Evidence of structural balance in spatial ecological networks

Journal:	*Ecography*
Manuscript ID	ECOG-02561.R1
Wiley - Manuscript type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Saiz, Hugo; Instituto Pirenaico de Ecologia, Grupo de Conservación de la Biodiversidad, Departamento de Biodiversidad y Restauración; Universite de Rennes 1, ECOBIO Gomez-Gardeñes, Jesus; Universidad de Zaragoza Facultad de Ciencias, Department of Condensed Matter Physics; Universidad de Zaragoza, Institute for Biocomputation and Physics of Complex Systems (BIFI) Nuche, Paloma; Instituto Pirenaico de Ecologia, Grupo de Conservación de la Biodiversidad, Departamento de Biodiversidad y Restauración Giron, Andrea; Universidad de Zaragoza Facultad de Ciencias, Department of Condensed Matter Physics Pueyo, Y.; Instituto Pirenaico de Ecologia, Grupo de Conservación de la Biodiversidad, Departamento de Biodiversidad y Restauración Lopez Alados , Inmaculada; Instituto Pirenaico de Ecologia, Grupo de Conservación de la Biodiversidad, Departamento de Biodiversidad y Restauración
Keywords:	Spatial ecological networks, Signed networks, Structural balance
Abstract:	Despite recent advances in applying networks to study ecological systems, most of the network datasets are built attending only to a single type of interaction between nodes, which can be an oversimplification. In the present work, we built ecological networks that had positive and negative links for multiple plant communities based on the local spatial association between species. Then, we evaluated whether those networks were in balance, a hypothesis commonly formulated for real signed graphs but never tested in systems other than social networks. Specifically, we quantified the global and the local structural balance in the networks. We found that plant community networks were more balanced than expected by chance, and that this pattern was due to a large number of balanced triads to the detriment of unbalanced ones. Furthermore, this pattern was consistent among all of the types of the plant communities examined, which suggests that configurations that promote structural balance might be common in ecological signed networks. We also found that almost all networks had some unbalanced components, which might be responsible for the adaptation of the system. Mechanisms behind these structure and possible applications for community ecology are discussed. Our results encourage testing structural balance in other ecological networks to confirm if it is a widespread architecture of natural systems.

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.
Frustrationcode.rar

Evidence of structural balance in spatial ecological networks

Hugo Saiz*[1,2], Jesús Gómez-Gardeñes[3,4], Paloma Nuche[1], Andrea Girón[3], Yolanda Pueyo[1], Concepción L. Alados[1]

[1] Grupo de Conservación de la Biodiversidad, Departamento de Biodiversidad y Restauración, Instituto Pirenaico de Ecología (CSIC), Av. Montañana, 1005, 50059, Zaragoza, Spain.

[2] UMR 6553 Ecobio, CNRS - University of Rennes 1, Av. du General Leclerc, 35042, Rennes Cedex, France.

[3] Department of Condensed Matter Physics, University of Zaragoza, E-50009 Zaragoza, Spain.

[4] Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, E-50018 Zaragoza, Spain.

*Corresponding author: Hugo Saiz. UMR 6553 Ecobio, CNRS - University of Rennes 1, Av. du General Leclerc, 35042, Rennes Cedex, France. Tel: +33.(0)2.23.23.63.99. email: saizhugo@gmail.com

Field Code Changed

Abstract

Despite recent advances in applying networks to study ecological systems, most of the network datasets are built attending only to a single type of interaction between nodes, which can be an oversimplification. In the present work, we built ecological networks that had positive and negative links for multiple plant communities based on the local spatial association between species. Then, we evaluated whether those networks were in balance, a hypothesis commonly formulated for real signed graphs but never tested in systems other than social networks. Specifically, we quantified the global and the local structural balance in the networks. We found that plant community networks were more balanced than expected by chance, and that this pattern was due to a large number of balanced triads to the detriment of unbalanced ones. Furthermore, this pattern was consistent among all of the types of the plant communities examined, which suggests that configurations that promote structural balance might be common in ecological signed networks. We also found that almost all networks had some unbalanced components, which might be responsible for the adaptation of the system. Mechanisms behind these structure and possible applications for community ecology are discussed. Our results encourage testing structural balance in other ecological networks to confirm if it is a widespread architecture of natural systems.

Keywords: spatial ecological networks, signed networks, structural balance.

Introduction

The use of network science to study ecological communities has become a rising trend in recent times (Heleno et al. 2014, Kissling and Schleuning 2015). Networks allow the analysis of the interaction patterns among the elements of complex systems (Albert and Barabási 2002, Newman 2003) and the role of the structural organization in the functioning of these systems (Boccaletti et al. 2006). Thus, it is not surprising that networks are nowadays a common tool to study the organization of biotic interactions in real ecosystems (Ings et al. 2009). However, despite of the advances achieved by applying networks to ecosystems, most of ~~the~~ ecological networks are usually built attending only to one particular type of interaction between species, which could frequently represent an oversimplification of the true functioning of real communities.

In ecological communities, living organisms can interact with others positively (*e.g.*, mutualistic and facilitative interactions) or negatively (*e.g.*, competitive and parasitic interactions); and commonly these types of interactions occur at the same time within a given community (Kéfi et al. 2012, 2015). In ecology, however, there have been few empirical examples that have considered multiple types of interactions within the same network, probably because of the high logistic effort involved in documenting all ~~of~~ the potential interactions present within a community (Melián et al. 2009, Pocock et al. 2012). Furthermore, the analysis of these networks presents some difficulties over the analysis of networks with only one type of interactions, being necessary the use of new methodologies or theoretical frameworks (like the use of multilayer networks, Mucha et al. 2010, Boccaletti et al. 2014, Kivelä et al. 2014). Networks which include positive and negative links are called signed networks and have been mostly considered theoretically (Harary and others 1953, Zaslavsky 1982, Traag and Bruggeman 2009) and used to study social networks (Leskovec et al. 2010, Szell and Thurner 2010, Szell et al. 2010, Facchetti et al. 2011), while being ignored in other contexts such as ecological systems.

Signed networks exhibit a property called structural balance, which is based on how nodes organize in subgroups within the network (Cartwright and Harary 1956). A signed network is said to be balanced if it can be partitioned into groups of nodes in such a way that (*i*) every pair of connected nodes within the same group share a positive link, and (*ii*) links between nodes within different groups have a negative sign (Doreian and Mrvar 2009). Structural balance is associated with the resilience of social networks because it prevents the appearance of conflicts that might disrupt the system (Cartwright and Harary 1956). However, real networks rarely organize in a perfectly balanced way (*i.e.*, some links do not fulfill the criterion for structural balance), and the deviation from perfect balance is called ‘frustration’ (Doreian and Mrvar 2009). Specifically, frustration appears when there are negative links between nodes within the same group, or when there are positive links between nodes from different groups (Traag and Bruggeman 2009). Previous studies have shown that social systems (both human and animals) are more balanced (*i.e.*, exhibit less frustration) than is expected by chance, and some have hypothesized that the evolution of the system encourages cooperation and promotes stability by removing these unbalanced links (Szell et al. 2010, Facchetti et al. 2011, Ilany et al. 2013, Estrada and Benzi 2014). However, the scarcity of real signed networks has hampered the evaluation of this hypothesis in other systems apart from social networks (but see Altafini et al. 2010 for a similar approach in metabolic networks).

