Universitat de les Illes Balears
Master in Physics of Complex Systems.

Master's Thesis:

Multiple options noisy voter model: application to European elections

Supervisor: José J. Ramasco   Co-Supervisor: Juan Fernández-Gracia

Student: Gianmarco G. Pisano

UNIVERSITAT DE LES ILLES BALEARS

MASTER THESIS

Multiple options noisy voter model: application to European elections

Author:
Gianmarco Giuseppe Pisano

Supervisor:
Dr. José Javier RAMASCO
Co-Supervisor:
Dr. Juan FERNÁNDEZ-GRACIA

A thesis submitted in fulfillment of the requirements for the degree of Master in Physics of Complex Systems

in the

IFISC-CSIC-UIB

September 24, 2017

Universitat de les Illes Balears

Abstract

IFISC-CSIC-UIB

Master in Physics of Complex Systems

Multiple options noisy voter model: application to European elections

by Gianmarco Giuseppe Pisano

In the following master thesis we are going to present a noisy voter model based on social influence and recurrent mobility. The state of the agents of the model is described through a ternary option $a = \{-1, 0, 1\}$, and interactions between elements holding opposite opinion are not permitted (-1 and 1). The aim is to observe if our model is capable to reproduce the time evoution of real general elections, and the Spanish political context is chosen for a confrontation: IU, PSOE, and PP will correspond respectively to the options (-1,0,1) of our model. During the time evolution the agents will move repeatedly between the municipality in which they live and the municipality in which they work (commuting network), and at each time all agents present in the same municipality will be considered reciprocally neighbours. The commuting network is constructed according with the spanish census data, and electoral results of year 2000 are used as initial state of the whole system. A preliminar statistical analysis of real elections show us that the three options considered can be divided into majoritarian (PSOE and PP) and minoritarian (IU).The original version of the model will not be able to capture the statistical features of real elections data, giving rise to stationary states in which the three options result to be statistically equivalent (on the contrary of what observed in real election data). However, a simple modification of the same model (well sustained theoretically) will permit us to reproduce most of the aspects of time evolution showed by real data, as the stationarity of the vote sharing standard deviations , the conservation of the shape of the vote sharing distrutions, and the time evolution of the average vote sharing of the minoritarian option, which in this modification keep staying minoritarian.

Ringraziamenti

L'anno appena trascorso è stato veramente importante, di scoperta e di crescita, e questa tesi ne rappresenta in un certo senso la "meta". Vorrei dunque ringraziare anzitutto chi mi ha permesso di fare questa esperienza: Grazie IFISC; mi hai dato molto più di quanto tu possa immaginare. Ma grazie anche per essere ogni giorno quel posto unico che sei. Grazie ad ogni singola persona che ogni giorno fa parte di questo grande "fenomeno collettivo". Per quanto riguarda la mia esperienza, vorrei rivolgere un ringraziamento particolare a Maxi, Marta, Raul, Roberta, Antonia, Damia e Pere. Infine, un grazie speciale rivolto ai miei tutori José e Juan per tutta la pazienza e la disponibilità mostratami durante la stesura di questo lavoro.

La crescita però non è stata solamente di natura accademica ma anche un continuo confronto con un'altra cultura e con tutto ciò che questo implica; di sicuro, coloro che più di tutti mi hanno permesso di stare a contatto con i diversi aspetti della stessa sono stati i miei amici e colleghi Alejandro, Álex, Anna, Eduardo, Joan Losa y Joan Pont, Luca, Miguel, Oriol, Nacho, Patrick, Victor; più di tutti gli altri, siete stati "la mia Spagna". Grazie Ciccios.

Vorrei anche ringraziare la mia famiglia: sembrerà scontato da dire, ma senza i loro insegnamenti non sarei la persona che sono adesso. Mi sono sforzato molto, ma non riesco a trovare le parole per descrivere quanto importanti siate per me. Grazie Mamma, Qruzzu, Stefano. Un grazie speciale anche a Sabrina, Alessio e Alessia, che si stanno "prendendo cura di loro in mia assenza".

Seppure lontani geograficamente, gli amici veri sanno comunque starti accanto, e così hanno fatto Johnny (Michele), Il Guascone (Daniele), Ciccio (Francesco), ed Émile. Vi voglio bene picciotti, spero le nostre strade tornino a coincidere il prima possibile.

Per concludere, vorrei velocemente menzionare e ringraziare alcune persone senza le quali quest'anno avrebbe avuto tutta un'altra piega: grazie Aurora, Juanjo, Ludovica, Ralf.

Non potevo immaginarmi anno migliore.

Agradecimientos

El año que acaba de pasar ha sido veramente importante, de descubrimiento y crecimiento, y este trabajo representa de unas formas "el destino" del año mismo. Por lo tanto, me gustaría agradecer en primer lugar quien me permitió de hacer esta exprreriencia: Gracias IFISC, me has dado mucho mas de lo que tu puedas imaginarte. Ademas, gracias para ser cada dia aquel lugar unico que eres. Gracias a cada singula persona que cada dia forma parte de ese grande "fenomeno colectivo". En lo relativo a mi experiencia, querría manifestar mi agradecimiento particular a Maxi, Marta, Raul, Roberta, Antonia, Damia y Pere. Por ultimo, un agradecimiento especial a mis tutores José y Juan, Por toda la paciencia y la disponibilidad que me demostraron durante la elaboraci'on de este trabajo.

El crecimiento pero no ha sido solamente de natura academica, sino tambien una continua confrontacion con otra cultura, y con todo lo que eso implica; definitivamente, los que mas que nadie me permitieron de estar a contacto con muchos aspectos de la misma fueron mis amigos y compañeros Alejandro, Álex, Anna, Eduardo, Joan Losa y Joan Pont, Luca, Miguel, Oriol, Nacho, Patrick, Victor; mas de alguien otro, habeis sido "mi España". Grazie Ciccios.

Querría tambien agradecer mi familia: parecerá obvio de decir, pero sin sus enseñamentos no sería la persona que de hecho soy. Me Esforzé mucho, pero no acabo de encontrar las palabras para describir los importantes que sois para mi. Grazie Mamma, Qruzzu, Stefano. Y tambien un gracias especial a Sabrina, Alessio y Alessia, que "cuidan de ellos" en mi absencia.

Aunque geograficamente lejos, los verdaderos amigos sabes estar cerca de ti de todas formas, y asì hicieron Johnny (Michele), El Gascón (Daniele), Ciccio (Francesco), y Émile. Os quiero chicos, y espero que nuestros caminos vualvan a coincidir lo cuanto antes.

En conclusión, querrìa mencionar rapidamente y agradecer unas personas sin las cuales este año hubiese tenido una dirección totalmente distinta: Gracias Aurora, Juanjo, Ludovica, Ralf.

No habría podido imaginarme año mejor.

Contents

Abstract **iii**

1 Introduction **1**
1.1 Opinion dynamics and voter model 1
1.2 Our aim 4

2 The model **5**
2.1 Recurrent mobility 6
2.2 Neighbours 6
2.3 Quantities 6
2.4 Model Dynamics 8
2.4.1 Some modifications 9
2.5 Update rules 10

3 Preliminar analysis **11**
3.1 Vote sharing distribution 11
3.2 Spatial correlations 13

4 Model results **17**
4.1 Time evolution of the average vote sharing 17
4.2 Stationary vote sharing distributions 18
4.3 Stationary spatial correlations 18

5 Some modifications to the model; introduction and results **23**

6 Conclusions and future work **29**

A Construction of the commuting network **33**

B Preliminar statistical discussion **37**

"Denti stretti e testa bassa, sempre."

Chapter 1

Introduction

Since their origins, the purpose of the *computational social sciences* has been to understand the role of different interaction mechanisms between (social) elements of a system, and how such interactions result to be the origin of collective (social) phenomena. To do this, a multitude of models have been proposed during the years, in the context of human dynamics [1-2], social spreading [3-4], crowd behavior [5], hierarchy formation [6], but also language[7-80], cultural[9], opinion dynamics[10-17] and so on. Most of such models are characterized by very simple interaction rules: in fact, even though they could appear to be not *realistic*, their simplicity permit us to assess the effects of single interaction mechanisms on the way collective phenomena are formed.
Beside the theoretical developments, an absolutely fundamental part of the work consists in the confrontation at different levels between the model predictions and real data, from qualitative statistics features to concrete evolutions of the social systems. In this context, the recent technological developments, including the new information technologies, allow for a detail tracking and registering of human interactions and opinion expression in real time. The new computer power available has largely contributed to efficiently store and to access easily a huge data body on social systems; the simultaneous occurrence of these events has brought us to the so-called Era of Big Data.

1.1 Opinion dynamics and voter model

As already stressed in [18] "Agreement is one of the most important aspects of social group dynamics. Everyday life presents many situations in which it is necessary for a group to reach shared decisions. Agreements makes a position stronger, and amplifies its impact on society. [...] In any mathematical model, opinion has to be a variable, or a set of variables, i.e., a collection of numbers. This may appear too reductive, thinking about the complexity of a person and of each individual position. Everyday life, on the contrary, indicates that people are sometimes confronted with a limited number of positions on a specific issue, which often are as few as two: right or left, Windows or Linux, buying or selling, etc. If opinions can be represented by numbers, the challenge is to find an adequate set of mathematical rules to describe the mechanisms responsible for their evolution and changes.". The branch of the computational social sciences which deal with quantitative descriptions of the agreement is Opinion Dynamics, and one of the most succesfull models developed and used in these years, together with an uncountable quantity of variants, is the well-known voter model.

Agent based models

The voter model, and all its variants, can be well contextualized in a larger class of models, the so called agent-based models. Such models are characterized by the presence of microscopic elements, *the agents* of the systems, and the state of everyone of them is defined assigning a value to a few number of variables. Very often, such variables can assume only discrete values, and sometimes this values are even binary. The state of every agent is updated due to the interaction with other agents of the system[1]. The interactions between agents take place according to the update rules of the model; in particular, each agent will interact only with its *neighbours*.
The way in which the neighbours of each agent are defined, characterizes the *topology* of the system; we can imagine different topologies, as for example regular squared or triangular lattices, uni-bi-tri dimensional or even more; or we can also dispose our agents in a more varied configuration, as for example the case of a random network, in which the number of neighbours for each agent is not constant, and it is not possible to recognize a regular structure. The topology is a fundamental feature of the system and it affects the macroscopic properties of the set of elements playing a role as important as the one played by the interaction rules between the elements.
Agent-based models have many advantages: if from the one hand they are very easy to implement for numerical studies and simulations, on the other hand they can be well contextualized and studied through analytical instruments developed in the field of statistical physics. Furthermore, as already stressed, this class of models permit to shed light on the relation between the microscopic properties of the individual elements of the system, the interaction rules between them, and the topological properties of the system itself, with the collective phenomena related to the macroscopic scale.

The original voter model

In its original formulation [10,11], the voter model consists of a set of microscopic agents disposed in an (infinite) integer lattice of dimension d[2] whose states are defined as binary variables, *the opinion of the agent i* $\sigma_i = \pm 1$. At each time step, an agent of the system and one of its neighbours are randomly selected, say i and j; the update of the state is developed through a process wich is known as *unconditional imitation*: the opinion of i will be simply imposed equal to the opinion of j, $\sigma_i = \sigma_j$.
At $t \to \infty$, and for $d \leq 2$, the simple version of the voter model always reach the stationary, ordered collective state in which one of the two option has totally disappeared in favour of the other one; such configuration is also known with the name of *total consensus*, and with the language of the statistical physics it can be seen as an absorbing state. Otherwise, it is possible to show that for $d \geq 3$ the two options coexist in the system for an undetermined time [3].

[1]Depending on the context, the agents are also called with other names, as for example particles, individuals etc.

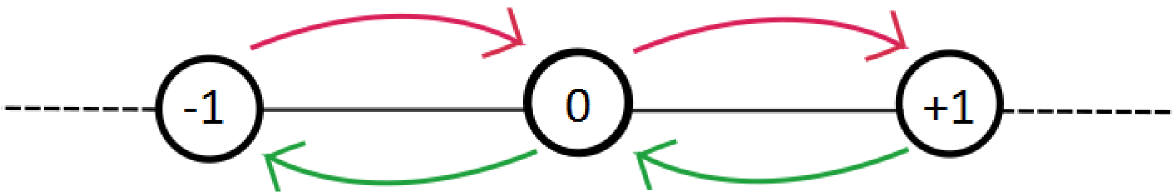
[2]The integer lattices are also known as cubic lattices. At each dimension d, the sites can be imagined as the points defined by integer d-plets in the correspondent euclidian space $\mathbf{R}^d$.

[3]What stated it is true, and can even be proved theoretically, in the case of infinite systems. Finite systems (the unique ones that can be simulated) reach always total consensus due to the finite-size-fluctuations.

Some variant versions of the voter model

We have seen how the key points that fix the collective properties of the system are its topology, the definition of the microscopic states of the agents, and the interaction rules. The modifications of such aspects will make the voter model more complicated, but also more varied. In this sense it is worth to cite the works of [12-13], in which a *noisy* voter model is studied: the modifications apported to the original version are related to the interaction mechanism between the agents and their neighbors: instead of the unconditional imitation of the neighbor's state, the agent involved in the temporal evolution will process an *imperfect imitation process*; this can be interpreted as it follows: the agent will copy the neighbor's state with probability λ, and it will assume the opposite one with probability $1 - \lambda$. The term *noisy* makes sense because in the original literature the model is studied with the wide formalism of the random walkers and stochastic processes, and the imperfect imitation is implemented trough the addition of a noise term. From a mathematical point of view the presence of the noise excludes the existence of infinite values of the correlation lenght, and this means that the collective states of *total consensus* are no longer absorbing states. So, in presence of noise, the voters will tend to organize themselves in spatial domains of size of the same order of the correlation lenght. The higher the value of the noisy term, and smaller will be the correlation length.
Another interesting modification to the original version of the voter model is the work proposed by F. Vazquez in 2003 [14], in which the state of each agent result to be ternary instead of binary: the three options represent leftist, centrist, and rightist ideologies. The leftist and the rightist are supposed to be so different that interaction between the two parts is forbidden (costrained interaction). Such modification gives rise to a wider collection of final absorbing states, but it is also responsible of a slower collective dynamics. During the elaboration of our work, the idea of a ternary opinion with costrained interaction will assume a foundamental role.
In conclusion of this section, just in order to understand how the different topologies can affect the collective properties of the system, we briefly present the results of the work developed by Castellano et al. (2003) [15] and continued in Vilone and Castellano(2004) [16]. Here the dynamics of the voter model is studied, contextualized to a small world network [18]. With such topology, what it is possible to observe in the case of a finite network is that the dynamics makes the system converge to an absolute consensus, as in case of the square lattice with $d \leq 2$. The interesting thing is that the transitory phase is characterized by the presence of a quasistationary state which corresponds to a *plateau* in the plotting of the fraction of active links of the network (links between agents with different opinions)[4] in function of the time. In correspondence of this plateau the correlation lenght assumes finite values, and this indicates the presence of clusters. Moreover, the works show that the temporal window in which the system stay in this quasistationary state is proportional to the number of elements of the network itself, so that, for $N \to \infty$, social consensus never will be reached, and the system will remain in the quasistationary state.

