

Índexs de similitud i papallones: quins fer servir? Consideracions sobre un cas pràctic

David Monteagudo-Sabaté

Travessera de les Corts, 289, 3r 2a; E-08014 Barcelona

davidmonteagudosabate@yahoo.es

Abstract. Similarity indexes and butterflies: Which must be used? Considerations about a practical case. Clusters obtained using four different similarity indexes, all of them usual in practical work, are compared. Dates proceed from butterfly species on 16 sampling stations with different altitudes and vegetation from two mountain ranges in the Sierra Madre del Sur (Mexico). Results show that Jaccard and Sorensen similarity indexes are better for similar comparative studies, with great number of dates and good lepidopterofaunas knowledge.

Resum. Es comparen els dendrogrames de similitud obtinguts amb la utilització de quatre índexs molt usats en la pràctica a partir de les espècies de papallones registrades en 16 estacions de mostreig, caracteritzades per diferents altituds i tipus de vegetació, en dues unitats muntanyoses de la Sierra Madre del Sur (Mèxic). Els resultats semblen indicar que els índexs més adequats per a estudis comparatius semblants, amb gran nombre de dades i bon nivell de coneixement de les lepidopterofaunes, són els índexs de Jaccard i de Sorensen.

Key words: Similarity indexes, butterflies, Sierra de Manantlán, Sierra de Atoyac, Sierra Madre del Sur, Mexico.

Introducció

Sovint, en biologia comparada, ens trobem amb la necessitat de mesurar en quin grau s'assemblen dos conjunts d'organismes (o dos organismes atenent als seus caràcters) per establir relacions entre ells i ordenar-los en classificacions, les quals es poden interpretar de diferents formes. La lògica d'aquestes comparacions es basa en l'existència d'elements compartits pels conjunts a comparar i en d'altres que no ho són. És en aquest punt on neixen els índexs de similitud. Amb la present contribució es pretén discutir certs aspectes de l'ús d'alguns dels índexs de similitud faunística més utilitzats sobre gran nombre de dades des de la consideració de la seva utilitat en el treball d'indagació biogeogràfica.

Encara que s'han assajat moltíssimes possibilitats, els índexs de similitud més habituals en la bibliografia són els índexs de Jaccard, Sorensen i Simpson i el coeficient Simple Matching. Per exemple, tots ells han estat utilitzats en el treball lepidopterològic de la Universitat Nacional Autònoma de Mèxic (UNAM): índex de Jaccard (Vargas-Fernández *et al.* 1992; Salinas-Gutiérrez 1999); índex de Sorensen (Díaz-Batres 1991; Monteagudo-Sabaté *et al.* 2001); índex de Simpson (Balcázar 1993);

coeficient Simple Matching (Salinas-Gutiérrez 1999). Tots aquests índexs, tal com es mostra en el present treball, tenen la interessant particularitat d'estar acotats entre 0 –valor mínim quan dos inventaris no comparteixen cap element– i 1 –valor màxim–. Malgrat que cadascun d'ells parteix d'una filosofia particular i que probablement el seu ús ve determinat per les necessitats de cada situació, sovint s'ha de prendre una decisió *a priori* sobre l'índex a utilitzar.

L'índex de similitud de Jaccard es correspon amb el quocient entre el nombre d'espècies compartides per dos inventaris i el nombre d'espècies total entre els dos inventaris a comparar. L'índex de Sorensen té un funcionament semblant a l'anterior, però dóna més pes a les espècies compartides. L'índex de Simpson ha estat reivindicat per alguns autors (Sánchez & López 1988) com el millor en l'ús pràctic, perquè redueix l'efecte de la mida relativa entre les faunes en establir la comparació de les espècies compartides respecte a la mida de la fauna menys rica. Per últim, el coeficient Simple Matching estableix les comparacions ingressant com a caràcters compartits no tan sols les presències comunes als dos inventaris a comparar, sinó també les absències compartides. La formulació d'aquests quatre coeficients es presenta en la taula 1.

Taula 1 Formulació matemàtica dels índexs de similitud estudiats.