Recent advances in the use of spatial networks provide a good milieu for evaluating the balance in real signed networks in ecology. Spatial networks use the data of spatial co-occurrences between species to address possible interactions between them (Araújo et al. 2011, Araújo and Rozenfeld 2014, Morales-Castilla et al. 2015). The spatial distribution of species arise from different mechanisms including biotic interactions, dispersal patterns and environmental heterogeneity (Pearson and Dawson 2003, Escudero et al. 2005). However, each mechanism operates differently at different spatial scales, with biotic interactions being recognized as one of the fundamental drivers of spatial pattern at local scales (Pearson and Dawson 2003, Araújo and Rozenfeld 2014). Thus, if the considered scale is fine enough, it is possible to approximate biotic interactions as one of the principal drivers of spatial pattern.

The relationship between biotic interactions and spatial pattern becomes more important for sessile organisms like plant species, which interact positively and negatively with other plants in their immediate neighboring. For example, plants can benefit from the microenvironments created by other species, which can present more favorable conditions and increase plants survivability (*i.e.*, facilitation through habitat amelioration, Bruno et al., 2003); and plants that are pollinated by animals can form clusters to attract more pollinators and improve plants reproductive success (*i.e.*, facilitation through “magnet species effect”, Johnson et al. 2003). On the other hand, plants can compete among them for space and resources (Tilman 1981); or impede the persistence of others plants by producing chemical compounds (*i.e.* allellopathy, Wardle et al. 2011). Most~~ of~~ plant interactions occur at a local scale and affect the spatial distribution of species (Wu et al. 1985), such ~~as ~~that it is possible to establish a relationship between the ways in which species associate in space and how they interact. Thus, when two species occur together in space more often than is expected by chance, a benefit for at least one of the species can be inferred from that positive association (+, 0/+ interaction); however, if two species co-occur in space less than expected, interference between them might be inferred (-/- interaction) (Tilman and Kareiva 1997, Tirado and Pugnaire 2005)~~(Tilman and Kareiva 1997, Tirado and I Pugnaire 2005)~~. Analysis of the spatial pattern among plant species has been proved to be a valuable tool to study ecological communities and has been successfully applied to study the ecological interactions within them (Cavieres et al. 2006, Raventós et al. 2010, Verdú and Valiente-Banuet 2011, Saiz and Alados 2012).

Field Code Changed

In the present study we built ecological signed networks based on the local spatial association between pairs of plant species, and evaluated the structural balance of these networks. For that, we built 31 networks based on data from several plant communities in Spain, including grasslands, shrublands and forests from Mediterranean and alpine environments. For each network we quantified properties associated with structural balance, such as network frustration and the proportion of each type of complete triads, which are fully connected subgraphs of three nodes within the network (Estrada 2011). While the study of network frustration informs about the balance of the whole system, the study of subgroups within networks (such as triads) provides information about the local organization of links within the system (Stouffer 2010, Szell et al. 2010)~~(Szell et al. 2010, Stouffer 2010)~~. We hypothesize that, if real signed networks promote~~s~~ the existence of balanced structures, ecological networks should present low frustration and the triads which promote balance should be preferentially selected over the triads which do not. Specifically, we predict that, independently of the type of plant community, plant spatial association networks would exhibit less frustration and more triads which promote balance than expected by chance. To our knowledge, this is the first time that one study investigates the presence of structural balance in an ecological network (but see Saiz et al. 2014 for an example of block-partitioning in an ecological network based on structural balance criterion).

Methods

Vegetation survey

We built plant spatial association networks based on vegetation data collected at four locations in Spain that differed in abiotic and biotic conditions (Table 1). To avoid possible confounding factors, all of the analyzed plant communities were in a good state of conservation and did not present significant levels of disturbance. At each location different sites were surveyed and one network was created for each site (ranging from 6 to 10 networks at each location). Vegetation was surveyed using the point-intercept method (Goodall 1952), and, to minimize ~~limitations to account~~overlooking rare species (van der Maarel and Franklin 2012), at each site we recorded a large number of points (3750 to 10000 points). Specifically, at each site linear transects were established perpendicular to the slope and, at points separated by 20-cm intervals in grasslands and shrublands, and at 40-cm intervals in forests, we documented the plants that were in contact with the transect line. For each species, the total number of occurrences (n_i) and their position along the transects were recorded without distinguishing between plants life stages.

Network construction

For each of the 31 sites, one signed network was built based on the data from all the transects. In those networks, plant species were the nodes (i, j) and the spatial association between each pair of species in the adjacency matrix **A** were the links (l_{ij}), which can have positive or negative values. Spatial association was calculated comparing the number of times that two species i and j appeared at the same point on the transects (*i.e.* number of real spatial co-occurrences; a_{ij}), and the expected number of co-occurrences based on the total occurrences of the two species ($e_{ij} = n_i/T \times n_j/T \times T$, *i.e.*, the probability of both species occurring at the same point multiplied by the total number of points surveyed in the site, T). If real spatial co-occurrences were significantly higher or lower than a 95% confidence interval of a Poisson distribution fitted with the expected co-occurrences, a positive or negative (l_{ij} = +1/-1) link between species was indicated, while if real co-occurrences were not different to the confidence interval a neutral association (l_{ij} = 0, absence of link) was indicated. As we only recorded the presence of species at each point, we could not evaluate the intra-specific association, and so, the diagonal of **A** was set to zero.

Although spatial association has been commonly employed as a proxy of real biotic interactions, its use is not exempt of criticism. For example, plants spatial patterns are influenced by factors other than biotic interactions, like plant dispersal strategies (Escudero et al. 2005), and environmental heterogeneity (Getzin et al. 2008), which could also lead ~~also~~ to non-~~-~~neutral association patterns. Nevertheless, the importance of each of these factors is dependent on the particular spatial scale considered in the study, with biotic interactions being dominant at local scales (Pearson and Dawson 2003, Araújo and Rozenfeld 2014, Morales-Castilla et al. 2015). Thus, we measured the spatial association in a similar scale to that in other works studying biotic interactions at community level (Cavieres et al. 2006, Raventós et al. 2010, Verdú and Valiente-Banuet 2011). Other possible limitation is that parasitic interactions result in a positive spatial association in ~~while~~ which one species ~~gets~~ benefits in detriment of the other (+/- interaction,~~)~~ ~~(~~Morales-Castilla et al. 2015). Therefore, differentiating between facilitation and parasitism in the network could not be possible. To avoid this possible source of error and considering that we only recorded one parasitic species in one of the sites, we decided to directly exclude that species from the network (but this could be more problematic in plant communities with higher presence of parasitic species). Taking into account these considerations, we believe that networks based on plant spatial associations at local scale are a valid method to approximate the structure of interactions in plant communities.

Computing structural balance

A partition of a network ~~in~~ is defined by a set of m ($1 \leq m \leq N$) groups (or modules) so that each ~~of the~~ node~~s~~ is assigned to one (and only one) group. The problem of evaluating the structural balance in a signed network consists in finding the partition that maximizes (minimizes) the number of positive (negative) links between the nodes belonging to the same group. This problem strongly resembles that of finding the community structure of a network, since in both problems the number of groups and their respective sizes are unknown. Below, we show how we searched for the partitions that were closest to structural balance.

