Title

Abstract

Introduction

**[P1 - Role of oceanic currents in long-distance dispersal]**

**[P2 - Examples of studies. in oceanic archipelagos ideal for testing - but lack of data and challenges]**

**[P4 - Also challenges to address Long-distance traits - some traits straightforward for vectors of zoochory or anemochory. For sea dispersal, thalassochorous traits proposed but, need experimentation]**

**[P5 - Littoral flora]**

**[P6 - Objective]**

1. Test whether there exist a relationship between archipelago patterns of littoral the flora and oceanic current connectivity
2. Identify central islands for the current connectivity network and explore whether they contain a higher percentage of littoral plants.

Methods

**Oceanic currents data**

Oceanic current data were sourced from the HYCOM (HYbrid Coordinate Ocean Model) dataset (www.hycom.org), which provides spatially and temporally resolved information on ocean currents (see Chassignet et al. 2007). Data retrieval was implemented using custom R scripts leveraging the get.hycom function from the “HMMoce” r package. This function facilitates the download of NetCDF files containing variables of interest—eastward (water\_u) and northward (water\_v) current velocities—from specified time intervals and geographic regions. These encompassed 26 years, from 1992 to 2018, and three distinct oceanic archipelagos: the Galapagos (-1.8 to 0.9°N, -92.0 to -88.7°W), Canaries (26.7 to 29.9°N, -18.8 to -12.6°W), and Azores (36.5 to 40.1°N, -32.2 to -23.4°W).

Using the extracted velocity components, sea direction (in degrees) and speed (in m/s) were computed. Direction was calculated using the atan2 function from the “raster” package, converted from radians to the 0-360 degrees range. We extracted this information every two days along the 26 years, yielding 4770 (+-1) days with ocean currents data for the analysis in each archipelago, at a resolution of 8km.

**Oceanic currents connectivity**

A conductance matrix, representing the ease of movement between geographic points due to ocean currents, was computed for each day using the flow.dispersion function from the “rWind" package, which uses the alrgorythm proposed by Muñoz et al. (2004) and their variation in Felicísimo et al. (2008) to calculate the movement cost from any starting cell to one of its 8 adjacent cells (Moore neighborhood) accounting for the directional and speed characteristics of ocean currents (see Fernández-López & Schliep 2018).

To assess connectivity between islands within each archipelago, cost matrices were derived from the conductance matrices: Using the costDistance function, the minimum cost path (conductance) required to travel between pairs of islands was computed. This step integrates geographic coordinates (latitude and longitude) of islands with the conductance values derived from the oceanic current data. We selected a coordinate point on each side of the islands to cover all orientations (North, South, West, East) from which we calculated the current connectivity paths to the rest of the islands of the archipelago. For large islands, we added one connection point every 40km (5 pixels of 8 km) to increase our sampling of points and therefore of potential shortest paths connecting with other islands. For instance, Isabela island in the Galápagos archipelago has a total of eight connection points (one in the North and South, and three in the East and West sides). For small islands with an area that occupies less than a pixel of the raters (<8km), a single connection point was assigned.

The marine currents can vary through time. Events such as seasonality and El nino o la nina can alter the connectivity, and plants may have windows of opportunity to disperse in periods with higher connectivity due to these conditions. To obtain this minimum cost that occurs among islands and still capture some variability (avoiding the effect of an anomalous extreme climatic event), we considered the 5% minimum values of cost distances within the temporal range of the data, and obtained their median as their distribution was non-parametric. This yielded a cost matrix of ocean currents, representing the median of the 5 % minimum costs connecting each pair of geographic points.

**Floristic data**

We selected species within the lowland habitat, defined as any area under 500m (ref), from floristic databases: ref for Galapagos, ref for Canaries, ref for Azores. We then grouped them into those classified as littoral and those that don't, non-littoral: 25 and 429 in Galapagos, 111 and 1122 in Canaries, and 24 and 121 in Azores (Supplement Table X-X’).