[4]The fraction of active links is selected as order parameter of the system. Working in connected graphs, this quantity will be equal to zero only when social consensus is reached.



2a

fig 1 A simple representation of the one-dimensional opinion space. Every agent can assume opinion states which are correspondent to discrete values on the line, correspondent. In our model we consider only three options, so that $\sigma = \{-1, 0, 1\}$, and transitions (represented by the arrows) are permitted only between adjacent opinions.

1.2 Our aim

The findings of the work developed by Keith T. Poole and Howard Rosenthal [19] show that most of opinion dynamics in a political context can be well described with the use of a metric space (let us call it *opinion space*) in which electors are placed in function of their opinions; an agent's displacement in such space corresponds to an agent's opinion change. Taking such results as a starting point, and considering also the electoral data modeling exercise of [17], in this thesis we are going to introduce a more elaborated model which takes into account the existence of more than two electoral options. We introduce a one dimensional space representing a left to right ordering of the parties, in which every elector is placed in correspondence of one of the options (see fig.1). Hence, at every time step they can move (change opinion) according to the adaptation-of-the-voter rules. To test the model in a simple scenario, we use data from Spanish general elections and reduce the voters' options to only three parties[5]. Basically, our aim is to observe whether the dynamic of our model is capable to reproduce results whose statistical features and temporal evolution are similar with the ones of the real elections data.

[5]The Spanish electoral system, as most of parliamentary democracies in Europe, is characterized by the coexistence of many parties. For the sake of convenience, we reduce the complexity to only three. However, we will discuss later the effect of including more parties in the model.

Chapter 2

The model

In this chapter we will explain in detail the different aspects of the model and we will describe the features of the system studied. We will introduce the idea of recurrent mobility (section 2.1) and we will define the neigbours of each agent (section (2.2)). In section 2.3 we will define some basical quantities, and in section 2.4 we will use them to introduce the rates at which agents change opinion, making the system evolve; finally (section 2.5) we will explain in detail the update rules we used to compute the model.
Our starting point for the new model is the one exposed in [17]: we deal with a Noisy voter model with recurrent mobility. The agents of the system represent electors, and they are spatially divided in municipalities. The American and the Spanish political situations are different, as well as the models used to the description of the two : for this reason, it is worth to stress the differences betwen the two cases.
Accordingly with the American two-party political system, in the previous model every agent was characterized by opinion $+1$ or -1. Electoral results of the year 2000 were used as initial data: for the construction of the model, only the elements who voted one of the two main options (Republican or Democratic party) were considered: said with other words, Democratic and Republican agents are the only elements considered by the model in [17]. All the voters that voted for other options (or did not vote at all) were not considered. Such approximation results to be reasonable in the American context, in which the Republican and Democratic parties are absolutely majoritarian.
The same statement does not hold in the Spanish context, in which we can recognize many more parties, different between them not only for their importance (electoral vote sharing) or political belief, but also for their sense of regional identity. In fact, the Spanish context is characterized by the presence of many parties that, altough they are not present in the whole national territory, in the passed elections they have even reached the majority of votes in their own belonging regions [1], and such circumstance makes not possible to ignore their presence in the description of the dynamics of the system.
What we do is to describe the system with a "3+1" parties mechanism: the agents of the system subjected to the dynamics can individually hold opinion $(+1)$, (0), or (-1), which according with the left-to-the-right ordering correspond with the IU, the PSOE, and the PP parties; during the dynamic the agents will have the possibility to change opinion, but in any case such opinion will be one of the three introduced above. However, as said before, the elements of these parties will not constitute the totality of the system: we introduce a "fourth option", say opinion $(\times)$, in which we include all the voters of all the other different options respect to the three ones cited

[1] As example of regional parties we can cite the CIU in the catalan region (nowadays known as PDeCAT), the PNV in the Basque Country, BNG in the Galicia region etc.

before[2]. These agents are not subjected to the dynamics of the system, so that they result to be constant in time. We can state that their presence inside the system can be seen as a set of exogenous variables which affect the evolution of the system itself. Electoral result of Spanish general elections of 2000 were used as initial conditions of the system.

2.1 Recurrent mobility

Basically, our agents are divided into cells, which represent the municipalities in which people live or work. Our agents are supposed to perform daily recurrent movements between the cell they live in and the cell in which they work (fig 2a): any agent is supposed to pass a fraction α of the day in the residence place, and a fraction $(1-\alpha)$ in the working cell. Recurrent mobility is foundamental in order to permit the opinion diffusion between different cells, and coherently with a statistical survey of the I.N.E. (Spanish National Statistical Institute) [20] we impose $\alpha = 0.5$, which corresponds to the situation in which any person spend 8 hours working, 8 hours in the residential place and 8 hours sleeping.

2.2 Neighbours

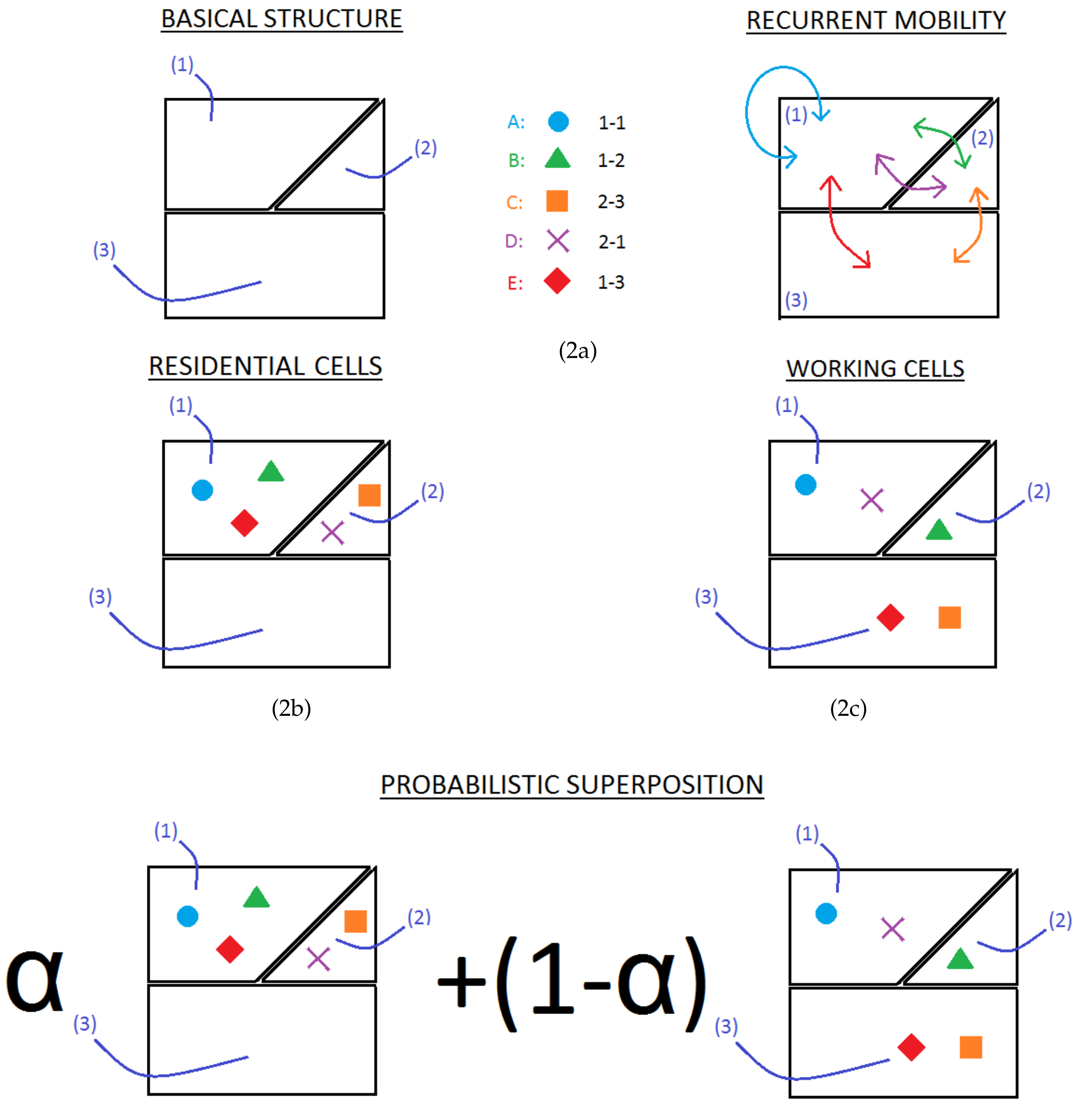
In our model, the neighbours of an agent at time t will be all the other agents present in the same cell at the same time : all the agents living in the cell i when the elements is in its residence cell, and all the workers in j when the agent is in its working cell[3] (fig 2b-c). The time scale at which the agents-recurrent-mobility takes place results to be smaller respect to the time scale at which elections occurs. Such consideration allows us to give an equivalent interpretation for the localization of the agents and the neighbours interaction: we can consider every element delocalized in both the cell i and j at any time; in such case, it will interact with the neighbours of the cell i with probability α, and with the agents of the cell j with probability $1-\alpha$ (fig 2d).
Having exposed the "kinetic rules" of the agents, now we are ready to introduce some quantities of interest.

2.3 Quantities

Let us start with some definitions: the recurrent mobility typical of the model suggests us to define the following commuting cells; every commuting cell is identified by a pair of ordered indeces ij. Basically, the agents belonging to such commuting cell will live in the municipality i and work in the municipality j. From this point of view, the quantity N_{ij} will indicate the number of people living in commuting cell ij. At the same way, we can also define the number of people holding opinion (-1), (0), or $(+1)$, belonging to such commuting cell, V_{ij}^{-}, V_{ij}^{0}, and V_{ij}^{+}. The criteria used for the construction of the whole commuting network starting from the electoral data are explained in detail in the appendix A.

[2] In the fourth option we will include also the voters who voted a null option and the ones who did not vote.

[3] We assume that all the agents of the system work simultaneously, and stay in their residential cell also all at the same time.



2d

FIGURE 2:**fig (2a)** Simple example of municipalities in which people live and work. On the center, for each agent, municipalities in which they live and work are specified. On the right, a basical representation of the recurrent mobility is given. **fig (2b-c)** Representations of the system during the free and the working time. During the free time, agents A-B-E are neighbours, staying all in the municipality 1, as well as agents C-D, located in municipality 2. During the working time we have agents A-D in the municipality 1, agent B in municipality 2, and agents C-E in municipality 3. We assume that all the agents of the system work simultaneously, and stay in their residence municipalities also all at the same time. **fig (2d)** Representation of the probabilistic interpretation of the state of the system. Due to the different time scales of the recurrent mobility and electoral votations, the system can be imagined to be in a state which is a superposition (with coefficients α and $1-\alpha$) of its configurations during the free and the working time.

We briefly introduce the following quantities:

$$N_i = \sum_j N_{ij} \quad \longrightarrow \quad \text{Number of people living in the municipality } i \tag{2.1}$$

$$N'_j = \sum_i N_{ij} \quad \longrightarrow \quad \text{Number of people working in the municipality } j \tag{2.2}$$

$$V_i^a = \sum_j V_{ij}^a \longrightarrow \text{Number of people holding opinion a=\{-1, 0,+1\}, living in the municipality } i \tag{2.3}$$

$$V_j'^a = \sum_i V_{ij}^a \longrightarrow \text{Number of people holding opinion a=\{-1, 0,+1\}, working in the municipality } j \tag{2.4}$$

2.4 Model Dynamics

Trough the quantities just exposed above, it is possible to define the probabilities p_{ij}^{-0}, p_{ij}^{0-}, p_{ij}^{0+}, p_{ij}^{+0}, with which the agents change opinion at time step t:

$$p_{ij}^{-0}(t) = \left[\alpha\left(\frac{V_i^{(0)}(t)}{N_i}\right) + (1-\alpha)\,\frac{V_j'^{(0)}(t)}{N'_j}\right] + D\eta_{ij}^{-0}(t) \tag{2.5}$$

$$p_{ij}^{0-}(t) = \left[\alpha\left(\frac{V_i^{(-)}(t)}{N_i}\right) + (1-\alpha)\,\frac{V_j'^{(-)}(t)}{N'_j}\right] + D\eta_{ij}^{0-}(t) \tag{2.6}$$

$$p_{ij}^{0+}(t) = \left[\alpha\left(\frac{V_i^{(+)}(t)}{N_i}\right) + (1-\alpha)\,\frac{V_j'^{(+)}(t)}{N'_j}\right] + D\eta_{ij}^{0+}(t) \tag{2.7}$$

$$p^{+0}(t)_{ij} = \left[\alpha\left(\frac{V_i^{(0)}(t)}{N_i}\right) + (1-\alpha)\,\frac{V_j'^{(0)}(t)}{N'_j}\right] + D\eta_{ij}^{+0}(t) \tag{2.8}$$

The equations above are very immediate to understand; in the original formulation of the voter model [10,11] the opinion of an agent is updated selecting randomly one of its neighbours, and copying unconditionally its state: we have already defined the neighbours of an element in section 2.2. Let us make the following example: an agent belonging to the commuting cell ij holding opinion (-1) will change (deterministically) its state due to the interacion with one of the agents holding opinion (0) which live in the municipality i or work in the municipality j. Actually, the probability to interact with one these agents in the cell i (j) is given by the fraction $V_i^{(0)}/N_i$ $(V_j'^{(0)}/N'_j)$. This permit us to understand the term inside the square brackets, where the two fractions are summed after being multiplied by the correspondent interaction probabilities α and $(1-\alpha)$. Furthermore, in order to consider imperfect imitation processes we add by hand the term $(D \cdot \eta_{ij}^{-0})$, where D is the noise parameter, which regulates the intensity of the imperfect imitation, and η_{ij}^{-0} is uncorrelated gaussian white noise with zero mean and unitary variance, that is:

$$\eta_{ij}^{ab}(t)\,\eta_{lm}^{cd}(s) = \delta_{ac}\;\delta_{bd}\;\delta_{il}\,\delta_{jm}\;\delta_{ts} \tag{2.9}$$

Equations (2.5)-(2.8) resume the whole dynamic of the noisy voter model with recurrent mobility: in particular, we stress how the generic quantity N_i is equal to:

$$N_i = V_i^{(-1)} + V_i^{(0)} + V_i^{(+1)} + V_i^{(\times)} \tag{2.10}$$

This is how the elements of the "fourth option" come inside the dynamic of the system: a commuting cell with a high presence of such elements in the related residential and working municipalities, will be characterized by a dynamic that will be slower respect to another commuting cell in correspondence of which $V_i^{(\times)}$ and $V_j'^{(\times)}$ are lower.
Another important point to stress is that, according with eq(2.5 - 2.8), the only permitted transitions are the ones such that the opinion of any agent change from a value to an adjacent one: similarly with the work in [14], no transitions are permitted from the state (-1) to $(+1)$ and viceversa. The reason why such constrain is applied is that the agents are supposed to change their opinion progressively.
The strenght of equations (2.5 - 2.8) is that *they permit a mesoscopic approach for the simulation of the model*.

