrupt some element, system, or other unit of analysis."

2.2 Correlation with other Rs

As shown in section 1.2, Resilience is often seen as the sum of contribution of four dimensions: Robustness, Rapidity, Redundancy and Resourcefulness.

An important issue to study in deep is the connection between each R, in particular Resourcefulness, and the other ones.

The matter has been drafted expanding the Resilience function to the third axis of resources, obtaining what happens to Resilience when resources and Resourcefulness increase; the reasoning is displayed in Figure 2-1.
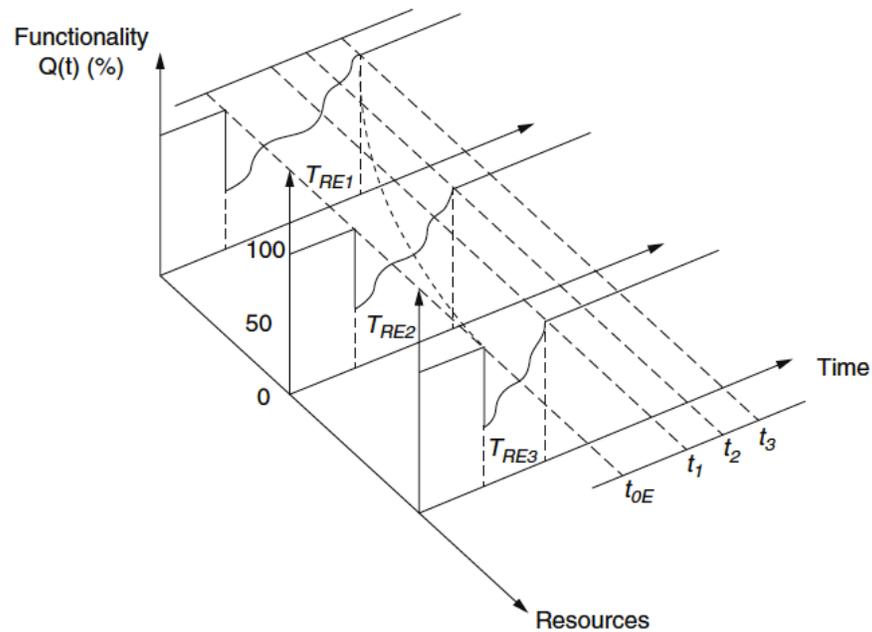
Figure 2-1: Expansion of Resilience on the axis of Resources (Bruneau and Reinhorn 2007)

It's clear that, the more are the resources available, the quicker will be the recovery phase and, therefore, the better will be the Rapidity: as noticed by Cimellaro (Cimellaro 2016), a theoretical infinite quantity of resources would take to a zero-time recovery.

Although Figure 2-1 does not consider that, Resourcefulness can also influence Robustness; particularly thinking about its side related to preparedness, it can be easily figured how Resourcefulness can reduce the loss of functionality after the event, in terms of human lives and material damage; for this reason, Figure 2-1 should be replaced by Figure 2-2.
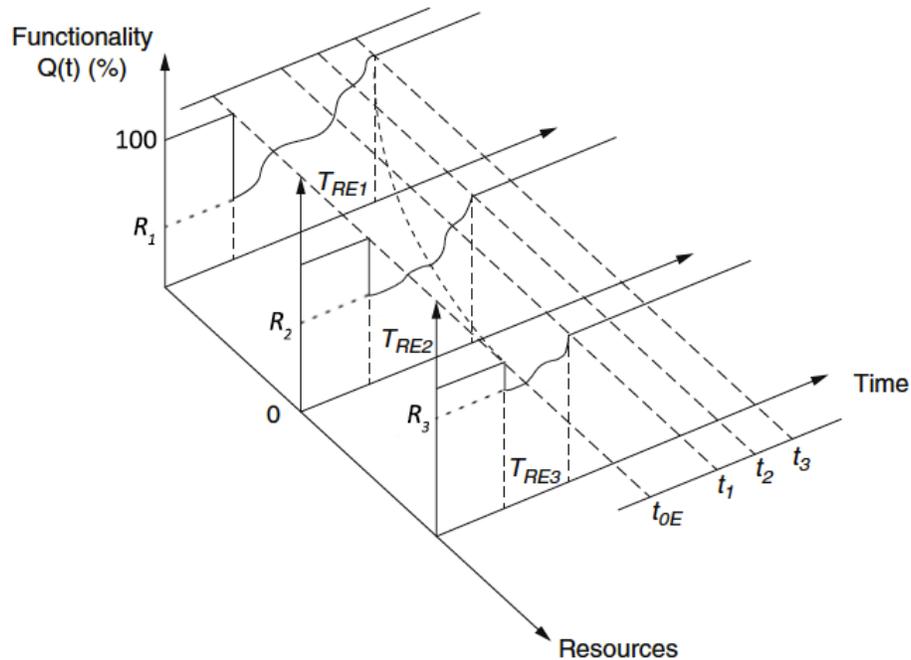
Figure 2-2: Figure 2-1 has been modified pointing out also intercorrelation between Resourcefulness and Robustness (R_1, R_2 and R_3).

Interconnection between Resourcefulness and Redundancy has been addressed by Cimellaro: "Resourcefulness and Redundancy are strongly interrelated. For example, resources, and resourcefulness, can create redundancies that did not exist previously" (Cimellaro 2016).

Following on from this statement, correlation between Resourcefulness and Redundancy seems to be one-way: enhancing Resourcefulness or adding resources can improve Redundancy, but enhancing Redundancy does not necessarily mean improving Resourcefulness. Further debate will be necessary on this issue.

We can sum up these interconnections by the diagram in Figure 2-3:
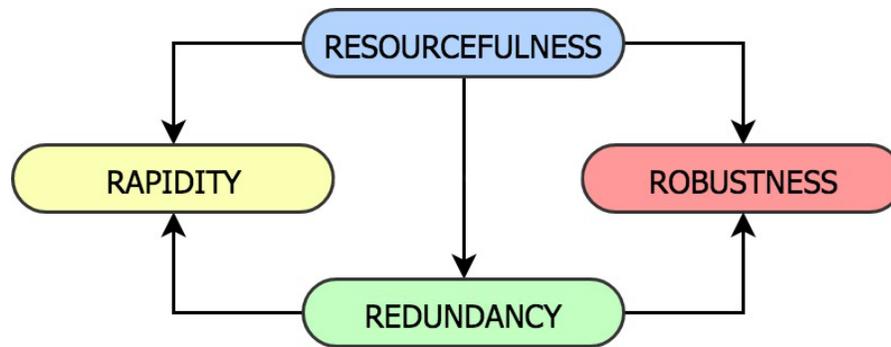
Figure 2-3: Interconnection between the 4 Rs of Resilience.

2.3 How to enhance Resourcefulness

A simple and intuitive example of an action to improve the Resourcefulness of a community are the drills: by simulating the event, each component of the community can try to play its role, identifying any problems and critical points, so that, if a real disaster happens, it is taught not only in theory, but also in practice on how it has to act.

Thus, enhancing Resilience through Resourcefulness can be achieved, for example, improving information sharing between the organizations involved in the incident response or teaching people what they have to do during and after the disaster.

Furthermore, many researchers (Perrow 2011) (Tierney 2008) (Drabek 2003) agree to consider decentralization of decision making an important factor that can enhance Resilience mainly through Redundancy and Resourcefulness: indeed, if decision makers are close to and familiar with the problems created by the event, they're more likely to know what the community needs and what kind of resources must be mobilized, both in the emergency and in the recovery phases; moreover, given the nature of the disaster-related tasks, relationships between organizations involved in emergency response require coordination, rather than hierarchical control over people and resources, networks, rather than hierarchies.

Weakness of a centralized system of decision making is evident looking at the case study of Kobe Earthquake in section 2.4.2.

A consequence of a decentralized disaster-response system is the birth of Emergent Multi-Organizational Networks (EMONS), that are a set of entities that "pursue repeated, enduring exchange relations with one another and, at the same time, lack a legitimate organizational authority to arbitrate and resolve disputes that may arise during the exchange" (Podolny and Page 1998).

The behaviour of the response system during the emergency phase following the WTC catastrophe has been compared (Tierney and Trainor 2004) to the more centralized way in which other political systems, as the Japanese one, usually manage emergencies, finding four main characteristics:

- Size: the sheer size of the EMON was commensurate to the severity of the event;
- Diversity: the network that developed to cope with the WTC attack was highly diverse, referring to the tasks responders were assigned to;
- Decentralization: decentralization of the response network was the key for a quick dissemination of information and an efficient coordination;
- Emergence: WTC attack stimulated the development of collective behaviour, regulated by emergent norms instead of institutionalized ones.

All these features are symptoms of an effective response system and are useful to enhance Resilience through Resourcefulness.

2.4 Case studies

In the following sections case studies will be shown and described to better frame the concept of resourcefulness, with examples in which communities pointed out a good resourcefulness or a lack of it.

2.4.1 Indian Ocean tsunami (2004)

A famous example of lack of Resourcefulness is the Indian Ocean tsunami, in December 2004: it took more than three hours to reach the coasts of Sri Lanka and India; nevertheless, the tsunami hit these countries with no warnings (death toll of more than 230,000 people), though they were sent to the coast guards above all from Honolulu, Hawaii. The reason was that there were neither communication infrastructures nor emergency planning, so that many people were surprised unaware by the catastrophe.

After that disaster, The GEOFON group of GFZ Potsdam completed a warning system in less than four years, training local scientists to operate and maintain the system.

The effect of this operation is an increase of Resourcefulness and, consequently, of Resilience.

2.4.2 Kobe earthquake (1995)

Another example of lack of Resourcefulness is the Kobe earthquake, in January 1995: in Japan, government officials usually live outside the city centre whereas offices are concentrated in the centre, so that, after the quake, only 7,300 members out of 15,000 employees were able to reach their offices (Rosenthal, Boin et al. 2001) and the emergency management system could not work.

Additionally, the Kobe earthquake brought the concept of volunteerism to Japan (alien to the Japanese culture until 1995: it did not exist, in fact, a Japanese word for "volunteerism"). The enormous amount, close to half a million, of volunteers produced more problems than solutions because they were not informed and coordinated by a special unit, so that they didn't know what to do and meant other people to get food and lodging for.

2.4.3 921 Taiwan earthquake (1999)

On September 21, 1999, Taiwan was hit by a shocking earthquake ($M_s = 6.6$), that produced a death toll of more than 2,350 and nearly 10,000 buildings collapsed[1].

A research (Lee and Loh 2003) pointed out the resourceful spirit showed by the community: people effectively cooperated to save lives and help each other; nevertheless, many began to criticize the slowness of the government, others praised the work of military forces and international aid teams.

An interesting study (Paton and Johnston 2017) analysed the contribution of the *Hakka spirit* to the behaviour of the community to respond to the disaster: Hakkas are Han (Chinese) people who have emigrated to many countries, among which Taiwan; their approach to natural disasters has been described as "the *spirit of the sturdy neck*, which means to hold on firmly despite extreme adversity, or to keep on doing something without regard to one's own strength" (Paton and Johnston 2017). This spirit had a great influence to the quick recovery of Tung Shih town, also pushed by commitment to the temple.

The government, for its part, seemed to have offered a good response, though its work was limited by the shortage of important material as local area maps and building blueprints and lack of earthquake preparedness was evident.

---

[1] Data reported by National Fire Agency, Ministry of the Interior R.O.C.

The spirit of the community, its confidence in the expertise of the policy, the availability of necessary resources to cope with the emergency phase and the preparedness to the event are all features amenable to the wider concept of Resourcefulness.

2.4.4 Queensland (Australia) floods (2010, 2011)

From 2010, a series of floods hit the region of Queensland, in the north-east of Australia, for a death toll of 44 and at least 200,000 people affected.

In the days following the floods a Mud Army born, armed with buckets and brooms, and the bureaucracy, for its part, shed much of its slow mechanisms to make the tasks of the rescuers easier.

World Bank sent a team of fifteen people to help and its director for the Pacific "expressed his admiration for the state's positive attitude and announced a partnership with Queensland to help teach other countries how to face down a life-changing disaster" (Allen 2013).

2.4.5 Chile earthquake (2010)

In 2010, a 8.8 M_w earthquake hit the coast of central Chile, followed by a tsunami that devastated many south-central coastal towns.

Preparedness to the event seems to have been effective: ONEMI (comparable to FEMA) Regional Emergency Management in Talca reported that most coastal communities practiced evacuating the coastal area during drills every month (Hinrichs, Jones et al. 2011).

The impact of the disaster was mitigated by the community spirit of the Chilean people: they shared resources as food and shelters with family and neighbourhood, so that in the town of Constitución the first outside supply was brought in by family members who drove from the city of Talca.

Moreover, Chile, during non-disaster times, uses to feed poor people storing food in "4x4" boxes: they are so called because they contain non-perishable items sufficient to feed four people for up to four days; in the post-quake phase, Chile expanded this system and used it to deliver food to the earthquake and tsunami impacted areas (SALAZAR).

A notable example of adaptability of the community to the mutated conditions following the disaster is the evacuation of the hospitals: two medical directors reached their hospitals less than 30 minutes following the quake but found the facilities completely evacuated. This was made possible because the staff and patients did not wait for orders from any centralized

command structure or the person in charge, but immediately helped patients out of the building.

It's doubtless a hazardous decision, but waiting for the arrival of the appointed unit could have been tragic consequences.

Unfortunately, someone took advantage of the temporary absence of the usual control systems: many supermarkets, pharmacies, liquor stores have been looted; furthermore, prison riots broke out in the prison of El Manzano in Concepción and in another one in Chillán, from which 203 prisoners escaped and burned seven houses close to the prison[2].

This last aspect highlights the importance of crime rate and of effectiveness of the law enforcement officers in the community Resourcefulness (and thus Resilience).

2.5 Theoretical principles of Resourcefulness

In this section theoretical principles of Resourcefulness will be set, *i.e.* the mathematical boundaries and conditions that the Resourcefulness dimension has to respect to represent a good and faithful expression of its conceptual definition, even according to the described examples.

2.5.1 The higher the better

Minimal possible value for Resourcefulness is obviously 0: it is therefore impossible that a community has less than the total absence of resources to cope with natural disasters.

On the other hand, it would not be correct to set an upper limit for Resourcefulness: theoretically it is always possible to add resources to a community and its response to hazardous events will be gradually better (in terms of losses and recovery time).

For this reason, the higher will be the community's Resourcefulness, the better will be its response; Resourcefulness varies, therefore, between 0 and $+\infty$:

$$RFS \in [0,+\infty] \tag{2.1}$$

---

[2] *El Mercurio*, 1 march 2010. Cuerpo C, p. C13

2.5.2 More resources higher Resourcefulness

Adding (or removing) resources necessarily means enhancing (decreasing) Resourcefulness: it would be obviously nonsense that an increase of a resources implies stagnating or decreasing Resourcefulness.

Graphically, if plot resource x – Resourcefulness is plotted, the graph will be necessarily monotonically strictly increasing:

$$RFS(x_2 > x_1) > RFS(x_1) \qquad\qquad (2.2)$$

2.5.3 Independency among communities

There's no reason why Resourcefulness of different communities should be interdependent; in other words, it must be possible to compute RFS of community c irrespective of the availability of data for other communities:

$$RFS_c \neq f(RFS_{d \neq c}) \qquad\qquad (2.3)$$

This theoretical principle will be very important and limiting in choosing normalizing and weighting methods to build the composite index.

3 Methodology

3.1 Resourcefulness as composite index

Since resourcefulness is such a difficult aspect of a community to quantify (as has been already explained in Chapter 2) and it involves many different features (such as preparedness, trust, sense of belonging, etc.), the best methodology seems to be the construction of a quantitative composite index, also by looking at many frameworks which quantify resilience in a similar way (Cutter, Ash et al. 2014).

As stated by Cimellaro, "one condition for quantitative indicators, in contrast to qualitative indicators, is that they have to be fully operationalized" (Cimellaro 2016).

Nevertheless, this is not possible in the matter of resourcefulness, which involves many subjective issues; actually, task of this thesis is exactly to convert qualitative indicators into quantitative metrics; however, should always be kept in mind that, "despite transferring qualitative indicators into quantitative metrics, the underlying information remains still subjective" (Cimellaro 2016).

3.2 Steps of the construction of the composite index

Considering what discussed by OECD (Commission 2008), the construction of a composite index must follow the steps below:

      1. Definition of the theoretical framework (*cf.* Chapter 2);

      2. Data selection (*cf.* Chapter 4);

      3. Imputation of missing data (*cf.* Chapter 3);

      4. Normalisation (*cf.* Chapter 3);

      5. Weight allocation (*cf.* Chapter 3);

      6. Aggregation (*cf.* Chapter 3);

      7. Uncertainty and sensitivity analysis (*cf.* Chapter 5).

3.3 Dealing with missing data

In common practice it's almost impossible to get a complete data set on which to carry out an analysis; missing data have, therefore, to be dealt with to make possible subsequent operations.

Obviously, the higher is the amount of missing data, the more the final results will be affected by the assumptions on how to impute them; it's therefore a good practice to delete cases or variables which are affected by an higher amount of missing data than a certain threshold. Remaining missing data can be imputed through single or multiple imputation; the difference is that multiple imputation method draws a certain number of imputations, which will be synthesized by the analyst (*e.g.* through arithmetic mean).

Before choosing the imputation method, it is important to analyse the missing data pattern, to find out if that is (Commission 2008):

- *Missing completely at random* (MCAR): missing values do not depend on the observed variable or any other variables in the data set;
- *Missing at random* (MAR): missing values do not depend on the observed variable but on other variables;
- *Not missing at random* (NMAR): missing values depend on the values themselves.

The great majority of the methods of imputation of missing data proposed in the state of the art require MAR or MCAR missing data pattern.

Unfortunately, there are no statistical tests to prove whether data pattern is NMAR or MCAR/MAR; nevertheless, statistical software like SPSS dispose of missing values analysis functions which can help for this purpose.

Many methods of imputation have been proposed in the state-of-the-art (Commission 2008); the main are the following:

- Unconditional mean imputation: missing data are substituted by the mean value of the corresponding variable across cases;
- Regression imputation: if a (or more) good regressor is available, missing data are imputed through regression; good regressor is chosen on the basis of a certain criterion (*e.g.* R^2);
- Expected maximisation imputation: it allows to impute missing data through an iterative maximum likelihood process.
- Markov Chain Monte Carlo (MCMC) imputation: it requires MCAR or MAR missing data pattern and it generates a sequence of random variables, iterating the process till covariance and mean vector reach convergence.

At first, in order to ensure that final results will not be too much influenced by simulated data, incomplete data set is transformed as follows:

1. Indicators with more than 75% missing data are excluded from the analysis;
2. Years with more than 50% missing data are excluded from the analysis.

The more permissive threshold for indicators is motivated by the observation that deleting an indicator is, from a theoretical point of view, much more drastic than deleting a case (*i.e.* a year); it has therefore to be done only in extreme cases (*e.g.* when less than 25% of data are available).

Remaining missing data are imputed following the following algorithm:

1. Each $x_i - x_j$ is plotted;
2. R^2 of each plot is computed;
3. If $R^2 \geq 0.5$, x_j is considered a good regressor for x_i.

R^2 is a unitless dimension (ranged between 0 and 1) which represents the reliability of a predicting model fitting starting data:

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{N}(y_i - \bar{y})^2} \tag{3.1}$$

Missing data of indicators for which have been found at least a good regressor are imputed through multiple imputation (5 imputations, for example), averaging the results.

Missing data of indicators for which have not been found a single good regressor are imputed through MCMC imputation method.

3.4 Outliers detection

The analysis can be affected by the presence of outliers data; they therefore needed to be detected to purify the data set from these abnormalities.

Outliers can be univariate, when the outlier data depends on only one variable (*e.g.* man high 250 cm), or multivariate, when the abnormality is not evident from only one variable, but it requires the combined observation of more than one to be detected (*e.g.* man high 180 cm, but aged 8 years old).

The most common technique to detect multivariate outliers is the Mahalanobis distance (Mahalanobis 1936):

$$D_M(\vec{x}) = \sqrt{(\vec{x}-\vec{\mu})^T S^{-1}(\vec{x}-\vec{\mu})} \qquad\qquad (3.2)$$

where $\vec{x}$ is a multivariate vector, $\vec{\mu}$ is the means-vector of data set and S is the covariance matrix.

Each subject is considered an outlier if its D_M exceed a critical value (square root of chi-squared, χ^2), which depends on the probability level set and the number of degrees of freedom: many texts (Tabachnick and Fidell 2007) suggest a prudential probability level $p < 0.001$, while the number of degrees of freedom is the number of variables; through these two values it can be obtained the critical value from a table.

Once detected outliers, it must be decided what to do, but there is not a right (or wrong) way to deal with them; if it's demonstrable that they are erroneous data, it will be legitimate to delete them and to continue the research without; in all other cases (the most), if the analyst thinks the research could be affected by the presence of the detected outliers, it could be a good idea to choose robust normalising and weighting methods.

3.5 Normalisation

Collected and simulated data must be normalised, in order to transform their measurement units (which usually are different between each indicator) into pure and dimensionless numbers, so that they can be aggregated.

Moreover, some sub-indicators have a positive influence on the phenomenon, others a negative one, so another purpose of the normalising phase is to fix a direct proportionality link between sub-indicators and the composite indicator.

Choice of the normalisation method is mainly steered by the theoretical principles that have to be respected by the Resourcefulness index (*cf.* 2.5).

Great part of the methods presented in the state of the art (Commission 2008) (Leys, Ley et al. 2013) provides a normalisation depending on a mathematical function that involves all the data; that implies that, if x_c is score of country c before normalisation and x_c^* is score of country c after normalisation, when x_c varies, all x_d^* ($d \neq c$) consequently vary.

For example, let's take into account the *Arithmetic Mean Reference* normalisation method: it is a normalisation method belonging to the greater group of *Distance to a reference methods*, which transform data as follows:

$$x_c^* = \frac{x_c}{ref} \tag{3.3}$$

This method compares the data to a reference index, that could be a mathematical operation (e.g. arithmetic mean, max) or an external or arbitrary benchmark.

It is clear that, when x_c increases (decreases), the arithmetic average (taken as ref) accordingly increases (decreases): therefore, even though x_d ($d \neq c$) has not changed, every x_d^* ($d \neq c$) will decrease (increase), because the denominator of the normalisation formula has increased (decreased).

Therefore, if x_c changes, not only RFS_c will change, but also every RFS_d ($d \neq c$): this is contrary to the Independency among communities principle (*cf.* 2.5.3).

The need for independency among communities implies that Resourcefulness of community A is computable even though data about community B are not available.

Therefore, during not only normalisation phase, but also every other step of the construction of the composite index, will must not be used arithmetical functions involving data that come from different communities; in other words, will not be filled out a $C \times N$ (*Communities × Variables*) data set, but each community will have an its own data set.

In order to allow the construction of weights and the computation of Resourcefulness over years, community's c data set will be a $Y \times N$ (*Years × Variables*) matrix.

Obviously, if RFS_{c,y^*} (Resourcefulness of country c on year y^*) wanted to be computed, data set will be limited to cases referring to years $y < y^*$.

A likely solution is to choose as normalising benchmark (for indicators which need to be normalised) an external value, called *Target Value TV*, considered as a good value for the given indicator (Cutter, Burton et al. 2010).

Each target values must be properly discussed for each variable: many indicators, especially those expressed by a percentage, may have as *TV* the maximum value which they can assume (*i.e.* 100%), others, although expressed by a percentage, must have as target a value which is not the unit (*cf.* 4.4.2), but a value which is considered optimal on the basis of the state of the art.

Normalising formula will therefore be the same shown in Equation 3.2, using *TV* as *Ref*.

The same method has been adopted in PEOPLES framework for Resilience evaluation of urban communities (Kammouh, Zamani-Noori et al. 2017), but in that case each normalised indicator cannot be higher than 1; that's the reason why, if raw indicator *x* is higher than *TV*, 1 is imposed instead of x/TV.