**Floristic connectivity**

We measured the Bray-Curtis Dissimilarity Index (Bray & Curtis, 1957) to measure the relative separation (high dissimilarity) or closeness (low dissimilarity) of floristic composition. It ranges from zero, when species are very similar among sites, to unity, when no single species is shared among them. We used this index because it takes into account double-zero attributes also called co-absences, which the absence of a species is also indicative of the floristic composition of the island (Greenacre, 2008; Todeschini et al., 2012), and it is commonly used in studies that compare community distances in oceanic islands (e.g. Castro-Urgal & Traveset, 2014; Florencio et al., 2013, Thuesen et al., 2011). We then computed two floristic distance matrices between islands, one for littoral and other for non-littoral species, for each archipelago.

**Island factors connectivity**

To gather further insights from the patterns we measure, we also explored how the floristic connectivity of littoral floras relate to geographic distance, island age and area. To compute geographic distances (km), we downloaded shapefile maps of the archipelagos from NOAA ([www.noaa.gov](http://www.noaa.gov)) and obtained the geographic centroid of each island with the QGIS software and the tool ‘geometry by expression’. These maps also contain the measure for island area (km2). For the island ages (Myr), we considered the mean between minimum and maximum emergence estimates from (ref) for the Galapagos. For the Canaries, we obtained estimated ages from Carracedo et al. 1998 and Anguita & Hernán 2000, and considered the mean for those islands with discrepancies between the two. For the Azores, we obtained the data from \_\_. For each archipelago, the pairwise geographic distances between island centroids forms a connectivity matrix. We computed Euclidean distances between island ages and areas using the function daisy from the “cluster” package to obtain the connectivity matrices for each.

**How does the structure of ocean currents connectivity networks within archipelagos correlate to that of the littoral flora? and to that of other island factors?**

Procrustes analysis is a statistical method used to compare the shapes of datasets (Legendre & Legendre 2012). It measures the degree of similarity between two matrices by finding an optimal superimposition through transformations such as translation, rotation, and scaling. Its flexibility and interpretability has made it widely employed in ecology to measure biogeographic patterns of floristic connectivity (Munoz et al. 2004, Legendre et al. 2012). The goal is to minimize the sum of squared differences between corresponding points in the matrices. After the transformation, the Procrustes distance is calculated as the sum of squared distances between the corresponding points in the transformed configuration and the target configuration. This distance quantifies the degree of similarity between the two sets of points.

To explore how the structural pattern of ocean currents connectivity networks matches that of the littoral flora, we analyzed the correlation between the floristic distance and the current cost matrices. Furthermore, we analyzed the correlation between the structure of the ocean current connectivity network with that formed by geographic, age and area distances. We used the “vegan” package (Oksanen et al., 2007) to perform the Procrustes analysis on Principal Coordinates Analysis of the distance matrices. To test significance, a permutation procedure is used by randomly permuting the points in one dataset 999 times and calculating Procrustes distances for each permutation (Peres-neto & Jackson 2001).

**Island centrality based on oceanic current connectivity**

The inverse of the current cost matrix is the matrix of ocean current connectivity in an archipelago. Within networked systems, graph theory provides centrality measures that determine the importance or influence of a node within a network (Delmas et al. 2019). Studies addressing landscape connectivity often analyze the centrality of ecological patches to identify those that are most critical in terms of connectivity, information flow, or influence within the network (Estrada & Bodin 2008, Bodin & Saura 2010, Treml & Kool 2018, Cecino et al. 2021). In our system, nodes correspond to specific islands (nodes) in the current network of the archipelago, whose link weights correspond to cost values computed from current velocity and direction. To approximate the islands’ roles as critical connectors of the current network, we computed In-degree, Out-degree, Closeness, and Betweenness centrality measures (Opsahl et al. 2010). Starting with a cost matrix representing distances for ocean currents, we inverted the weights to derive connectivity values (). In-degree and out-degree centrality were calculated using these inverted weights to identify islands that act as major sinks and sources in the network, respectively. For closeness and betweenness centrality, we used the original distance-based weights, ensuring that higher closeness centrality values indicate shorter average distances to other islands and higher betweenness centrality values reflect nodes frequently serving as bridges on shortest paths.