2.4.1 Some modifications

As already stressed, the results of our model will be compared with the Spanish electoral results of different years (presented in detail in the next chapter). A peculiarity which emerges from the election results is that the parties considered in our model dynamics can be separated in two different groups. In fact, differently from the other two parties, the IU cover the role of "minoritarian option". Such particular aspect will not rise in the model results (chapter 4), so, in order to have a better correspondence with real data, we shall make some modifications to the original model. In particular, what we will do is to modify the probability with which the agents change opinion from 0 to (-1); to do this, we multiply the original probability by a factor $\gamma \in [0, 1]$.

$$p_{ij}^{0-} \longrightarrow p_{ij}^{0-'} = \gamma \cdot p_{ij}^{0-}$$

A wider background which can justify such choice will be given in chapter 5, together with the results of the modified model.

2.5 Update rules

By definition, the Monte Carlo step of a model is the set of updates the simulation must do in order to reproduce a single update of the whole system. Basically, the concept behind the Monte Carlo step is that "every element of the system must have the opportunity to change its state, at least once". In the previous section we have defined the probabilities p_{ij}^{ab} $(a, b = -1, 0, 1)$, and we have also concluded the section saying that such probabilities permit a mesoscopic approach for the simulation of the model; let us proceed by order.
Let us consider a particular commuting cell ij. Equations (2.5-2.8) permit us to know the probability with which a single agent of the cell ij in the state a change its opinion in b. Now, in the hypotesis in which the states of the elements of the same commuting cell will be updated synchrounously, the changes inside the commuting cell will happen at the same time, so that they will be independent between them. Such independence can be used for the mesoscopic update of the commuting cell ij: instead of computing the change of a single agent at any step, we can compute the total number of negative elements becoming zero (and viceversa), $\Delta_{-\to 0}$ ($\Delta_{0\to -}$), and the number of zero elements becoming positive $\Delta_{0\to +}$ (analogously, we calculate $\Delta_{+\to 0}$) generating random numbers in accordance with the following Binomial distributions $B(V_{ij}^{-}, p_{ij}^{-0})$, $B(V_{ij}^{0}, p_{ij}^{0-})$ $B(V_{ij}^{0}, p_{ij}^{0+})$, and $B(V_{ij}^{+}, p_{ij}^{+0})$.
In conclusion, any Monte Carlo step with wich we update the state of our system is made in accordance with the following rules:

1. Select all the commuting cells of the network in a random order.

2. For each commuting cell selected, calculate the probabilities r_{ij}^{ab}, and generate the quantities $\Delta_{a\to b}$ as random variables following the binomial distributions $B(V_{ij}^{a}, p_{ij}^{ab})$. Calculate the variations ΔV_{ij}^{a} .

It is important to stress how the whole update procedure result to be neither synchrounous, nor asynchronous: if from the one hand the state of the agents of a single commuting cell are updated with the hypotesis of synchronicity, on the other hand any change ΔV_{ij} will affect the calculation of the rates on the successive commuting cells.

Chapter 3

Preliminar analysis

Before showing the model results, we briefly expose the statistical features of the Spanish general elections from the year 2000 to the year 2008. The Spanish electoral system, as most of parliamentary democracies in Europe, is characterized by the co-existence of many parties. For the sake of convenience, we reduce the complexity to only three. The parties chosen with which we identify the three options considered in the model are the IU, the PSOE, and PP, which correspond respectively to the options (-1), (0), and (1). This choice is supported by many reasons: these three parties (in particular the PSOE and the PP) result to be the biggest political entities in the election considered. Moreover, all of these three options are well spread all around the national territory. Finally, the political identities are coherent with the order inside the opinion space. We will focus our attention on the vote sharing distributions (section 3.1), and on the spatial correlations (section 3.2).

3.1 Vote sharing distribution

The vote sharing for the option a and the municipality i is defined as:

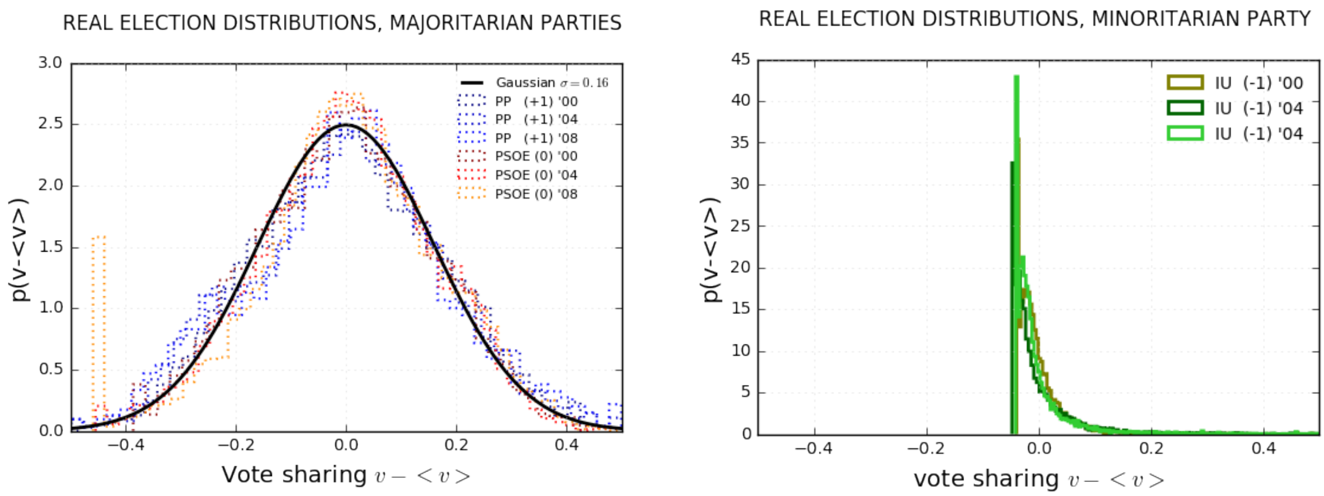
$$v_i^{(a)} = \frac{V_i^{(a)}}{V_i^{(tot)}} \quad ; \quad V_i^{(tot)} = V_i^{(-)} + V_i^{(0)} + V_i^{(+)}. \tag{3.1}$$

It is important to stress that they are calculated putting on the denominator $V_i^{(tot)}$, the sum of voters related to the three dynamic options studied by the model. In this sense, the vote sharings considered here differ by the "real" ones, in the sense that they are not considered respect to the total number of people having the right to vote[1].

On the table below we report the average vote sharing the three parties obtained during different elections, toghether with the correspondent standard deviations:

Year	IU $V_i^{(-)}$	$\sigma^{(-)}$	PSOE $V_i^{(0)}$	$\sigma^{(0)}$	PP $V_i^{(+)}$	$\sigma^{(+)}$
2000	0.039	0.0462	0.386	0.1589	0.575	0.1669
2004	0.047	0.0697	0.454	0.1556	0.499	0.1759
2008	0.041	0.0646	0.459	0.1720	0.500	0.1807

[1] The vote sharings defined in eq(11) would have been equal to the real ones if, instead of $V_i^{(tot)}$ we would have put N_i in the denominator of eq (11), that is the total number of people(voters) present in the cell i.



(3a) (3b)

FIGURE 3: plots of vote sharing distibutions related to general Spanish elections of the years 2000, 2004 and 2008. **fig (3a)** electoral results related to PSOE and PP are reported here. Is it possible to observe how the distributions appear gaussian-like.**fig (3b)** electoral results for the IU are reported here. We observe that such distributions are characterized by a sharp peak on the left, which is coherent with the fact that in many municipalities the IU option was not voted at all

A quick look to the values reported on the table permit us to understand that the vote sharings obtained by each party during different elections, as well as the related standard deviations, are more or less constant in time. The PSOE and PP parties (options 0 and 1) are fairly similar, meanwhile the IU party (option -1) plays the role of minoritarian option. From now on, we will refer to the first two parties calling them "majoritarian" options, meanwhile the IU party will be called "minoritarian" option. The difference between the parties can be appreciated plotting the vote sharing distributions. We do this in figure 3(a-b), where the vote sharing distributions are shownn in function of $v^a - \langle v^a \rangle$. [2] The curves are obtained as normalized histograms. In fig 3a we appreciate the fact that the vote sharing distributions related to the majoritarian options appear to be gaussian-like, although with some rather visible deviations in municipalities where one of the two options is not well implemented. These results are coherent[3] with the work developed in [21].
The same gaussian-like behavior does no hold for the minoritarian option, whose

[2] The tranlation term $-\langle v^a \rangle$ will shift the distribution so that the value of the distribution $f(\langle v^a \rangle)$ will be plotted in correspondence of the null value on the x-axis. This permit to better recognize symmetry properties of the distributions obtained.

[3] Our definition of vote sharing result to be slightly different from the normally used one. For each municipality i the two vote sharings are related by the relation $v_{r,i} = v_i \cdot (V^{(tot)_i}/N_i)$. So, the transformation which links the two distributions is not properly a dilation, because the dilation factor $(V_i^{(tot)}/N_i)$ is not constant, and in principle there is no reason to think that gaussianity-like must be conserved. By the way, the dilation factor it is not that far to be constant: for each municipality i, it represents the ratio between the voters who voted one of the three dynamic options and the total number of voters. Considering options that are majoritarian even respect to the global, real, vote-sharing, it makes sense to assume that the dilation factor assumes values that are approximately equals. Coherently with this, we observe gaussian-like distributions, even if we are not considering the real vote-sharings. In fact, it is possible to show analitically that the dilation of a gaussian-like function is also a gaussian-like function

plots(figure 3b) are characterized by the presence of strong peaks on the left side, which corresponds to municipalities in which the IU was not voted at all (see fig. 5).

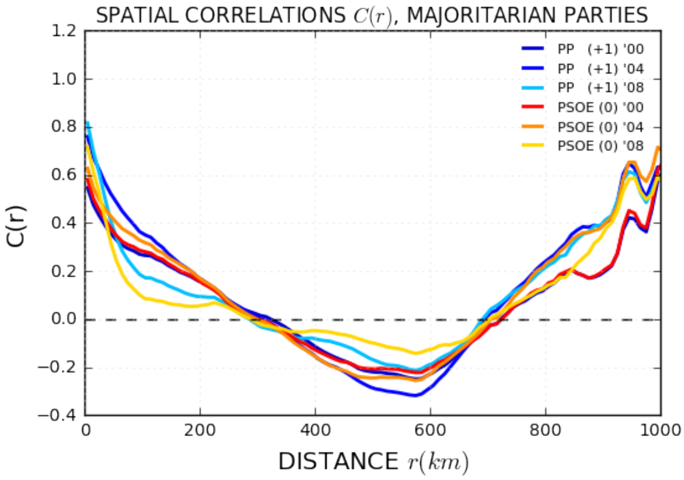
3.2 Spatial correlations

Spatial correlations in function of the distance r are calculated as:

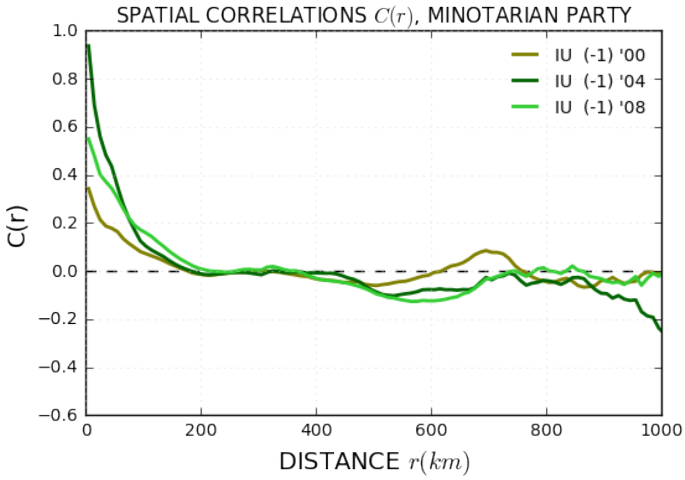
$$C^a(r) = \frac{\langle v_i^a v_j^a \rangle_{d(i,j)=r} - \langle v^a \rangle^2}{(\sigma^a)^2} \tag{3.2}$$

The first term on the numerator represents an average made considering all the pairs of municipalities (i, j) whose distance[4] is r. The terms $\langle v^a \rangle^2$ and $(\sigma^a)^2$ are referred to the whole set of vote sharings obtained by the option a in the different municipalities.
Correlation results are reported in Figure 4(a-d). In order to consider only results supported by a good statistics behind, data are shown only for an interval of $r \in [0 - 1000]km$. A deeper discussion on the motivation of such choice, together with other important details regarding the data analysis, can be found in appendix B.
Similarly with what happened for the vote sharing distributions, the minoritarian option (IU) shows a different behavior respect to the majoritarian ones (PSOE and PP). The decay of the IU correlation appears logarithmic (fig 4d), as already observed in other cases [22,23], and it reaches approximately the zero value in correspondence of $r \approx 200km$ (fig 4b). On the other hand, majoritarian options are charactherized by a decay which appears linear and not logarithmic (fig 4a), and they reach the zero value at $r \approx 300\,km$.
The plots of $C^a(r)$ provide with us important informations. The high positive values of $C^a(r)$ in correspondence of small values of the distance r followed by a systematic decay indicate the presence of domains, geographical areas in which the vote sharing for a particular party tend to assume similar values. The value of r at which $C(r)$ reaches the zero value has an important physical sense: it indicates the characteristic size of the domains. Moreover, the negative values of the correlation functions in the range between 300 and 700 km suggest us that most of the pairs used for the calculations of $C(r)$ (eq 3.2) belong to different clusters, characterized by different dominant opinions. It is interesting to observe how such characteristic sizes result to be constant in time. Such domains can be observed in detail in fig. 5, where the vote sharing maps of the different elections are reported. It is important to stress how a brief sight at these maps permit us to appreciate that most of the IU domains are located in the same regions in which PSOE domains are present too. This consideration fortify our basic assumption according to which the IU agents change their opinion only due to the interaction with agents holding opinion (0) (PSOE voters). The confrontation of maps belonging to different elections give us another interesting information: independently from the year considered, we can generally recognize a north-to-south opinion alignment: The PP option result to be stronger in the north regions, meanwhile the PSOE presence is dominant in the southern part(Andalucia region). In the center, we can recognize also a treshold line (in white) in which no one of the two options is dominant.