In our framework, on the contrary, principle of *The higher the better* has to be applied (*cf.* 2.3.1); therefore, x/TV will be used even though it is higher than 1.

Value of *TV* has to be discussed on case by case basis.

In order to allow the application of the weighting method that will be chosen, the need for a z-scores transformation will be displayed in section 3.6.2.

Z-scores method transforms data as follows:

$$x_y^* = \frac{x_y - \mu(x)}{\sigma(x)}$$

$$(3.4)$$

This technique transforms a data set with variance σ^2 and mean μ to a standardised set with variance 1 and mean equals to 0.

3.6 Weights allocation

To aggregate normalised indicators, each of them must be assigned a weight, that is a measure of its contribution to the phenomenon studied. Several weighting methods exist, but they can be summarized in three groups:

- Equal weighting: the simplest weighting technique, it gives the same weight to all variables;
- Statistical methods: they arise statistical techniques such as Principal Components Analysis (PCA) or Factor Analysis (FA), which assigns weights based on the correlation among indicators;

- Participatory methods: they assign weights based on the opinion of a group of people (experts, politicians, citizens): main techniques are Budget Allocation Process (BAP) and Analytic Hierarchy Process (AHP).

3.6.1 Equal Weighting method

Equal weighting method assigns the same weight to all indicators, so that the w_i weight of variable i will be:

$$w_i = \frac{1}{N} \tag{3.5}$$

where N is the total number of variables.

OECD (Nicoletti, Scarpetta et al. 1999) claimed that "if variables are grouped into dimensions and those are further aggregated into the composite, then applying equal weighting to the variables may imply an unequal weighting of the dimension (the dimensions grouping the larger number of variables will have higher weight)".

3.6.2 PCA/FA weighting method

PCA/FA based weighting method assigns weights to correct the overlap of information contained by the indicators; that means that weights assigned by this technique will not represent a coefficient of importance.

The algorithm exploits multivariate analysis techniques to allocate weight.

Multivariate data analysis is a set of statistical techniques which aims to analyse data sets involving more than one variable. It's especially useful to detect correlation among variables and so grouping correlated dimensions to reduce the number of variables.

Indeed, a possible trouble that can affect the reliability of a composite indicator is that one of more of its dimensions are highly-correlated, so that it will contain redundant information.

Therefore, many techniques have been developed to investigate the underlying structure of a multi-variables phenomenon.

Principal Components Analysis (PCA)

Principal Components Analysis is a multivariate technique, the object of which is "to explain the variance of the observed data through a few linear combinations of the original data" (Commission 2008).

It was firstly proposed by Pearson (Pearson 1901) and then developed by Hotelling (Hotelling 1933).

The main assumptions in this method are:

- Sufficient number of cases; many different rules of thumb have been proposed in the state of the art, all based on the cases/variables ratio: 10:1 (Commission 2008), 3:1 (Commission 2008), 5:1 (Bryant and Yarnold 1995), *etc.*;
- No outliers: OECD suggests Mahalanobis distance to reveal them (Commission 2008).

The principle is that, starting from a $x_1, x_2, ..., x_N$ variables (indicators), much of their variation can be explained by variables, called Principal Components, $Y_1, Y_2, ..., Y_N$, that are a linear combination of them and are uncorrelated, among which can be chosen a smaller number Q representing a sufficient amount of the original cumulative variance:

$$
\begin{aligned}
Y_1 &= a_{11}x_1 + a_{12}x_2 + ... + a_{1N}x_N \\
Y_2 &= a_{21}x_1 + a_{22}x_2 + ... + a_{2N}x_N \\
&... \\
Y_Q &= a_{Q1}x_1 + a_{Q2}x_2 + ... + a_{QN}x_N
\end{aligned}
\tag{3.6}
$$

ignoring $Y_{Q+1}, Y_{Q+2}, ..., Y_N$, considered negligible by the user with regards to their contribution to the cumulative variance.

Principal components are mutually uncorrelated, *i.e.* they are orthogonal:

$$
\text{cov}(Y_i, Y_j) = 0
\tag{3.7}
$$

The aim of this method is to select Q and choose the *component loadings* a_{ij}.

First step is to calculate the covariance matrix S:

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{N1} & \cdots & s_{NN} \end{bmatrix} \tag{3.8}$$

where

$$s_{ij} = \mathrm{cov}(x_i, x_j) \tag{3.9}$$

It must be noticed that S is symmetric, because $s_{ij} = s_{ji}$.

Standardizing initial data x_N (using z-scores normalisation method, *cf.* 3.5), S will be the Correlation Matrix (P), *i.e.* a matrix in which coefficients represent correlation among indicators (Pearson 1895): the higher will be the correlation between two indicators, the most they will contain similar information.

$$P = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1N} \\ \vdots & \ddots & \vdots \\ \rho_{N1} & \cdots & \rho_{NN} \end{bmatrix} \tag{3.10}$$

where

$$\rho_{ij} = \mathrm{corr}(x_i, x_j) = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_{x_i} \sigma_{x_j}} \tag{3.11}$$

is the Pearson's correlation coefficient.

As argued for S, P is symmetric too, so its eigenvalues and eigenvectors can be calculated. Each eigenvalue, solution of $\det(P - \lambda I) = 0$, will represent the percentage of variance (of the original data) involved; they have to be ordered in a decreasing scale, so that it's possible to select a group of them for which the cumulative variance is considered sufficient to represent the original data with no excessive information loss.

Many criterions on how to select eigenvalues have been purposed:

1. Kaiser criterion: select all the eigenvalues greater than 1;
2. Variance based criteria: select a sufficient number of eigenvalues to explain the 80% (or 90%) of the total variation;
3. Scree plot criteria: plot principal components on the x axis and eigenvalues on the y; select all the eigenvalues before the abrupt slop change.

Once selected, the next step is to multiply each eigenvector (from $\det(P - \lambda I) = 0$) by the square root of the corresponding eigenvalue to obtain the *Component Loadings Matrix A*. So,

$$Y = AX \qquad (3.12)$$

In order to frame the question from a geometric point of view, let, for the sake of simplicity, x_1 and x_2 be the only two variables involved in a statistical analysis, so that there will be a $\mathbb{R}^2$ (2-dimensional) space, but this reasoning should be extended to a n-dimensional ($\mathbb{R}^n$) space; data concerning the C competitors can be then represented in Figure 3-1.
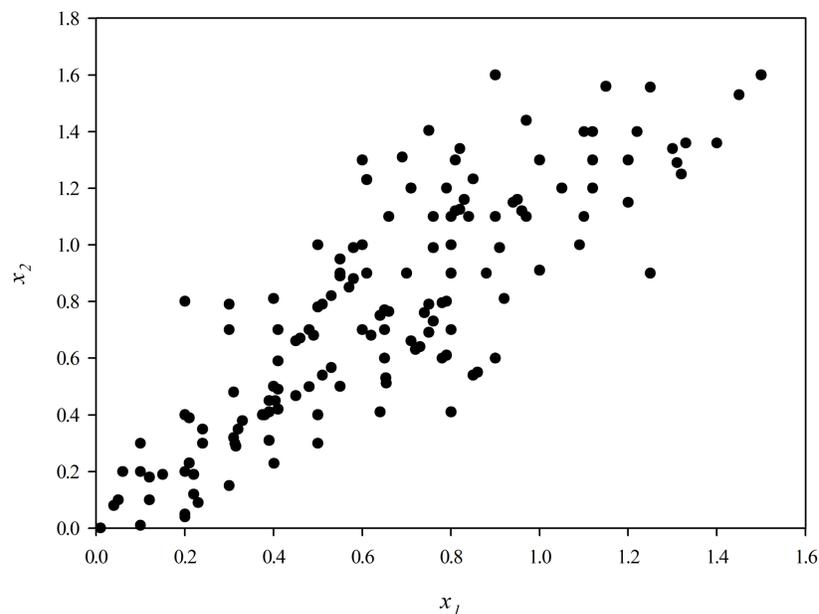

Figure 3-1: Hypothetical data distribution in a $\mathbb{R}^2$ space (*i.e.* with only two indicators)

A vector, *i.e.* the first principal component, can be identified, centred on the average and oriented in order to minimize the sum of the squared distances from the points to the line or, in other words, to maximize their variance (*i.e.* the eigenvalues of P, Correlation Matrix). Since the space is 2-dimensional, there will be a second principal component, orthogonal to the first, that will explain all the remaining variance.

Principal components can be observed as vectors in Figure 3-2, and they are the geometric meaning of eigenvectors of matrix P.

The most variance will be explained by the first principal component, the less will be the information loss neglecting the second component, reducing the question to a 1-dimensional space.
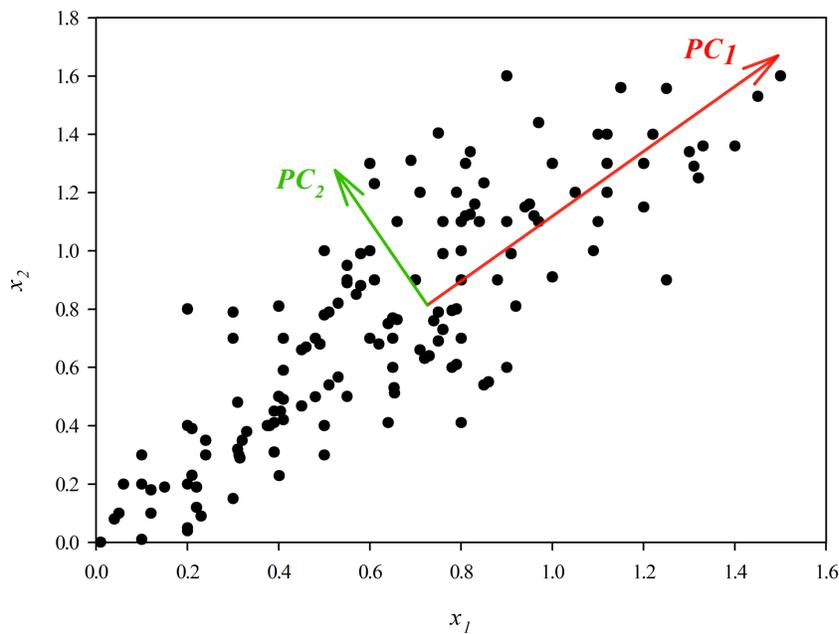

Figure 3-2: Vector representation of the two principal components of the distribution displayed in Figure 3-1; it can be seen that the first principal components explain much more variance than the second one.

For example, if the second principal component (the green one in Figure 3-2) was neglected, hypothetical distribution displayed in Figure 3-1 and Figure 3-2 would be treated as the one which would be obtained if every point was projected on the first principal component, as displayed in Figure 3-3.
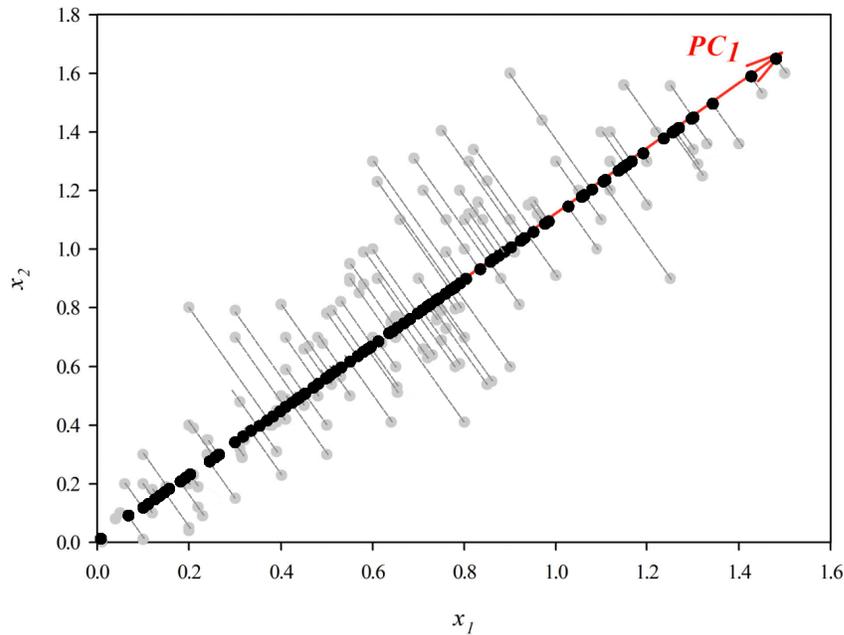
Figure 3-3: Projection of data shown in Figure 3-1 and Figure 3-2 on the first principal component; neglecting the second principal component implies loosing information related to the variance explained by that.

It's clear that principal components have a different meaning from the original variables: since they will take a quota of one variable and a quota of the another one, they will represent new different dimensions. For example, if the original axes are "weight" on the x, "height" on the y and the data distribution is approximately like the one in the Figure 3-1, the new axes, *i.e.* the principal components, will represent, respectively, "size" and "shape" scores of the data: specimens with high coefficients on the first and the second p.c. will be high and heavy, specimens with high coefficient on the second p.c. and low on the first one will be high and thin, and so on.

Getting back to the indicators issue, principal components are new indicators summing up the N variables to $Q < N$ new ones, which have a different meaning from the original ones and have to be aggregated to form the composite index.

Factor Analysis (FA)

Factor Analysis is another multivariate analysis technique, similar to PCA and thus usually confused with. There are no agreements about which of the two is preferable and whether the

methods produce radically different results, but there's no doubt they share common ground: as PCA, FA aims to express a N-variables problem by a smaller Q number of dimensions. Nevertheless, the theoretical principle from which FA comes from is that an underlying causal model exists, whereas PCA is simply a dimensions reduction technique based on linear data combinations; FA's assumption is therefore that original data x_i (with zero mean and unit variance) can be written as products of *factor loadings* α_{ij} and *common factors* F_j, plus an unobserved stochastic error e_i, where $i = 1, 2, ..., N, j = 1, 2, ..., Q$:

$$
\begin{aligned}
x_1 &= a_{11}F_1 + a_{12}F_2 + \cdots + a_{1Q}F_Q + e_1 \\
x_2 &= a_{21}F_1 + a_{22}F_2 + \cdots + a_{2Q}F_Q + e_N \\
&\cdots \\
x_N &= a_{N1}F_1 + a_{N2}F_2 + \cdots + a_{NQ}F_Q + e_1
\end{aligned}
\tag{3.13}
$$

Several approaches exist to extract factors, and principal components are usually used for this purpose; in that case the first part of the algorithm is the same as that shown for PCA. Once selected the factors, *i.e.* the principal components, the factor loadings matrix has to be rotated to enhance its interpretability; the most used rotation technique is the *varimax* rotation (Kaiser 1958), the effect of which is that factor loadings (which are in the range $[-1,1]$) will be closer to 1 or 0 or -1, depopulating the two intermediate regions.

Varimax rotates factors to maximize the sum of the variances of the squared factor loadings, saving however their orthogonality, *i.e.* their uncorrelation.

Thus, will be maximized the quantity:

$$
V = \sum_{f=1}^{Q} V_f = \sum_{f=1}^{Q} \left[\sum_{i=1}^{N} \alpha_{if}^4 - \frac{1}{N} \left(\sum_{i=1}^{N} \alpha_{if}^2 \right)^2 \right]
\tag{3.14}
$$

Therefore, the rotate factor loadings matrix will be:

$$
F_{VARIMAX} = \arg\max_{F} \sum_{f=1}^{Q} \left[\sum_{i=1}^{N} \alpha_{if}^4 - \frac{1}{N} \left(\sum_{i=1}^{N} \alpha_{if}^2 \right)^2 \right]
\tag{3.15}
$$

After the rotation, each factor will have only few variables with large factor loadings, simplifying its interpretability.

Since varimax rotation doesn't change the amount of total variance explained by each factor, the analytical solutions obtained with or without the rotation will be basically the same.

Nicoletti (Nicoletti, Scarpetta et al. 1999) suggested grouping variables in intermediate indicators, composed by the ones that match, in each factor, the most of variance, *i.e.* the ones with the highest factor loadings.

For examples, if rotated factor loadings matrix is that in Table 3-1, the best way to create intermediate indicators is:

$$ii_i = 0.27x_3 + 0.17x_5 + 0.23x_6$$
$$ii_2 = 0.66x_3 \qquad\qquad (3.16)$$
$$ii_3 = 0.27x_1 + 0.17x_4$$

Table 3-1: Example of factor loadings

Indicator	Factor loadings			Scaled squared factor loadings		
	F_1	F_2	F_3	F_1	F_2	F_3
x_1	0.01	0.07	0.55	0.00	0.00	0.27
x_2	0.13	0.02	0.08	0.01	0.00	0.01
x_3	0.85	0.98	0.04	0.27	0.66	0.00
x_4	0.21	0.11	0.44	0.02	0.01	0.17
x_5	0.67	0.08	0.00	0.17	0.00	0.00
x_6	0.79	0.20	0.01	0.23	0.03	0.00
Explained Variance	1.86	1.02	0.58			
Explained Variance/Var.tot.	54%	30%	17%			

It must be noticed that 6 original variables have been reduced to only 3 intermediate indicators, whose meaning has to be identified by the researcher.

Intermediate indicators can be then aggregated on the basis of how much variance each one explains on the total variance:

$$CI = 0.54\ ii_1 + 0.30\ ii_2 + 0.17\ ii_3 \qquad\qquad (3.17)$$

A logical consequence of this method is that it's sensitive to data update or change: new observations or corrections will force the researcher to repeat the analysis to recalculate the weights; moreover, imagining the presence of any outliers in the Figure 3-1, it's easy to understand the need for a prior analysis to detect them and to purify the original data.

3.6.3 Participatory methods

Budget Allocation method (BAP)

Budget Allocation is a participatory method: a group of people considered experts is selected and they have to allocate a *P* points budget to the indicators, depending on the contribution they think the dimensions have to the composite indicator.

The obvious matter is how to select the group of experts and OECD User Guide (Commission 2008) argued that this method manifests troubles when the indicators are more than 10, because of cognitive stress given to the experts.

Analytic Hierarchy Process (AHP)

AHP is a method included in the group of techniques for multi-attribute decision making (Wind and Saaty 1980).

Main principles are:

- Weights based on pair-wise comparison among variables;
- Comparison conducted by experts;
- Weights as trade-off coefficients.

Expert will develop a pair-wise comparison choosing a value in a scale (usually $0 \div 9$); the higher is the value, the most the expert thinks variable x_1 is more important than x_2. Points are put in a *Comparison Matrix A*, where, obviously, $A_{ij} = A_{ji}^{-1}$; for example:

	x_1	x_2	x_3	x_4	x_5
x_1	1.00	3.00	0.50	0.50	3.00
x_2	0.33	1.00	0.20	0.33	2.00
x_3	2.00	5.00	1.00	6.00	0.20
x_4	2.00	3.00	0.17	1.00	4.00
x_5	0.33	0.50	5.00	0.25	1.00

This kind of comparison is subject to consistency issues: indeed, who makes the comparison could prefer x_1 to x_2, x_2 to x_3 but x_3 to x_1; in order to solve this problem, weights are computed by an eigenvector algorithm, iterating the process and verifying every time the size of a Consistency Ratio, until this reaches a value considered reasonable.

3.6.4 When to use what

A table to answer "when to use what" for weighting methods, depending on the aim of the research, has been purposed about Ecological Indicators construction (Gan, Fernandez et al. 2017), but it can be generalized and displayed in Table 3-2:

Table 3-2: Suitability of each weighting method for different purposes and scales (Gan, Fernandez et al. 2017). REC, OK and NO mean, respectively, recommended, applicable and not to be used.

Weighting	Purposes		Scales			
	Assessing state/ prediction	Comparison	Spatial scales		Temporal scales	
			Coarse	Fine	Long	Short
EW	REC	OK	REC	REC	REC	REC
PCA/FA	REC	OK	REC	OK	OK	REC
BAP	REC	OK	NO	REC	REC	REC
AHP	REC	OK	NO	REC	REC	REC

It can be useful to comment on the reason why participatory weighting methods have to be preferred at fine spatial scales, whereas statistical methods must be chosen when the spatial scale is coarse: weights allocated by public participation methods might be influenced by local conditions when spatial scale is too large (Gan, Fernandez et al. 2017).

On the other hand, when the spatial scale is finer, *e.g.* sub-national, opinion of politicians and government officials should be taken into account (Van de Kerk and Manuel 2008).

Discussion about spatial scale for Resourcefulness composite index will be deepened in section 0.

3.6.5 Weighting method for Resourcefulness composite index

First of all, equal weighting method is excluded mainly for two reasons:

1. As observed in section 3.6.1, equal weighting applies unequal weighting if sub-indicators are aggregated in dimensions before being aggregated in the final composite index: although, at present, this is not provided for, it is not excluded that the framework may be changed in the future, maybe after having selected a complete and exhaustive list of indicators, each one precisely located in a dimension;

2. Equal weighting method does not allow to prevent overlapping information among indicators.

At this point, the fundamental choice about the weighting allocating strategy is between data-driven approaches and opinion based methods.

PEOPLES framework, for example, allocates weights basing on an interdependency matrix, filled out by an expert (or a group of experts) who assigns 1 if he thinks indicator in the row depends on the indicator in the column; that is done to "prevent possible overlap among the indicators" (Kammouh, Zamani-Noori et al. 2017).

The principle to allocate weights in Resourcefulness framework has to be same: indeed, indicators could contain information already contained in other indicators (they are correlated) and, if this overlap is not removed, the final composite index will be affected by this.

Nevertheless, opinion-based method used in PEOPLES framework does not seem to be suitable in our case; reasons are the following ones:

1. Indicators in PEOPLES framework are mainly statistic data representing tangible dimensions, so that it is possible to select a (or more) expert to evaluate the interdependency among indicators; for example, an economist could have an authoritative opinion about the interdependency between *Income* and *Occupation*, or an environmental scientist between *Air quality* and *Water quality*; anyway, the state-of-the-art about Resilience and its dimensions is rich and it can be used to fill out interdependency matrix.

   On the contrary, Resourcefulness's dimensions are often intangible; it would be a stretch to claim that it could be selected a (or more) expert to evaluate the interdependency among indicators;

2. Resourcefulness is an inherent feature of communities and it must not change to the changing of people's views;

3. Opinion (of experts or public) about weights may change across communities, so that the choice about subjective weights would imply the issue of experts (or public) of what community have to be engaged and why.

Now the specific method to assign objective weights has to be chosen.

As described above, framework's goal is to prevent overlap among indicators; it's therefore a good idea to assign higher weights to those indicators which do not seem to be highly correlated with other indicators, *i.e.* do not contain information already contained in other indicators.

The most proper method to assign weights is therefore one which refers to principal components analysis; it is necessary, however, to remember that weights assigned through this method are not coefficients of importance (*cf.* paragraph 2.2.5), but factors that correct overlapping information and maximize variance of starting data explained by the composite index.

After having carried out the principal components analysis accordingly to what described in section 3.6.2, Q principal components are selected following both Kaiser criterion and a variance-based criterion; scree plot criterion has been excluded because considered impractical for the purposes of the research; all eigenvalues greater than 1 are selected and total variance explained has to be at least equal to 75%.

Once selected Q principal components, factor loadings of reduced matrix [D*] can be rotated through *varimax* rotation to enhance the interpretability of factor loadings (*cf.* 3.6.2):

$$D^* = \arg\max_D \sum_{j=1}^{Q} \left[\sum_{i=1}^{N} D_{ij}^4 - \frac{1}{N} \left(\sum_{i=1}^{N} D_{ij}^2 \right)^2 \right] \tag{3.18}$$

Previous operation is iterative, so the number of iterations must be enough to fill out the rotated matrix D^*.