*Island in and out degrees - pulls and sinks of oceanic currents*

Degree centrality () is a simple count of the total number of connections linked to a vertex. The weighted In-degree of a node in an ecological network measures the sum of weights (or strengths) of all edges connected to that node. It reflects the overall importance or influence of a node based on the weighted interactions it has with other nodes. On the other hand, the weighted Out-degree measures the sum of weights of all outgoing edges from that node. It indicates the total contribution or output of resources or influence from the node to other nodes in the network. We note them as and :

,

For instance, in a mutualistic network where species exchange resources (e.g., pollination network), the weighted of a plant species represents the total amount of resources (e.g., pollen) it receives from its pollinator species (incoming edges), and the would be the total amount of resources (e.g., pollen) it provides to its plant partners (outgoing edges). In a network of current connectivity, where , of an island accounts for the sum of incoming currents with their associated strengths and for the outcoming ones. We consider the former as “sinks” for oceanic currents in the archipelago and the later as “pulls”.

*Island closeness - isolation based on oceanic currents*

A path is a series of steps that go from one node to another. Closeness centrality () of a node is the inverse of the sum of distances to all the other vertices in a graph. For a node in a network with nodes:

,

where is the shortest path distance between nodes and . Thus, the more central a node is, the closer it is to all other nodes, making it potentially influential in terms of information dissemination or resource transfer. Within the matrix of island distances based on costs to current flow, Closeness centrality provides a quantitative measure of how centrally located and accessible an island is within the current connectivity network in the archipelago.

*Island betweenness - hubs of oceanic currents*

For every pair of vertices in a connected graph, there exists at least one shortest path between the vertices such that the sum of the weights of the edges (for weighted graphs) is minimized. Betweenness centrality () is defined as the number of shortest paths that pass through the node divided by the total number of shortest paths between all pairs of nodes. For a node in a network with nodes:

,

where is the total number of shortest paths from node to node , and is the number of those paths that pass through node .

A node with high betweenness centrality acts as a bridge or bottleneck in the network, facilitating or controlling communication and interactions between other nodes. Within a network of ocean currents in an archipelago, islands with high betweenness centrality may serve as ecological corridors or hubs for species dispersal. Central metrics describe similar aspects and are usually correlated somehow. We study the correlation of these central measures computing their pearson correlation between the islands of each archipelago.

**Does island centrality in the ocean connectivity network affect its percentage of littoral flora?**

To investigate the relationship between the proportion of littoral plants and the centrality measures across different archipelagos (Azores, Canaries, and Galapagos), we applied a generalized linear model using the “glmmTMB” package. Our response variable, the proportion of littoral plants, was modeled using a beta distribution with a logit link function, suitable for proportional data:

for which

Where is a given island, and “” refers to one of the predictor centrality variables: either , , or . was log-transformed and centered to address its large variance across archipelagos. In this case, represents the baseline logit proportion of littoral plants for Azores. ​ and represent how the logit proportion of littoral plants changes for Canaries and Galapagos compared to Azores, respectively. ​ and ​ represent how the effect of in\_degree on the logit proportion of littoral plants varies with Canaries and Galapagos, respectively. All models underwent diagnostic assessments for uniformity, dispersion, homoscedasticity, and identification of outliers, employing the “DHARMa” package (Hartig 2022).

Results

**Correlation between littoral flora and oceanic currents connectivity**

The correlation between the network structures of the ocean current and littoral flora connectivity differed between islands. The correlation was marginally significant in the Galapagos (ProcCor = 0.55, p = 0.057) and also significantly positive in the Canaries (ProcCor = 0.87, p = 0.002), whereas in Azores it was not significant (ProcCor = 0.15, p = 0.919).