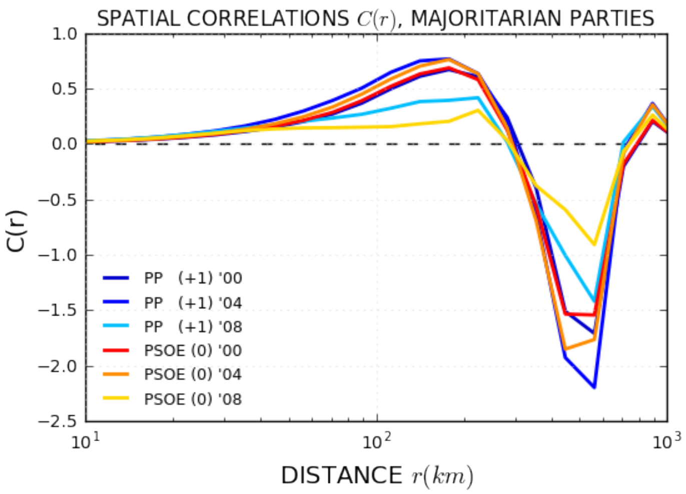
[4]The distance between two municipalities is defined as the distance between the centroids of the two geographical area, and it is calculated using the Vincenty's formula.



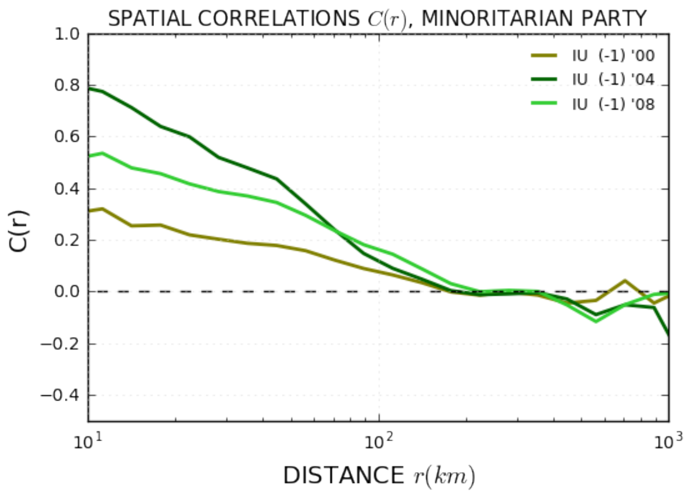
(4a)



(4b)

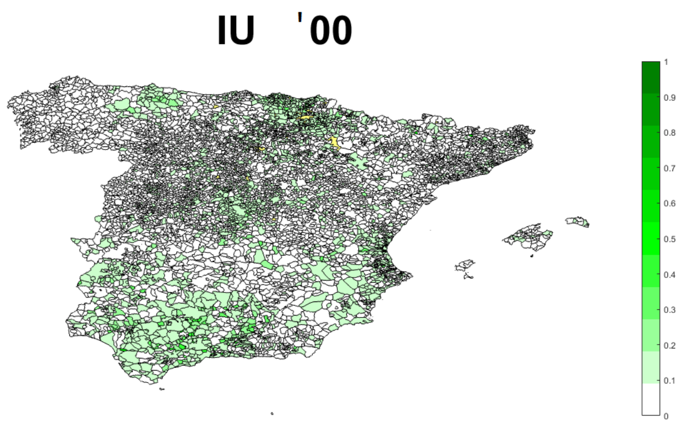


(4c)

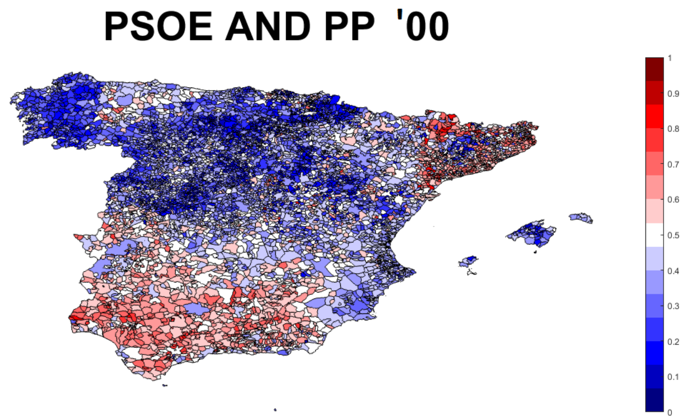


(4d)

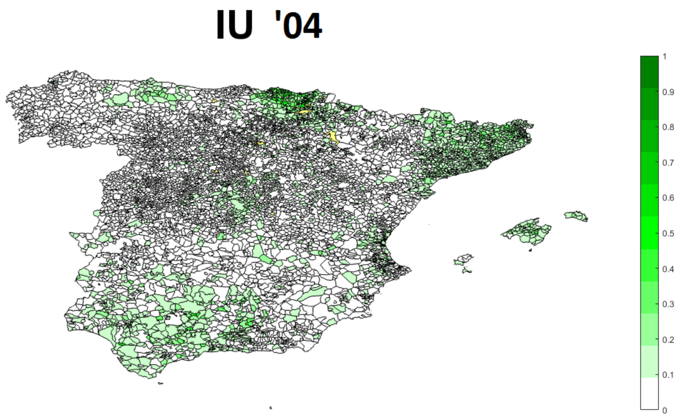
FIGURE 4:Plots of correlations in function of distance, related to general Spanish elections of the years 2000, 2004 and 2008. **fig (4a)** PSOE and PP correlations are reported here. Is it possible to observe how the decay appear linear, reaching a value equal to zero approximately at $r \approx 300\,km$. **fig (4b)** IU correlations are reported here. It is possible to observe how the the three curves reach the zero value for $r \approx 200\,km$. **fig (4c)** PSOE and PP correlations are reported here. Differently from fig 4a, logarithmic bins have been used. **fig (4d)** IU correlations, calculated using logarithmic bins.



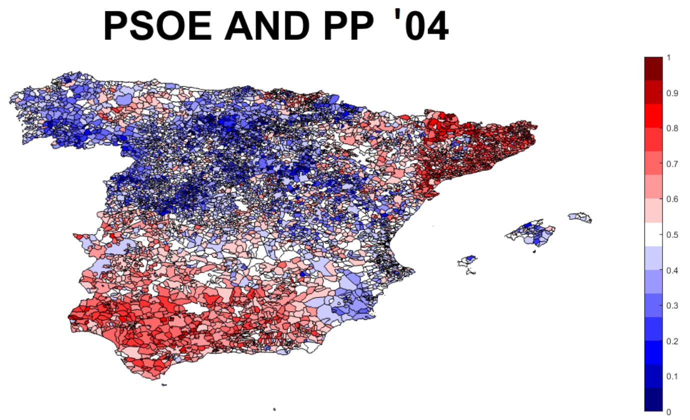
(5a)



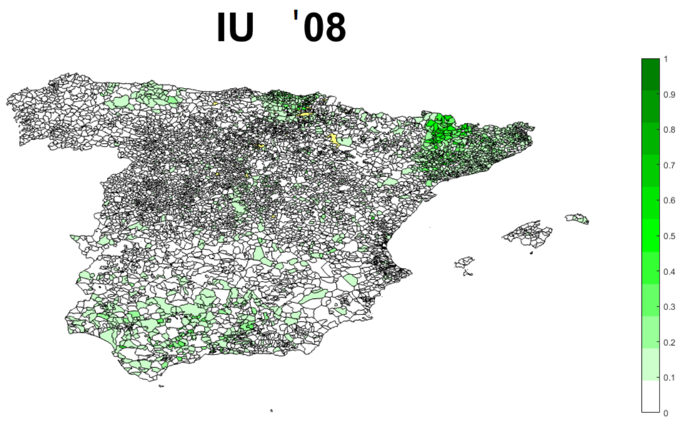
(5b)



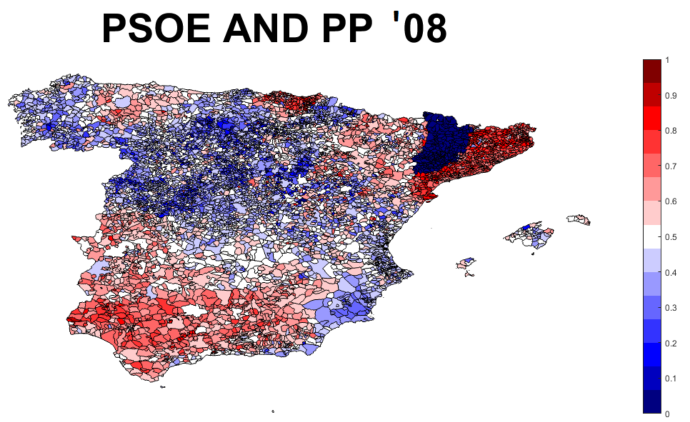
(5c)



(5d)



(5e)



(5f)

FIGURE 5: Vote sharing maps for different elections and different parties. The IU data are plotted separately: due to its nature of minoritarian party, its vote sharing does not affect significantly the PSOE and PP distributions, so that these two options can be plotted togheter and interpreted as two opposite choices. Red and blue colours indicate respectively a PSOE or a PP majority. White colour indicates that the two parties obtained both a vote sharing around $v = 0.5$. The values in such colorbars are referred respect to the PSOE vote sharings $v^{(-)}$ normalized respect the sum of PSOE and PP vote sharings:

$$v^{(-)}_{maps} = v^{(-)}/(v^{(-)} + v^{(+)})$$

fig (5a) IU vote sharings of 2000 **fig (5b)** PSOE and PP vote sharings of 2000
fig (5c) IU vote sharings of 2004 **fig (5d)** PSOE and PP vote sharings of 2004
fig (5e) IU vote sharings of 2008 **fig (5f)** PSOE and PP vote sharings of 2008

Chapter 4

Model results

In the previous chapter we have presented different statistical features of real Spanish elections. Coming back and better analizing the plots and tables presented, we can notice an important fact: from a statistical point of view, the different general elections studies are pretty similar between them. In section 1.2 we stated that our aim is to observe whether our model is capable to reproduce results statistically comparable to real elections. On the basis of what showed in chapter 3, this purpose coincides to observe whether, for some values of the noise term D, our model admits the existence of stationary states which conserve the statistical properties of the initial state.
In this chapter we will show the model results with different representative values of the noise parameter $D = \{0\ ,\ 0.20\ ,\ 0.40\ ,\ 0.60\}$. In section 4.1 we will focus our attention on the time evolution of the average-vote-sharings and the related standard deviations. In section 4.2 we will report the vote sharing distributions related to the stationary states of the system. Finally In section 4.3 we will report the spatial correlation plots, again referred to the stationary states.

4.1 Time evolution of the average vote sharing

For any value of $D \neq 0$ the time evolutions of the average vote sharings are characterized by the same behavior : all the plots (fig. 6a-c-e-g) made for different values of D show that, even if the system starts from some initial conditions in which it is possible to distinguish a minoritarian option (IU) and two majoritarian ones (PSOE, PP), the dynamics is such that the system always converge to a stationary state in which $v^a \approx v = 1/3 \quad \forall a$, so that at the stationary the three options result to be statistically equivalent[1]. In accordance with this, we conclude that a value of D such that the statistical properties of the initial state are preserved does not exists. The discrepancies with real data can be fixed apporting some modifications at the original model, as we will observe in Chapter 5.
Even if the convergence takes place $\forall D$, we can still observe some aspects which depend on the value of the noise parameter, as for example the time needed by the system to reach the stationary state. The higher D, the faster will be the convergence. Also the standard deviations of the stationary vote sharing distributions show a dependence on D. As we can observe in fig.6b-d-f-h, they increase with increasing D.

[1]As it is possible to observe by the plots in fig.6, it is also true that for increasing values of D the PSOE tends to assume an average vote sharing slightly greater respect the vote sharings of the other parties. It could be that such behavior is related to the fact that the PSOE is the only party of the system which can interact with all the other ones. However, being interested in obtaining the preservation of the initial statical features, we will not investigate further the reasons of such behavior.

A special case is given by $D = 0$: Even if the average vote sharings assume values fairly similar to the initial ones (fig.6a), the standard deviations of the majoritarian parties decay strongly (fig.6b). Having observed that the system always reach a stationary state, in the next sections we will report only the stationary-vote sharing distributions.