Finally, weights are computed as follows:

$$w_i = \frac{\sum_{i=1}^{N}\sum_{j=1}^{Q} D_{ij}^{*2}\cdot\lambda_j}{\sum_{j=1}^{Q} D_{ij}^{*2}\cdot\lambda_j} \qquad\qquad (3.19)$$

To conclude, it is important to note that using $Y \times N$ data sets implies both community-specific and year-specific: different communities will assign different weight to the same indicator and this is also consistent with the principle of independency among communities (*cf.* 2.5.3); moreover, each community will assign different weight to the same indicator every year: this is due to a process of refinement, for which it stands to reason that variation of weights decreases over the years, as number of cases (*i.e.* years) increases.

3.7 Aggregation method

Last step in constructing a synthetic indicator is to choose an aggregation technique; three main methods are purposed in the state-of-the-art (Commission 2008):

- Additive aggregation;
- Geometric aggregation;
- Non-compensatory multi-criteria approach (MCA).

Non-compensatory multi-criteria approach will not be shown because it is effective only in analysis targeted to comparison among competitors (Gan, Fernandez et al. 2017).

Additive aggregation

It's the simplest aggregation method and it's based on the following equation:

$$CI = \sum_{i=1}^{N} x_i^* w_i \qquad\qquad (3.20)$$

It's easy to understand that the main issue correlated to this method is that weights assume the meaning of trade-off coefficients: a country could obtain a satisfying CI although a bad score in many variables if it can rely on adequately high score in other indicators.

Competitors could therefore adopt a compensatory logic for which it's acceptable a lack in a variable if another one is sufficiently strong.

For example, if country A's score is 25, 25, 25, 25 in the four sub-indicators considered and country B's one is 1, 1, 1, 97, supposing just for simplicity equal weighting, A and B will obtain the same $CI = 25$ score, although A manifest better balance in the sub-indicators.

Geometric aggregation

Geometric aggregation solves part of the full-compensability manifested by additive aggregation method:

$$CI = \prod_{i=1}^{N} x_i^{*w_i} \qquad (3.21)$$

Resuming the previous examples, country A would obtain again 25, but B would see its score decreasing from 25 to only 3.13. Moreover, B would enhance its composite indicator much more improving variables in which it lacks than improving its "warhorse".

3.7.1 When to use what

A table to answer "when to use what" for aggregation methods, depending on the aim of the research, has been purposed about Ecological Indicators construction (Gan, Fernandez et al. 2017), but it can be generalized and displayed in Table 3-3:

Table 3-3: Suitability of each aggregation method for different purposes and scales (Gan, Fernandez et al. 2017). REC, OK and NO mean, respectively, recommended, applicable and not to be used.

Aggregation	Purposes		Scales			
	Assessing state/ prediction	Comparison	Spatial scales		Temporal scales	
			Coarse	Fine	Long	Short
Linear	REC	OK	REC	REC	REC	REC
Geometric	REC	OK	REC	REC	REC	REC

Since the purpose of the framework is not to compare communities among them, but assess resourcefulness (or predicting its future evolution) of a community (also considering section 2.5.3), there are no limitations in choosing between linear and geometric aggregation method.

3.7.2 Aggregation method for Resourcefulness composite index

Case of Resourcefulness seems to be one of those in which one indicator can compensate for others' weakness: for example, as described in section 2.4.3, during the 921 Taiwan Earthquake (1999), *Hakka spirit* partly mitigated the lack of preparedness to such an event. More specifically, the idea of compensability is actually inherent in the concept of Resourcefulness: indeed, communities marked by a lack in some kind of resources can be able to cope with natural disasters better than one could expect tapping into more intangible resources.

There's no doubt, therefore, that the most proper aggregation method for Resourcefulness composite index is the additive one.

3.8 Algorithm

The mathematical procedure is given below:

Given a $Y \times N$ (Y years, N variables) raw matrix of data $[X]$, x_{yj} will be the indicator j score of community c on year y.

If $[X]$ is not complete, it is transformed as follows:

1. Indicators with more than 75% missing data are excluded from the analysis;
2. Years with more than 50% missing data are excluded from the analysis.

Remaining missing data are imputed through the following method:

1. Each $x_j - x_{i \neq j}$ is plotted;
2. R^2 of each plot is obtained;
3. If $R^2 \geq 0.5$, x_i is considered a good regressor for x_j; otherwise x_i is not considered a good regressor for x_j;
4. Missing data of indicators which have at least one good regressor are imputed through multiple linear regression, taking the average of the different imputations;
5. Missing data of indicators which don't have good regressors are imputed through Monte-Carlo simulation.

6. Complete data matrix $[X]$ is obtained.

If needed, x_{yj} is divided by its corresponding *TV*.

$[X^*]$, normalized matrix of data, is obtained accordingly to the z-scores method:

$$x_{yj}^* = \frac{x_{yj} - \frac{1}{y}\sum_{k=1}^{y} x_{kj}}{\sqrt{\frac{1}{y}\left(\sum_{k=1}^{y}(x_{kj} - \frac{1}{y}\sum_{k=1}^{y} x_{kj})\right)}}$$

(3.22)

Matrix [P] of Bravais-Pearson correlation coefficients between the normalized starting data can now be computed; each element ρ_{ij} of [P] will be obtained as follows:

$$\rho_{ij} = \mathrm{corr}\left(x_i^*, x_j^*\right)$$

(3.23)

$$\rho_{ij} = \frac{\mathrm{cov}\left(x_i^*, x_j^*\right)}{\sigma_{x_i^*}\sigma_{x_j^*}}$$

(3.24)

$$\rho_{ij} = \frac{\sum_{k=1}^{N}\left(x_{ik}^* - \overline{x_i^*}\right)\left(x_{jk}^* - \overline{x_j^*}\right)}{\sqrt{\sum_{w=1}^{N}\left(x_{iw}^* - \overline{x_i^*}\right)^2 \sum_{z=1}^{N}\left(x_{iz}^* - \overline{x_i^*}\right)^2}}$$

(3.25)

Where

$$i, j \in \mathbb{N} : 0 < i, j < N$$

(3.26)

They have to be now calculated eigenvalues and eigenvectors of matrix [P]: have to be found, in other words, coefficients λ and vectors d such that they satisfy the following equation:

$$Pd = \lambda d$$

(3.27)

Eigenvalues λ are organised in a vector $\{\lambda\}$, eigenvectors d by column in a matrix [D].

$Q < N$ principal components are selected following the Kaiser criterion:

$$Q \in \mathbb{N} : Q < N, \; \lambda_i \geq 1 \; \forall i \leq Q \tag{3.28}$$

Once selected Q principal components, factor loadings of reduced matrix $[D^*]$ can be rotated through *varimax* rotation (if rotation is chosen, computational cost could be high):

$$D^* = \arg\max_D \sum_{j=1}^{Q} \left[\sum_{i=1}^{N} D_{ij}^4 - \frac{1}{N} \left(\sum_{i=1}^{N} D_{ij}^2 \right)^2 \right] \tag{3.29}$$

Previous operation is iterative, so the number of iterations must be enough to fill out the rotated matrix D^*.

Weights that have to be assigned to each variable will be:

$$w_j = \frac{\sum_{i=1}^{N} \sum_{j=1}^{Q} D_{ij}^{*2}}{\sum_{j=1}^{Q} D_{ij}^{*2}} \tag{3.30}$$

Finally, Resourcefulness will be:

$$RFS_{c,y} = \sum_{i=1}^{Q} x_{yj} \cdot w_j \tag{3.31}$$

Flow chart of the algorithm is displayed in Figure 3-4 and MATLAB spreadsheet of the is given in Appendix A.
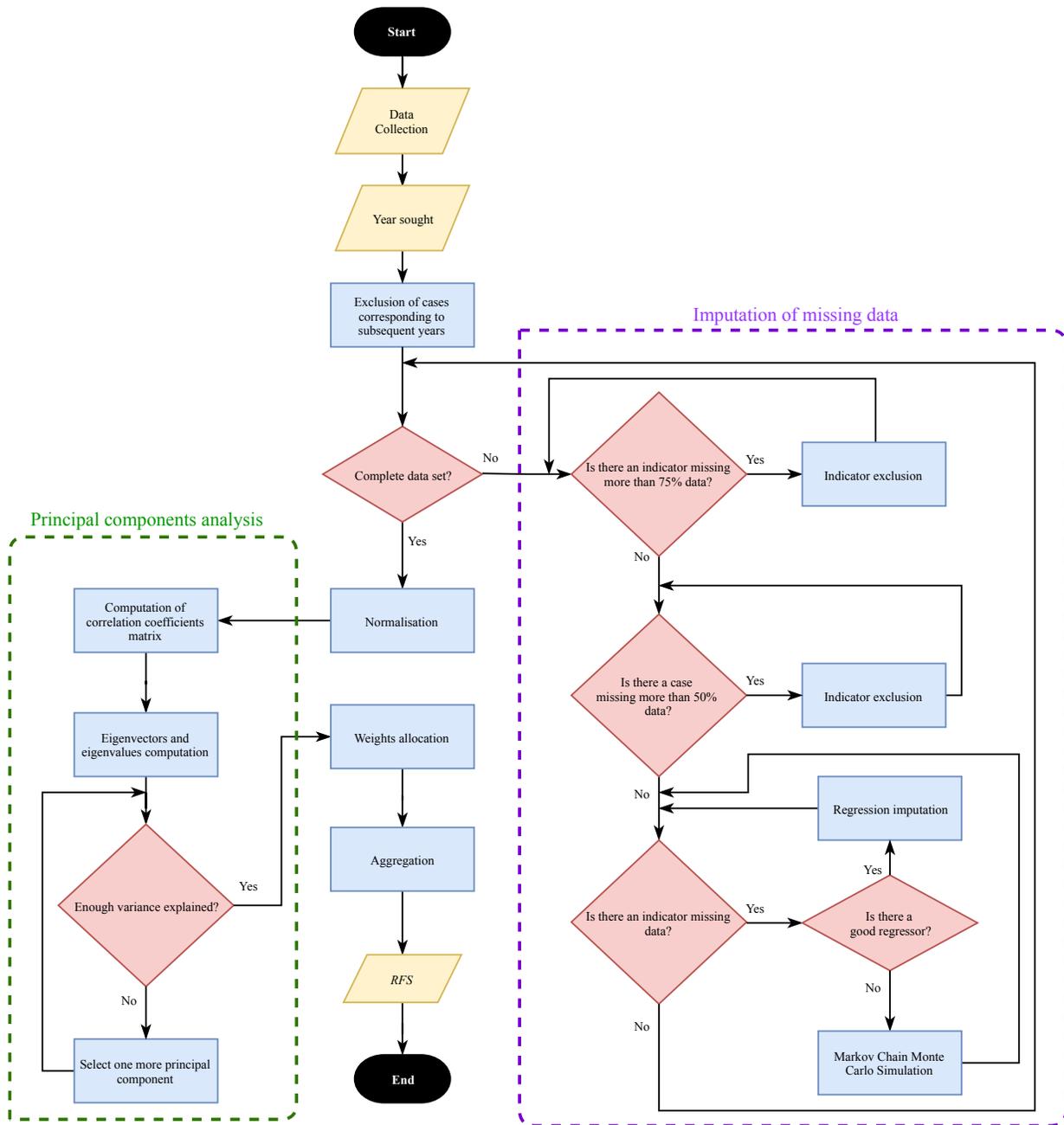
Figure 3-4: Flow chart of the algorithm developed to compute Resourcefulness composite indicator

3.9 Resourcefulness composite index software

The algorithm can be used to develop an executable software file to compute *RFS* composite index of a given community.

A code has been written by using of Visual Basic .NET programming language, with all the benefits of the Framework .NET (interoperability among .NET programming languages, such as C#, F# etc., common libraries, and so on).

It could be easily implemented in software in the field of Resilience written in a compatible programming language, or used as external executable file to compute *RFS* index.

Steps of the algorithm described in section 3.8 provided for in the code are displayed in Table 5-3.

Table 3-4: Steps of the developed algorithm provided and not provided for in the software

Steps provided	**Steps not provided**
- Years/indicators elimination on the basis of the filling percentage - Imputation of missing data through regression; - Normalisation; - Weights allocation through PCA; - Aggregation	- Imputation of missing data through MCMC simulation; - Indicators transformation

The code allows the user to import .xls, .xlsx, .csv, .txt files containing starting data set to the main form (Figure 3-6); as displayed in Table 3-4, transformation of indicators polarity is currently still not provided for.

User is forced to follow the algorithm developed to impute missing data and, more generally, to follow the order of the whole algorithm; indeed, he cannot compute weights if missing data have not been yet imputed, he cannot impute missing data before having deleted any cases/variables (if needed), and so on.

At first, he has to access to a form (Figure 3-6) in which years and indicators filling percentage are displayed in two tables, setting as to be removed those which are below the threshold described in section 3.3; user could also be free to not use default thresholds and to custom them.
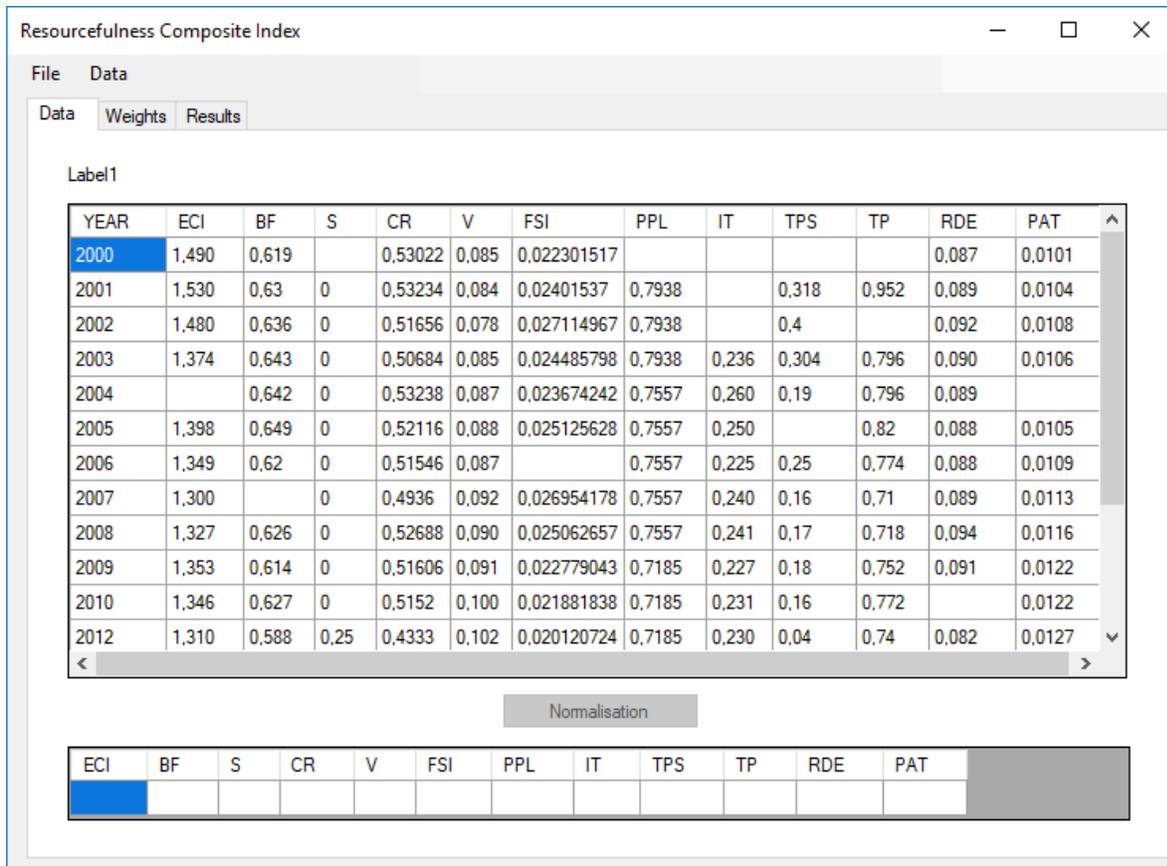
Figure 3-6: Main form of the software
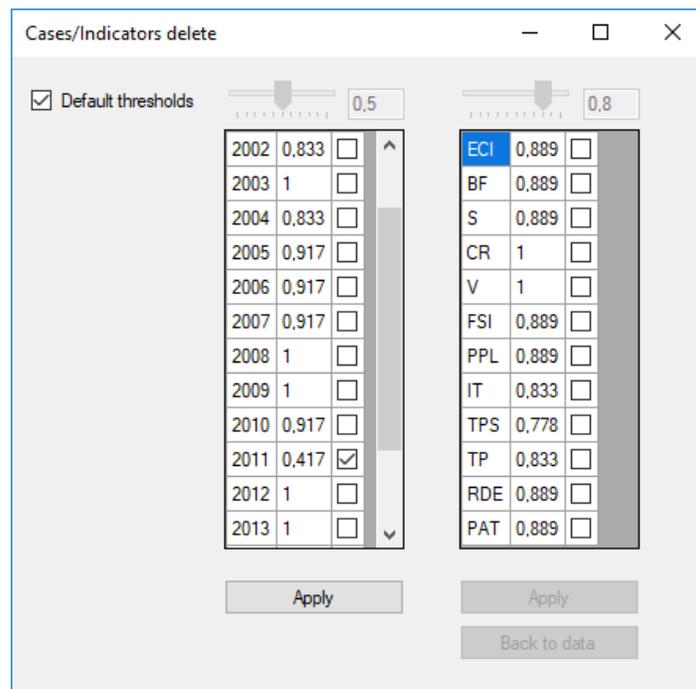
Figure 3-5: Form in which the user has to delete, if needed, cases and variables with insufficient amount of data; threshold can be changed manually

If any missing data remain to be imputed, user has to access to a second form (Figure 3-7), in which the software looks for the best regressor for each indicator, plotting R^2 of each variable with all the others; automatically, best regressor for each indicator is considered the one to which corresponds the highest R^2 coefficient and it's displayed in a table on the left of the form. Here again, user is allowed the freedom to not use the automatic R^2 threshold and to move it by using a scroll bar and to not follow automatic regressors selection and to set them manually. Clicking on "Regression" button will impute missing data.
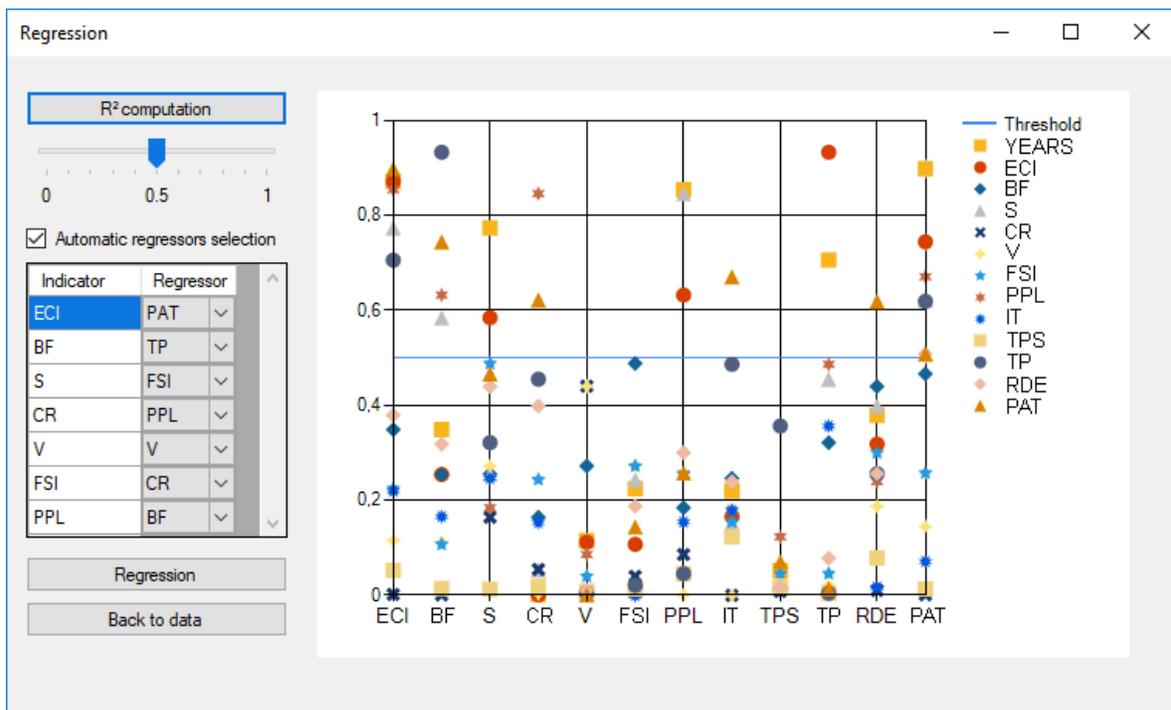

Figure 3-7: Form to look for best regressor for each indicator on the basis of R^2 coefficient; threshold above which indicators are considered good regressor for the given indicator can be set manually.

At this point, completely filled data set has to be normalised, setting manually the *Target Values* and then clicking on "Normalisation" (Figure 3-8).
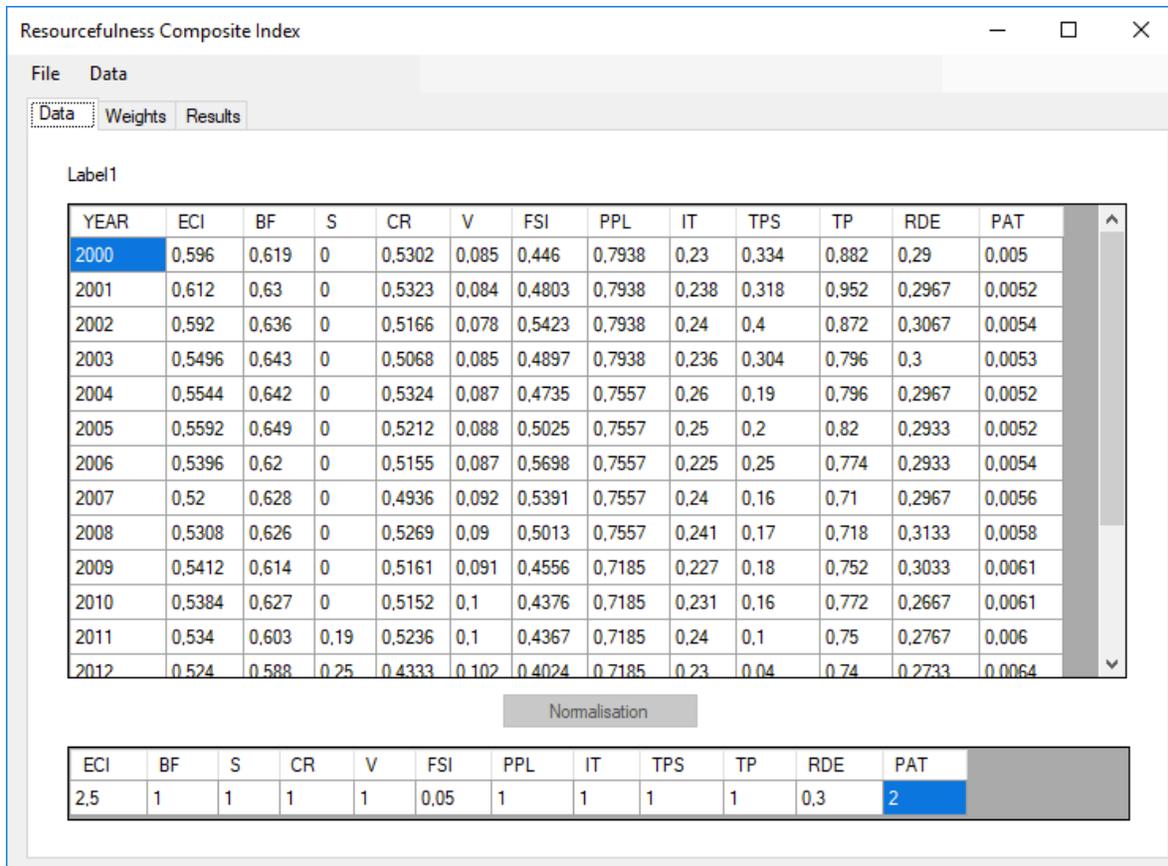
Figure 3-8: After having imputed missing data, normalisation is allowed in the main form

The user can now move to the weights allocation tab page (Figure 3-9), in which he can use Principal Components Analysis (which is automatically set) or he can set manually the weight for each indicator; in that case, the code will ensure that the sum of weights is equal to 1.