First, we exploited the resemblance between the problem of finding structurally balanced partitions and that of community detection. Specifically, we used Modularity Optimization, the most commonly used framework in studies of community detection (Fortunato 2010). For unsigned networks, Modularity Optimization allows the clustering of nodes based on the link densities within and between communities. The partitions of a network that exhibit high modularity present dense connections between nodes within the same module and sparse connections between nodes in different modules (Newman 2006). Mathematically, being $\mathbf{A} = \{A_{ij}\}$ the Adjacency matrix of a network, ~~the~~ modularity is given by:

$$Q = \frac{1}{4L} \sum_i \sum_j \left(A_{ij} - \frac{k_i k_j}{2L} \right) \delta(C_i, C_j), \qquad (1)$$

where $k_i = \sum_j A_{ij}$ is the degree of node i and L is the total number of links in the network $2L = \sum_i k_i = \sum_i \sum_j A_{ij}$. In addition, C_i denotes the module that contains node i, and the function $\delta(x,y)$ is the Kronecker delta function defined as $\delta(x,y) = 1$ when $x = y$, and $\delta(x,y) = 0$ otherwise.

For signed networks, Gómez and collaborators (Gómez et al. 2009) proposed a generalization of Eq. (1) that takes into account both types of links:

$$Q = \frac{1}{2L^+ + 2L^-} \sum_i \sum_j \left[A_{ij} - \left(\frac{k_i^+ k_j^+}{2L^+} - \frac{k_i^- k_j^-}{2L^-} \right) \right] \delta(C_i, C_j), \qquad (2)$$

where $k_i^+ = \sum_j A_{ij}^+$ and $k_i^- = \sum_j A_{ij}^-$ are the number of positive and negative links on node i (where $\mathbf{A}^+$ and $\mathbf{A}^-$ are the Adjacency matrices that contain those positive and negative links, respectively), and L^+ and L^- are the total number of positive and negative links in the network.

To find the partition that maximize Eq. (2), we used the deterministic algorithm in Xue et al. (2014). However, because the modularity function focuses on maximizing the density of positive links within a module rather than minimizing the number of negative links between nodes within the same group, the partition obtained is not the one that yields the maximum structural balance. Therefore, we used the partition obtained by maximizing Eq. (2) as the initial condition for another optimization method which maximized the positive and minimized the negative connections within modules.

Given a partition of a signed network, such as the one obtained by maximizing Eq. (2), we quantified the extent to which this partition departed from perfect structural balance, which can be done by counting the number of links that contribute to frustration, *i.e.*, the number of positive links between groups and the number of negative links within groups.

Namely, consider a ~~network~~ partition of a network that is described by an Adjacency matrix **A**. The frustration of this partition is:

$$F = \sum_i \sum_j \left(\left[-a \cancel{A_{ij}} A_{ij}^{-} + \epsilon \left(1 - \left| A_{ij} \right| \cancel{A_{ij}} \right) \right] \delta\left(C_i, C_j \right) + b A_{ij}^{+} \left(1 - \delta\left(C_i, C_j \right) \right) \right), \qquad (3)$$

where a, b and ϵ are constants. The first term in the double sum aims at penalizing both the existence of negative links (weighted with a) and the absence of a positive links (weighted with ϵ) within a group. In addition, the second term (weighted with b) penalizes the existence of positive links between different groups

The goal is to find the partition that minimizes Eq. (3) and, to this aim, by implementing a Simul~~t~~ated Annealing technique (Kirkpatrick et al. 1983), we find the partition that minimizes function F starting from the partition that maximizes Q. The combination of those two optimization processes allowed us to reduce considerably the computational time of the partitions obtained compared to those of other methods (Traag and Bruggeman 2009). Those improvements occur because the partition obtained by maximizing Q is very close to the final one and the computational cost associated with it is much smaller than is the one associated with minimizing F.

We also calculated the proportion of each type of fully connected subgroup of three nodes within the network (*i.e.* triads, Estrada 2011). In undirected signed networks, there can be four types of complete triads that differ based on the sign of the three links: (+ + +), (+ + -), (+ - -) and (- - -). The proportions of each type of triad have been used to study the structural balance of signed graphs because they relate to the local balance within the system (Szell et al. 2010). Specifically, (+ + +) and (+ - -) are considered balanced triads because they allow the partitioning of nodes fulfilling the structural balance criterion (*sensu*, nodes can be divided in groups such that links within a group are positive while links between groups are negative); while (+ + -) and (- - -) are considered unbalanced because they do not fulfill that criterion (although it has been proposed that (- - -) triad can be considered balanced based on a 'relaxed' structural balance criterion Kulakowski 2006). In our case, triads represent the possible spatial association between groups of three plant species (Figure 1). The (+ + +) triad corresponds to a group of three species which appear associated in space, and the (+ - -) triad corresponds to a pair of species which associate between them while segregate to another different one. On the other hand, the (+ + -) triad stands for two species which segregate between them but both associate to another species, and the (- - -) represents three species which segregate from each other.

Network parameters and analysis

For each signed network, we calculated the size (S), connectance of positive ($C^+ = 2L^+/(S(S-1)$, diagonal of the matrix is always zero) and negative links ($C^- = 2L^-/(S(S-1))$), and ratio between positive and negative links ($R = (L^+ - L^-)/(L^+ + L^-)$). To evaluate structural balance, we also calculated system frustration as the proportion of unbalanced links in the network ($F_L = F/L$), and the proportion of each type of triad. As differences in connectivity patterns might lead to differences in the mechanisms behind networks organization (Newman

2002), we used a Pearson test to assess whether there were significant correlations between networks parameters and F_L and the proportion of each type of triads (Sokal and Rohlf 1995). To assess whether real networks were in balance, for each real network, we simulated 10000 networks that had the same degree distribution as did the original graph. Furthermore, those null networks presented the same positive and negative degree sequences so that the number of positive and negative links, k_i^+ and k_i^-, that each node *i* shared with the rest of the system remained the same as it was in the real network. To that end, we generalized the method for randomizing unsigned networks introduced by ~~(~~Maslov and Sneppen (2002) so that it could be applied to signed networks. The generalized method starts with the original (real) graph and proceeds as an iterative process. At each step, two pairs (*i*,*j*) and (*l*,*m*) of connected nodes are chosen randomly. In our case, those two links had to have the same sign. Then, partners are switched and, after deleting the previous two links, we set (*i*,*m*) and (*l*,*j*) as the new edges (with the same sign as the previous ones) provided neither of the new links are already present in the network. By repeating this step a number of 4*L* times we generate a null network in which any kind of structural correlation in the original graph is removed.

We tested the balance in real networks in two ways. First, we compared the real *F* and proportions of each type of triad against a simulated 95% confidence interval. Values were significantly higher than expected if real values were higher than 95% of the simulated values, and significantly lower than expected if values were lower than 95% of the simulated values, which allowed us to assess how common were balanced plant communities. Second, we compared the same real values to the mean of simulated values using a linear-model ANOVA for repeated measures that included location, type of network (real vs. simulated) and the interaction of the two as explanatory variables. Thus, it was possible to test whether real plant communities were more balanced than expected, and whether this balance differed significantly among the types of plant communities.

Results

Plant-plant spatial association networks

All of the parameters in the plant spatial association networks exhibited a wide range of values (Table 2). In general, positive links were more common than were negative links, but there were networks in which the reverse was true (8 of 31 networks, 26%, presented $R < 0$). Furthermore, there was no significant correlation between any of the network parameters and F_L, which suggests that unbalanced links arise from mechanisms unassociated with network connectivity. On the other hand, the correlations between the proportions of triads of each type and R were significant. This result was expected as (Leskovec et al. 2010) showed that, in a complete network, the probability of finding a triad of a given type is correlated to the proportion of positive and negative links in the system.