For the non-littoral flora, the correlation was significant in the Canaries (ProcCor = 0.82, p = 0.008) and non-significant in the Galapagos (ProcCor = 0.49, p = 0.122) and Azores (ProcCor = 0.38, p = 0.452) (Fig. 2*A*).

The network structure of littoral flora connectivity was also significantly correlated with the connectivity by geographic distance in Galapagos (ProcCor = 0.62, p = 0.015) and Canaries (ProcCor = 0.88, p = 0.002), but not for Azores (ProcCor = 0.16, p = 0.918). The correlation was significant only in Canaries for non-littoral flora (ProcCor = 0.79, p = 0.015) (ProcCor = 0.54, p = 0.070 in Galapagos; ProcCor = 0.43, p = 0.337 in Azores) (Fig. 2*A*).

A significant positive correlation between littoral flora connectivity and the connectivity pattern by island ages was only in Galapagos (ProcCor = 0.58, p = 0.018) (ProcCor = 0.28, p = 0.690 in Canaries; ProcCor = 0.20, p = 0.733 in Azores). We did not detect significant correlations for non-litoral floras (ProcCor = 0.33, p = 0.301 in Galapagos; ProcCor = 0.23, p = 0.786 in Canaries; ProcCor = 0.20, p = 0.686 in Azores) (Fig. 2*A*).

Non-significant correlations were detected between neither littoral and non-littoral flora connectivities and the connectivity pattern by island area in any archipelago (ProcCor = 0.38 and 0.44, p = 0.174 and 115 in Galapagos; ProcCor = 0.17 and 0.13, p = 0.876 and 0.907 in Canaries; ProcCor = 0.11 and 0.26, p = 0.924 and 0.566 in Azores for littoral and non-littoral floras, respectively) (Fig. 2*A*).

**Correlation between oceanic currents and island factors connectivity**

The correlation between the network structures of the ocean current and geographic connectivity was strongly positive and significant in the three archipelagos (ProcCor = 0.92, p = 0.001 in Galapagos; ProcCor = 0.98, p = 0.001 in Canaries; ProcCor = 0.98, p =0.001 in Azores). With island age connectivity, it was significant for Galapagos (ProcCor = 0.62, p = 0.003) and Azores (ProcCor = 0.69, p = 0.039), but not for Canaries (ProcCor = 0.19, p = 0.671). We detected no significant correlation with island area connectivity ProcCor = 0.18, p = 0.697 in Galapagos; ProcCor = 0.12, p = 0.847 in Canaries; ProcCor = 0.09, p =0.860 in Azores) (Fig. 2*B*).

**Island centrality and percentage of littoral species**

The mean percentage of littoral species was highest in the Azores with 16.51% (SD = 1.49), followed by the Canaries with 14.74% (SD = 6.59) and Galapagos with 9.75% (SD = 2.51) (Table 1). The islands with the highest and lowest percentages were max = , min = in Azores, max = min = in Canaries, and max = , min = in Galapagos (Table SX).

The Canaries showed an of 0.56 (SD = 1.48), whereas it was much lower in the Galapagos ( **=** 0.03, SD = 0.02) and the Azores ( **=** 0.01, SD = 0.02). These results were identical for  in the three archipelagos. The mean Closeness centrality in the Galapagos was = 0.17 (SD = 0.04), twice as big as the Canaries (= 0.09, SD = 0.03) and Azores (= 0.06, SD = 0.02). Lastly, the mean Betweenness centrality was highest in Azores (= 13.56, SD = 11.20), followed by the Galapagos (= 10.17, SD = 10.18) and the Canaries (= 7.43, SD = 7.72).

Islands with the highest centrality measures were Santa Cruz ( and ) and Santa Fe ( and ) in Galapagos; Fuerteventura (), Lanzarote () and Tenerife ( and ) in Canaries; and Pico (, ), Faial (), and Sao Jorge () in Azores (Table SX).

**Effect of island centrality on % of littoral species**

In all models, the archipelago significantly influenced the proportion of littoral plants (model : χ²(2) = 40.96, p < 0.001; model : χ²(2) = 26.17, p < 0.001 , model : χ²(2) = 21.76, p < 0.001) (Table 2 and SX).