Weights will be plotted as bars on the right of the form.
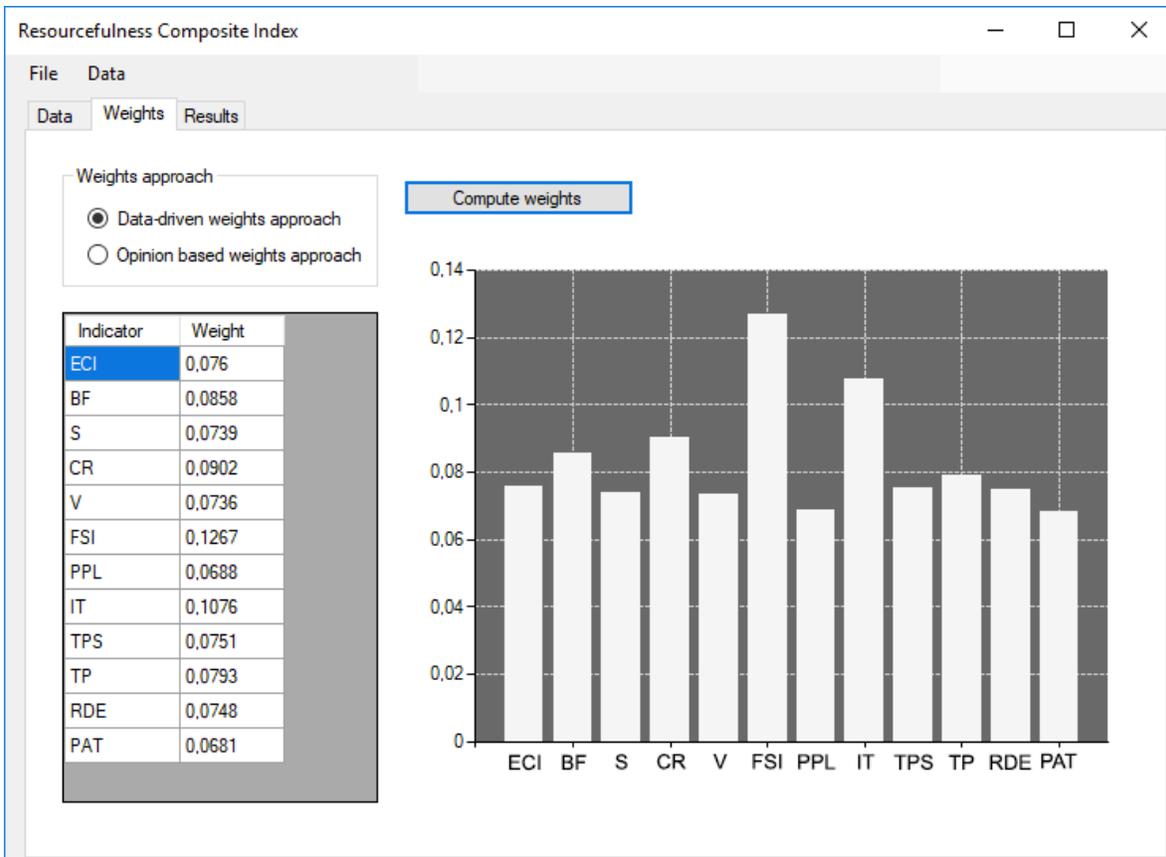
Figure 3-9: By using of Principal Components Analysis, the code allocates weights to each indicator, displayed as numbers in the table on the left and as bars in the plot on the right; user can also set weights manually.

After having allocated weights, in the Results tab page data can be aggregated and the results are obtained: *RFS* composite index is displayed and area plot too (Figure 3-10), compared to the ideal area plot, but this issue will be better dealt with in section 5.3.1.
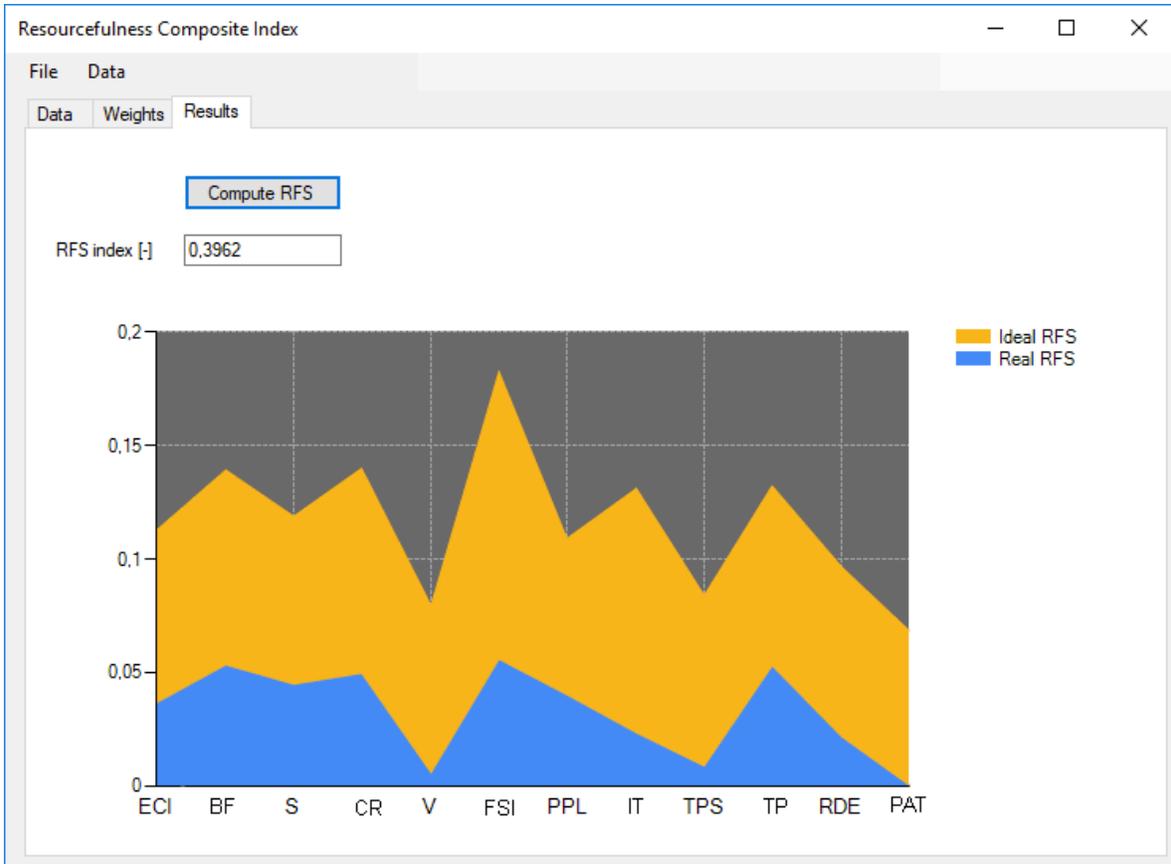
Figure 3-10: After having imputed missing data, normalised data and allocated weights, data can be aggregated and results are obtained; *RFS* is displayed as number and as comparison between real RFS area plot and ideal RFS area plot (*cf.* section 5.3.1)

4 Indicators selection

Selection of indicators will be discussed in the present chapter mainly following two purposes:

1. Covering as much as possible the whole set of features which are part of Resourcefulness, taking into account the possible different facets of each dimension;
2. Ensuring equity among different dimensions of Resourcefulness, *i.e.* avoiding as much as possible theoretical overlap among sub-indicators.

In this respect, dimensions of Resourcefulness can be summed up as follows:

- **Political-economic**: support provided by the economic and political structure to the emergency management system, referring only to Resourcefulness' own issues;
- **Preparedness**: disaster-preparedness of individual citizens and the whole community, intended as knowledge about the risk and what it has to be done in a disaster's aftermath;
- **Trust**: ability of a community to cope with natural disasters as a cohesive unit, tapping into its trust resources (*e.g.* interpersonal trust, trust in the government);
- **Creativity**: ability of a community to take smart and not obvious decisions during the emergency, getting as result a mitigation of losses.

Also in PEOPLES framework sub-indicators are aggregated in higher dimensions; that is done both as a proper step of the theoretical algorithm to compute Resilience and to highlight any weaknesses of the community.

In our case, as will be seen in the following paragraph, some indicators have one foot in a dimension and one in another one; it would be therefore far-fetched to use such a classification in the constructing process of the composite index, whereas remains valid its usage to pointing out flaws.

4.1 Political-economic indicators

4.1.1 Economic Complexity

Dependence of economy on few products can slow down, or even stop, the recovery following a natural disaster, especially when the sector that has been most hit is just the leading one.

That often happens when the economy of the considered community is driven by a product (or few products) belonging to the primary sector: the collapse in the availability of a raw material, or difficulties in obtaining it, can cause an enormous variation of its prices, so that importers move towards other markets; consequent economic crisis can be very hard to solve, despite attempts of intervention by the institutions.

A well described example (Cutter, Emrich et al. 2006) is the case of Bayou La Batre, Alabama: based almost exclusively on seafood like shrimps and oysters, its economy has been devastated, on August 2005, by Hurricane Katrina, which destroyed more than half of the shrimp boats fleet; seafood economy has therefore moved to imported catches, against low prices of which Bayou La Batre cannot compete.

For these reasons, it is appropriate to introduce in the Resourcefulness framework an indicator whose value is inversely proportional to the risk that the economy of the community collapses after a natural disaster.

A good source is the *Economic Complexity Index*, developed by the Observatory of Economic Complexity, which measures every year the knowledge intensity of the products exported by each country through a eigenvalues methodology using as data the quantity (in USD) of a product p that a country c exports.

Numerical range of Economic Complexity Index is $(-\infty, +\infty)$, but its value is mostly between -2.5 and +2.5.

4.1.1 Bureaucracy Flexibility

Red tape can constitute a tricky obstacle to overcome in the aftermath of a natural disaster.

For example, following a natural hazard event almost always the number of requests for exemptions from buildings regulations significantly increases (Arendt and Alesch 2014).

It's widespread in the Asian Southeast the combination of high corruption index and rigid red tape: in the Philippines, for example in the aftermath of Typhoon Ketsana, "the rigid

regulatory requirements attending housing and land development contribute significantly to greater costs and increased inaccessibility of low-cost housing" (Ballano 2017).

At a finer community-scale, as described in the case study of Chile Earthquake 2010 (*cf.* 1.3.4), many times hospitals need a readjustment of their bureaucratic procedures to allow an efficient intervention despite the challenges that a natural disaster throws.

With regards to country-communities, bureaucracy flexibility can be evaluated through the *Index of Economic Freedom*, developed by the Heritage Foundation: it gives a numerical value to the freedom of citizens to work, produce, consume and invest in any way they please. It is obtained averaging (with equal weights) Rule of Law, Government Size, Regulatory Efficiency and Open Markets (each one evaluated between 0 and 100), so that its numerical range is $[0, 100]$.

4.1.1 Fragility

Exposure to natural disasters can threaten the global stability of the community, in terms of institutions and relationships with other communities, often even enhancing the risk of civil conflict; this cause-and-effect may also been reversed, so that fragility can much worse the community's response to a disaster, so that the relationship between vulnerability and fragility can be seen as two-way. The aetiology is rarely easy to prove, but this subject has been well debated (Nel and Righarts 2008); moreover, civil conflicts in states hit by natural disaster seem to take longer (Eastin 2016).

After the Indian Tsunami in 2004, civil conflict in Sri Lanka between the government and the secessionists of Liberation Tigers of Tamil Eelam (between which was granted a cease-fire), took new strength after the disaster (Billon and Waizenegger 2007).

How to quantify community's fragility would be an interesting issue, but at country-scale a good way is to use the *Fragile States Index*, a composite index developed by The Fund For Peace, which uses cohesion, economic, political and social indicators; it also implements indicators like demographic pressures, refugees and external intervention, obtaining then the global score by summing up the scores of each indicator.

Considering that each indicator is ranged between 0 and 10 and that there are 12 indicators, the numerical range of Fragile States Index is [0,120].

The index has, though, negative polarity, since it allocates high global scores to the most fragile countries; it's therefore necessary to transform the index into a positive-polarity indicator by reversing it. So, its numerical range will be [0.00833,1].

As for Safety Rate, the borderline case (practically only ideal) of 0 score in Fragile States Index, which would report arithmetical error, is solved by imposing 1 before the transformation.

4.1.1 Mitigation Spending

Although it disregards how much efficiently they are spent, it is clear that communities that allocate more funds to mitigate future hazards are more likely to suffer less damage than other communities: for example, benefit-cost ratio (BCR) of 4000 mitigation investments has been quantified equal to 4 by FEMA (Council 2005).

The forward-seeing ability of the institutions to allocate smart and sufficient funds can much improve the response in respect of disasters, whereas it has been argued that relief funds do not mitigate future damage (Healy and Malhotra 2009).

Despite the aptness of Mitigation Spending as Resourcefulness indicator, it is faced with inherent difficulties about its definition; indeed, funds allocated to mitigate future disasters usually come from different sources: they can be provided by the local, state or federal government, by higher institutions (*e.g.* the European Union), by volunteer organizations and so on; moreover, it's rare that mitigation funds are rigorously tracked[3].

Nevertheless, Mitigation Spending is measured as the sum of the funds allocated to cope with natural disasters as percentage of GDP: its numerical range is [0,1].

4.1.1 Safety Rate

Described example of 2010 Chile Earthquake (*cf.* 1.3.4) well shows the importance of crime rate in the response of a community coping with a natural hazard event. In fact, not only crime episodes after a disaster can cause damage in terms of economy or human lives, but also it may lead to a waste of resources which would be targeted to the recovery, *e.g.* law enforcement officers.

Another example are the looting episodes following Nepal's 2015 earthquake, after which many damaged sacred temples were looted to steal sacred art items (Yates and Mackenzie 2018).

---

[3] The Pew Charitable Trusts, *Natural Disaster Mitigation Spending Not Comprehensively Tracked*, Web, Accessed on December 14, 2018

After the earthquake of Bam, Iran (2003), although people's reports should not be accepted uncritically, looting was consistent both in damaged houses and on rescued dead bodies (Ibrion, Parsizadeh et al. 2015).

Although it's not easy to find a numerical indicator for safety rate, we can refer to its opposite crime rate, subsequently reversed.

As crime rate index can be chosen the yearly number of reported violent crimes, numerical range of which is $[0, +\infty)$; the latter has to be reversed to obtain the Safety Rate, ranged between $[0, 1]$. The borderline case (in practice only ideal) of 0 reported violent crimes, which would report arithmetical error after the transformation, is solved by imposing 1 report violent crime.

4.1.1 Participation in Public Life

Citizens' participation in public life of the community is important both in the mitigation and in the recovery phase; it stands to reason that people actively interested in public life are likely to better know the emergency plan and the risks which they are exposed to; they will also efficiently interact with other members of the community and with emergency managers during the emergency and the recovery phases.

Importance of community participation and opportunity for institutions to take measures to enhance it has been well discussed by World Health Organization (Organization 2002).

For example, during the Japan's March 2011 earthquake, social capital has allowed to rescue many people and to offer aid in terms of financial and non-financial resources (Aldrich and Meyer 2015).

Van Krieken (Van Krieken, Kulatunga et al. 2017) has stated that "empowering the community and maximizing the community's participation at the local level […] means tapping into the community's resourcefulness"; even the issue of what kind of strategies have to be adopted to empower the community has been explored (Kieffer 1984) (Paton and Long 1996).

Measurement of participation in public life is made, in this framework, through the percentage of population voting at the last presidential elections; it might seem a rough methodology to quantify it, but other aspects of social life are already contained in other indicators (Disaster Preparedness, Interpersonal Trust, Trust in the Political System, Trust in the Police, Patriotism), so that it guarantees a safe measurement of interest of the people in the interaction with public life.

Years in which elections did not take place are filled with the data referring to the most recent year in which elections took place.

Its numerical range is, therefore, [0,1].

4.2 Preparedness indicators

4.2.1 Smartphone Penetration

A crucial issue during the emergency phase following a natural disaster is the handling of information flow: an efficient and clear sharing of information and directives can considerably improve the response of the whole system.

In the last years the main medium to share information is the smartphone and emergency managers, above all of the most developed countries, are promoting its usage[4], and often of wearable technologies, to mitigate the aftermath of a natural disaster.

For example, during Hurricane Harvey flooding, many rescuers communicate through Twitter with people who were not still rescued to locate them or to inform them about what to do[5].

The most obvious and logical measurement of this issue is the smartphone penetration, *i.e.* the percentage of smartphone owners and users.

Obviously, data about smartphone penetration are not available (or always equal to 0) before smartphone started to spread.

Its numerical range is $[0, 1]$.

4.2.1 Disaster Preparedness

Disaster preparedness is a key issue for Resilience nowadays, because it has been widely observed that a good emergency plan, together with the knowledge of every actor of which is the role he has to play, can much mitigate the immediate aftermath of a natural disaster (Paton and Long 1996).

Preparedness is often at the base of the difference in disaster response: for example, even to disaster preparedness has been attributed the radically different behaviour of Chile and Haiti dealing with 2010 earthquakes [6].

---

[4] Federal Emergency Management Agency (FEMA) has developed a smartphone app to receive real-time alerts, to learn emergency safety tips, to locate open emergency shelters and to find disaster recovery centers.

[5] Nace T., December 15, 2017. *How Technology is Advancing Emergency Response and Survival During Natural Disasters*, Forbes, Web. Accessed on December 11, 2018.

[6] *Earthquake preparedness: what the United States can learn from the 2010 Chilean and Haitian earthquakes*: hearing before the Ad Hoc Subcommittee on State, Local, and Private Sector Preparedness and Integration of the Committee on Homeland

The most logical way to quantify disaster preparedness is the percentage of population which have an emergency plan; numerical range is, therefore, [0,1].

4.2.2 Emergency Kit Preparedness

Complementary to the previous Disaster Preparedness indicator can be seen the practical foresight of each household, in terms of prediction of what would be necessary to survive until the arrival of the rescuers if a disaster happens.

Emergency kit is one of the factors (at personal, household and community level) that American Red Cross consider as key steps for emergency preparedness[7].

Emergency Kit Preparedness is quantified as percentage of population which have prepared an adequate emergency kit; its numerical range is, therefore, [0,1].

4.3 Trust indicators

4.3.1 Safety perception

Although the described importance of the real safety rate following a natural disaster, at least an equal role is played by safety perception.

Indeed, frequently they don't go hand in hand and their discrepancy can be exacerbated during an emergency; testament to this are the several studies about the false perception of crime following a disaster[8] (Nogami 2015).

Safety perception can, obviously, affect the interpersonal trust, the trust in the institutions and, consequently, the global response to the emergency situation.

It's therefore a good idea to implement an indicator expression of the safety perception in the calculation of Resourcefulness.

Safety Perception will be measured as percentage of people thinking that crime is lower than previous year, so its numerical range will be [0, 1].

Necessity for setting a target value, although *SP* is a percentage indicator, will have to be discussed in the future.

At a finer urban-community scale it's also available a Safety Index developed by Numbeo[9].

Security and Governmental Affairs, United States Senate, One Hundred Eleventh Congress, second session, September 20, 2010

[7] American Red Cross, 2018, available at: www.redcross.org (accessed December 13, 2018)

[8] Ross J., Lowery W., September 1, 2017. *Looting rumours and fear of crime often exaggerated after natural disasters*, The Washington Post, Web. Accessed on December 12, 2018.

[9] https://www.numbeo.com/crime/rankings.jsp, accessed on December 14, 2018

4.3.1 Volunteering

Volunteering, intended as volunteers and economic aid, can be very effective to mitigate the effects of a natural disaster.

In the Mexico City 1985 earthquake's aftermath, the absence of official rescue service was tempered by the action of emergent volunteer groups (Whittaker, McLennan et al. 2015), which saved many people who were trapped under the debris; the major of these groups became an specialised organization in the following years, offering its contribution in the days after L'Aquila earthquake (2009) and Haiti earthquake (2010).

That has to be accompanied by a smart management of volunteers to use them efficiently and the habit to do it; otherwise, volunteering can actually be counterproductive (*cf.* Kobe Earthquake, 1.3.4). Moreover, it often happens that professional rescuers are not able to reach the disaster zone because of the traffic jams caused by civilians who were willing to offer their support (Liu and Robinson 2013).

The simplest and most logical measurement of this indicator is the percentage of population volunteering, taking for granted that communities in which volunteering is more widespread are also the ones which are better organised to manage the volunteers during and emergency situation.

An interesting alternative, for country-scale communities, is the World Giving Index, a composite index developed by Charities Aid Foundation[10] based on volunteering time, propensity to donate money and to help a stranger.

They both have numerical range [0,1].

Necessity for setting a target value, although V is a percentage indicator, will have to be discussed in the future.

4.3.1 Interpersonal Trust

The role of social networks in the emergency management has been widely discussed (Carlin, Love et al. 2014).

It has been, for example, highlighted the importance of Vietnamese community as social network in the immediate Hurricane Katrina's aftermath in New Orleans East (Patterson, Weil et al. 2010): they actively supported other members of the community housing many evacuated and taking pictures to them to assure their loved ones that they were fine; they

---

[10] https://www.cafonline.org/about-us/publications/2018-publications/caf-world-giving-index-2018, accessed on December 15, 2018

keep acting as a community during the recovery as well and they have been able to repair many of the damaged buildings in about a half-year.

On the contrary, the rest of the city of New Orleans "showed little, if any broad scale cooperative activity" (Aldrich 2010), so that recovery of Kobe (after 1995 earthquake) and Tamil Nadu (after 2004 tsunami) have been quicker and more efficient, despite material disadvantages and thanks to their social ties and trust.

It's therefore clear the reason why trust indicators have to be implemented in the Resourcefulness framework. Interpersonal trust can be quantified as the percentage of population which thinks that people can be trusted; its numerical range is [0,1].

Necessity for setting a target value, although *IT* is a percentage indicator, will have to be discussed in the future.

4.3.2 Trust in the Political System

The second trust-sphere indicator is the trust in the political system; indeed, it is reasonable to assume that a community with a good confidence in the decision-makers is likely to act accordingly to received directives; it will be more likely that decision-making system and population form a cohesive and consistent unit coping with the challenges imposed by the natural hazardous event.

Trust in the Political System is measured through the percentage of population which trusts in government, and it's ranged [0,1].

Necessity for setting a target value, although *TPS* is a percentage indicator, will have to be discussed in the future.

4.3.3 Trust in the Police

Similarly to what explained about Interpersonal Trust and, above all, Trust in the Political System, trust in the police and generally law enforcement officers (*i.e.* those which manage emergency and take decisions at fine scale in the disastrous event's aftermath) can much improve the global community's response in regards to the natural disaster.

For example, "The citizen has to trust the police to abandon home and leave his or her property to a relatively unknown place" (Busa, Musacchio et al. 2015).

It is quantified as the percentage of people thinking that police can be trusted and its numerical range is [0,1].

Necessity for setting a target value, although *TPS* is a percentage indicator, will have to be discussed in the future.

4.3.4 Patriotism

Patriotism and sense of belonging to its own community can raise the morale and provide the right determination to better manage the emergency and to start a quick and efficient recovery. It seems to be the case of *Hakka spirit* described in the case study of 921 Taiwan earthquake (*cf.* 1.3.4): Hakka's keen sense of belonging have them allowed to mitigate the damage (in terms of economic and human loss) despite the lack of "mainstream" resources.

It can be seen as the last trust-sphere indicator and it is quantified as the percentage of people who are proud to belong to their community.

Its numerical range is, therefore, [0,1].

Necessity for setting a target value, although *TPS* is a percentage indicator, will have to be discussed in the future.