Structural balance in plant-plant networks

Plant spatial association networks were significantly more balanced than expected when compared against random networks (*i.e.* exhibited significantly less frustration, F-value = 38.41, $p<0.001$, Figure 2). Linear-model ANOVA did not reveal any significant effect of location or of the interaction between location and type of network, which suggests that in general, the ~~interactions~~ associations in plant communities were organized in a way that promoted structural balance. Specifically, real networks had 4% more balanced links than expected (F_L^{real} - $F_{L\ \cancel{sim}}^{random}$ = -0.043) and 21 of 31 networks (68%) were significantly more balanced than expected. Furthermore, none of the real networks exhibited more frustration than expected (Figure 2).

Formatted: Superscript

Formatted: Superscript

The high balance in real networks was ~~because of~~due to the high abundance of balanced triads (+ + +, + - -) to the detriment of unbalanced triads (+ + -, - - -). Specifically, (+ + +) and (+ - -

) triads were significantly more common (+ + +, *F*-value = 61.69, $p < 0.001$; + - -, *F*-value = 59.76, $p < 0.001$), and (+ + -) and (- - -) triads were significantly less common than expected (+ + -, *F*-value = 30, $p < 0.001$; - - -, *F*-value = 79.64, $p < 0.001$, Figure 3). Location had a significant effect on the proportion of (+ + -) and (+ - -) triads (+ + -, *F*-value = 8.46, $p < 0.001$; + - -, *F*-value = 9.19, $p < 0.001$, Figure 3); however, as the proportion of each type of triad is correlated to the proportion of positive and negative links in a network, this result might have been influenced by the differences in the *R* values at each location (Table 2, networks in Monegros exhibited high *R* value, together with high proportion of positive links, (+ + +) and (+ + -) triads, and low proportion of (+ - -) and (- - -) triads, Figure 4). Furthermore, 26 of 31 real networks exhibited at least one type of triad that differed significantly from the expected (65% of the networks presented more (+ + +) triads than expected; 58% presented less (+ + -) triads; 65% presented more (+ - -) triads; and 74% presented ~~more~~ less (- - -) triads), and no networks had triads that exhibited the reverse pattern (*i.e.*, balanced triads were always more common than or as common as expected, while unbalanced triads were always less common than or as common as expected).

Discussion

Studies of structural balance in real signed networks have shown that signed interactions are not organized randomly, but networks can be divided into blocks such that links among nodes within a block are positives while links between groups are negatives (Cartwright and Harary 1956). We found that one ecological system, the network formed by the biotic interactions among plant species derived from their spatial association, also presented this organization in different types of plant communities. This result suggests that organizations which promote structural balance might be the norm in plant communities and an interesting structure to evaluate in real networks with signed links.

Structural balance has been mostly tested in social networks. Social theory posits that links involved in conflict relationships decrease as the frequency of interactions among social groups increase (Pettigrew and Tropp 2006), and hypothesizes that balance in social networks increases over time (Doreian and Krackhardt 2001, Marvel et al. 2011). Our study expands this theory to another type of real system, and presents balanced structures as a common configuration for real signed networks. Mature ecological networks have been proposed as examples of real systems that are robust and efficient (Arditi et al. 2005, Estrada 2006). Ecologists suggest that ecological communities present common patterns (Sugihara et al. 1989), converge to common functional structures (Fukami et al. 2005), and increase their stability throughout succession (Neutel et al. 2007); resulting in common network architectures that increase the robustness of natural communities. Those architectures can change depending on the type of interaction considered, with nested patterns associated to mutualistic systems, and the presence of modules associated to herbivory networks (Thébault and Fontaine 2010). Thus, an architecture that promotes structural balance is possible to be a robust organization for networks with positive and negative links.

An important question in the study of systems that have positive and negative links is what the mechanisms that underlie the emergence of structural balance are. In sociology, structural balance is the result of general processes in which actors change their social arrangements to reduce unbalance. Usually, those processes operate at two levels: node level, including specific decisions by individual actors; and system level, presenting group behaviors that constrain individual decisions to minimize tensions within social groups (Hummon and Doreian 2003). In ecological systems these two levels also exist. At species level processes such as niche partitioning (*i.e.*, species differentiate to occupy different niches and minimize competition, Wisheu 1998) can play a major role in determining the interactions in the system (*e.g.*, by promoting positive interactions between the species that exploit different niches, Verdú and Valiente-Banuet 2011). On the other hand, at the system level, evolution constrains the possible interaction patterns in the network by selecting the species and interactions that are more successful each time (*e.g.*, by selecting the traits responsible of the "forbidden links" in the community, Bascompte and Jordano 2007). In fact, species traits and phylogeny have a major role in ecological networks organization (Vázquez et al. 2009, Verdú and Valiente-Banuet 2011, Maglianesi et al. 2014, Borrelli et al. 2015), which suggests that species characteristics and diversification processes shape interactions in ecological communities. Specifically in our case, plant spatial association networks represent the different types of vegetation patches which can be found in the ecosystem. Results support the idea that multi-specific patches are common in plant communities (Tirado and Pugnaire 2005, Cavieres et al. 2006). Studies on facilitation have revealed that positive interactions between plants are not specific, but many species benefit from the direct facilitation of the so-called 'nurse' species (Maestre and Cortina 2005, Verdú et al. 2008). Thus, this results in a positive association between several species coexisting under the same nurse. Furthermore, this species will also segregate from the species segregating from the nurses, which explain the high proportion of (+ - -) triads. On the other hand, communities showed a low number of unbalanced triads. These triads represent groups of three species competing for the same space, either competing among all of them (- - -), or competing for the space provided by a 'nurse' (+ + -). Competitive models predict that best competitors will outcompete other species, relegating them to the gaps where it is not present and eventually excluding them from the community (Tilman and Kareiva 1997, Chesson 2000). Therefore, in our plant communities the low proportion of unbalance triads could be the result of this exclusion by better competitors (although this exclusion is not total, as almost all networks presented some unbalanced links). The inclusion of intra-specific associations and differences in the ontogenetic stages of plant species could provide valuable information to link the results of local spatial networks to theoretical models on spatial coexistence.

Our study unveils new possibilities in the study of ecological systems. First, it is necessary to study explicitly the link between structural balance and the robustness of the network. Several studies on the dynamic of ecological networks have evaluated the role that positive and negative links have on the stability of the system, but always considering links at the whole system level (May 1972, Allesina and Tang 2012, Mougi and Kondoh 2012). It is possible that the particular organization of the links (*e.g.* minimizing the frustration) could play an important role in the robustness of ecological systems. On the other hand, studies in biotic interactions have suggested that including weighted and asymmetric interactions (the common case in real ecosystems) can have a significant effect in the structure of natural communities (Vázquez et al. 2012, Lin et al. 2012). Therefore, including strength and directions would provide a more realistic information about the actual role of balanced organizations. Furthermore, our results might encourage other researchers to evaluate structural balance in other biological systems (*e.g.*, ecological networks that have multiple types of interactions, Kéfi et al. 2015; biochemical and neural networks that have activation-inhibition, Guelzim et al. 2002, Takahashi et al. 2010) to asses if it is a general architecture in signed networks. If so, it would be interesting to identify the functional traits related to balanced groups at species level, which would help in identifying the traits which play a major role maintaining the structure of the ecological networks (*e.g.* the importance of resistant plants in stressful environments, Saiz et al. 2014). Moreover, the analysis of species involved in unbalanced links could provide valuable information about the adaptability of the system, as it has been proposed that balanced structures are less likely to change (Doreian and Krackhardt 2001), and is through the unbalanced links that the system is able to adapt (Ilany et al. 2013). On the other hand, the study of structural balance throughout ecological succession or perturbation gradients could identify the role of ecological mechanisms in the structural assemblage of real systems. If structural balance is a general architecture in ecological signed networks, then a set of fundamental rules could be defined to explain the structure and dynamics of real ecosystems and therefore, further the development of a general theory for community ecology.