4.2 Stationary vote sharing distributions

The stationary vote sharing distributions of the three options becomes indistinguishable between them for any value of $D \neq 0$. This is another evidence of the fact that the model gives rise to a dynamic such that the three options eventually become equivalent. We have already observed how the width of the distributions (the standard deviations) increase with increasing D. Different behaviors are showed in fig. 7 a-d; in each figure we plot also a referential gaussian distribution with σ equal to the standard deviations of the correspondent vote sharings at the stationary regime. As we can see, for small values of D (fig. 7a) the vote sharing distributions result to be sharper respect to the referential gaussian. However, higher values of D make the distribution wider, but destroy the symmetry: the effect is such that in each municipality the vote sharings will tend to be more unbalanced in favour of a single option. As consequence of this, in correspondence of small and high values of v the stationary distributions will be higher respect the referential gaussian, but it will be lower for intermediate values (fig. 7 b-d). Higher will be D, higher will be the unbalance. What said is confirmed by the vote sharing maps. In fig. 9 a-b we report the maps referred to the IU and PSOE options. A quick look at such maps can confirm what stated above even if, according with the result reported in fig. 6 g-h and fig. 7d, the two parties are characterized by a slightly different behavior. Even if not reported, we expect that the PP map will be characterized by a behavior similar to the IU map.

4.3 Stationary spatial correlations

In correspondence of the stationary state, and for every value of D, we observe that the spatial correlation function $C(r)$ is constantly equal to zero. This means that differently from the real geopolitical situation, in the stationary states of the model it is not possible to identify any kind of spatial domain. This results disagree with the results obtained in [17], contextualized at the american elections, in which initial correlations grow for small values of D. Further simulations based on the elimination of the minoritarian option can prove that such difference is not due to the presence of the third option, and should be caused by different topological properties of the two commuting networks. We will not go further in this discussion, but two pictures which stress the differences between the American and the Spanish commuting networks can be found at the end of appendix A. A quick look at the vote sharing maps (fig. 9 a-b) will confirm the absence of clusters in the system.

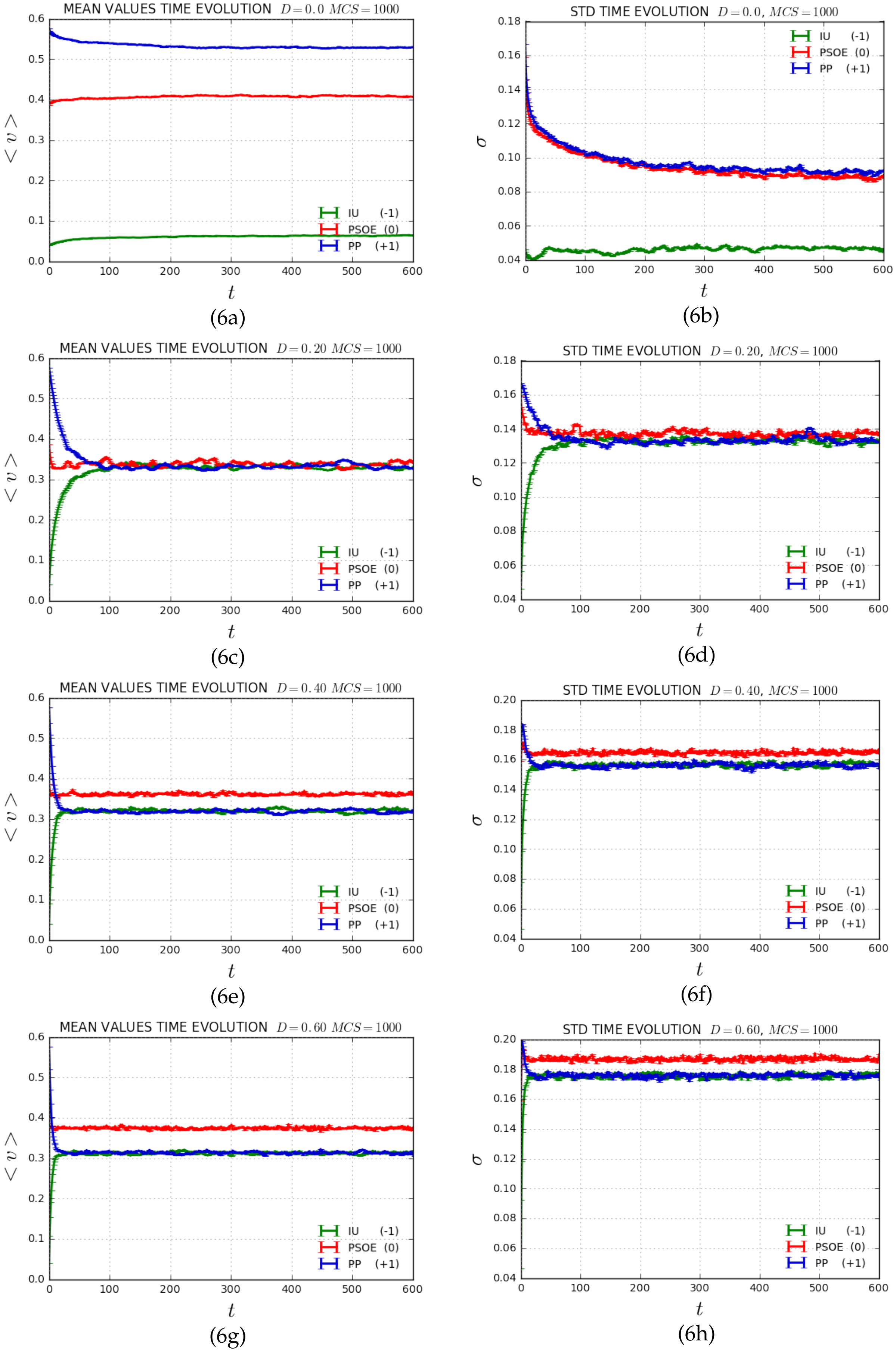
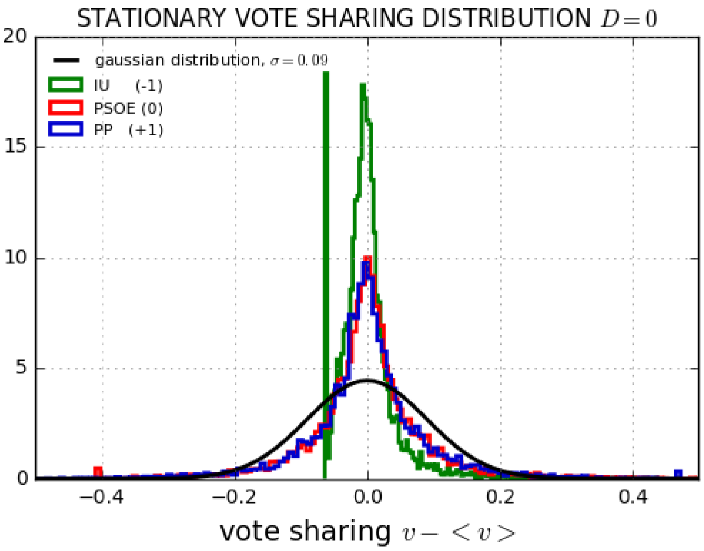
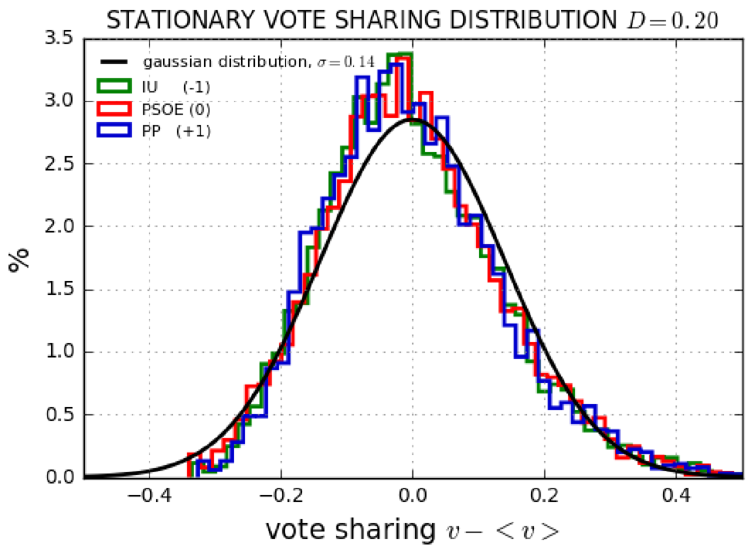


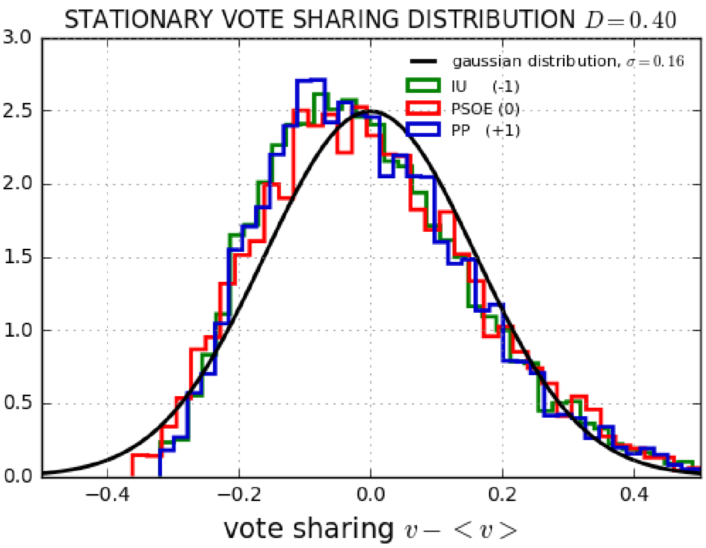
FIGURE 6:Time evolution plots of the average vote sharing (left column) and standard deviation (right column) of the different parties, for different values of the parameter D. It is possible to observe that, for each value of D the system converges to a stationary state . Bigger will be value of D, and faster will be the convergence to the stationary state. **fig (6a-b)** Time evolution with $D = 0$ **fig (6c-d)**Time evolution with $D = 0.20$ **fig (6e-f)** Time evolution with $D = 0.40$, **fig (6g-h)** Time evolution with $D = 0.60$



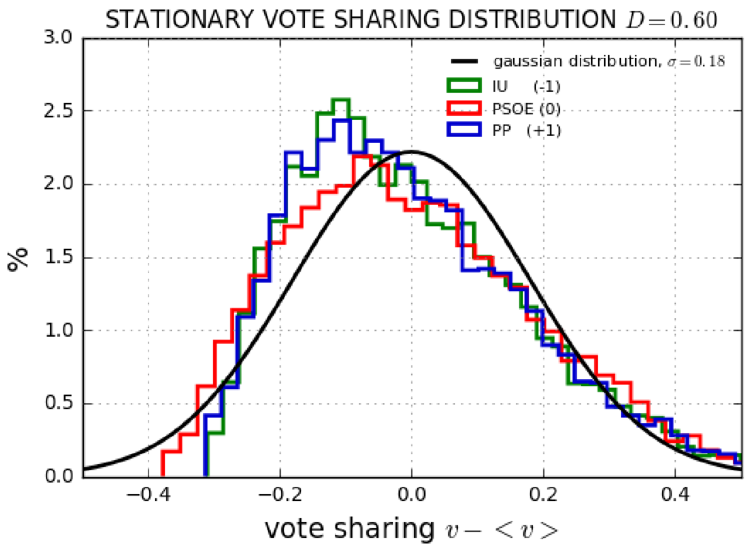
(A) (7a)



(B) (7b)

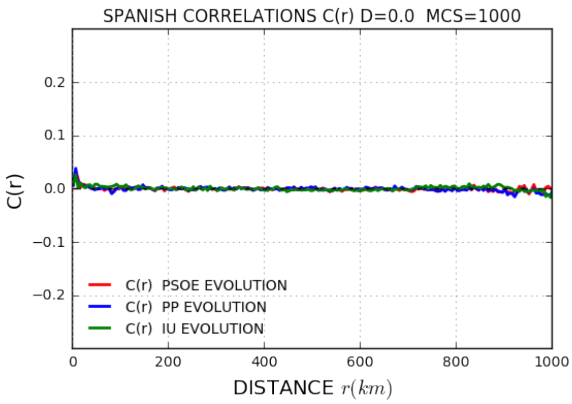


(C) (7c)

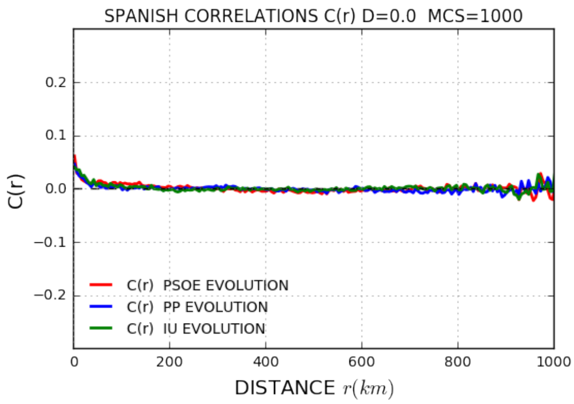


(D) (7d)

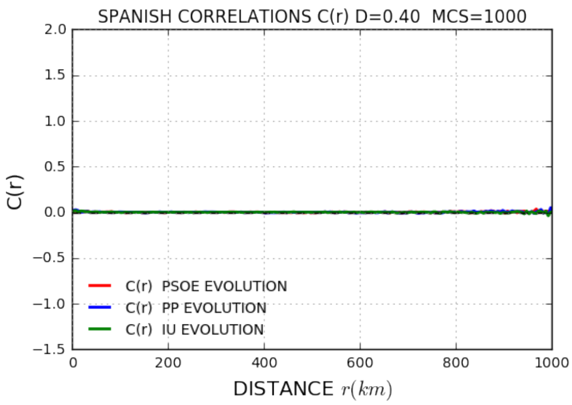
FIGURE 7:Plots of the stationary vote sharing distributions for different values D. The curves assume different behaviors respect the referential gaussian in function of the value of the parameter. **fig (7a)**For $D = 0$ (and in general small values) the real distributions are more sharped respect the referential gaussian. **fig (7b-c-d)**For $D = 0.20$ and higher values, real distributions lose their simmetry: the higher D, higher the unbalance of the distributions.