4.4 Creativity indicators

4.4.1 Patent Applications

The need to implement creativity indicators in the Resourcefulness framework is due to the necessity to explain the ability of a community to understand that, for example, a wooden table can be used as a raft during a flood.

Indeed, if the importance of planning has been already explored and quantified (*cf.* 2.2.9, 2.2.10, 2.2.11), during an emergency it is also fundamental the ability to improvise (Kreps 1990) (Tierney 2002), so that training programmes for improvisation in emergency management have been developed (Mendonça and Fiedrich 2006).

Improvisation can be seen as an expression of creativity, intended as ability to adapt to the new situation consequent to the hazardous event, to find a new shape and to take creative decisions to mitigate the event's aftermath and recovery.

For country-scale communities, creativity could be quantified through the Global Creativity Index, developed by the Martin Prosperity Institute: it aims to quantify creativity through indicators like tertiary education rate, number of employees in creative occupations, research and development expenditure, number of patent applications, tolerance and so on.

Global Creativity Index is unfortunately available only for 2011 and 2015 and it is therefore impossible to efficiently use it; nevertheless, till there will be enough data, its indicators can be used to replace it.

First chosen indicator is Patent Applications, measured as number of patent applications for 1,000 people; its numerical range is $[0, \infty)$.

4.4.2 Research and Development Expenditure

The second creativity indicator is Research and Development Expenditure, measured as percentage of GDP expended to fund research and development.

Although it is expressed by a percentage, a target value has to be set; indeed, 100% of GDP expended on research and development is neither a possible nor a good value.

3% can be taken as target value as suggested by the European Council, which set that GDP percentage as target by the year 2010[11].

Its numerical range is $[0,1]$.

4.5 Summary

Table 4-1 sums up described indicators, with their symbols, measure and justification.

---

[11] https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:R_%26_D_intensity, accessed on December 13, 2018

Table 4-1: Summary of selected indicators.

Dim.	Indicator	Symbol	Measure	Justification
Political-economic	Economic Complexity	*ECI*	*Economic Complexity Index ÷ TV*	(Cutter, Emrich et al. 2006)
	Bureaucracy Flexibility	*BF*	*Economic Freedom Index ÷ TV*	(Ballano 2017)
	Fragility	*FSI*	*(Fragile States Index)$^{-1}$ ÷ TV*	(Nel and Righarts 2008)
	Mitigation Spending	*MS*	*% GDP allocated by the community to cope with disasters ÷ TV*	(Council 2005)
	Safety Rate	*SR*	*(Reported violent crime rate per 100,000 people)$^{-1}$ ÷ TV*	(Yates and Mackenzie 2018)
	Participation in public life	*PPL*	*% turn-out at last presidential election*	(Organization 2002)
Preparedness	Smartphone penetration	*S*	*% population having and using a smartphone*	(Palen, Anderson et al. 2010)
	Disaster Preparedness	*FDP*	*% population reporting having a family emergency plan*	(Paton and Johnston 2017)
	Emergency Kit Preparedness	*EKP*	*% population reporting having adequate emergency kits*	American Red Cross (2018)
Trust	Safety Perception	*SP*	*% population thinking crime is less than previous year*	(Nogami 2015)
	Volunteering	*V*	*Average volunteering hours per week ÷ TV*	(Whittaker, McLennan et al. 2015)
	Interpersonal Trust	*IT*	*% population thinking others can be trusted*	(Carlin, Love et al. 2014)
	Trust in the political system	*TPS*	*% population thinking government can be trusted*	(Carlin, Love et al. 2014)
	Trust in the police	*TP*	*% population thinking police can be trusted*	(Carlin, Love et al. 2014)
	Patriotism	*P*	*% population proud to belong to the community*	(Lee and Loh 2003)
Creativity	Patent applications	*PAT*	*Patent applications per 1,000 people ÷ TV*	(Kreps 1990)
	Research and development expenditure	*RDE*	*% GDP invested in research and development ÷ TV*	(Kreps 1990)

4.6 Community scale

It's fundamental to set the right community scale for which Resourcefulness is intended as representative.

In this respect, PEOPLES framework "is based on basic community organizational units at a local (*i.e.* neighbourhoods, villages, towns or cities) and regional scale (*i.e.* counties/parishes, regions, or states)" (Cimellaro 2016).

For simplicity sake, we can adopt the same spatial subdivision used for PEOPLES framework, shown in Figure 4-1.
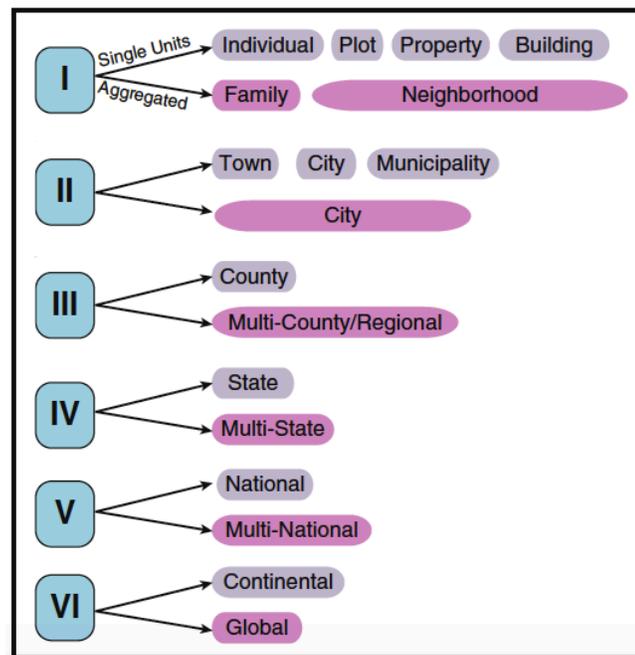

Figure 4-1: Community scale for community resilience framework (Cimellaro 2016)

Cimellaro (Cimellaro 2016) claimed that "on a small scale, when considering critical infrastructures, the focus is mainly on technological aspects. On a larger scale, when considering an entire community, the focus is broadened to include the interplay of multiple systems such as human, environmental, and others which together add up to ensure the healthy functioning of a society".

Since we are handling the latter issue, it does not seem to be proper to talk about Resourcefulness referring to too small communities such as infrastructures. Indeed, it stands to reason that, from the human factor point of view, the response of a small community in a

natural disaster's aftermath will be similar to the one of the larger community in which it is included; it is likely that it is possible to observe slight differences among components of the community (*e.g.* hospitals, *cf.* 1.3.4, Chile Earthquake, 2010), but it's obvious that the behaviour of small communities (*e.g.* infrastructures) will be strongly influenced by the one of the larger community (*e.g.* town or actually the state).

Moreover, when considering town-level communities, it's clear that the response of the city will significantly depend on the characteristics of the state; for example, it does not make sense to talk about government fragility or research and development expenditure at town-level.

For that reason, even when town or cities are considered, some indicators have to be considered referring to the state in which they are. Table 4-2 shows what level of community has to be referred to each indicator when an analysis on a certain community level is being carried on.

Level I and VI are not considered, because it does not seem to be meaningful to talk about Resourcefulness at family/neighbourhood level, just like at continental/global level.

Everything said so far is independent of the ease of collecting data at each community level. Many of the chosen indicators are based on surveys, which almost always are conducted at national level; for example, it's unlikely they are available data about *FDP* or *PAT* at city level. Moreover, measurement method for economic complexity is based on a composite index, *Economic Complexity Index*, which is computed by OEC at country-level, making it impossible in practice its usage at a scale smaller than IV.

Table 4-2: Community scale to which each indicator has to be referred for each community scale analysis

Indicator	Community scale					
	I	II	III	IV	V	VI
ECI		II	III	IV	V	
BF		IV/V	IV/V	IV/V	V	
S		II	III	IV	V	
SR		II	III	IV	V	
SP		II	III	IV	V	
V		IV/V	IV/V	IV/V	V	
FSI		IV/V	IV/V	IV/V	V	
PPL		II	III	IV	V	
FDP		II	III	IV	V	
EKP		II	III	IV	V	
MS		IV/V	IV/V	IV/V	V	
IT		II	III	IV	V	
TPS		II	III	IV	V	
TP		II	III	IV	V	
P		II	III	IV	V	
PAT		II	III	IV	V	
RDE		IV/V	IV/V	IV/V	V	

5 Case studies

5.1 Motivation

After having discussed every step of the theoretical framework and developed an algorithm to compute Resourcefulness of a community, it is fundamental to test a posteriori its consistency with the theoretical principles and its good behaviour when parameters change. For that reason, it seems to be a good idea to choose communities for which it is possible to collect a sufficient amount of data, in order to make possible the application of the algorithm as faithfully as possible compared to what has been explained at section 3.8.

It has been already described the reason (*cf.* 4.6) why, in this development phase of the algorithm, it seems to be better to focus the analysis on country-level communities.

The first country chosen is, therefore, the United States; after certain attempts, in fact, they have proven to be the country with the highest number of available and retrievable data.

In order to get a confirmation of results obtained for the first case study, an analysis on a second case study is carried out: Italy is chosen for this purpose.

It's fundamental to make clear that, at the current state of development of the framework, Resourcefulness of communities are not comparable: it would be a conceptual mistake to state that community A is more resourceful than community B because $RFS_A > RFS_B$.

Indicators, in fact, are widely based on surveys, frequently conducted nationwide by a national institute, so that almost never they are comparable to similar surveys conducted in a different community, because they probably have been carried out through different questions and on different statistical samples.

For that reason, in the following paragraphs *RFS* of United States and Italy will never be plotted in the same graph, in order to not induce the reader to think that communities are comparable; only plots in which data coming from analysis on United States and Italy will be plotted together are those whose purpose is to check the good behaviour of the algorithm.

In the following paragraphs data sources will be described and the most crucial steps of the algorithm (*e.g.* imputation of missing data) and the results will be discussed; an analysis will be also carried out to verify the proper functioning and any weaknesses.

5.2 Data collection

5.2.1 Data sources

Table 5-1 and Table 5-2 shows sources from which have been collected the available data:

Table 5-1: Summary of indicators used for the case study of United States; each indicator is followed by its symbol, its dimension and the source from which data have been collected.

Dim.	Indicator	Symb.	Source
Political-economic	Economic Complexity	ECI	https://atlas.media.mit.edu/en/
	Bureaucracy Flexibility	BF	https://www.heritage.org/index/
	Fragility	FSI	http://fundforpeace.org/fsi/data/
	Mitigation Spending	MS	? (*cf.* 4.1.1)
	Safety Rate	SR	https://www.statista.com/statistics/191219/reported-violent-crime-rate-in-the-usa-since-1990/
	Participation in public life	PPL	https://www.fairvote.org/voter_turnout#voter_turnout_101
Preparedness	Smartphone penetration	S	https://www.statista.com/statistics/201183/forecast-of-smartphone-penetration-in-the-us/
	Disaster Preparedness	FDP	https://ncdp.columbia.edu/
	Emergency Kit Preparedness	EKP	https://ncdp.columbia.edu/
Trust	Safety Perception	SP	https://www.statista.com/statistics/205525/public-perception-of-trend-in-crime-problem-in-the-usa/
	Volunteering	V	https://www.statista.com/statistics/189295/percentage-of-population-volunteering-in-the-united-states-since-2003/
	Interpersonal Trust	IT	https://gssdataexplorer.norc.org/variables/441/vshow
	Trust in the political system	TPS	http://www.people-press.org/2017/12/14/public-trust-in-government-1958-2017/
	Trust in the police	TP	https://news.gallup.com/poll/213869/confidence-police-back-historical-average.aspx
	Patriotism	P	https://news.gallup.com/poll/236420/record-low-extremely-proud-americans.aspx?utm_source=twitterbutton&utm_medium=twitter&utm_campaign=sharing
Creativity	Patent applications	PAT	https://data.worldbank.org/indicator/IP.PAT.NRES?locations=US&view=chart
	Research and development expenditure	RDE	https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?display=graph

Table 5-2: Summary of indicators used for the case study of Italy; each indicator is followed by its symbol, its dimension and the source from which data have been collected.

Dim.	Indicator	Symb.	Source
Political-economic	Economic Complexity	ECI	https://atlas.media.mit.edu/en/
	Bureaucracy Flexibility	BF	https://www.heritage.org/index/
	Fragility	FSI	http://fundforpeace.org/fsi/data/
	Mitigation Spending	MS	n/a
	Safety Rate	SR	https://www.statista.com/statistics/191219/reported-violent-crime-rate-in-the-usa-since-1990/
	Participation in public life	PPL	https://www.tgcom24.mediaset.it/politica/infografica/l-andamento-storico-dell-affluenza-alle-urne_1001472-2018.shtml
Preparedness	Smartphone penetration	S	https://www.statista.com/statistics/201183/forecast-of-smartphone-penetration-in-the-us/
	Disaster Preparedness	FDP	n/a
	Emergency Kit Preparedness	EKP	n/a
Trust	Safety Perception	SP	https://www.istat.it/it/files//2018/06/EN_Fear_of_crime.pdf
	Volunteering	V	https://www.lastampa.it/2012/12/04/blogs/datablog/il-volontariato-in-italia-basWoxRZc2U9svassRt6TO/pagina.html
	Interpersonal Trust	IT	https://www.statista.com/statistics/641012/level-of-interpersonal-trust-italy/
	Trust in the political system	TPS	http://www.realinstitutoelcano.org/wps/portal/rielcano_en/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/europa/ari39-2018-toygur-guide-to-understanding-italy-2018-elections-and-beyond
	Trust in the police	TP	https://www.statista.com/statistics/579685/public-trust-in-state-police-italy/
	Patriotism	P	n/a
Creativity	Patent applications	PAT	https://www.statista.com/statistics/412674/european-patent-applications-from-italy/
	Research and development expenditure	RDE	https://www.statista.com/statistics/420976/gross-domestic-expenditure-on-research-and-development-gdp-italy/

5.2.2 Dealing with missing data

United States' starting data matrix was 29.4% filled, whereas Italy's one was 18.3% filled.

Following the algorithm described at section 3.3, *Mitigation Spending* indicator has been excluded from the analysis of United States; analysis of Italy has required the exclusion of five indicators: *Mitigation Spending*, *Safety Perception*, *Family Disaster Preparedness, Emergency Kit Preparedness* and *Patriotism.*.

Resulting data matrices had many cases with less than 50% data filled; 30 cases have, therefore, been excluded from the analysis of both case studies; they are, in particular, years from 1960 to 1989 (USA) and from 1970 to 1999, so that data matrices contain years from 1990 to 2017 (USA) and from 2000 to 2017 (Italy).

[X] is thus respectively a 28×16 matrix and a 18×12 matrix, with missing values summarised in Figure 5-1.
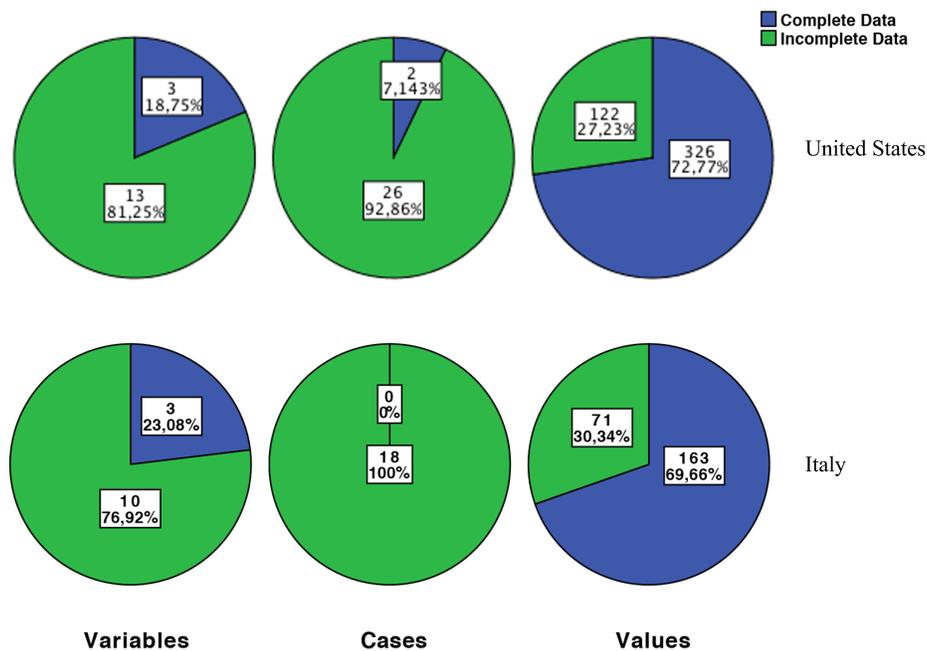
Figure 5-1: Summary of missing data after having excluded indicators with less than 25% filled data and cases with less than 50% filled data.

Missing values can be analysed through missing value patterns using software *SPSS*; patterns are shown in Figure 5-2 and Figure 5-3.
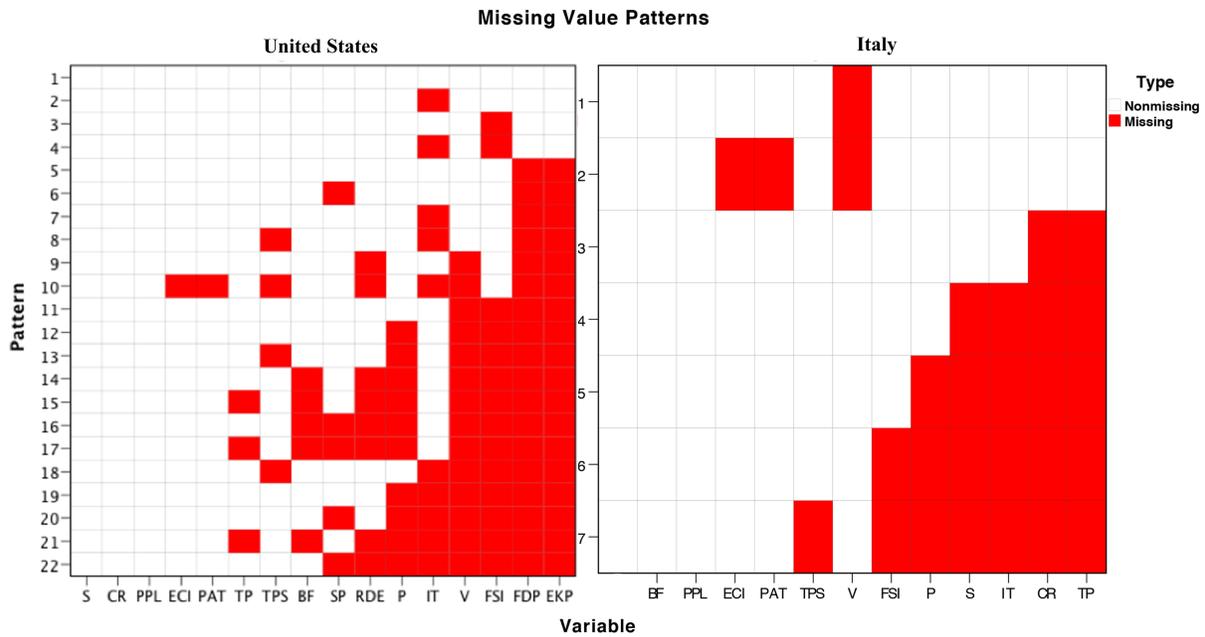
Figure 5-2: Missing value patterns obtained from software SPSS; each pattern represents a group of cases with missing values for the same indicators: for example, United States' pattern 6 represent a group of cases with missing values for indicator *SP*, *FDP* and *EKP*. Patterns are used to reveal monotonicity and to check whether data set is NMAR, MAR or MCAR.
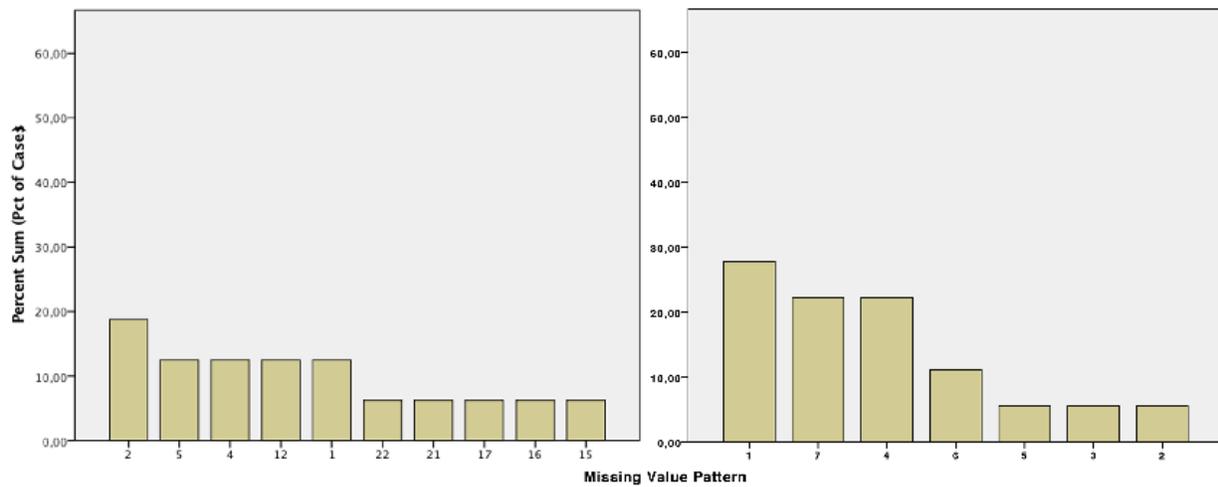

Figure 5-3: Percentage frequency of the most 10 most frequently occurring patterns displayed in Figure 5-2.

Each pattern is intended as a group of cases with the same pattern of incomplete and complete data. For example, with regards to United States, Pattern 1 represents cases which have no missing values, while Pattern 8 represents cases that have missing values on *TPS* (*Trust in*

the Political System), *IT* (*Interpersonal Trust*), *FDP* (*Family Disaster Preparedness*) and *EKP* (*Emergency Kit Preparedness*)[1].

Variables and patterns are ordered to reveal monotonicity where it exists. Specifically, variables are ordered from left to right in increasing order of missing values. Patterns are then sorted first by the last variable (non-missing values first, then missing values), then by the second to last variable, and so on, working from right to left. This reveals whether the data structure is NMAR, MAR or MCAR (*cf.* 3.3). If the data are monotone, then all missing cells and non-missing cells in the chart will be contiguous; that is, there will be no "islands" of non-missing cells in the lower right portion of the chart and no "islands" of missing cells in the upper-left portion of the chart[2].

If United States are considered, Figure 5-shows a border-line behaviour, but this is mainly due to the fact that *FDP* and *EKP* are taken form the same survey, so that when one misses, the other one misses too; even though *FDP* and *EKP* are deleted from the plot, it still does not point out a completely non-monotone behaviour; nevertheless, Figure 5- shows that most frequent patterns (2, 5, 4, 12 and 1: total equal to about 70% of cases) are not those which would point out a monotonicity.

It seems therefore to be proper to use multiple imputation or MCMC to impute remaining missing data.

Speaking of Italy, Figure 5-seems to show a monotone behaviour; nevertheless, the scarcity of cases has to be considered. Therefore, multiple imputation or MCMC will be anyway used to impute missing data.

Algorithm described in section 3.8 has been applied.