Acknowledgements

HS is partially supported by the EU FP7-ICT-2011-9 No. 600654 DIVERSIFY Project; JGG is supported by the Spanish MINECO through the Ramón y Cajal program and Projects FIS2011-25167 and FIS2012-38266-C02-01, the European Commission through FET IP projects MULTIPLEX (Grant No. 317532) and PLEXMATH (Grant No. 317614), the Comunidad de Aragón (FENOL group); YP is supported by Project CGL2012-37508; and CLA is supported by CGL2011-27259. We are very grateful of the dedicated work of many technicians, students, volunteers and researchers which helped in the vegetation data surveys, ~~and~~ Cendrine Mony, Carlos Melian and Bruce MacWhirter for their suggestions in early drafts of the manuscript, and three anonymous reviewers and the subject editor for their useful suggestions.

Bibliography

Albert, R. and Barabási, A.-L. 2002. Statistical mechanics of complex networks. - Rev. Mod. Phys. 74: 47–97.

Allesina, S. and Tang, S. 2012. Stability criteria for complex ecosystems. - Nature 483: 205–208.

Altafini, C. et al. 2010. Determining the distance to monotonicity of a biological network: a graph-theoretical approach. - IET Syst. Biol. 4: 223–235.

Araújo, M. B. and Rozenfeld, A. 2014. The geographic scaling of biotic interactions. - Ecography 37: 406–415.

Formatted: Bibliography, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Araújo, M. B. et al. 2011. Using species co-occurrence networks to assess the impacts of climate change. - Ecography 34: 897–908.

Arditi, R. et al. 2005. Rheagogies: Modelling non-trophic effects in food webs. - Ecol. Complex. 2: 249–258.

Bascompte, J. and Jordano, P. 2007. Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. - Annu. Rev. Ecol. Evol. Syst. 38: 567–593.

Boccaletti, S. et al. 2006. Complex networks: Structure and dynamics. - Phys. Rep. 424: 175–308.

Borrelli, J. J. et al. 2015. Selection on stability across ecological scales. - Trends Ecol. Evol. 30: 417–425.

Bruno, J. F. et al. 2003. Inclusion of facilitation into ecological theory. - Trends Ecol. Evol. 18: 119–125.

Cartwright, D. and Harary, F. 1956. Structural balance: a generalization of Heider's theory. - Psychol. Rev. 63: 277.

Cavieres, L. A. et al. 2006. Positive interactions between alpine plant species and the nurse cushion plant Laretia acaulis do not increase with elevation in the Andes of central Chile. - New Phytol. 169: 59–69.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. - Annu. Rev. Ecol. Syst.: 343–366.

Doreian, P. and Krackhardt, D. 2001. Pre-transitive balance mechanisms for signed networks. - J. Math. Sociol. 25: 43–67.

Doreian, P. and Mrvar, A. 2009. Partitioning signed social networks. - Soc. Netw. 31: 1–11.

Escudero, A. et al. 2005. Spatial pattern and neighbour effects on Helianthemum squamatum seedlings in a Mediterranean gypsum community. - J. Veg. Sci. 16: 383–390.

Estrada, E. 2006. Network robustness to targeted attacks. The interplay of expansibility and degree distribution. - Eur. Phys. J. B - Condens. Matter Complex Syst. 52: 563–574.

Estrada, E. 2011. The structure of complex networks: theory and applications. - Oxford University Press.

Estrada, E. and Benzi, M. 2014. Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. - Phys. Rev. E 90: 42802.

Facchetti, G. et al. 2011. Computing global structural balance in large-scale signed social networks. - Proc. Natl. Acad. Sci. 108: 20953–20958.

Fortunato, S. 2010. Community detection in graphs. - Phys. Rep. 486: 75–174.

Fukami, T. et al. 2005. Species divergence and trait convergence in experimental plant community assembly. - Ecol. Lett. 8: 1283–1290.

Getzin, S. et al. 2008. Heterogeneity influences spatial patterns and demographics in forest stands. - J. Ecol. 96: 807–820.

Gómez, S. et al. 2009. Analysis of community structure in networks of correlated data. - Phys. Rev. E 80: 16114.

Goodall, D. W. 1952. Some considerations in the use of point quadrats for the analysis of vegetation. - Aust. J. Biol. Sci. 5: 1–41.

Guelzim, N. et al. 2002. Topological and causal structure of the yeast transcriptional regulatory network. - Nat. Genet. 31: 60–63.

Harary, F. and others 1953. On the notion of balance of a signed graph. - Mich. Math. J. 2: 143–146.

Heleno, R. et al. 2014. Ecological networks: delving into the architecture of biodiversity. - Biol. Lett. 10: 20131000.

Hummon, N. P. and Doreian, P. 2003. Some dynamics of social balance processes: bringing Heider back into balance theory. - Soc. Netw. 25: 17–49.

Ilany, A. et al. 2013. Structural balance in the social networks of a wild mammal. - Anim. Behav. 85: 1397–1405.

Ings, T. C. et al. 2009. Review: Ecological networks–beyond food webs. - J. Anim. Ecol. 78: 253–269.

Johnson, S. D. et al. 2003. Pollination success in a deceptive orchid is enhanced by co-occurring rewarding magnet plants. - Ecology 84: 2919–2927.

Kéfi, S. et al. 2012. More than a meal… integrating non-feeding interactions into food webs. - Ecol. Lett. 15: 291–300.

Kéfi, S. et al. 2015. Network structure beyond food webs: mapping non-trophic and trophic interactions on Chilean rocky shores. - Ecology 96: 291–303.

Kirkpatrick, S. et al. 1983. Optimization by Simulated Annealing. - Science 220: 671–680.

Kissling, W. D. and Schleuning, M. 2015. Multispecies interactions across trophic levels at macroscales: retrospective and future directions. - Ecography 38: 346–357.

Kulakowski, K. 2006. Some recent attempts to simulate the Heider balance problem. - ArXiv Prepr. Physics0612197 in press.

Leskovec, J. et al. 2010. Predicting Positive and Negative Links in Online Social Networks. - Proc. 19th Int. Conf. World Wide Web: 641–650.

Lin, Y. et al. 2012. Differences between symmetric and asymmetric facilitation matter: exploring the interplay between modes of positive and negative plant interactions. - J. Ecol. 100: 1482–1491.

Maestre, F. T. and Cortina, J. 2005. Remnant shrubs in Mediterranean semi-arid steppes: effects of shrub size, abiotic factors and species identity on understorey richness and occurrence. - Acta Oecologica 27: 161–169.

Maglianesi, M. A. et al. 2014. Morphological traits determine specialization and resource use in plant–hummingbird networks in the neotropics. - Ecology 95: 3325–3334.

Marvel, S. A. et al. 2011. Continuous-time model of structural balance. - Proc. Natl. Acad. Sci. 108: 1771–1776.

Maslov, S. and Sneppen, K. 2002. Specificity and stability in topology of protein networks. - Science 296: 910–913.

May, R. M. 1972. Will a Large Complex System be Stable? - Nature 238: 413–414.

Melián, C. J. et al. 2009. Diversity in a complex ecological network with two interaction types. - Oikos 118: 122–130.

Morales-Castilla, I. et al. 2015. Inferring biotic interactions from proxies. - Trends Ecol. Evol. 30: 347–356.

Mougi, A. and Kondoh, M. 2012. Diversity of Interaction Types and Ecological Community Stability. - Science 337: 349–351.