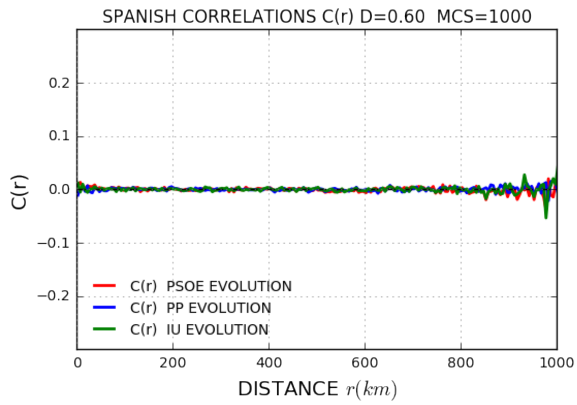
(8a)



(8b)

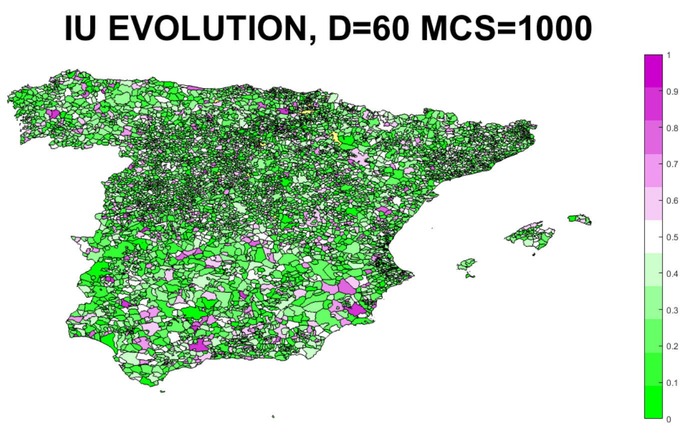


(8c)

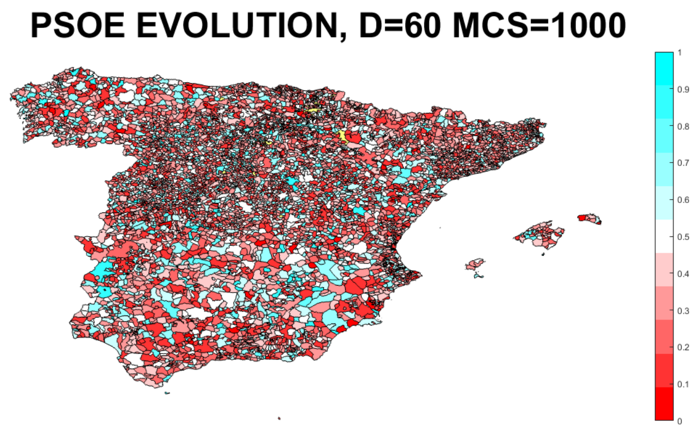


(8d)

FIGURE 8: Plots of the spatial correlation for different values D. Basically, correlations are destroyed for any value of D. This in accordance with the fact that it is not possible to identify any spatial domain in the spatial distribution of the vote sharings **fig (8a)** $D = 0$ **fig (8b)** $D = 0.3$ **fig (8c)** $D = 0.8$ **fig (8d)** $D = 1$



(9a)



(9b)

FIGURE 9:Plots of the IU and PSOE vote sharing maps at the stationary state, with $D = 0.60$. According with the correspondent vote sharing distributions (fig. 7d) the major part of the municipalities are characterized by low values of vote sharing, even if some peaks can be found all over the map. On the other hand, the PSOE map result to be more unbalanced in favor of higher value of v. Furthermore, coherently with the correspondent correlation functions, in both maps it is not possible to identify any cluster. **fig (9a)** IU vote sharing map **fig (9b)** PSOE vote sharing map

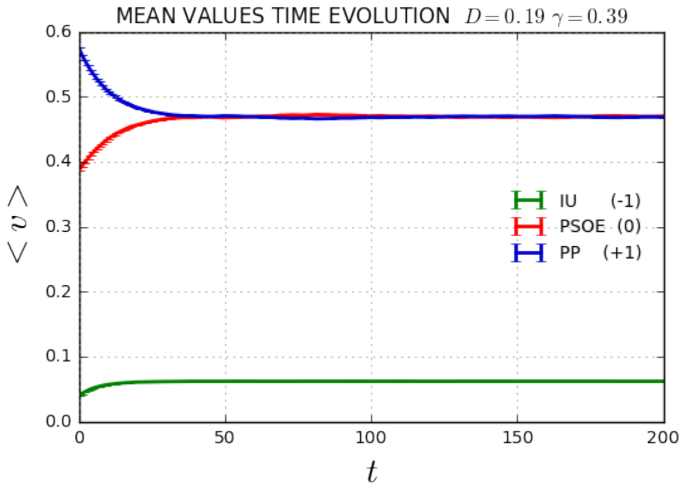
Chapter 5

Some modifications to the model; introduction and results

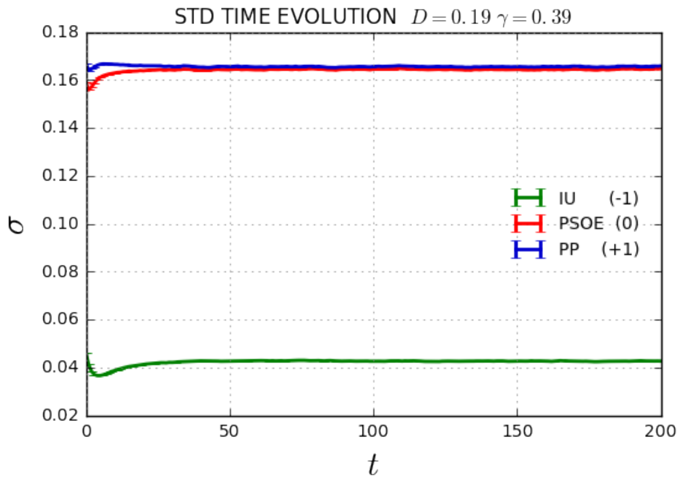
We stressed many times the fact that, according with real election data, it is possible to distinguish between the IU, which cover the role of minoritarian party, and the PP and PSOE options which result to be majoritarian. This concrete fact is not reproduced by the dynamics of the original model (chapter 4). To make the IU option stay minoritarian, what we do is to reduce the probability at which the agents in the state (0) assume opinion (-1) due to the interactions with the agents holding the minoritarian opinion. The new probability is obtained multiplying the old one by a factor $\gamma \in [0, 1]$

$$r_{ij}^{0-} \longrightarrow r_{ij}^{0-'} = \gamma \cdot r_{ij}^{0-} \tag{5.1}$$

Such choice must not appear simply as a forcing make by hand. The role of minoritarian options as been already studied by Gargiulo et al. in [24], and the reduction of the value of the transition probability is coherent with the collective behavior typical of the people holding extreme ideologies: they tend to cluster themselves, so that interaction with people holding different ideas result to be not so common, as well as persuading the latter to adopt the extreme position. Again, more than a forcing, the modification of the rate can be seen as an enrichment of the model, including in it also this aspect of the social opinion dynamic.
Differently from what we have seen so far, this modification allow the system to preserve its initial condition in correspondence of the values $D = 0.19$ and $\gamma = 0.39$. The two parameters affect, both, the stationary values of the average vote sharing and the standard deviations. In particular, γ strong influences the stationary value of the quantities related to the minoritarian options. The higher γ, the higher $\langle v_{st}^{(-)} \rangle$ and $\sigma_{st}^{(-)}$. On the other hand, the higher γ and the lower will be $\sigma_{st}^{(0,+)}$, the stationary standard deviation related to the majoritarian parties. The dependence on D of the same quantities is different: this time, all the quantities cited before ($\langle v_s t(-) \rangle$, $\sigma_{st}^{(-,0.+)}$) tend to increase with increasing D. In particular, in correspondence of the pair of values mentioned before, the system evolution conserves the initial values of the standard deviations, meanwhile $\langle v_{st}^{(-)} \rangle$ assume a value which is perfectly comparable with the original one.
The time evolution plots of the average vote sharings show that the average vote sharings related to the majoritarian parties tend to become equal, meanwhile the average vote sharing of the minoritarian option assume a stationary value which is fairly equal to the initial one, staying minoritarian even during the time evolution (fig. 10 a). Similarly, the standard deviations of all the parties weakly oscillate



(10 a)



(10 b)

FIGURE 10:Plots of the time evolution plots of the average vote sharings and standard deviations of the different parties, in correspondence fo $D = 0.19$ and $\gamma = 0.39$, averaged over 100 repetitions. **fig (10a)** Basically, the time evolution plot of the average vote sharing shows that the majoritarian parties tend to assume the same mean value, whereas the IU assume a stationary value which is fairly equal to the initial one. **fig (10b)** The evolution of the standard deviations conserve the initial values, and this is valid for all the three parties

around the correspondent intial values $\forall t$ (fig. 10b). The majoritarian vote sharing distributions continue to be gaussian-like, and the vote sharing distribution of the minoritarian option keep being characterized by a strong peak on the left, correspondent with the fact that in many municipalities $v^- = 0$ (fig. 11 a-b). Furthermore, differently from the original version of the model, correlations get not destroyed totally (fig. 12 a-b-c), and in the maps it is possible to recognize domains with characteristic size of $r \approx 200$ (fig. 13 a-b). Again, such result is different respect the one obtained in the study of the american case in [17], where at the stationary state even the domains conserve their original size, and correlation functions do not decay. On the other hand, this version of the model has all the characteristics we were expecting for, mantaining the initial state of the system stationary.

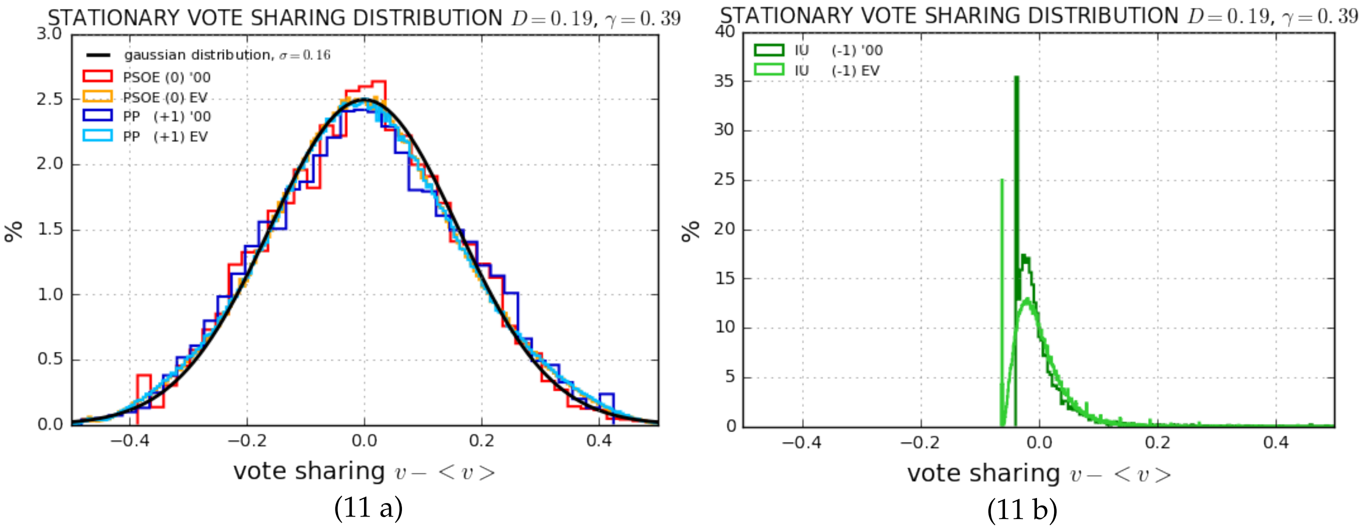
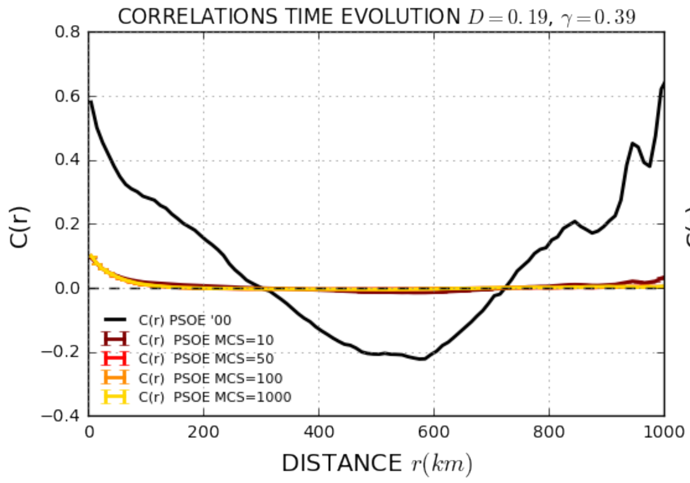
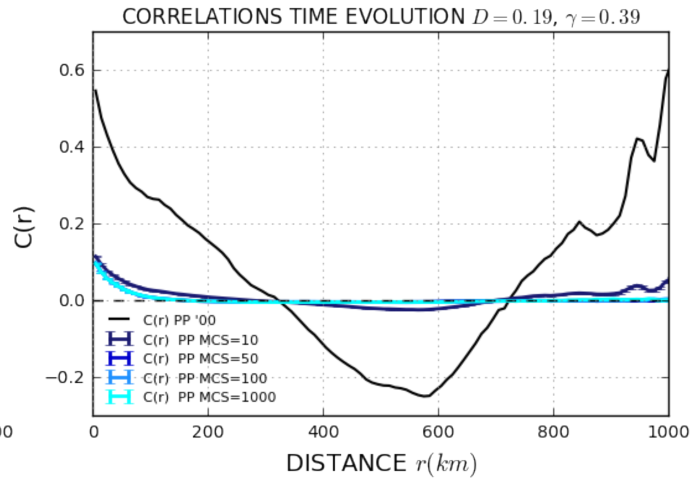


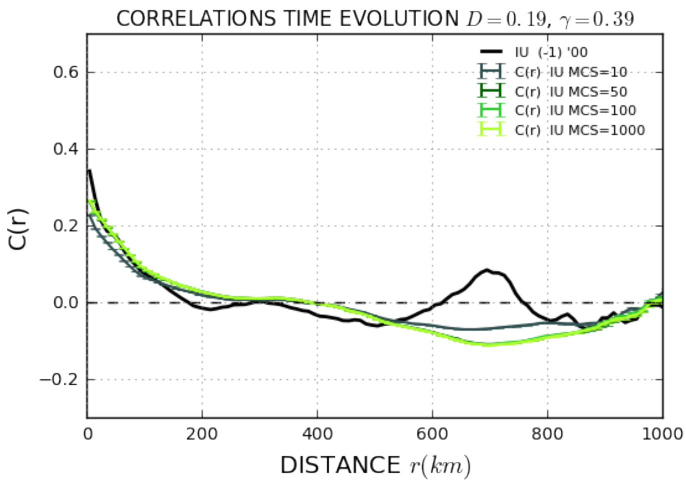
FIGURE 11:Plots of the stationary vote sharing distributions in correspondence of $D = 0.19$ and $\gamma = 0.39$, together with the correspondent initial vote sharing distributions (general Spanish elections ’00). The distributions are averaged over 100 repetitions. **fig (11a)** Here the vote sharing distributions of the majoritarian options are reported. A comparison with the initial distributions, and with a gaussian distribution with $\sigma = 0.16$,permit us to conclude that they tend to be gaussian like. **fig (11b)** Vote sharing distribution of the minoritarian option. Coherently with the correspondent initial distribution, the stationary one result to be not gaussian like, and it is characterized by a peak (lower than the original one) which indicates that, according with the simulation, the IU keep not being voted in many municipalities