Figure 5- displays the analysis carried out on each variable to find at least a good regressor which would be used to impute missing data:

---

[1] A data set can potentially have 2[number of variables] patterns; however, only 22 patterns are represented.
[2]   https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/tutorials/mi_patterns_telco_patterns.html, accessed on December 28, 2018
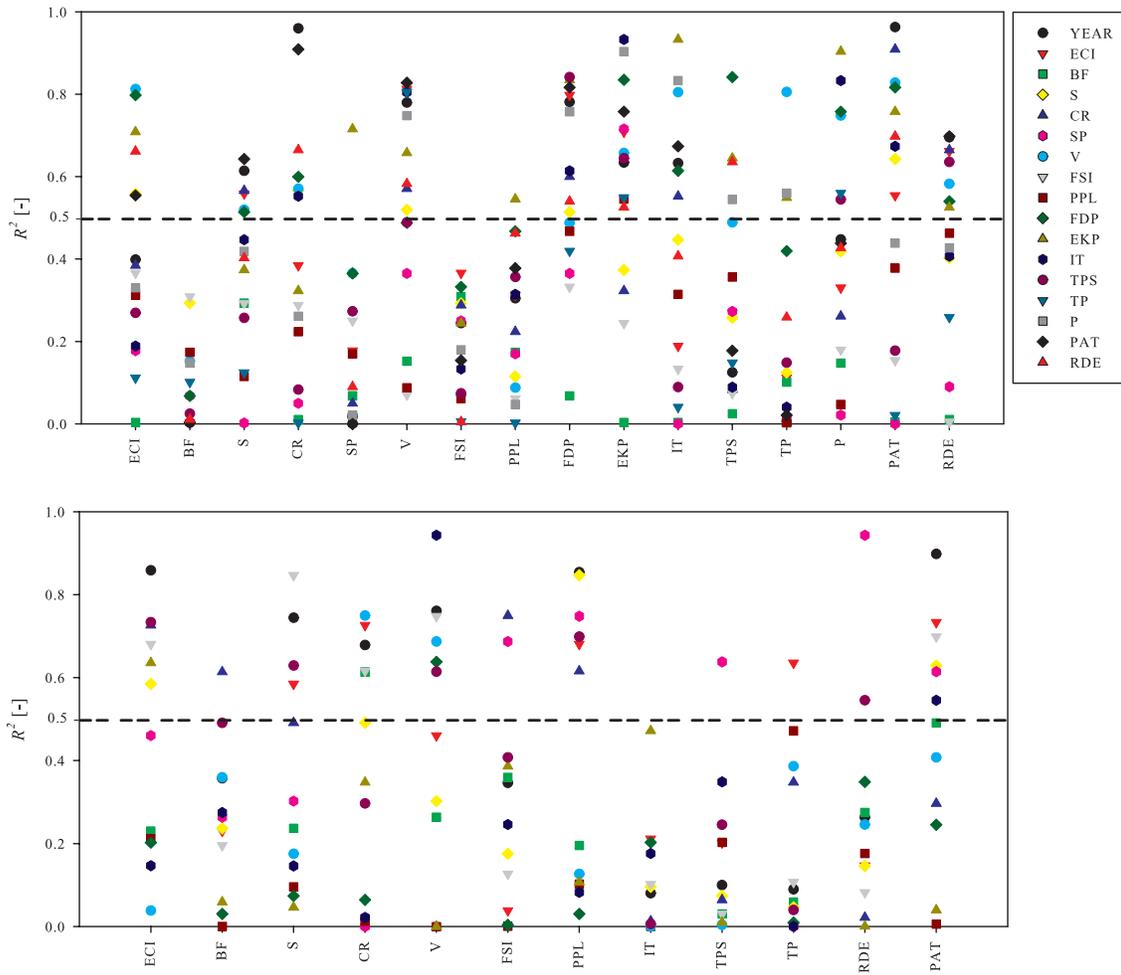
Figure 5-4: Selection of good regressors for each indicator; each symbol represents the R^2 of the plot if the indicator on the x-axis is plotted versus the indicator represented by the symbol in the legend; if the symbol lies above the threshold line (corresponding $R^2 = 0.5$), then the indicator represented by the symbol is considered a good indicator for the indicator on the x-axis; otherwise, it is not considered a good regressor.

Each symbol in the plot represents the R^2 (which can be read on the y axis) of the plot if the indicator on the x-axis is plotted versus the indicator corresponding to the symbol (which can be read in the legend on the right); if the symbol lies above the threshold line ($R^2 = 0.5$), then the indicator in the legend is considered a good regressor for the indicator on the x-axis; if the symbol lies below the threshold line, then the indicator in the legend is not considered a good regressor for the indicator on the x-axis.

As shown by the Figure 5-, the great majority (87.5% for USA, 91.7% for Italy) of variables has at least a good regressor; missing values of these indicators have been imputed through

multiple imputation: data have been imputed 5 times and the arithmetical average has been taken as simulated value.

Remaining indicators, *i.e. Bureaucracy Flexibility* and *Fragile States Index* (USA) and *Volunteering* (Italy), for which has not been found even a good regressor, have been imputed through a Monte-Carlo simulation.

Then, having done that, a simulated complete [*X*] data matrix has been obtained.

5.3 Findings

5.3.1 Results

Spider plot displayed in Figure 5- shows the USA data set referred to year 2017.
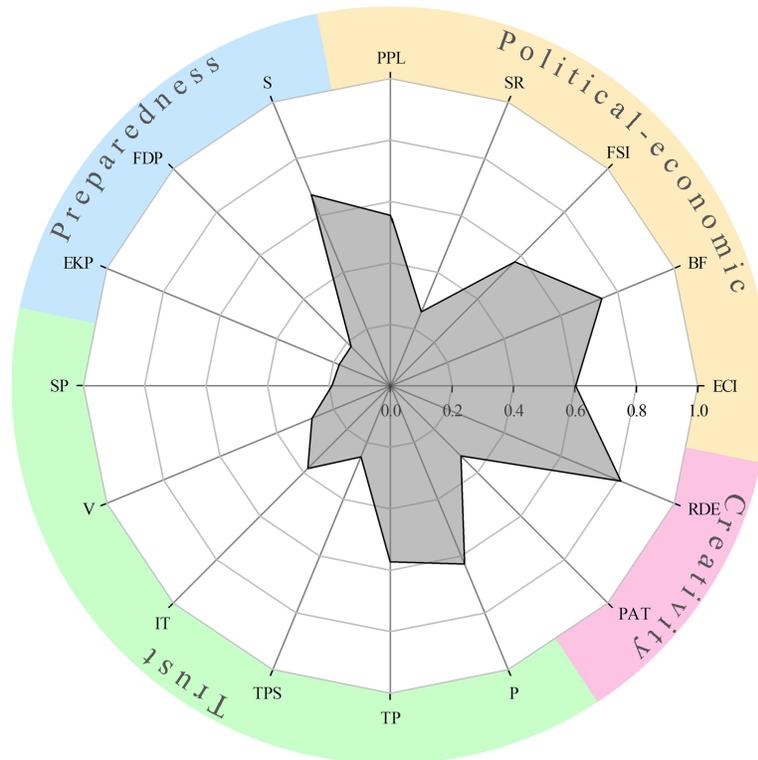
Figure 5-5: Spider plot of USA referred to year 2017. It should be noted that weaknesses in some indicators can be due to the chosen Target Value. Hypothetical completely filled spider plot refers to the ideal community which has a score in each indicator equal to the *TV*.

Figure 5-4 points out certain weaknesses, above all at trust and preparedness level.

It should be considered that this depends on the choice of Target Values; it is therefore fundamental to further debate that issue, finding the justification of each *TV* in the state of the art.

The analysis has returned a value for Resourcefulness of United States in 2017 equal to:

$$RFS_{USA,2017} = 0.4605 \qquad (5.1)$$

and a value for Resourcefulness of Italy in 2017 equal to:

$$RFS_{ITA,2017} = 0.3954 \qquad (5.2)$$

It stands to reason to expect that, if it was possible to carry out the analysis on other countries, value obtained for United States is a rather high Resourcefulness score.

One might think that Figure 5- displays real $RFS_{USA,2017}$ versus the ideal Resourcefulness value: in this respect, grey area should be 46.05% of the entire white spider plot; nevertheless, spider plots do not work in that way, because they progressively distort areas from the centre. Therefore, spider plot of Italy has not been plotted.

In order to get a plot in which it is possible to see the difference between the real RFS of a community and its ideal value, so that proportions are respected, Figure 5-6 and Figure 5-have been developed.
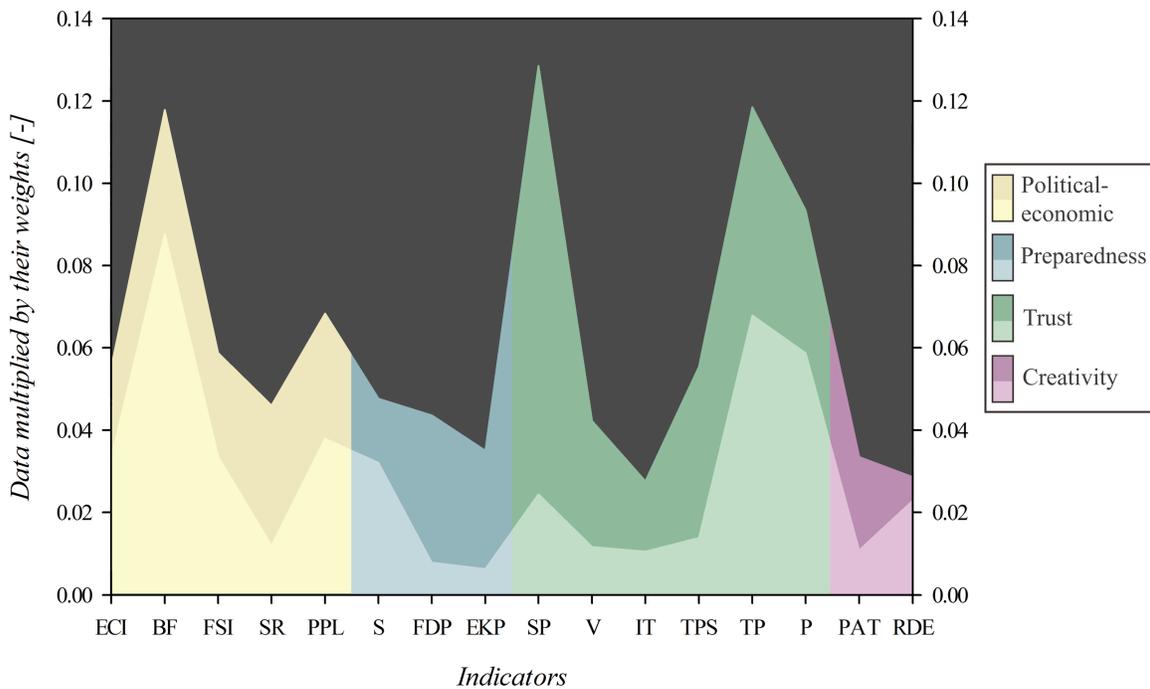

Figure 5-6: Real data are plotted with light colours, whereas dark colours refer to the ideal values; the entire coloured area is equal to 1 (*i.e.* 100%, ideal *RFS*), whereas the light coloured area is equal to 0.4605 (*i.e.* 46.05%, real *RFS*_{USA,2017}).
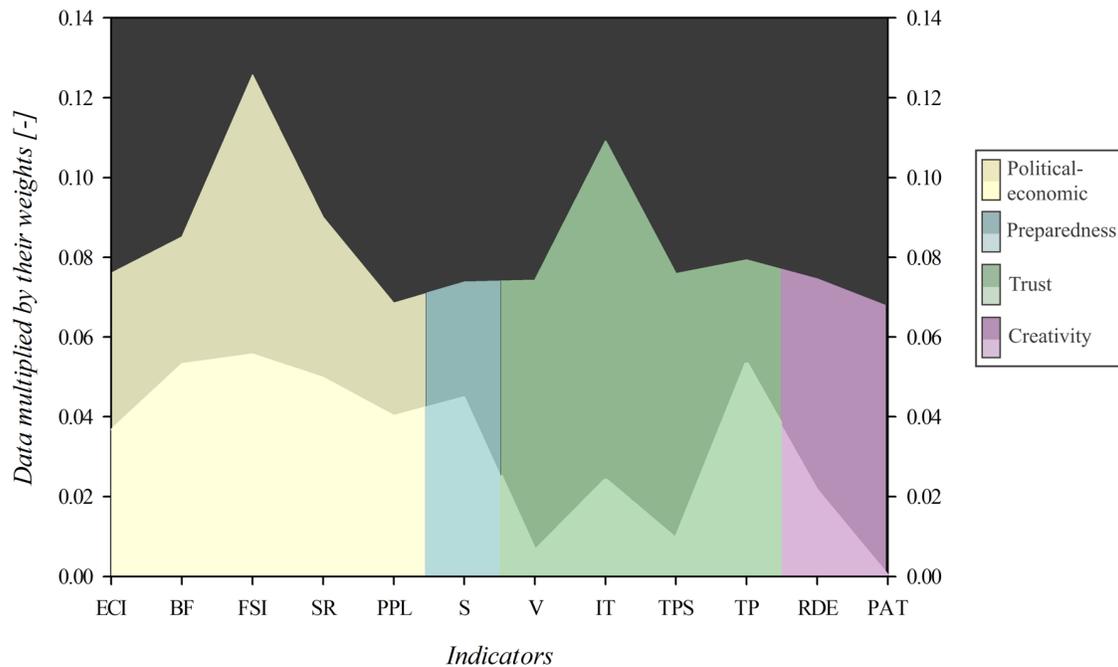
Figure 5-7: Real data are plotted with light colours, whereas dark colours refer to the ideal values; the entire coloured area is equal to 1 (*i.e.* 100%, ideal *RFS*), whereas the light coloured area is equal to 0.3954 (*i.e.* 39.54%, real $RFS_{ITA,2017}$).

For each indicator, the real value is displayed (multiplied by its weight) with the light version of the colour corresponding to the dimension of the considered indicator; the darkest colour refers to the ideal value of the considered indicator, so that the entire coloured area (the light one plus the dark) corresponds to the ideal *RFS* value (*i.e.* equal to 1, 100%), whereas the light coloured area (only the light one) corresponds to the real $RFS_{USA,2017}$ and $RFS_{ITA,2017}$ value (*i.e.* equal to, respectively, 0.4605, 46.05% and 0.3954, 39.54%).

For that reason, spider plots can be potentially used to compare data sets among communities, but they are not effective if a visual representation of the real Resourcefulness versus the ideal one is sought; in that case, area plots have to be used.

Nevertheless, at the current state of the research in the field of resourcefulness, the final value returned by the analysis is far less important than the study on the behaviour of the algorithm. First of all, in order to evaluate the correct functioning of the algorithm, consistency between *RFS* across years and evolution of indicators can be checked.

Before doing that, it should be discussed till what year it is possible to compute *RFS* with sufficient accuracy; indeed, as described in section 3.6.2, Principal Components Analysis should have at its disposal a sufficient amount of cases to be performed with precision;

nevertheless, none of the rules of thumb proposed can be satisfied ($[X]_{USA}$ matrix is 28×16, with a cases/variables ratio equal to 1.75, whereas $[X]_{ITA}$ is 18×12, with a cases/variables ratio equal to 1.50), so that it is good to keep in mind that results are certainly affected by lack of cases.

Moreover, as shown in 4.2.1, *Smartphone Penetration* indicator cannot have data before about 2010, for the simple reason that before that year smartphones did not exist.

For those reasons, it seems to be a good choice not to compute *RFS* before 2010 for USA and before 2011 for Italy, so that any further analysis will be limited to years between 2010 (2011) and 2017.

RFS of USA between 2010 and 2017 is shown in Figure 5-; *RFS* of Italy between 2011 and 2017 is shown in Figure 5-8.
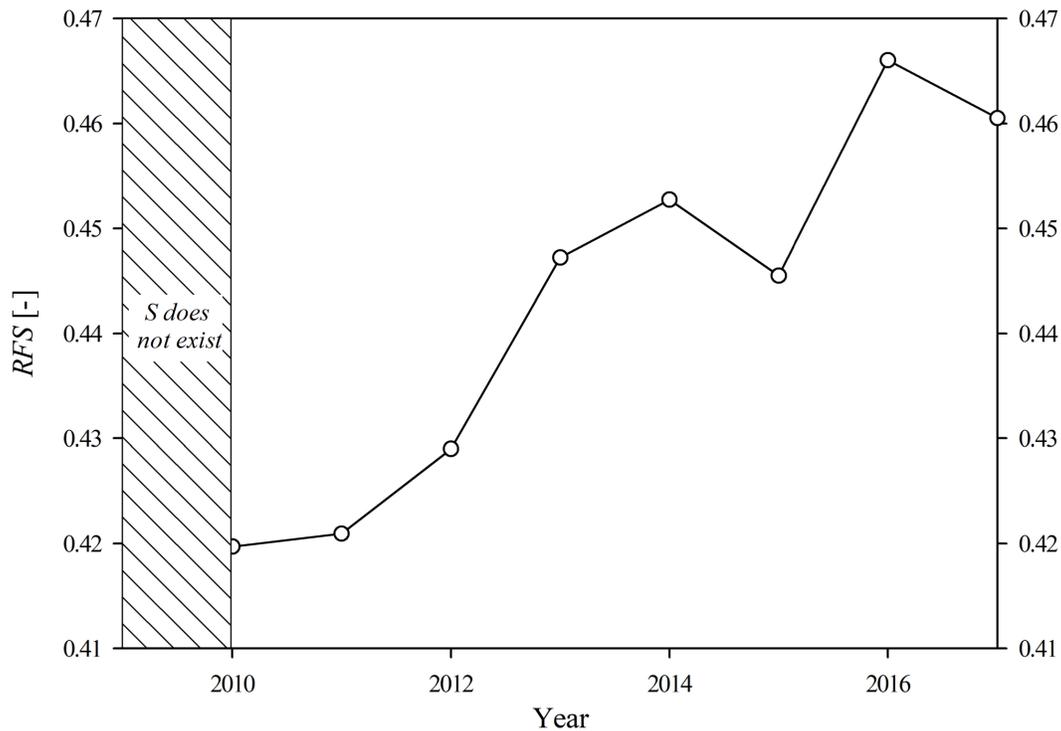

Figure 5-7: Evolution of *RFS* of United States over the years. Years before 2010 are not computed because *Smartphone Penetration* indicator does not exist and because cases are too few to obtain meaningful results from the Principal Components Analysis.
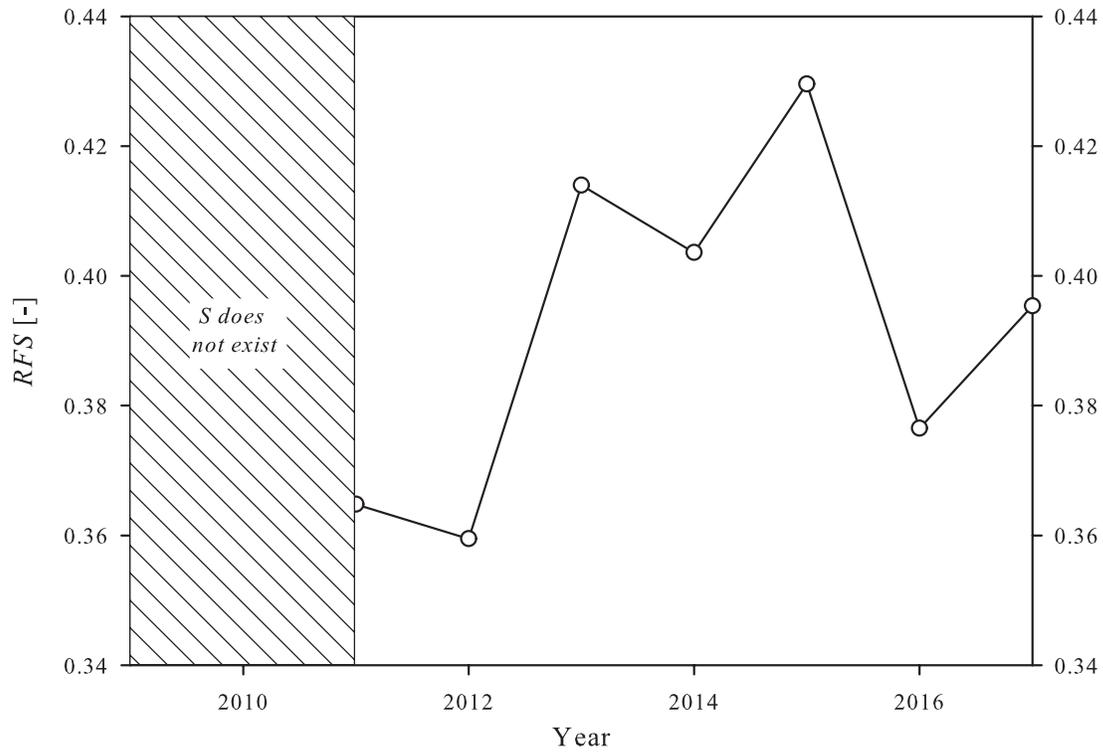
Figure 5-8: Evolution of *RFS* of Italy over the years. Years before 2011 are not computed because *Smartphone Penetration* indicator does not exist and because cases are too few to obtain meaningful results from the Principal Components Analysis.

Evolution of *RFS* over the years must be consistent with the evolution of indicators; in order to make possible this confrontation, data are compared to those referred respectively to year 2010 and to year 2011; it is then taken the average of the percentage increases of indicators. They are shown in Figure 5-9 and Figure 5-10.
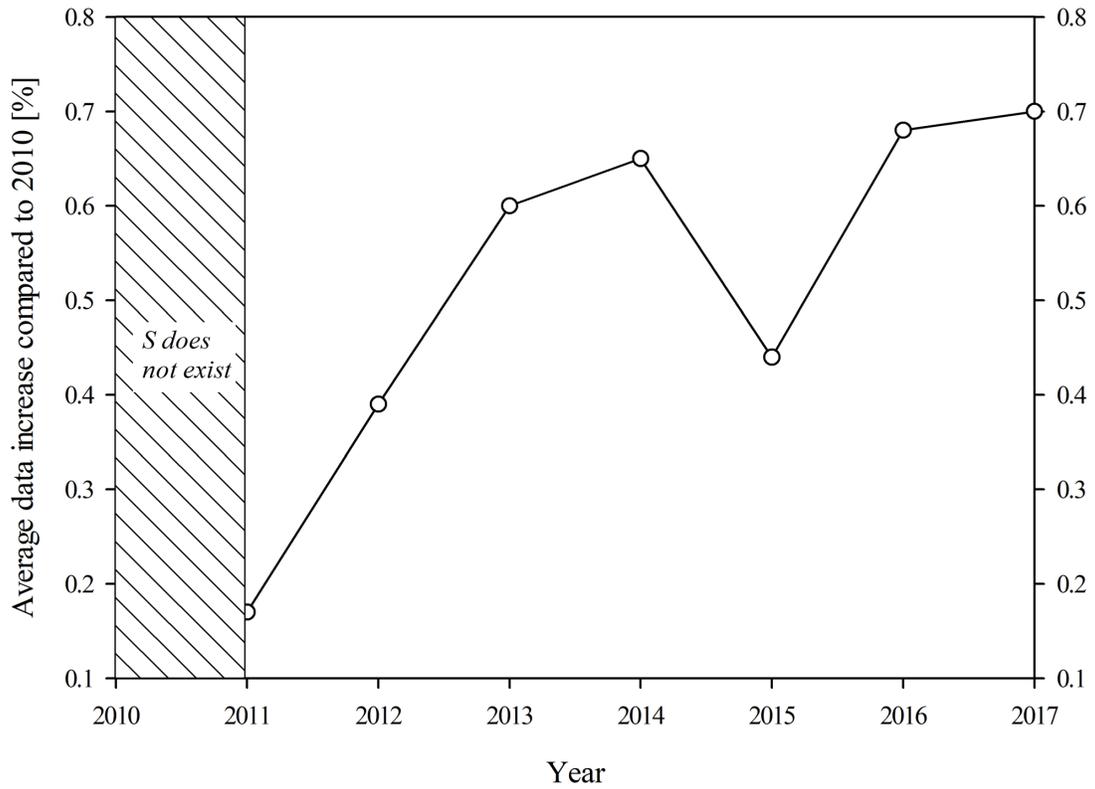
Figure 5-9: Evolution of data (averaged over the indicators) of United States compared to data referred to year 2010.
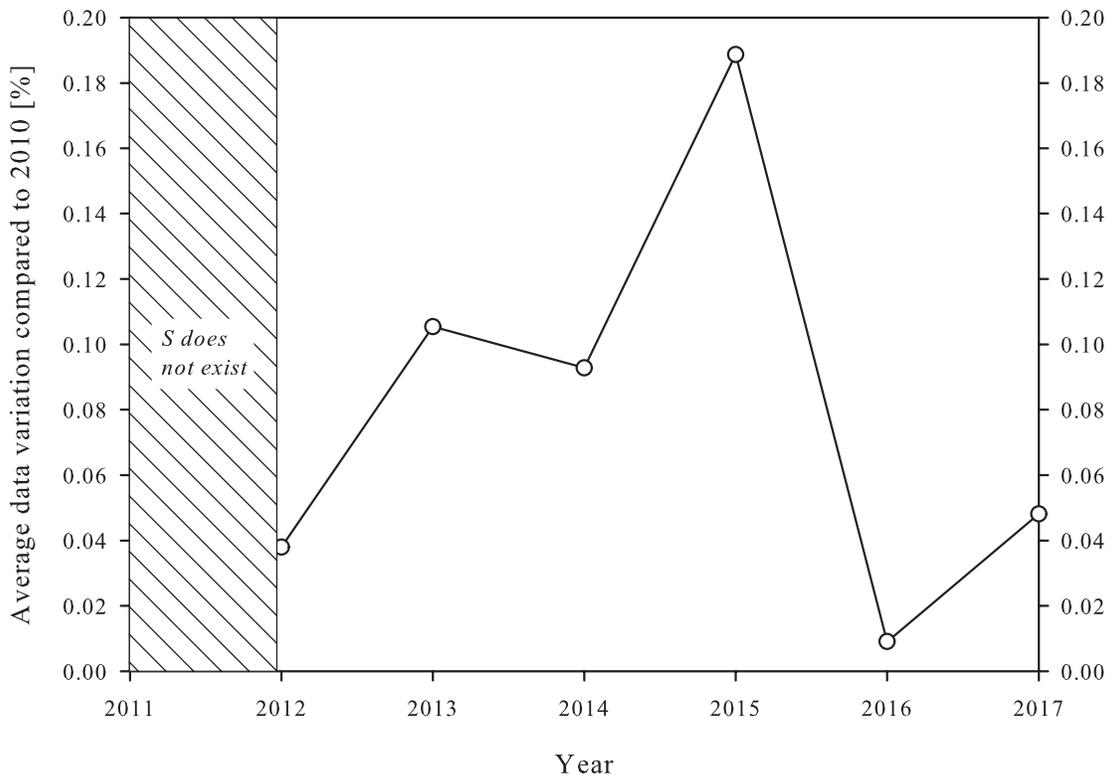

Figure 5-10: Evolution of data (averaged over the indicators) of Italy compared to data referred to year 2011.