Neutel, A.-M. et al. 2007. Reconciling complexity with stability in naturally assembling food webs. - Nature 449: 599–602.

Newman, M. E. J. 2002. Assortative Mixing in Networks. - Phys. Rev. Lett. 89: 208701.

Newman, M. E. 2003. The structure and function of complex networks. - SIAM Rev. 45: 167–256.

Newman, M. E. J. 2006. Modularity and community structure in networks. - Proc. Natl. Acad. Sci. 103: 8577–8582.

Ninyerola, M. et al. 2005. Atlas Climático Digital de la Península Ibérica. - Universidad Autónoma de Barcelona.

Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? - Glob. Ecol. Biogeogr. 12: 361–371.

Pettigrew, T. F. and Tropp, L. R. 2006. A meta-analytic test of intergroup contact theory. - J. Pers. Soc. Psychol. 90: 751–783.

Pocock, M. J. et al. 2012. The robustness and restoration of a network of ecological networks. - Science 335: 973–977.

Raventós, J. et al. 2010. Evidence for the spatial segregation hypothesis: a test with nine-year survivorship data in a Mediterranean shrubland. - Ecology 91: 2110–2120.

Saiz, H. and Alados, C. L. 2012. Changes in Semi-Arid Plant Species Associations along a Livestock Grazing Gradient. - PLoS ONE 7: e40551.

Saiz, H. et al. 2014. Plant–plant spatial association networks in gypsophilous communities: the influence of aridity and grazing and the role of gypsophytes in its structure. - Web Ecol. 14: 39–49.

Sokal, R. R. and Rohlf, F. J. 1995. Biometry: the principles and practice ofstatistics in biological research. - State Univ. New Yorkat Stony Brook N. Y. NY in press.

Stouffer, D. B. 2010. Scaling from individuals to networks in food webs. - Funct. Ecol. 24: 44–51.

Sugihara, G. et al. 1989. Scale invariance in food web properties. - Science 245: 48–52.

Szell, M. and Thurner, S. 2010. Measuring social dynamics in a massive multiplayer online game. - Soc. Netw. 32: 313–329.

Szell, M. et al. 2010. Multirelational organization of large-scale social networks in an online world. - Proc. Natl. Acad. Sci. 107: 13636–13641.

Takahashi, N. et al. 2010. Circuit topology for synchronizing neurons in spontaneously active networks. - Proc. Natl. Acad. Sci. 107: 10244–10249.

Thébault, E. and Fontaine, C. 2010. Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. - Science 329: 853–856.

Tilman, D. 1981. Resource competition and community structure. - Monogr. Popul. Biol. 17: 1–296.

Tilman, D. and Kareiva, P. M. 1997. Spatial ecology: the role of space in population dynamics and interspecific interactions. - Princeton University Press.

Tirado, R. and Pugnaire, F. 2005. Community structure and positive interactions in constraining environments. - Oikos 111: 437–444.

Traag, V. A. and Bruggeman, J. 2009. Community detection in networks with positive and negative links. - Phys. Rev. E 80: 36115.

van der Maarel, E. and Franklin, J. 2012. Vegetation ecology. - John Wiley & Sons.

Vázquez, D. P. et al. 2009. Evaluating multiple determinants of the structure of plant–animal mutualistic networks. - Ecology 90: 2039–2046.

Vázquez, D. P. et al. 2012. The strength of plant–pollinator interactions. - Ecology 93: 719–725.

Verdú, M. and Valiente-Banuet, A. 2011. The relative contribution of abundance and phylogeny to the structure of plant facilitation networks. - Oikos 120: 1351–1356.

Verdú, M. et al. 2008. The Nested Assembly of Plant Facilitation Networks Prevents Species Extinctions. - Am. Nat. 172: 751–760.

Wardle, D. A. et al. 2011. The ecosystem and evolutionary contexts of allelopathy. - Trends Ecol. Evol. 26: 655–662.

Wisheu, I. C. 1998. How Organisms Partition Habitats: Different Types of Community Organization Can Produce Identical Patterns. - Oikos 83: 246–258.

Wu, H.-I. et al. 1985. Ecological field theory: A spatial analysis of resource interference among plants. - Ecol. Model. 29: 215–243.

Xue, R.-R. et al. 2014. Community Detecting in Signed Networks based on Modularity. - Math. Methods Sci. Mech.

Zaslavsky, T. 1982. Signed graphs. - Discrete Appl. Math. 4: 47–74.

Figure 1. Ecological meaning of triads in plant spatial association networks. Triads represent the spatial relationships among three different species. (a) Three species appear spatially associated in space, creating a multi-specific vegetation patch. (b) Two species segregate between them, but both positively associate to the third one. (c) Two species associate in space while segregate to the third one. (d) Three species spatially segregate from each other. It is important to note that (a) and (c) represented balanced triads, while (b) and (d) represent unbalanced ones.

Figure 2. Proportion of unbalanced links in plant-plant spatial association networks. (Left). F_L^{Real}: proportion of unbalanced links in real network; F_L^{Random}: proportion of unbalanced links in random network. Values in the *x*-axis represent the ratio between F_L^{Real} and F_L^{Random}, and grey area represents the 95% confidence interval built according to random networks. Values lower than the confidence interval indicate that $F_L^{Real} < F_L^{Random}$, and values higher than confidence interval indicate $F_L^{Real} > F_L^{Random}$. Numbers in the *y*-axis represent the 31 networks, sorted by increasing F_L^{Real}/F_L^{Random}. (Right). Difference between the mean proportion of unbalanced links in real and random networks. $N^{Real} = 31$. Error bars represent the standard error of the difference between means. Real values are significantly lower than random values with a *p*-value < 0.001.

Figure 3. Mean proportions of each type of triads in all the plant-plant spatial association networks. $N^{Real} = 31$. For each type of triad, real and random values are presented. Error bars represent the standard error of the differences between means. Real and random values are significantly different with a p-value < 0.001.

Figure 4. Mean proportions of each type of triads in the plant-plant spatial association networks at four locations. $N^{Real} = 6$ for Monegros; $N^{Real} = 7$ for Cabo de Gata; $N^{Real} = 10$ for Guara; $N^{Real} = 8$ for Ordesa-Monte Perdido. For each type of triad and location, real and random values are presented. Error bars represent the standard error of the differences between means at each location. At all locations, real and random values are significantly different with a p-value < 0.001. There is a significant effect of location for (+ + -) and (+ - -) triads ($p < 0.001$).

Table 1. Characteristics of the locations of the vegetation surveys in Spain.

Plant community, dominant plant community; Temperature, mean annual temperature; Precipitation, mean total annual precipitation. One network was built for each site based on the pooling of the data from all transects. Each location had +50% species unique to that location. Climatic data were obtained from ~~(~~Ninyerola et al. (2005).

Table 2. Plant-plant spatial association network indices for each location in Spain and their correlations with the proportion of unbalanced links and with each type of triad.

S, network size; C^+, connectance of positive links; C^-, connectance of negative links; *R*, ratio of positive and negative links; F_L, proportion of unbalanced links; + + +, proportion of (+ + +) triad; + + -, proportion of (+ + -) triad; + - -, proportion of (+ - -) triad; - - -, proportion of (- - -) triad. Values are mean ± standard error. Significance of correlations is *n.s.* > 0.05, * < 0.05, ** < 0.01, *** < 0.001.