Índex	Formulació	Notació
Jaccard	$J = \frac{s}{(n_1 + n_2 - s)}$	s = nombre d'espècies compartides n ₁ , n ₂ = nombre d'espècies dels inventaris en comparació
Sorensen	$QS = \frac{2s}{(n_1 + n_2)}$	s = nombre d'espècies compartides n ₁ , n ₂ = nombre d'espècies dels inventaris en comparació
Simpson	$S = \frac{s}{n_2}$	s = nombre d'espècies compartides n ₂ = nombre d'espècies de l'inventari menys ric
Simple Matching	$SM = \frac{(a + d)}{(a + b + c + d)}$	a = nombre de presències compartides b, c = nombre de presències en un sol inventari d = nombre d'absències compartides

Material i mètodes

Les dades procedeixen del treball de camp del grup de lepidopterologia del Museu de Zoologia Alfonso L. Herrera, de la Facultat de Ciències (MZFC) de la UNAM. Es va estudiar la distribució altitudinal de les faunes de papallones (Lepidoptera: Papilionoidea) en dues unitats fisiogràfiques pertanyents a la Sierra Madre del Sur (Mèxic): la Sierra de Manantlán (Jalisco i Colima) i la Sierra de Atoyac de Álvarez

(Guerrero). A partir de les dades es va desenvolupar un estudi comparatiu, possible perquè els estudis originals (Vargas-Fernández *et al.* 1992; Vargas-Fernández *et al.* 1999) s'havien fet amb unes tècniques i un esforç similars, a partir d'un gran volum de dades –exemplars i espècies registrades–, i els havia dut a terme el mateix grup de treball per reconèixer els patrons altitudinals de riquesa per al grup a la Sierra Madre del Sur (Monteagudo-Sabaté *et al.* 2001). Les dades de registres es poden consultar en aquestes publicacions.

En total es va comparar la distribució de 414 espècies de papallones, representades per més de 37.000 registres, en 16 estacions de mostreig (que funcionen en el present estudi com a unitats geogràfiques operatives, OGU –Murguía & Rojas 2001–) ubicades en diferents altituds en el conjunt de les dues serralades. La taula 2 resumeix les característiques d'aquestes estacions. Les espècies registrades en les diferents OGU es van incloure en una matriu de dades de presències (1) i absències (0), a partir de la qual es van calcular els valors de similitud entre totes les estacions fent servir els índexs de Jaccard, Sorensen, Simpson i Simple Matching. Posteriorment es van representar gràficament els dendrogrames de similitud entre totes les OGU obtinguts a partir de l'ús dels diferents índexs, agrupant-les segons el mètode del lligament mitjà no ponderat (UPGMA), que és el que més respecta les relacions inicials entre les OGU (Crisci & López-Armengol 1983). Per al càlcul dels índexs de similitud i l'obtenció dels cladogrames d'àrees es va utilitzar el programa informàtic Ntsys ver. 1.80 (Applied Biostatistics Inc. 1993, 1994).

Taula 2 Caracterització de les OGU que s'han fet servir en el present estudi. Acrònims utilitzats: **ATO**, Sierra de Atoyac de Álvarez; **MAN**, Sierra de Manantlán; **BTS**, Bosc tropical subcaducifoli; **Ca**, Pertorbació per cafetars; **BMM**, Bosc mesòfil de muntanya; **BQ**, Bosc de *Quercus*; **BL**, Bosc de lauràcies; **BPQ**, Bosc de *Pinus* i *Quercus*; **BTC**, Bosc tropical caducifoli. Totes les espècies estudiades pertanyen al grup dels Papilionoidea.

Clau	Localitat	Sierra	Altitud (m)	Vegetació	Espècies registrades
LP	Rincón de Las Parotas	ATO	300	BTS/Ca	186
RS	Río Santiago	ATO	680	BTS/Ca	220
PL	Puente de Los Lugardo	ATO	800-900	BTS/Ca	191
EF	El Faisanal	ATO	1.250	BTS/BMM	161
ND	Nueva Delhi	ATO	1.350-1.450	BMM/Ca	135
LR	Los Retrocesos	ATO	1.580-1.650	BMM/Ca	115
LG	La Golondrina	ATO	1.800	BMM	74
ED	El Descanso	ATO	2.000	BQ/BMM/BL	35
EI	El Iris	ATO	1.950-2.150	BQ/BMM/BL	40
PG	Puerto del Gallo	ATO	2.450-2.500	BMM/BPQ	73
AD	Agua Dulce	MAN	250	BTS	179
PT	Platanarillos	MAN	350	BTS	202
LC	La Calera	MAN	650	BTS	240
ZE	Zenzontla	MAN	800	BTC	154
AH	Ahuacapán	MAN	900	BTC	198
LM	Los Mazos	MAN	1.600-1.750	BMM	143

Es discuteixen els resultats obtinguts a partir del treball amb els diferents índexs de similitud per concloure quins van oferir resultats més lògics en aquest estudi a partir del coneixement previ de què es disposa sobre la distribució de les faunes.