Comparison between Figure 5- and Figure 5-9 and between Figure 5- and Figure 5-10 well point out the consistency between *RFS* and average data increase; this relationship is even better shown in Figure 5-.
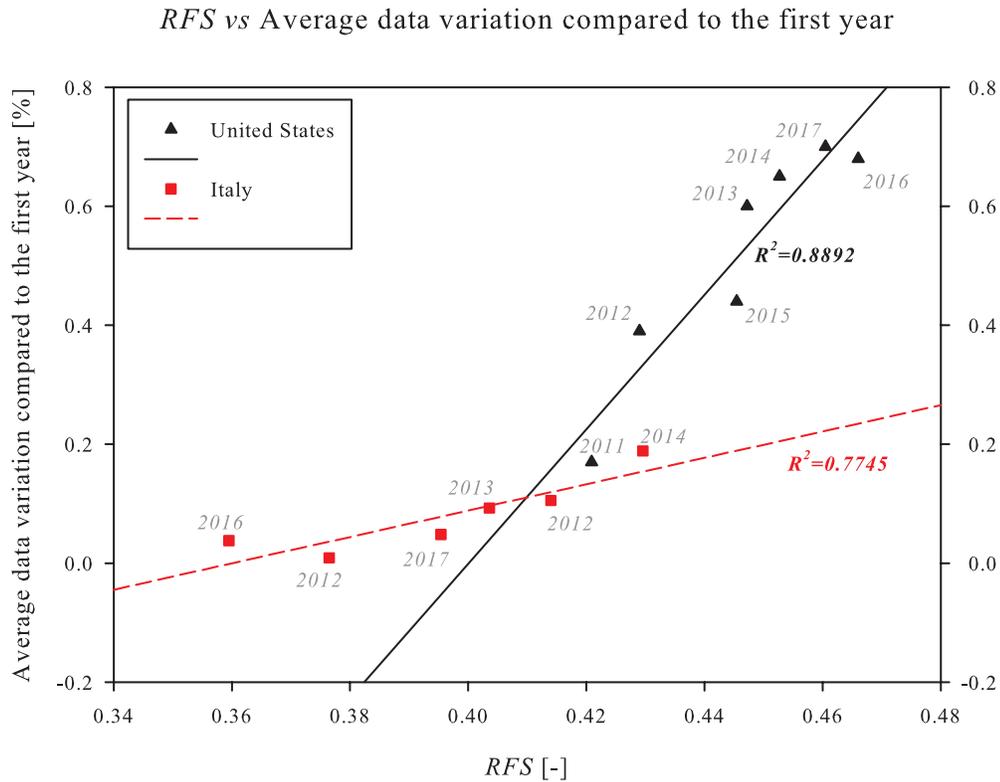

Figure 5-11: *RFS* over the years *vs* percentage average data variation compared to 2010.

5.3.2 Weights

The most crucial step of the algorithm and the one to which has been devoted the great majority of the debate about the construction of the composite index is the allocation of weights.

Weights returned by the analysis carried out on 2017 are shown in Table 5-3, Table 5-4 and Figure 5-12.

Following the developed algorithm, weights are the only factors which can suggest what is convenient to invest on to enhance Resourcefulness: indeed, there is no specific convenience of investing on weak indicators, since linear aggregation method (and not geometric aggregation) has been chosen, which allows full compensability; obviously, the higher is the weight corresponding to an indicator, the more is convenient for USA to invest on the considered indicator to enhance Resourcefulness.
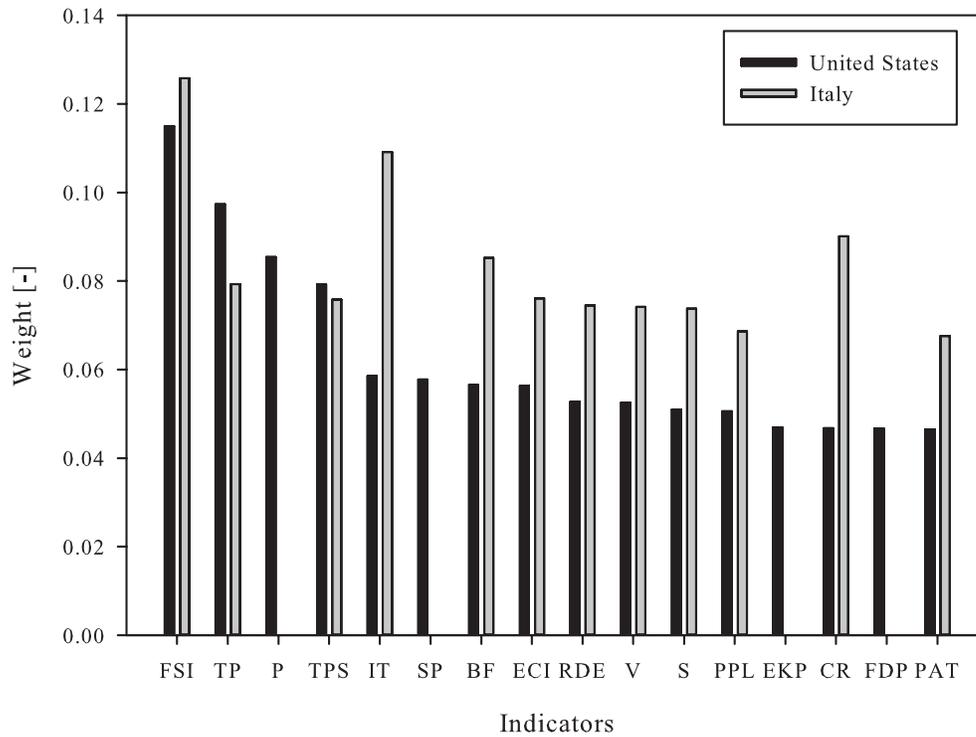
Figure 5-12: Indicators' weights

Table 5-3: USA – indicators' weights

Indicator	Weight [-]
FSI	0.115
TP	0.098
P	0.086
TPS	0.08
IT	0.059
SP	0.058
BF	0.057
ECI	0.057
RDE	0.053
V	0.053
S	0.051
PPL	0.051
EKP	0.047
SR	0.047
FDP	0.047
PAT	0.047

Table 5-4: Italy – indicators' weights

Indicator	Weight [-]
FSI	0.126
IT	0.109
SR	0.091
BF	0.085
TP	0.079
ECI	0.076
TPS	0.076
RDE	0.075
V	0.074
S	0.074
PPL	0.069
PAT	0.068

As described in section 3.6.5, chosen method to allocate weights implies not only community-specific weights, but also year-specific weights: weights are therefore subject to a process of refinement, as a result of which weights change every year.

It seems reasonable to expect that weights' variation is higher in the first years of the analysis and it lessens over the years, following the process of refinement explained in section 3.6.5. Evolution of weights can be seen in Figure 5-13 and Figure 5-14.
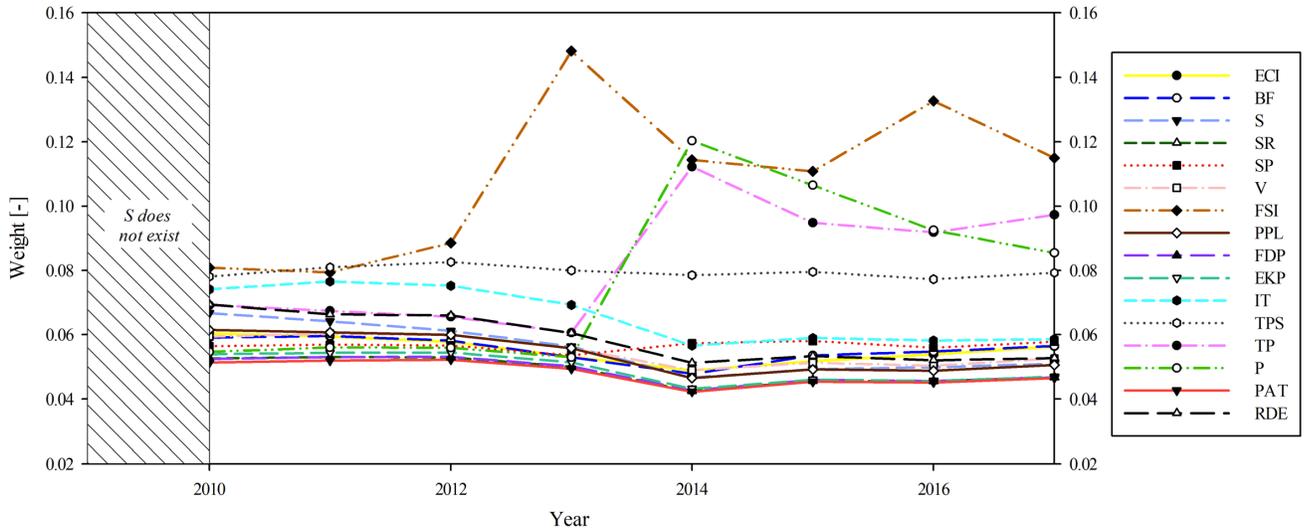

**Figure 5-13**: United States - evolution of weight of each indicator over the years
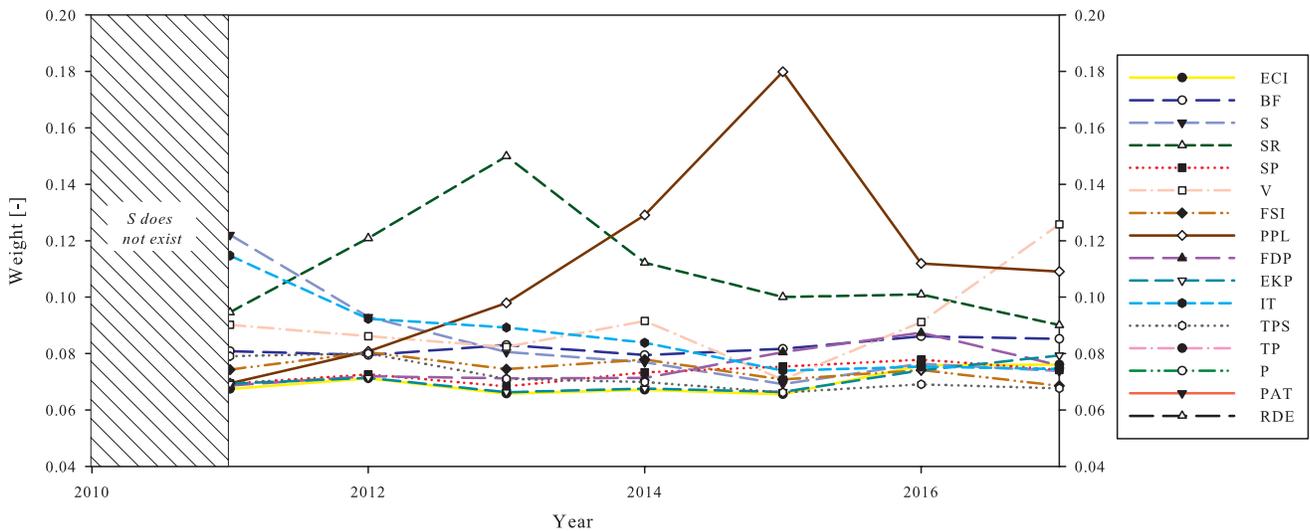
Figure 5-14: Italy - Evolution of weight of each indicator over the years.

As visible in Figure 5-13 and Figure 5-14 and better displayed in Figure 5-15, does not seem to be verified what has been assumed above; indeed, weights' variation does not lessen over years.
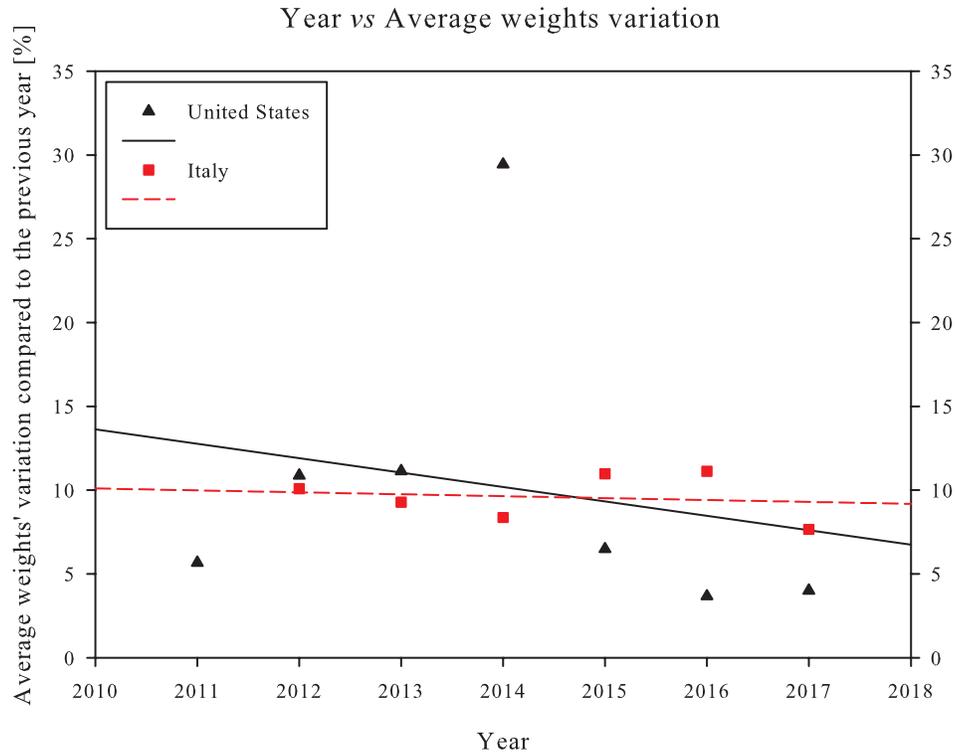

Figure 5-15: Year *vs* average weight's variation compared to the previous year

There can be multiple reasons:

1. Criterion used to select the number of principal components has selected 4 principal components in 2010, 2012, 2012 and 2013 for USA; instead, since 2013 the algorithm has selected 3 principal components, so that weights suffer an obvious surge; in the analysis carried out on Italy 4 principal components have been selected in 2011, 3 principal components have been selected in the following years;

2. As observed in section 5.3.1, none of the cases/variables ratios suggested by OECD is satisfied; it, therefore, stands to reason that analysis is affected by the low amount of cases.

At this point, consistency between obtained weights and the goal on the basis of which has been developed the weights assigning method must be checked (*cf.* 3.6.5).

Indeed, weighting method has been chosen with the main purpose of preventing information overlap among indicators; for that reason, the algorithm allocates lower weights to those indicators which show a high correlation coefficient with other indicators (*i.e.* a meaningful overlap of information) and higher weights to those whose information is not yet contained in other indicators.

Figure 5-16 displays relationship between average correlation coefficient (taken as absolute values), between each indicator and the other ones, and weight, with reference to year 2017.
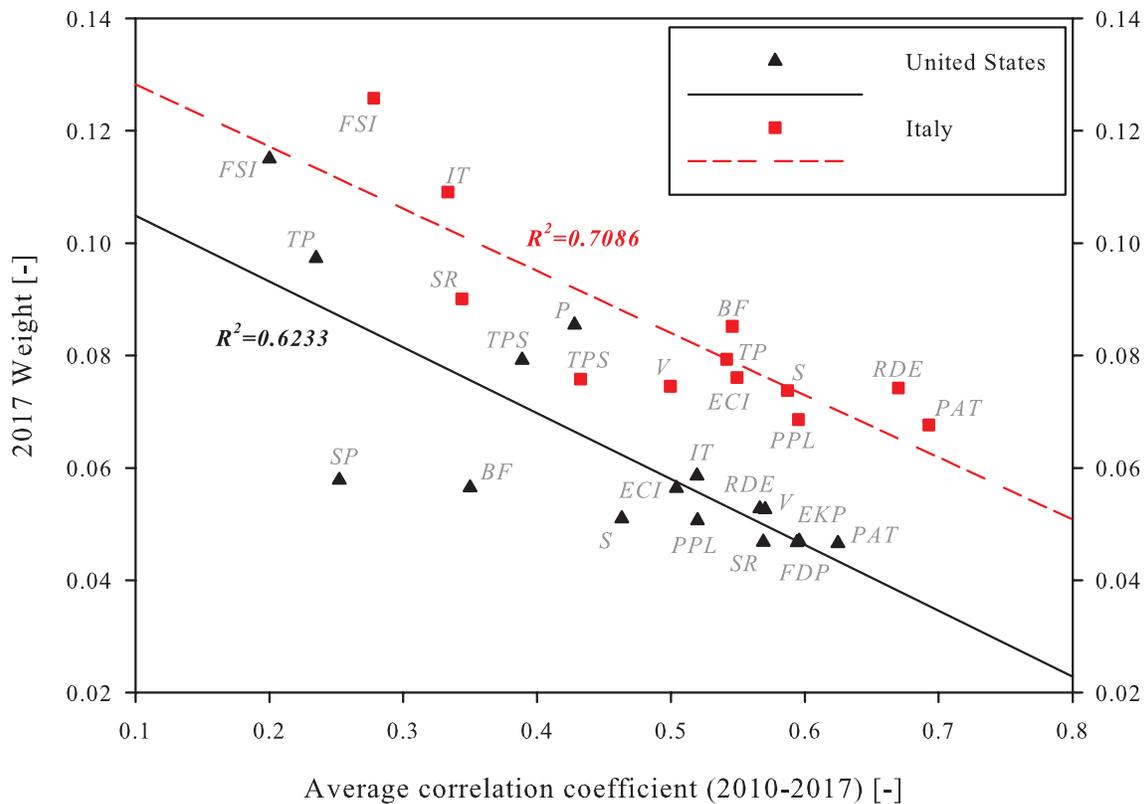

Figure 5-16: Correlation coefficient of each indicator averaged over the years 2010-2017 *vs* weight of each indicator referred to analysis carried out for year 2017.

Average correlation coefficient is computed averaging, for each year involved in the analysis (in this case since 2010 to 2017 for United States, since 2011 to 2017 for Italy), the correlation coefficients between the considered indicator and the other ones: *Y* averages are so obtained for each indicator. Finally, average of the *Y* averages is taken.

For example, if 2013 of USA is considered, average correlation coefficient between each indicator and the others is separately computed for 2010, 2011, 2012 and 2013; lastly, average of averages obtained for those four years is taken.

That is done because the algorithm uses all the cases since the first year to the considered year to compute weights and *RFS*; for that reason, if, for example, 2013 is considered, would not be correct taking into account average correlation coefficients referred only to 2013.

As displayed by the Figure 5-16, there is a good relationship between average correlation coefficient (averaged over all years) and weight, so that when an indicator points out a high correlation coefficient with the other ones it is assigned a low weight and *vice versa*.

For that reason, objective that was purposed in Section 3.6.5 can be considered satisfied.

Moreover, Figure 5-17 is obtained in the same way as Figure 5-16, repeating the process for 2011 (averaging, therefore, over 2010 and 2011), for 2012 (averaging, therefore, over 2010, 2011 and 2012) and so on.
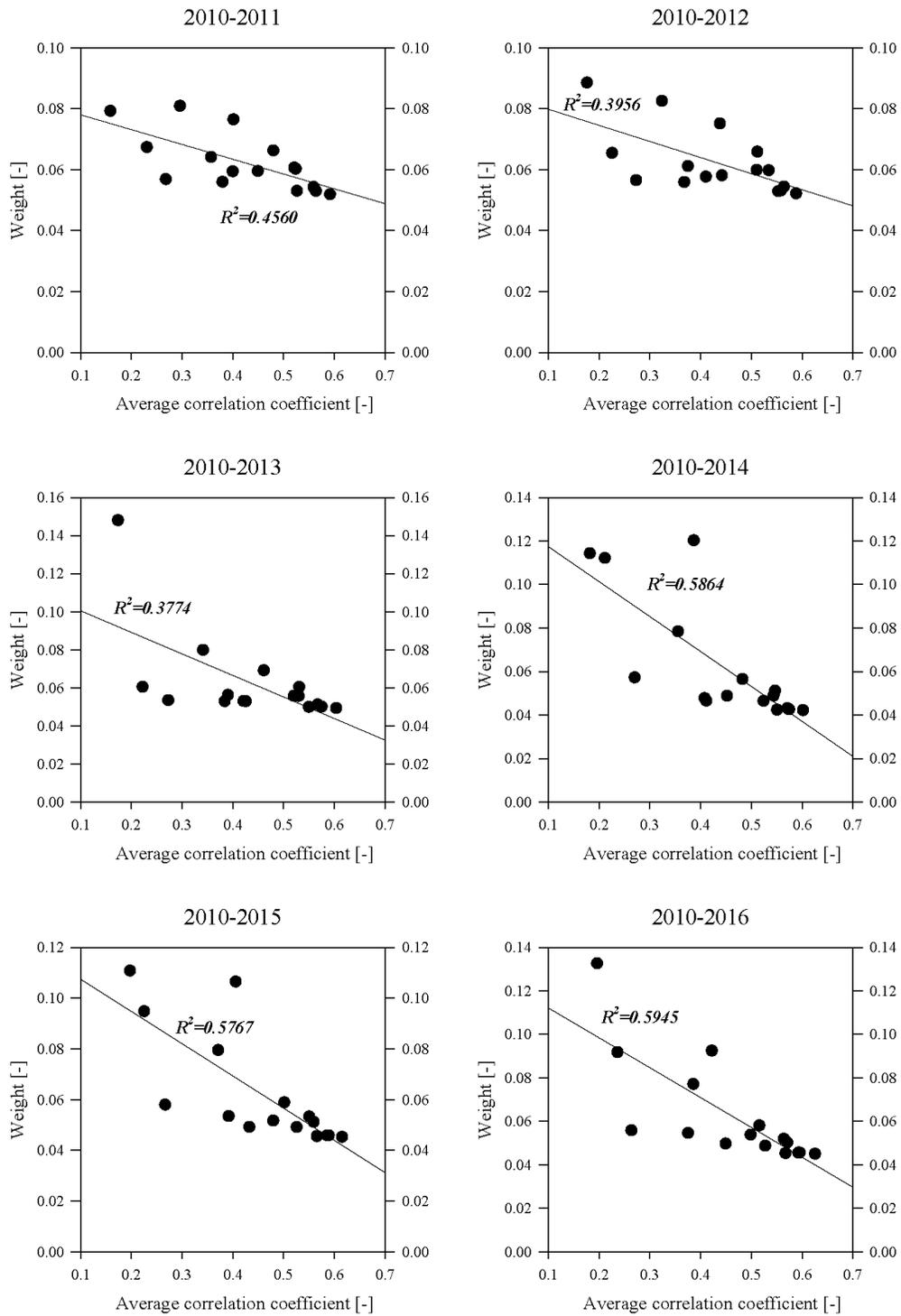
Figure 5-17: USA - What has been explained about Figure 5- is repeated averaging correlation coefficient over the years 2010-2011 and plotting weights obtained for year 2011, averaging correlation coefficient over the years 2010-2012 and plotting weights obtained for year 2012, and so on.