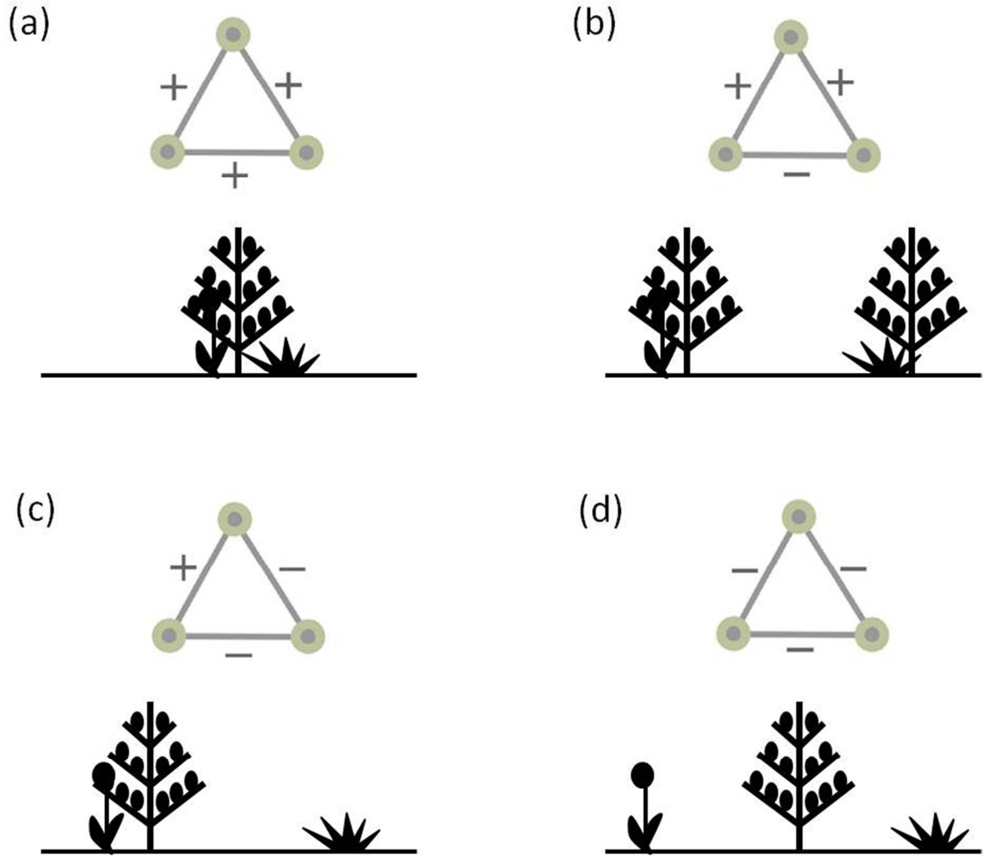

Figure 1. Ecological meaning of triads in plant spatial association networks. Triads represent the spatial relationships among three different species. (a) Three species appear spatially associated in space, creating a multi-specific vegetation patch. (b) Two species segregate between them, but both positively associate to the third one. (c) Two species associate in space while segregate to the third one. (d) Three species spatially segregate from each other. It is important to note that (a) and (c) represented balanced triads, while (b) and (d) represent unbalanced ones.
159x139mm (150 x 150 DPI)

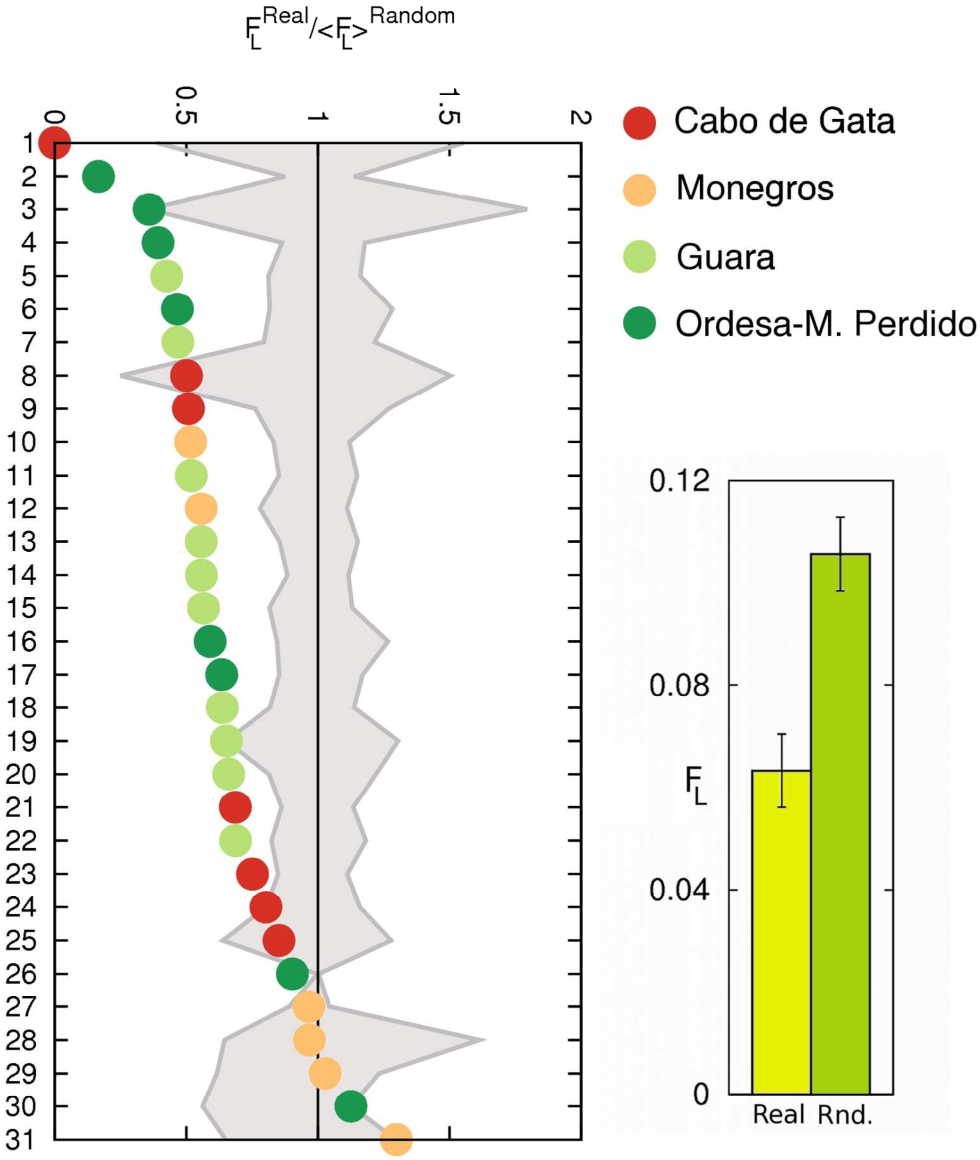

Figure 2. Proportion of unbalanced links in plant-plant spatial association networks. (Left). FLReal: proportion of unbalanced links in real network; FLRandom: proportion of unbalanced links in random network. Values in the x-axis represent the ratio between FLReal and FLRandom, and grey area represents the 95% confidence interval built according to random networks. Values lower than the confidence interval indicate that FLReal < FLRandom, and values higher than confidence interval indicate FLReal > FLRandom. Numbers in the y-axis represent the 31 networks, sorted by increasing FLReal/FLRandom. (Right). Difference between the mean proportion of unbalanced links in real and random networks. NReal = 31. Error bars represent the standard error of the difference between means. Real values are significantly lower than random values with a p-value < 0.001.