Resultats i discussió

Dissortadment, com manifesten Nelson & Platnick (1981), la biologia comparada, com qualsevol altra ciència, ha d'adoptar mètodes sense saber-ne *a priori* l'eficàcia, i, encara que aquests mètodes es podrien avaluar segons les solucions obtingudes, no serien necessaris si ja les coneguèssim. Això és perfectament aplicable a l'ús dels diferents coeficients de similitud.

En principi, s'espera que els arbres obtinguts representin una realitat lògica, i que defineixin la major similitud i els agrupaments segons la unitat fisiogràfica (Sierra) a la qual pertanyen els inventaris o bé segons altituds i tipus de vegetació similars. En el cas present, tenint en compte la pertinença de les dues serralades a la mateixa unitat fisiogràfica, Sierra Madre del Sur (Ferrusquía-Villafranca 1998), i la seva relativa proximitat, aquesta segona possibilitat sembla la més lògica, i, de fet, amb cap dels quatre índexs utilitzats es va definir un agrupament principal de les OGU segons les dues serralades.

En efecte, el factor altitudinal ha estat considerat sovint com el més important en la composició de les biotes en àrees tropicals, com es mostra des del clàssic estudi de Humboldt (1805) sobre el Chimborazo, a l'Equador, on descriu la gran diferenciació que s'estableix entre les flors pertanyents als diferents pisos altitudinals, assimilables amb comunitats de latituds ben definides, i on fa comparacions entre les comunitats vegetals en elevació en importants muntanyes arreu del món. Amb aquesta base, Grisebach (1876) assaja una primera classificació ecològica del territori mexicà destacant la zonació altitudinal en regions calenta, temperada i freda.

La biota mexicana actual és el resultat de complexos processos, a causa de l'enrevessada història tectònica, l'actual conformació orogràfica del territori i la ubicació entre dues grans regions biogeogràfiques (Graham 1998). El sud de Mèxic és una àrea de transició amb una biota composta per elements neàrtics procedents del nord del continent americà, elements d'origen neotropical provinents del sud i processos evolutius interns que donen lloc a tàxons d'origen propi i gran nombre de formes autòctones (Halffter 1978; Rzedowski 1998). La distribució altitudinal dels tàxons mostra una major afinitat de les faunes d'elevacions superiors amb la regió neàrtica i de les altituds baixes amb la regió neotropical. Llorente (1984) i Halffter (1987) destaquen, tanmateix, a partir de l'anàlisi d'entomofaunes, l'existència de tres pisos altitudinals, corresponents als descrits a continuació.

En àrees muntanyoses aïllades, com és el cas de la Sierra Madre del Sur, els patrons de riquesa biològica i d'endemisme són ben definits (Peterson *et al.* 1993). Les zones baixes, corresponents als boscos tropicals, presenten una major diversitat, amb un màxim que es pot establir entre els 500 i 1.000 m d'altitud, mentre que les elevacions

intermèdies, caracteritzades per la presència de taques aïllades de bosc mesòfil de muntanya, es caracteritzen pel major nombre d'endemismes entre els seus tàxons. A causa de l'antiguitat de la Sierra Madre del Sur, podem distingir en aquestes altituds mitjanes (1.200-1.800 m) l'existència d'alguns tàxons vicaris entre les dues serralades comparades, a nivell específic o subs específic (Llorente & Escalante 1992). En les elevacions superiors es dona un major grau d'aïllament, i una gran proporció d'endemismes, però sobre una fauna molt més pobre.