(12a)

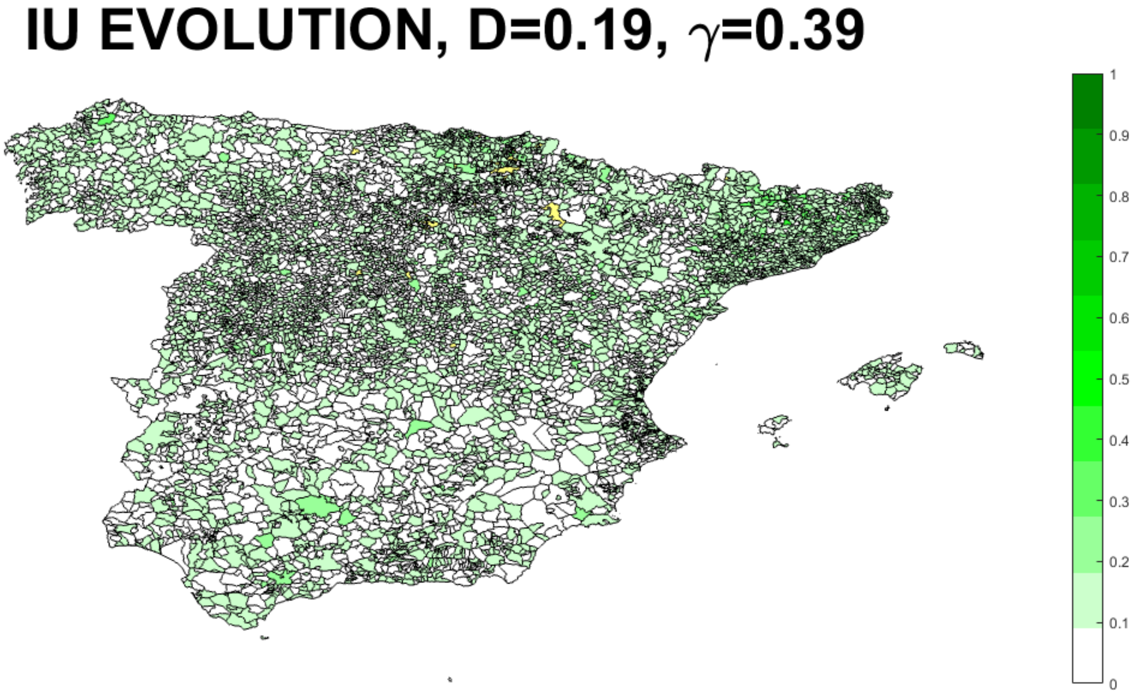


(12b)

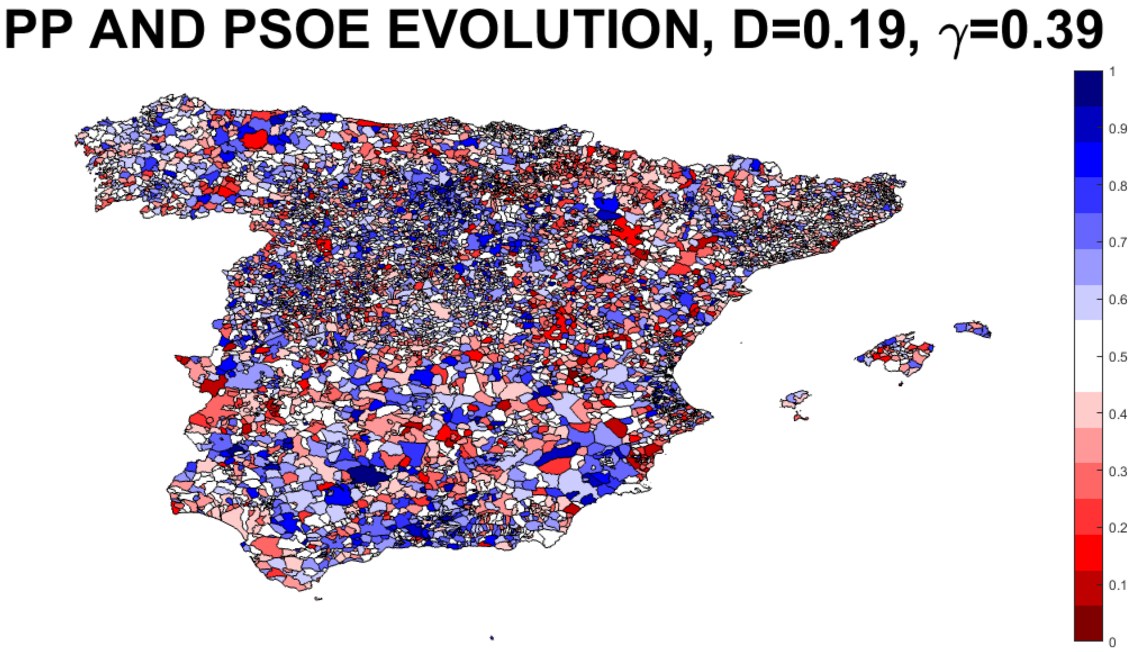


(12c)

FIGURE 12:Time evolution plot of the spatial correlation functions of the three electoral options, in correspondence of $D = 0.19$ and $\gamma = 0.39$. Results averaged over 100 repetitions. It is possible to observe that at the stationary states, even if lower respect the initial values, the correlations do not vanish completely. **fig (12a) fig (12b) fig (12c)**



(13 a)



(13 b)

FIGURE 13: Maps of the vote sharings of the three parties after an evolution of 100 MCS, with $D = 0.19$ and $\gamma = 0, 39$. **fig (13 a)**The IU keep remaining a minoritarian option, even if it result to be more spread all over tha Spain territory. **fig (13 b)**Again, Red and blue colours indicate respectively a PSOE or a PP majority. White colour indicates that the two parties obtained both a vote sharing around $v = 0.5$. The values in such colorbars are referred respect to the PSOE vote sharings $v^{(-)}$ normalized respect the sum of PSOE and PP vote sharings:

$$v^{(-)}_{maps} = v^{(-)}/(v^{(-)} + v^{(+)})$$

. In the picture, even if smaller, it is still possible to recognize the presence of clusters.

Chapter 6

Conclusions and future work

During this work we have proposed a noisy voter model based on social influence and recurrent mobility (SIRM). The state of the agents has been defined through a ternary options $a = \{-1, 0, 1\}$, and such states were ordered on a line (opinion space). The agents have been supposed to change opinion gradually, so that they could only assume a new opinion which (in the opinion space) was adjacent to the original one. Said with other words, no interactions has been admitted between agents holding opinion (-1) and agents holding opinion (+1).

The aim of our work has been to observe whether our model was capable to reproduce the time evolution of real general elections, and the Spanish political context has been chosen in order to do the confrontations.The political parties IU, PSOE, and PP have played the role of the three (dynamical) options of the model, having been identified respectively with the options (-1), (0), and (+1). Moreover all the other political entities has been inglobed inside a fourth (static) option (X), whose role was to slow down the interactions between agents holding one of the three dynamical opinions. No transitions have been admitted from the dynamical options and the static one. (and viceversa).

A key point of the model is the recurrent mobility of the agents: every agent has been supposed to make a recurrent mobility between the municipality in which it lived and the municipality in which it worked. At each time, all the agents present in the same cell (municipality) have been considered neighbours between them. The commuting network has been constructed coherently with the Spanish census data of 2001, and the results of Spanish general elections have been used as initial conditions of our system.

The statistical analysis of the Spanish general elections of the years 2000, 2004, and 2008 has showed that electoral results are more or less stationary from a statistical point of view, and this has been exactly the dynamic we were proposed to reproduce with our model. By the way, the basical version of the model did not captured the stationarity of the initial conditions for any value of the noise parameter D. In particular, we have observed that, differently from real election data, in which it is possible to distinguish between minoritarian and majoritarian parties (characterized by different statistical features), the stationary states admitted by our model did not permit such distinction.

So, we have introduced a second parameter γ (bounding parameter), whose role has been to mitigate the interactions between the agents supporting to the minoritarian option (-1) and agents supporting the opinion (0), which was majoritarian. We have observed that in correspondence of some particular values of the noisy and the bounding parameter ($D = 0.19$, and $\gamma = 0.39$) the new model reproduces a dynamical evolution according with which the statistical properties of the initial state are

conserved. The only quantities which change significantly are the correlations functions: they decay until they do not reach a stationary value different from zero, so that it is alway possible to recognize domains in the vote sharing (even if smaller than the original ones). This behavior is different respect the results obtained by a similar model applied to the american, bipartitic, context [17]. However,many trials done during the preparation of this master thesis showed that this discrepancy is not caused by the presence of a third party, so, going by exclusion, it could be that such differences are caused by different properties of the initial conditions or by different topologic properties of the commuting networks.
From the one hand the model proposed result to be a good mixing between many ideas proposed separately (in particular [14], [17], [19], [24]). On the other hand, the results obtained, and the concordances with real data can be interpreted as a step further toward the comprehension, the description and the formalization of the wide topic of the opinion dynamics. Many other steps can be done in this direction: it could be possible, for example, to further extend the model considering more electoral options, possibly of diffent natures (majoritarian and minoritarian, bounded in regional area or extended to the whole national territory), and observe if the topological properties of the opinion space, together with an appropriate arrangement of the parties inside of it, permit still to reproduce behaviors similar to the real ones.

References

[1]González, M. C., Hidalgo, C. A., & Barabási, A. L. (2008). Understanding individual human mobility patterns. Nature, 453(7196), 779-782.
[2]Balcan, D., Colizza, V., Gonç alves, B., Hu, H., Ramasco, J. J., & Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences, 106(51), 21484-21489.
[3]Gleeson, J. P., Cellai, D., Onnela, J. P., Porter, M. A., & Reed-Tsochas, F. (2014). A simple generative model of collective online behavior. Proceedings of the National Academy of Sciences, 111(29), 10411-10415.
[4] Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. ICWSM, 133, 89-96.
[5] Helbing, D. (2001). Traffic and related self-driven many-particle systems. Reviews of modern physics, 73(4), 1067.
[6]Bonabeau, E., Theraulaz, G., & Deneubourg, J. L. (1995). Phase diagram of a model of self-organizing hierarchies. Physica A: Statistical Mechanics and its Applications, 217(3), 373-392.
[7]Castelló, X., Eguíluz, V. M., & San Miguel, M. (2006). Ordering dynamics with two non-excluding options: bilingualism in language competition. New Journal of Physics, 8(12), 308.
[8]Baronchelli, A., Loreto, V., & Tria, F. (2012). Language dynamics.
[9]Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. Journal of conflict resolution, 41(2), 203-226.
[10]Clifford, P., & Sudbury, A. (1973). A model for spatial conflict. Biometrika, 60(3), 581-588.
[11]Holley, R. A., & Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. The annals of probability, 643-663.
[12]Scheucher, M., & Spohn, H. (1988). A soluble kinetic model for spinodal decomposition. Journal of statistical physics, 53(1), 279-294.
[13]Granovsky, B. L., & Madras, N. (1995). The noisy voter model. Stochastic Processes and their applications, 55(1), 23-43.
[14] Vazquez, F., Krapivsky, P. L., & Redner, S. (2003). Constrained opinion dynamics: Freezing and slow evolution. Journal of Physics A: Mathematical and General, 36(3), L61.
[15] Castellano, C., Vilone, D., & Vespignani, A. (2003). Incomplete ordering of the voter model on small-world networks. EPL (Europhysics Letters), 63(1), 153.
[16]Vilone, D., & Castellano, C. (2004). Solution of voter model dynamics on annealed small-world networks. Physical Review E, 69(1), 016109.
[17] Fernández-Gracia, J., Suchecki, K., Ramasco, J. J., San Miguel, M., & Eguíluz, V. M. (2014). Is the voter model a model for voters?. Physical review letters, 112(15), 158701.
[18] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of'small-world'networks. nature, 393(6684), 440

[19] Poole, K. T., & Rosenthal, H. L. (2011). Ideology and congress (Vol. 1). Transaction Publishers.
[20] http://www.ine.es/prensa/eet_prensa.htm.
[21] Klimek, P., Yegorov, Y., Hanel, R., & Thurner, S. (2012). Statistical detection of systematic election irregularities. Proceedings of the National Academy of Sciences, 109(41), 16469-16473.
[22] Borghesi, C., & Bouchaud, J. P. (2010). Spatial correlations in vote statistics: a diffusive field model for decision-making. The European Physical Journal B-Condensed Matter and Complex Systems, 75(3), 395-404.
[23] Borghesi, C., Raynal, J. C., & Bouchaud, J. P. (2012). Election turnout statistics in many countries: similarities, differences, and a diffusive field model for decision-making. PloS one, 7(5), e36289.
[24]Gargiulo, F., & Mazzoni, A. (2008). Can extremism guarantee pluralism?. arXiv preprint arXiv:0803.3879.