170x204mm (250 x 250 DPI)

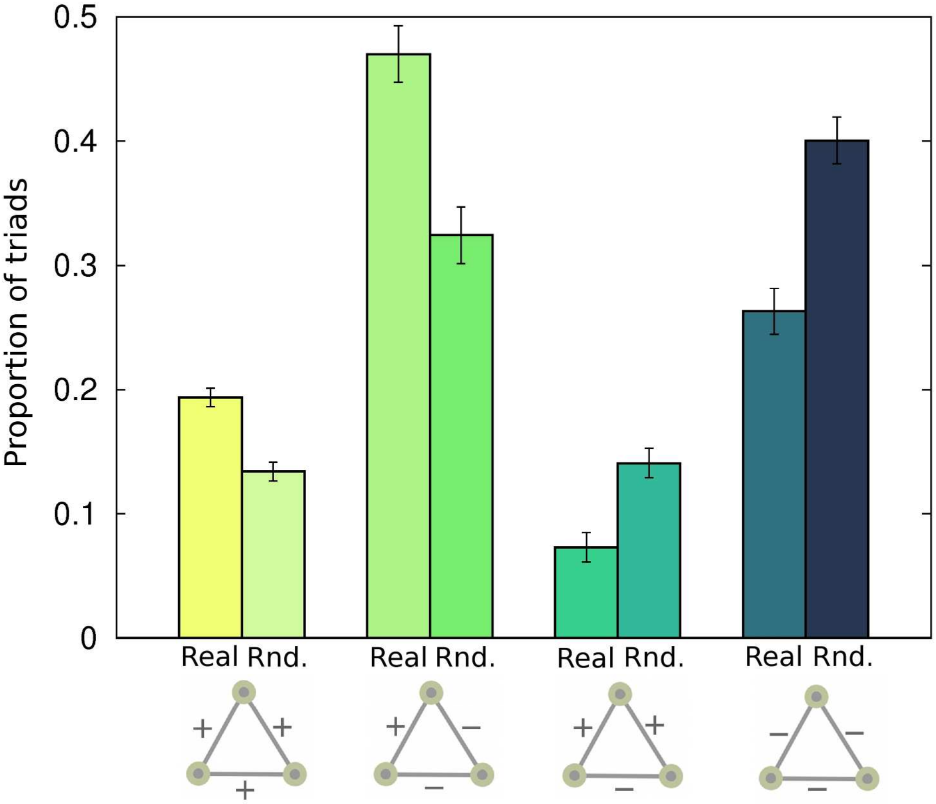

Figure 3. Mean proportions of each type of triads in all the plant-plant spatial association networks. NReal = 31. For each type of triad, real and random values are presented. Error bars represent the standard error of the differences between means. Real and random values are significantly different with a p-value < 0.001.
254x212mm (250 x 250 DPI)

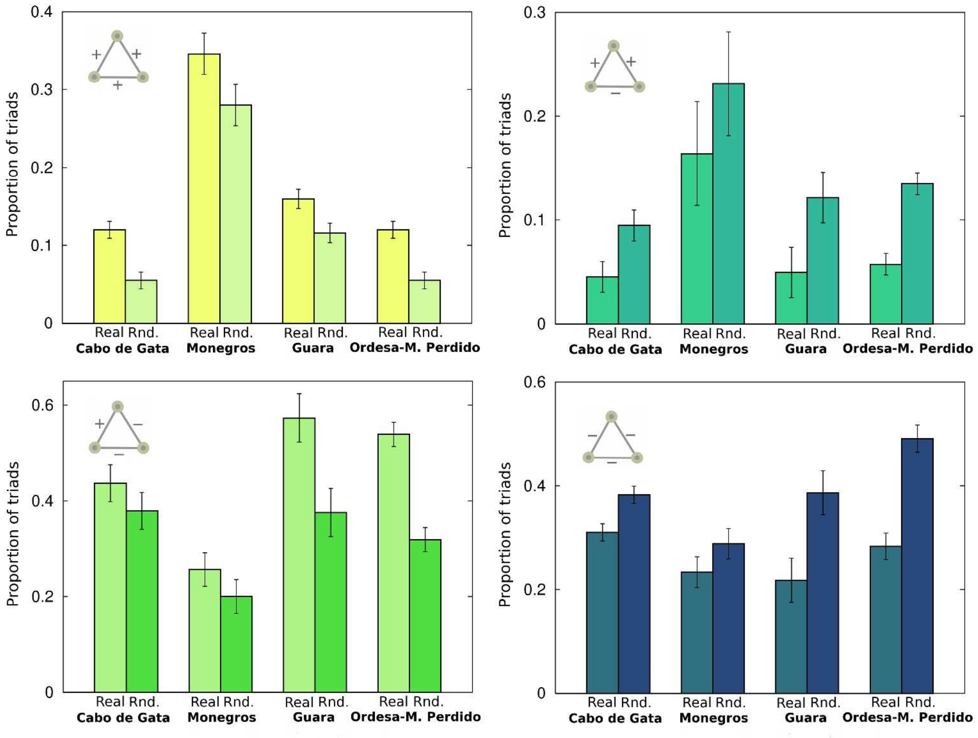

Figure 4. Mean proportions of each type of triads in the plant-plant spatial association networks at four locations. NReal = 6 for Monegros; NReal = 7 for Cabo de Gata; NReal = 10 for Guara; NReal = 8 for Ordesa-Monte Perdido. For each type of triad and location, real and random values are presented. Error bars represent the standard error of the differences between means at each location. At all locations, real and random values are significantly different with a p-value < 0.001. There is a significant effect of location for (+ + -) and (+ - -) triads (p < 0.001).
507x382mm (250 x 250 DPI)

Table 1.

Location	Coordinates	Plant communities	Temperature (°C)	Precipitation (mm)	Number of sites surveyed (transects per site)	Transect lenght (interval)
Cabo de Gata-Nijar Natural Park	36.77 N -2.11 W	Alpha-steppe Open shrubland	24	328	7 (3-4)	500 m (20 cm)
Monegros	41.65 N -0.71 W	Steppe Low shrubland	21	360	6 (6)	250 m (20 cm)
Sierra de Guara	42.27 N 0.18 W	Forest	17	927	10 (3)	500 m (40 cm)
Ordesa-Monte Perdido National Park	42.63 N -0.11 W	Alpine grassland	11	1485	8 (3)	250 m (20 cm)

Table 2.

		S	C^+	C^-	R
Location					
Cabo de Gata-Nijar Natural Park		74.9 ± 9.6	0.039 ± 0.002	0.024 ± 0.01	0.26 ± 0.05
Monegros		44.2 ± 6.9	0.066 ± 0.01	0.034 ± 0.01	0.35 ± 0.02
Sierra de Guara		60.9 ± 6.4	0.037 ± 0.002	0.023 ± 0.002	0.06 ± 0.11
Ordesa-Monte Perdido National Park		77.3 ± 6	0.058 ± 0.01	0.048 ± 0.01	0.11 ± 0.04
Pearson correlation					
F_L	R^2	0.04	0.046	0.006	0.084
	p-value	*n.s.*	*n.s.*	*n.s.*	*n.s.*
+ + +	R^2	-0.017	0.063	-0.215	0.764
	p-value	*n.s.*	*n.s.*	0.008**	<0.001***
+ + -	R^2	-0.203	-0.071	-0.272	0.201
	p-value	0.012*	*n.s.*	0.003**	0.011*
+ - -	R^2	0.154	0.045	0.365	-0.177
	p-value	0.029*	*n.s.*	<0.001***	0.018*
- - -	R^2	-0.002	-0.127	0.021	-0.533
	p-value	*n.s.*	0.049*	*n.s.*	<0.001***