As visible in Figure 5-17 and better displayed in Figure 5-18, correlation between average correlation coefficient and weight gets better over the years.

Figure 5-18 displays years *vs R^2* (*i.e.* the reliability of the regression line of each plot) of plots in Figure 5-17.
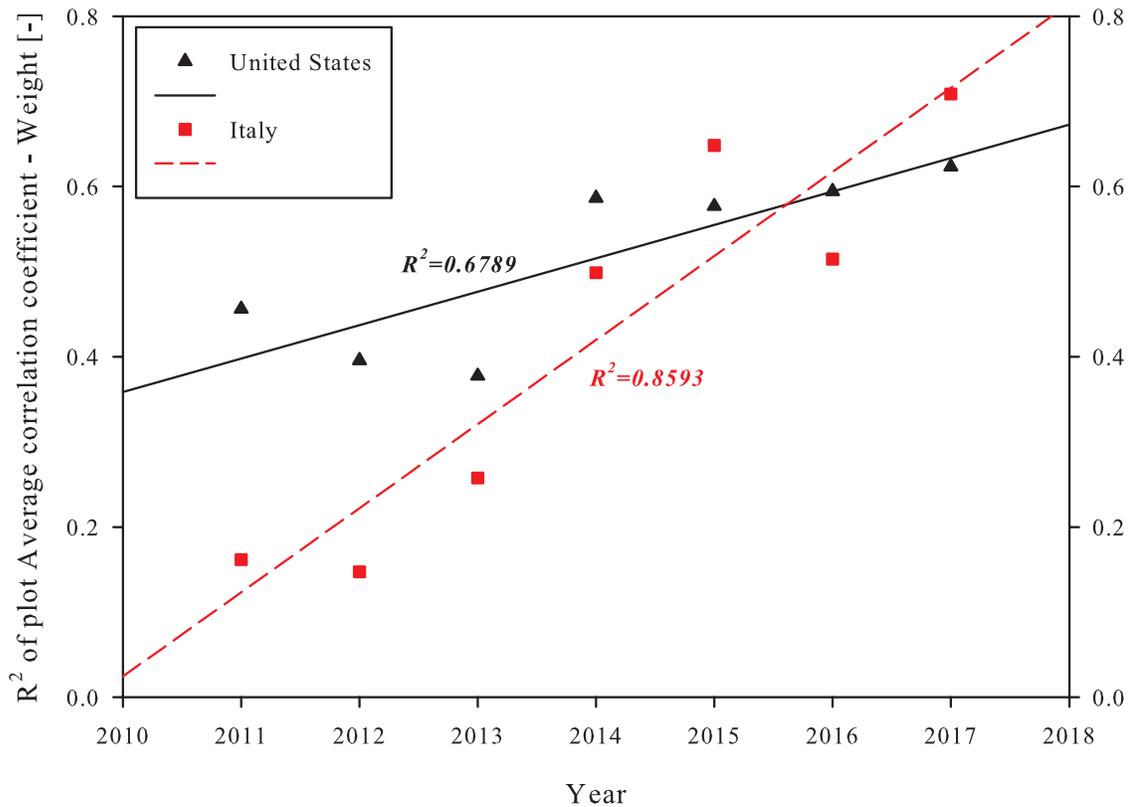
Figure 5-2: Years are plotted versus the R^2 of plots represented in Figure 5- and Figure 5-.

On the other hand, irrespective of what kind of methodology has been chosen to assign weights, it stands to reason to think that a good algorithm allocates the highest weights to those indicators which, if removed from the analysis, change the most the results.

That can be checked removing variables one at a time and comparing the consequent *RFS* to that obtained taking into account all indicators.

Results are displayed in Figure 5-19.
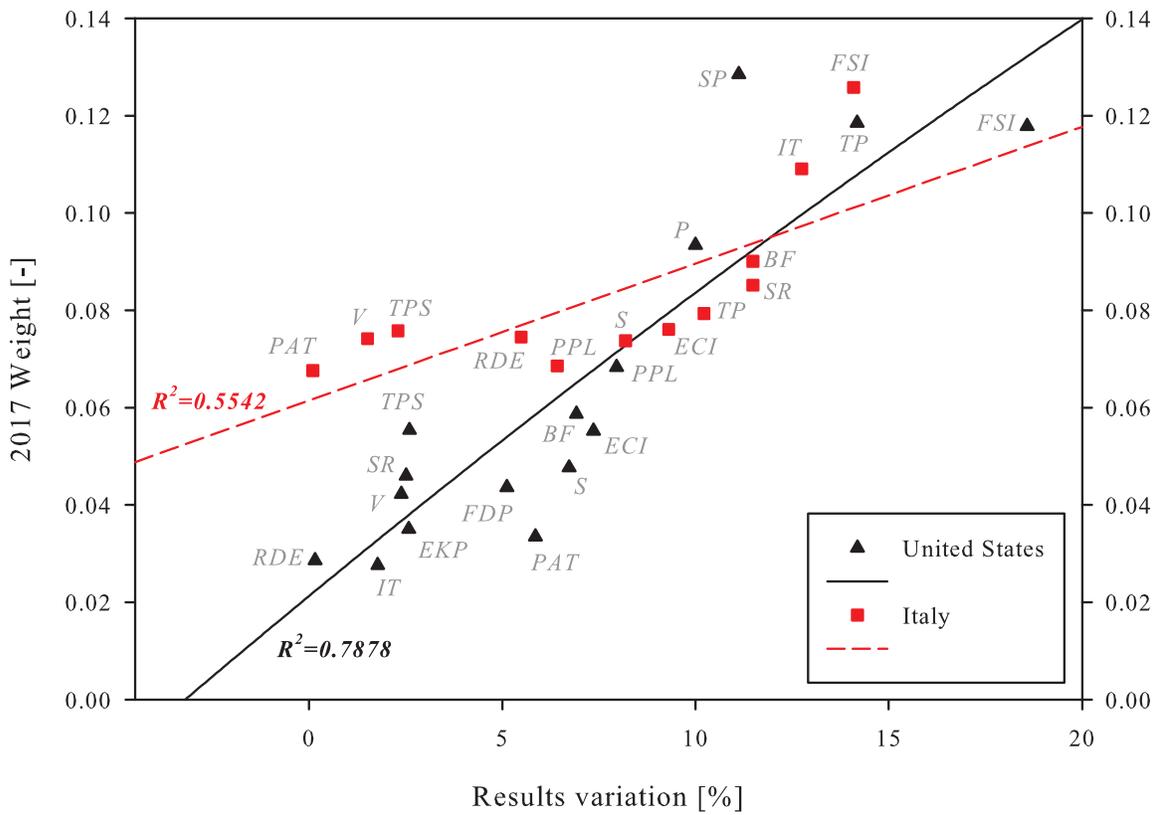
Figure 5-19: Each indicator is excluded from the analysis (on year 2017) one at a time and *RFS* percentage variation of each case is plotted versus the weight (referred to the complete analysis) of the excluded indicator.

Figure 5-19 points out also a good relationship ($R^2 = 0.7113$) between weight assigned and variation of results if the indicator is not taken into account into the analysis in the analysis carried out on United States; it seems to be weaker in the analysis on Italy ($R^2 = 0.5542$), but this could be due to the lack of cases. Further investigation is required on this issue.

6 Conclusion

6.1 Findings

The present thesis must be intended as the first attempt in the state of the art of emergency management to frame the issue of resourcefulness from a mathematical point of view.

To do this, it has been before necessary to analyse the state of the art and to give an accurate definition of the concept of resourcefulness; that has also allowed to set the principles that have to be respected, such as the interval in which resourcefulness is ranged, and so on.

It is then wondered how to quantify such an apparently unquantifiable dimension as the resourcefulness, agreeing (also by taking as an example the PEOPLES framework) that the best way is the construction of a composite index, which aggregates different variables, each one indicator of a particular characteristic of the community.

For that reason a methodology (discussing each step) has been developed to compute RFS of a given community, imputing missing data and normalising them, allocating weights to each indicator (through a data-driven approach) and, finally, aggregating them.

A list of indicators has been then drawn up, even though that has not to be intended as exhaustive, but rather as a good starting point.

The algorithm has been applied to two case studies, *i.e.* the USA and Italy, getting a good response: a sensitivity analysis has been carried out on the obtained results and it showed consistence between data evolution and resulting RFS; moreover, weights satisfy the target with which the algorithm to compute them has been developed.

Nevertheless, lack of cases surely affects the analysis, so that it has not been possible to find a reduction of weights' variation over the years; the process of refinement to which weights should be subjected was not, therefore, discernible.

However, main issue about the framework is another one: United States have been selected as the community for the case study mainly considering the amount of data that it was possible to collect; indeed, availability of data about chosen indicators for other countries is so scarce that the analysis would have been almost impossible to carry out.

Moreover, even though data were available, comparison between different communities would be, in theory, improper, as noticed in section 5.1.

Likely solutions to these problems will be proposed in section 6.2.

However, this work has given satisfactory and original results, paving the way for future developments on mathematical quantification of Resourcefulness and other three Rs.

6.2 Future developments

Surely, further work is needed on this matter; first of all, problems shown in section 6.1 have to be solved.

A likely solution to solve the lack of cases to carry out a reliable Principal Components Analysis is to decrease the number of indicators (*i.e.* increasing ratio cases/variables); further debates on selection of indicators and case studies are, therefore, needed, in order to identify the ones which can be deleted. Alternatively, the possibility of using the division of indicators in dimensions (economic-political, preparedness, trust, creativity), by taking as an example the PEOPLES framework, may be discussed; by doing that, the number of indicators involved in any step decreases, whereas the amount of cases remains unchanged.

Comparability among communities may be achieved, on the one hand, by proposing a consistent survey-set, discussing the questions that have to be asked for each indicator and the kind of statistical sample; on the other hand, further debates on selection of indicators may be aimed to the reduction of indicators based on surveys.

Future works may also focus on the evaluation of Resourcefulness at finer spatial scales: for example, it could be assessed the feasibility of a case study involving all US states; indeed, case study carried out on USA at Chapter 5 obviously does not take into account the differences that exist among different states (*i.e.* at finer spatial scale). It stands to reason that many of the selected indicators are available at state-level, so that proposed case study can maybe be conducted without modifying the group of indicators.

Results obtained from this thesis pave the way for the usage of the proposed framework as support of resilience frameworks (*e.g.* PEOPLES): after having evaluated the Resilience of a community, computation of Resourcefulness may help to detect any weaknesses and to direct policymaking.

In this respect, the way in which that is done must be thoroughly discussed, and evaluation of other three Rs (above all Redundancy), and so their comparison, must be explored with the same care.

Appendix A

Below is given the MATLAB spreadsheet used to compute Resourcefulness of United States on a selected year.

The name of the .xlsx spreadsheet from which data are taken and the interval must be inserted in the section named "Data acquisition".

The user has to insert, specifically:

- Name of the file, number of the sheet and interval where data have to be taken; data have to be ordered in a matrix with cases in the rows and indicators in the columns;
- Name of the file, number of the sheet and cell where the first year (*i.e.* the first case) of the data set is;
- Name of the file, number of the sheet and interval where *TV* (target values) of each indicator is; they have to be ordered in a row.

The user has also to choose whether factor loadings have to be rotated or not.

Obviously, the spreadsheet can be used to compute Resourcefulness for other communities, but in that case has to be considered what has been discussed in section 6.1.

```matlab
%%%%%%%%% DATA ACQUISITION %%%%%%%%%

DATA=xlsread('USA.xlsx',2,'C5:R32');
firstyear=xlsread('USA.xlsx',2,'B5');
TARGETVALUES=xlsread('USA.xlsx',2,'C35:R35');


%%%%%%%%%%%%%% INPUT %%%%%%%%%%%%%%%%

year=input('What year do you want to compute?');
rotation=input('Do you want to rotate factors? (1=yes, 0=no)');
year=year-firstyear+1;




SizeDATA=size(DATA);
RowsDATA=SizeDATA(1,1);
ColsDATA=SizeDATA(1,2);

threshold=0.75;

y=RowsDATA;
while y>year
      DATA(y,:)=[];
      y=y-1;
end

SizeDATA=size(DATA);
RowsDATA=SizeDATA(1,1);



%%%%%%%%%%%%% NORMALISATION %%%%%%%%%%%

for i=1:ColsDATA
    DATA(:,i)=DATA(:,i)./TARGETVALUES(i);
end

ZSCORE=zscore(DATA);



%%%%%%%%% WEIGHTS ALLOCATION %%%%%%%%%

COVZSCORE=cov(ZSCORE);
[D,L]=eig(COVZSCORE);
Lambda=flipud(diag(L));
Lambda=Lambda';
D=fliplr(D);
k=1;
PCSELSize=0;

SUMLambda=0;
```

```matlab
while k<ColsDATA
       if SUMLambda<threshold*sum(Lambda)
          SUMLambda=SUMLambda+Lambda(k);
          PCSELSize=PCSELSize+1;
       end
       k=k+1;
end


if rotation==1
   D=rotatefactors(D(:,1:PCSELSize), 'Method','varimax', 'Maxit', 1500);
end


Dsq=D.^2;

som=zeros(ColsDATA,1);

for i=1:ColsDATA
    k=1;
    for j=1:PCSELSize
        som(i)=som(i)+Dsq(i,j)*Lambda(k);
        k=k+1;
    end
end

W=zeros(ColsDATA,1);

for i=1:ColsDATA
    W(i)=sum(som)./som(i,:);
end

w=zeros(ColsDATA,1);

for i=1:ColsDATA
    w(i)=W(i)./sum(W);
end


%%%%%%%%%%%%% AGGREGATION %%%%%%%%%%%%%%

RFS=0;
for i=1:ColsDATA
        RFS=RFS+DATA(RowsDATA,i)*w(i);
end
```

References

Aldrich, D. P. (2010). "Fixing recovery: social capital in post-crisis resilience."

Aldrich, D. P. and M. A. Meyer (2015). "Social capital and community resilience." American Behavioral Scientist **59**(2): 254-269.

Allen, M. (2013). "Deconstructing'resilience'in the Aftermath of disasters in Australia." Senshu Social Capital Review **4**: 45-64.

Arendt, L. A. and D. J. Alesch (2014). Long-term community recovery from natural disasters, CRC Press.

Ballano, V. O. (2017). Law, Normative Pluralism, and Post-Disaster Recovery, Springer.

Berkeley, A., et al. (2010). "A framework for establishing critical infrastructure resilience goals." Final Report and Recommendations by the Council, National Infrastructure Advisory Council.

Billon, P. L. and A. Waizenegger (2007). "Peace in the wake of disaster? Secessionist conflicts and the 2004 Indian Ocean tsunami." Transactions of the Institute of British Geographers **32**(3): 411-427.

Brown, K. (2015). Resilience, development and global change, Routledge.

Bruneau, M., et al. (2003). "A framework to quantitatively assess and enhance the seismic resilience of communities." Earthquake spectra **19**(4): 733-752.

Bruneau, M. and A. Reinhorn (2007). "Exploring the concept of seismic resilience for acute care facilities." Earthquake spectra **23**(1): 41-62.

Bryant, F. B. and P. R. Yarnold (1995). "Principal-components analysis and exploratory and confirmatory factor analysis."

Busa, M. G., et al. (2015). Trust-building through social media communications in disaster management. Proceedings of the 24th International Conference on World Wide Web, ACM.

Carlin, R. E., et al. (2014). "Trust shaken: Earthquake damage, state capacity, and interpersonal trust in comparative perspective." Comparative Politics **46**(4): 419-453.

Cimellaro, G. P. (2016). Urban resilience for emergency response and recovery, Springer.

Cimellaro, G. P., et al. (2010). "Seismic resilience of a hospital system." Structure and Infrastructure Engineering **6**(1-2): 127-144.

Commission, J. R. C.-E. (2008). Handbook on constructing composite indicators: methodology and user guide, OECD publishing.

Council, M. M. (2005). Natural Hazard Mitigation Saves: An Independent Study to Assess the Future Savings from Mitigation Activities, Volumes 1 and 2, Report to US Congress on behalf of the National Institute of Building Sciences, Washington, DC.

Cutter, S. L., et al. (2014). "The geographies of community disaster resilience." Global environmental change 29: 65-77.

Cutter, S. L., et al. (2010). "Disaster resilience indicators for benchmarking baseline conditions." Journal of Homeland Security and Emergency Management 7(1).

Cutter, S. L., et al. (2006). "The long road home: Race, class, and recovery from Hurricane Katrina." Environment: Science and Policy for Sustainable Development 48(2): 8-20.

Drabek, T. E. (2003). Strategies for coordinating disaster responses, Institute of Behavior Sciences Boulder, CO.

Eastin, J. (2016). "Fuel to the fire: Natural disasters and the duration of civil conflict." International interactions 42(2): 322-349.

Gan, X., et al. (2017). "When to use what: Methods for weighting and aggregating sustainability indicators." Ecological Indicators 81: 491-502.

Healy, A. and N. Malhotra (2009). "Myopic voters and natural disaster policy." American Political Science Review 103(3): 387-406.

Hinrichs, R., et al. (2011). "Report on the 2010 Chilean earthquake and tsunami response." US Geological Survey Open-File Report 1053.

Hotelling, H. (1933). "Analysis of a complex of statistical variables into principal components." Journal of educational psychology 24(6): 417.

Ibrion, M., et al. (2015). "Handling of dead people after two large earthquake disasters in Iran: Tabas 1978 and Bam 2003–Survivors' perspectives, beliefs, funerary rituals, resilience and risk." International journal of disaster risk reduction 11: 60-77.

Kaiser, H. F. (1958). "The varimax criterion for analytic rotation in factor analysis." Psychometrika 23(3): 187-200.

Kammouh, O., et al. (2017). "Resilience Evaluation of Urban Communities Based on Peoples Framework." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, under review.

Kieffer, C. H. (1984). "Citizen empowerment: A developmental perspective." Prevention in human services 3(2-3): 9-36.

Kreps, G. (1990). "Organizing for Emergency Management." Emergency Management: Principles and Practices for Local Government: 86-99.

Lee, G. and C.-H. Loh (2003). Human and Institutional Perspective of the 921 Earthquake in Taiwan: Lessons Learned.

Leys, C., et al. (2013). "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median." Journal of Experimental Social Psychology 49(4): 764-766.

Liu, C. and P. Robinson (2013). "Better organisation of volunteers in disaster settings is needed: lessons for all from China." Australian and New Zealand journal of public health 37(6): 595-596.

MacKinnon, D. and K. D. Derickson (2013). "From resilience to resourcefulness: A critique of resilience policy and activism." Progress in Human Geography 37(2): 253-270.

Mahalanobis, P. C. (1936). On the generalized distance in statistics, National Institute of Science of India.

Mendonça, D. and F. Fiedrich (2006). "Training for improvisation in emergency management: Opportunities and limits for information technology." International Journal of Emergency Management 3(4): 348-363.

Nel, P. and M. Righarts (2008). "Natural disasters and the risk of violent civil conflict." International Studies Quarterly 52(1): 159-185.

Nicoletti, G., et al. (1999). "Summary indicators of product market regulation with an extension to employment protection legislation."

Nogami, T. (2015). "The myth of increased crime in Japan: a false perception of crime frequency in post-disaster situations." International journal of disaster risk reduction 13: 301-306.

Organization, W. H. (2002). "Community participation in local health and sustainable development: Approaches and techniques."

Palen, L., et al. (2010). A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters. Proceedings of the 2010 ACM-BCS visions of computer science conference, British Computer Society.

Paton, D. and D. Johnston (2017). Disaster resilience: an integrated approach, Charles C Thomas Publisher.

Paton, D. and N. Long (1996). Psychological aspects of disasters: Impact, coping, and intervention, Dunmore Press.

Patterson, O., et al. (2010). "The role of community in disaster response: conceptual models." Population Research and Policy Review **29**(2): 127-141.

Pearson, K. (1895). "Note on regression and inheritance in the case of two parents." Proceedings of the Royal Society of London **58**: 240-242.

Pearson, K. (1901). "LIII. On lines and planes of closest fit to systems of points in space." The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2**(11): 559-572.

Perrow, C. (2011). The next catastrophe: Reducing our vulnerabilities to natural, industrial, and terrorist disasters, Princeton University Press.

Podolny, J. M. and K. L. Page (1998). "Network forms of organization." Annual review of sociology **24**(1): 57-76.

Rosenthal, U., et al. (2001). Managing crises: Threats, dilemmas, opportunities, Charles C Thomas Publisher.

SALAZAR, K. "REPORT ON THE 2010 CHILEAN EARTHQUAKE AND TSUNAMI RESPONSE."

Tabachnick, B. G. and L. S. Fidell (2007). Using multivariate statistics, Allyn & Bacon/Pearson Education.

Tierney, K. (2008). Structure and process in the study of disaster resilience. 14th world conference on earthquake engineering, Beijing, China.

Tierney, K. and J. Trainor (2004). "Networks and resilience in the World Trade Center disaster." MCEER: Research Progress and Accomplishments 2003-2004: 157-172.

Tierney, K. J. (2002). "Lessons learned from research on group and organizational responses to disasters." Countering Terrorism: Lessons Learned from Natural and Technological Disasters. Academy of Sciences, February.

Van de Kerk, G. and A. R. Manuel (2008). "A comprehensive index for a sustainable society: The SSI—the Sustainable Society Index." Ecological Economics **66**(2-3): 228-242.

Van Krieken, T., et al. (2017). Importance of community participation in disaster recovery. 13th IPGRC 2017 Full Conference Proceedings, University of Salford.

Whittaker, J., et al. (2015). "A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges." International journal of disaster risk reduction **13**: 358-368.

Wind, Y. and T. L. Saaty (1980). "Marketing applications of the analytic hierarchy process." Management science **26**(7): 641-658.

Yates, D. and S. Mackenzie (2018). "Heritage, Crisis, and Community Crime Prevention in Nepal." International Journal of Cultural Property **25**(2): 203-221.

Çağnan, Z., et al. (2006). "Post-earthquake restoration planning for Los Angeles electric power." Earthquake spectra **22**(3): 589-608.