Appendix A

Construction of the commuting network

The dynamics of the model is based on the existence of the commuting network. The vote sharing data that we use as initial values of the model are divided per municipalities: no information about the places in which people work is included in such data. So, in order to construct the commuting network, we mix the vote sharing data with the Spanish census data of the year 2001. In those data we find specified the population of each municipality and the number of individuals W_{ij} residing in municipality i and working in municipality j . Let us consider for example the municipality of Golosalvo, (located in the province of Albacete) (code 02036), whose vote sharing and the worker network data are reported in the table below:

Residence mun. i	Total voters V_i	IU $V_i^{(-)}$	PSOE $V_i^{(0)}$	PP $V_i^{(+)}$	Other $V_i^{(\times)}$
02036	102	0	37	47	18

Municipality i	Working municipality j	N. of workers W_{ij}
02036	02003	1
02036	02024	2
02036	02034	2
02036	02036	22
02036	02045	1
02036	08019	2

Total number of workers W_i: 30 Total number of voters V_i: 102

As we can observe, the total number of workers is different respect to the total number of voters: this is a consequence of the fact that not all the voters are registered workers (freelancers, retirees, students) , and not all the workers have the right to vote (workers with foreign nationality).
To construct the commuting network, we start from a set of "empty commuting cells", in which we will dispose the voters steb-by-step. Each one of them, in accordance with the electoral data and even before to be displaced in the spatial structure, is characterized by one of the four possible opinions. For each municipality i, we use the following rules:

- If the total number of workers W_i of the municipality i is greater or equal to the total number of voters V_i, $W_i \geq V_i$, proceed directly to the next step. If it is not the case ($W_i < V_i$) , calculate the difference $\Delta = V_i - W_i$, and increase by hand the number of people living and working in the municipality i, W_{ii}, of such quantity[1], in order to reduce the situation to the case $W_i = V_i$.

- Repeat iteratively for $N = V_i$ steps the following:

a) Calculate the probability distribution:

$$P(n,j) = \frac{W'_{ij}(n)}{W'_i(n)} \qquad \text{with} \qquad W'_{ij}(n=1) = W_{ij} \qquad W'_i(n=1) = W_i \quad \text{(A.1)}$$

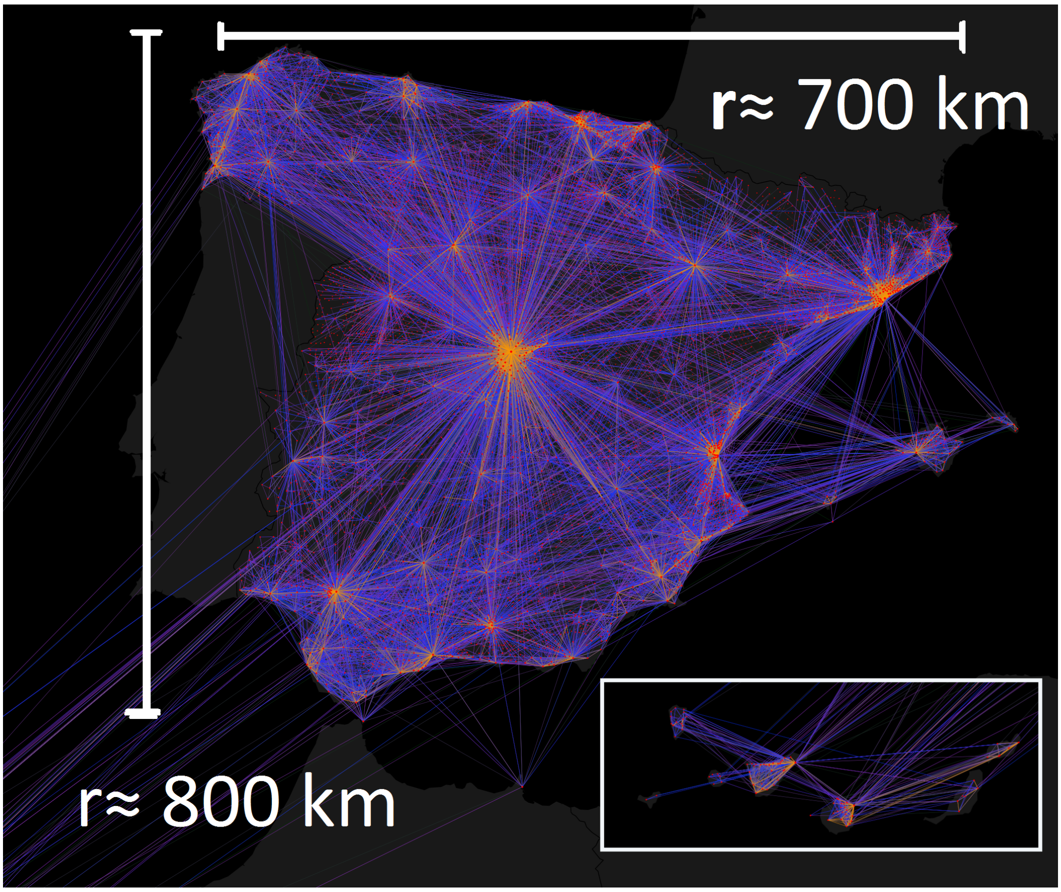
where n is the $n-$th step.

b) Select randomly one of the commuting cells ij,according with the probability distribution $P(n,j)$. Put one of the voters inside the selected commuting cell. at the next step we will have:

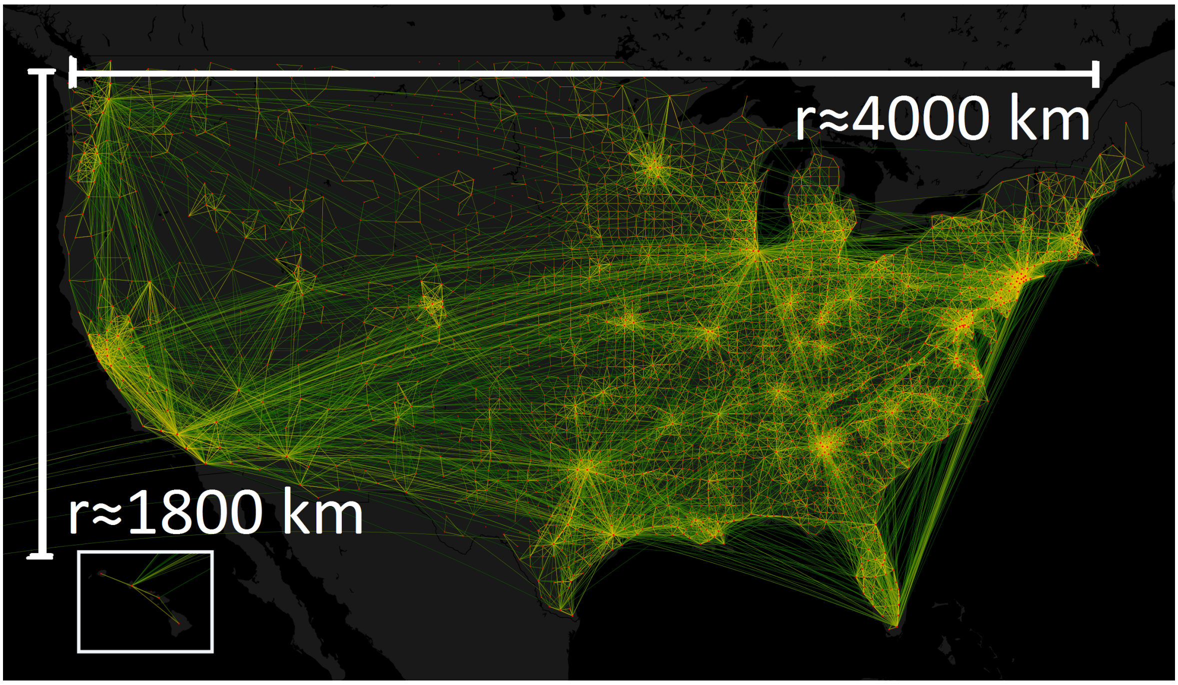
$$W'_{ij}(n+1) \, = \, W'_{ij}(n) \, - 1 \qquad W'_i(n+1) \, = \, W'_i(n) \, - 1 \qquad \text{(A.2)}$$

The resulting commuting network is reported In fig. 14a, Where only The 20% of commuting fluxes is shown. A confrontation with the American commuting network (fig 14b) say us that the two network seem to follow two different topologies: meanwhile the American network remind somehow a regular lattice (apart from some important regional areas which attract workers all over the cointry), in the Spanish one it is not possible to recognize any regular spatial pattern. The reason of such difference could be the different spatial scales which characterize the two states.

[1] The fact that the "missing workers" are added to the commuting cell ii is supported by the assumption that the majority of such "missing workers" is formed by retirees, and over-18 students that do not work still.



(14 a)



(14 b)

FIGURE 14: Representation of the Spanish and the American commuting maps. In each figures, only the 20% strongest links are shown. The two networks show a fundamental difference between them: The american one remind in some way a regular lattice, even though it is possible to recognize the presence of important regional parts which attract workers from all over the country. Otherwise, in the Spanish case it is not possible to recognize any kind of spatial regulat pattern. The reason of such difference in the commuting network could be the different spatial scales which characterize the two countries (stressed in the picture). **fig (14 a)** Spanish commuting map. **fig (14 b)** American commuting map.

Appendix B

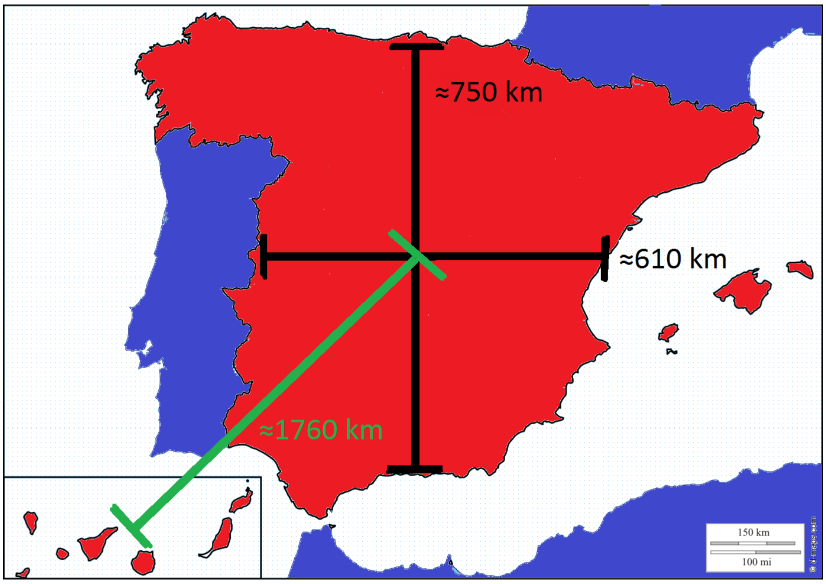
Preliminar statistical discussion

The calculation of spatial correlation is made through eq(12) which we report here again for the sake of convenience.

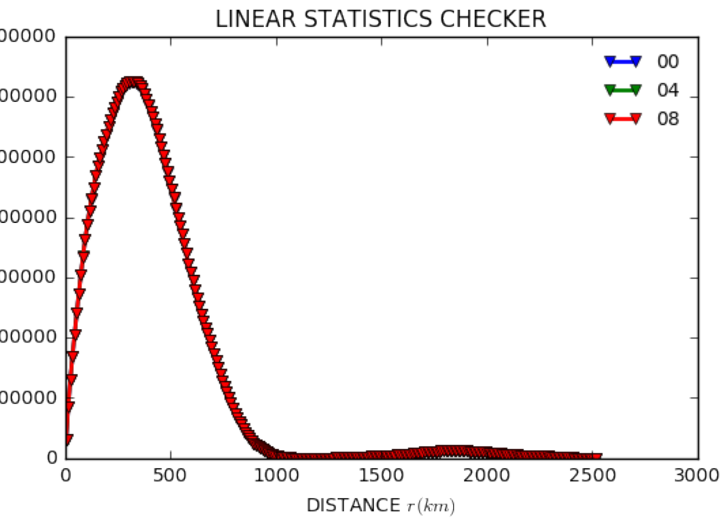
$$C^a(r) = \frac{\langle v_i^a v_j^a \rangle_{d(i,j)=r} - \langle v^a \rangle^2}{(\sigma^a)^2}$$

We have already stressed the fact that the first term on the numerator represents an average made considering all the pairs of municipalities (i, j) whose distance is exactly r. However, the distance r is a continuous variable, so that, in order to make some statistics calculations with equation (12), we divide the distance values in intervals, and we perform the average over the pairs belonging to the same distance interval. To do such division, we choose linear bins of 10 km of width, and base-10 exponential bins where the exponents change at steps of 0.1. We want to understand in which distance intervals we have enough pairs of municipalities to do reasonable statistics.

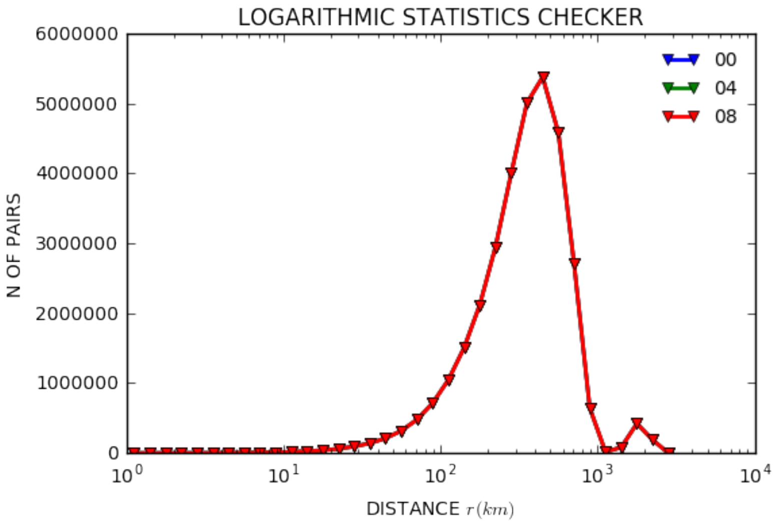
As we can see in Fig.(2b-c) the range in which some stastistics appear meaningful is the interval of 0-1000 $(0 - 10^3)$ km. A secundary peak is present in correspondence of $d \approx 1800(10^{3.25})km$, which approximately corresponds with distance between the canary islands and the city of Madrid, which in Fig (2a) is located more or less in correspondence of the center of the black cross. In fact, this peak is constituted by the pairs of municipalities such that one of the two is located in the canary islands, meanwhile the other one of the pair stays in the rest of Spain; according with we stated before, such pairs will be excluded by the statistical analysis.



(2a)



(2b)



(2c)

FIGURE 2:**fig 2a** An image of Spain, in which some characteristic distances are reported.**fig 2b (2c)** Number of pairs of municipalities in function of the distance, data divided in linear (Base-10-exponential) bins. We observe two peaks: the first one collects most of couples made selecting both municipalities in the Spanish peninsula, or Balearic islands. the second one in constituted by the pairs formed selecting one of the two municipalities in the canary island, the other one in the rest of Spain.