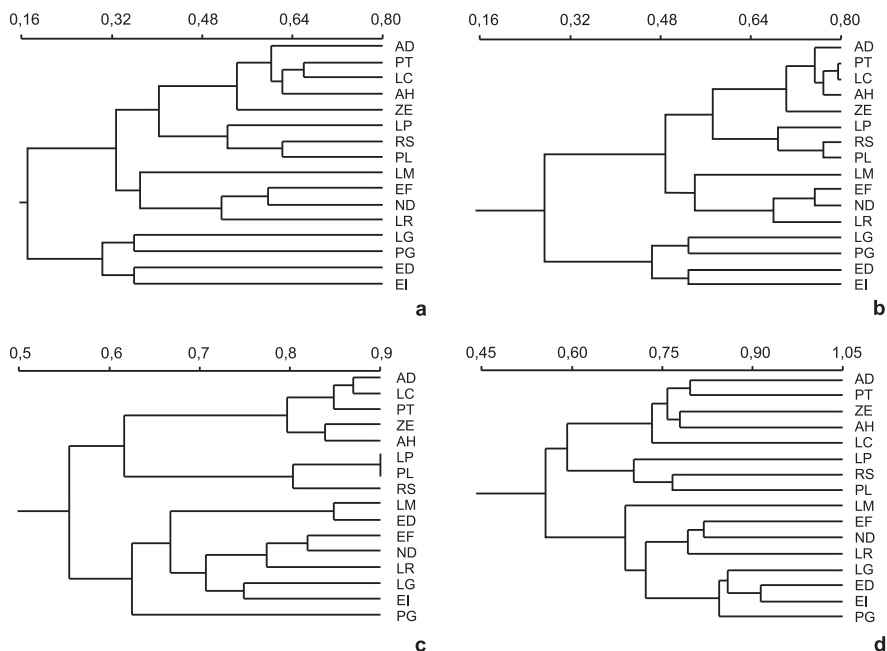


Fig. 1 Dendrogrames de similitud entre les localitats mostrejades, obtinguts amb l'aplicació dels diferents índexs de similitud estudiats: **a**, Índex de Jaccard; **b**, Índex de Sorensen; **c**, Índex de Simpson; **d**, Coeficient Simple Matching. Les claus en lletres que fan referència a les localitats són les que s'indiquen a la taula 2.

En la figura 1 es presenten els dendrogrames de similitud obtinguts amb els diferents índexs utilitzats en el present estudi. L'índex de Sorensen es va fer servir en el treball original i en la presentació dels resultats (Monteagudo-Sabaté 2000; Monteagudo-Sabaté *et al.* 2001) perquè és un dels més habituals en la bibliografia i perquè, en donar major pes a les espècies compartides, contribueix a la definició dels agrupaments en representar gràficament les relacions entre els inventaris en un dendrograma de similitud. Les estacions s'agrupen definint tres pisos altitudinals, a partir d'una

interpretació topològica de l'arbre resultant: un pis inferior (aproximadament fins als 1.000 o 1.200 m), on els inventaris s'agrupen segons la unitat a què pertanyen i per altituds contigües; un pis intermedi (aproximadament entre els 1.200 i els 1.800 m), on es mostra més relació entre les OGU de les dues unitats situades en aquest pis que amb les d'altituds inferiors, i, per últim, un pis superior (a partir de 1.800 m), només definit a la Sierra de Atoyac de Álvarez en l'anàlisi, amb molt poca relació amb els altres i entre els inventaris propis.

Els resultats dels agrupaments segons l'índex de Jaccard es mostren iguals als anteriors; tan sols els valors de similitud són menors. Segons la pràctica pedagògica, l'autor conclou que aquest índex, considerat com el quocient entre les espècies compartides i les totals (és a dir, la intersecció entre la unió de dos conjunts), sembla correspondre's amb la valoració més intuïtiva del concepte de similitud en ecologia i biogeografia.

L'ús de l'índex de Simpson, reivindicat per altres autors (Sánchez & López 1988), va oferir tan sols dos agrupaments, perquè les elevacions superiors s'uneixen amb les intermèdies. És lògic en comparar respecte a la mida de les faunes menors, pel fet que en les elevacions superiors hi ha faunes empobrides, on destaca, sobretot, la disminució del nombre d'espècies, més que no pas l'existència d'espècies exclusives. Així, la situació es complica per la influència de la diversitat de les mostres a comparar, i les definicions són menys lògiques; per exemple, l'extraordinària relació entre LM i ED només es deu a pobresa d'aquesta darrera estació. No es considera correcte per al cas, ja que hi ha una important diferenciació altitudinal de la biota en territori mexicà, per raons ecològiques i, sobretot, històriques, que certs autors consideren divisió biogeogràfica principal.