*Island In-degree*

The ANOVA table indicated a significant joint effect of archipelago and closeness centrality on littoral plant proportions (χ² = 6.70, p = 0.035), indicating that the effect of island varies among archipelagos (Table SX). Analyzing the individual coefficients, the interaction with archipelago showed a trend towards significance for the Canaries (β = 0.144, p = 0.103), indicating a potential effect where a one-unit increase in centered log-transformed in-degree centrality could correspond to a 14.4% increase in the odds of having littoral plants. This was non-significant for the Galapagos (β = -0.21, p = 0.257) and Azores (β = -0.01, p = 0.864) (Table 2).

*Island Closeness*

The ANOVA table indicated that the interaction between island and archipelago was significant (χ²(2) = 16.41, p < 0.001) (Table SX), indicating that the effect of varies among archipelagos. For the Canaries, a one standard deviation increase in results in a logit change of 1.075, corresponding to a substantial increase in the proportion of littoral plants, with an odds ratio of approximately 2.929. This means there is a 192.9% increase in the odds of having littoral plants for each standard deviation increase in . For the Galapagos and Azores, the interaction term was not significant (β = -0.046, p = 0.862; β = 0.084, p = 0.729) (Table 2).

*Island Betweenness*

The ANOVA table indicated that neither the main effect of (χ² = 0.117, p = 0.733) nor its interaction with archipelago (χ² = 0.675, p = 0.714) were statistically significant (Table SX). Consistent with the ANOVA results, the model summary indicated that neither the effect of itself (β = -0.002, p = 0.842) nor the interaction with Galapagos (β = -0.01, p = 0.693) or Canaries (β = 0.01, p = 0.585) was significant (Table 2).

Discussion

References

Anguita, F., & Hernán, F. (2000). The Canary Islands origin: a unifying model. *Journal of Volcanology and Geothermal Research*, *103*(1-4), 1-26.

Carracedo, J. C., Day, S., Guillou, H., Badiola, E. R., Cañas, J. A., & Torrado, F. P. (1998). Hotspot volcanism close to a passive continental margin: the Canary Islands. *Geological magazine*, *135*(5), 591-604.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., ... & Bleck, R. (2007). The HYCOM (hybrid coordinate ocean model) data assimilative system. *Journal of Marine Systems*, *65*(1-4), 60-83.

Delmas, E., Besson, M., Brice, M. H., Burkle, L. A., Dalla Riva, G. V., Fortin, M. J., ... & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, *94*(1), 16-36.

Oksanen, J., Kindt, R., Legendre, P., O’Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. A. S. S. (2007). The vegan package. *Community ecology package*, *10*(631-637), 719.

Legendre, P., & Legendre, L. (2012). *Numerical ecology* (Vol. 24). Elsevier.

Bodin, Ö., & Saura, S. (2010). Ranking individual habitat patches as connectivity providers: Integrating network analysis and patch removal experiments. Ecological Modelling, 221(19), 2393-2405.

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological monographs*, *75*(4), 435-450.

Munoz, J., Felicisimo, A. M., Cabezas, F., Burgaz, A. R., & Martinez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, *304*(5674), 1144-1147.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social networks, 32(3), 245-251.

Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*, *129*, 169-178.

Estrada, E., & Bodin, Ö. (2008). Using network centrality measures to manage landscape connectivity. Ecological Applications, 18(7), 1810-1825.

Treml, E. A., & Kool, J. (2018). Networks for quantifying and analysing seascape connectivity. Seascape ecology, 293-318.

**Figures**

|  |
| --- |
| **Figure 1.** **Network representation of ocean current connectivity in each archipelago and the percent of littoral species in each island.** (A) Galapagos, (B) Canaries, and (C) Azores. Edge width and transparency corresponds to the connectivity weights ( between island pairs. These are log transformed for Canaries and square root transformed for Azores to facilitate readability. |

|  |
| --- |
| **Figure 2. Procrustes correlations between structures of connectivity networks within archipelagos.** (A) Procrustes correlations between island connectivity by floristic composition (Bray-Curtis DI) and island connectivity by ocean currents and island factors. (B) Procrustes correlations between island connectivity by ocean currents and island factors. Black asterisks indicate the level of statistical significance. The solid point indicates marginally significant (p = 0.055 in this case). |

|  |
| --- |
| Figure 3. **Pearson correlation between island centrality measures for the ocean current connectivity network in each archipelago.** In and Out-degree are calculated from connectivity weights ( and approximate the role of islands as sinks or pulls of currents. Closeness and betweenness are calculated on the direct weights (distances) and approximate the role of islands as being more accessible and acting as hubs or bridges in the currents network, respectively. |
|  |

|  |
| --- |
| Figure 4. |

**Tables**

| Table 1. Summary statistics for the percentage of littoral plants and centrality within the ocean currents network in each archipelago. | | | | | |
| --- | --- | --- | --- | --- | --- |
| Variable | Statistic | Galapagos |  | Canaries | Azores |
| % littoral plants | Mean | 9.75 |  | 14.74 | 16.51 |
|  | SD | 2.51 |  | 6.59 | 1.49 |
|  | Mean | 0.03 |  | 0.56 | 0.01 |
|  | SD | 0.02 |  | 1.48 | 0.02 |
|  | Mean | 0.03 |  | 0.56 | 0.01 |
|  | SD | 0.02 |  | 1.48 | 0.02 |
|  | Mean | 0.17 |  | 0.09 | 0.06 |
|  | SD | 0.04 |  | 0.03 | 0.02 |
|  | Mean | 10.17 |  | 7.43 | 13.56 |
|  | SD | 10.18 |  | 7.72 | 11.20 |

| Table 2. Coefficients summary the for the Generalized Linear Mixed Models fitted as: | | | | |
| --- | --- | --- | --- | --- |
| **In-Degree ()** |  |  |  |  |
| Predictor | Estimate | SE | z value | p value |
| (Intercept) | -1.62 | 0.08 | -19.11 | **2.29E-81** |
| archipelago[Canaries] | -0.19 | 0.12 | -1.56 | 0.120 |
| archipelago[Galapagos] | -0.51 | 0.13 | -3.83 | **<0.001** |
|  | -0.01 | 0.08 | -0.17 | 0.864 |
| archipelago[Canaries]\* | 0.14 | 0.09 | 1.63 | 0.103 |
| archipelago[Galapagos]\* | -0.21 | 0.18 | -1.13 | 0.257 |
| **Closeness ()** |  |  |  |  |
| Predictor | Estimate | SE | z value | p value |
| (Intercept) | -1.53 | 0.25 | -6.13 | **<0.001** |
| archipelago[Canaries] | 0.11 | 0.27 | 0.41 | **0.678** |
| archipelago[Galapagos] | -0.73 | 0.28 | -2.63 | **0.009** |
|  | 0.08 | 0.24 | 0.35 | 0.729 |
| archipelago[Canaries]\* | 1.07 | 0.35 | 3.05 | **0.002** |
| archipelago[Galapagos]\* | -0.05 | 0.27 | -0.17 | 0.862 |
| **Betweenness ()** |  |  |  |  |
| Predictor | Estimate | SE | z value | p value |
| (Intercept) | -1.58 | 0.14 | -10.90 | **<0.001** |
| archipelago[Canaries] | -0.29 | 0.21 | -1.37 | 0.171 |
| archipelago[Galapagos] | -0.56 | 0.20 | -2.83 | **0.005** |
|  | -0.002 | 0.01 | -0.20 | 0.842 |
| archipelago[Canaries]\* | 0.01 | 0.02 | 0.55 | 0.585 |
| archipelago[Galapagos]\* | -0.01 | 0.01 | -0.40 | 0.693 |
| was transformed to logit scale.  was centered by subtracting the mean of the variable from each value, and scaled by dividing the centered values by the SD. | | | | |