Per últim, l'ús del coeficient Simple Matching va modificar les relacions individuals entre els inventaris de manera menys lògica, i proposa una major relació entre les faunes d'elevacions intermèdies i superiors, basada en les absències compartides, que adquireixen massa importància en definir les relacions entre els inventaris de les elevacions superiors, més pobres. L'inventari LM resulta el més diferenciat de la resta, cosa que no hauria de ser, perquè s'hi troba un gran nombre d'espècies comunes amb les d'altituds similars de la Sierra de Atoyac de Álvarez. Les agrupacions no són tan clares. El fet d'assumir com a criteri a favor la doble absència implica necessàriament una definició *a priori* de l'espai mostral, i no sembla recomanable en classificacions ecològiques i biogeogràfiques, si no és que es disposés de molts més inventaris i aquests representessin, per exemple, la totalitat d'espècies d'un territori més gran o més continu.

Conclusions

A la vista dels resultats, sembla que els índexs de Sorensen i Jaccard són els més recomanables en la pràctica per a aquests tipus d'estudis, on es disposa de gran nombre de dades i d'un bon coneixement de les faunes, com és el cas de les papallones en les

dues àrees treballades, malgrat que cal reconèixer sempre rere la utilització d'un índex una postura filosòfica. L'índex de Jaccard sembla el que es correspon de manera més intuïtiva amb el concepte de similitud biòtica.

Agraïments

Aquesta contribució neix de resultats no publicats de la Tesi de Mestratge de l'autor a la UNAM. Així, en primer lloc, vull donar les gràcies als companys i amics de la Facultat de Biologia de la UNAM, que van proporcionar les dades i condicions d'investigació per a la realització dels treballs originals. També he d'esmentar el conveni de col·laboració internacional entre la Universitat de Barcelona i la UNAM, que va fer possible la meua estada a Mèxic. El treball posterior i la publicació no s'haurien dut a terme sense el suport constant de Carmen Cervantes, la millor companya que es pot tenir.

Referències bibliogràfiques

- Balcázar, M. A. 1993. Butterflies of Pedernales, Michoacán, Mexico, with notes on seasonality and faunistic affinities (Lepidoptera: Papilionoidea and Hesperioidea). *Tropical Lepidoptera*, 4(2): 93-105.
- Crisci, J. V. & López-Armengol, M. F. 1983. *Introducción a la Teoría y Práctica de la Taxonomía Numérica*. 132 pp. Secretaría General de la Organización de los Estados Americanos. Programa regional de Desarrollo Científico y Tecnológico. Serie de Biología. Monografía 26. Washington D. C.
- Díaz-Batres, M. E. 1991. Estudio ecológico de los lepidópteros Hesperioidea de la Reserva de la Biosfera «La Michilia» Durango, México. *Folia ent. mex.*, 81: 325-333.
- Ferrusquía-Villafranca, I. 1998. Geología de México: una sinopsis. In: *Diversidad biológica de México. Orígenes y distribución* (Ramamoorthy, T. P., Bye, R., Lot, A. & Fa, J. comp.): 3-108. Instituto de Biología. Universidad Nacional Autónoma de México. México D. F.
- Graham, A. 1998. Factores históricos de la diversidad biológica de México. In: *Diversidad biológica de México. Orígenes y distribución* (Ramamoorthy, T. P., Bye, R., Lot, A. & Fa, J. comp.): 109-127. Instituto de Biología. Universidad Nacional Autónoma de México. México D. F.
- Grisebach, A. 1876. A Ramírez, J. 1889. *La Vegetación de México*. 271 pp. Secretaría de Fomento. México D. F.
- Halffter, G. 1978. El mesoamericano, un nuevo patrón de dispersión de la zona de transición mexicana. Descripción y análisis de un grupo ejemplo. *Folia ent. Mex.*, 39-40: 219-226.
- Halffter, G. 1987. Biogeography of the montane entomofauna of México and Central America. *A. Rev. Ent.*, 32: 95-114.
- Humboldt, A. von. 1805. *Essai sur la géographie des plantes; accompagné d'un Tableau physique des régions équinoxiales*. Paris.
- Llorente, J. 1984. Sinopsis sistemática y biogeográfica de los Dismorphiinae de México con especial referencia al género *Enantia* Hübner (Lepidoptera: Pieridae). *Folia ent. Mex.*, 58: 1-207.
- Llorente, J. & Escalante, P. 1992. Insular biogeography of submontane humid forests in Mexico. In: *Biogeography of Mesoamerica* (Darwin, A.S.P. & Welden, A.L. eds.): 139-146. E.O. Painter Printing Company. Florida.

- Monteagudo-Sabaté, D. 2000. *Estudio comparativo de la diversidad de las faunas de Papilionoidea (Lepidoptera) de las sierras de Atoyac de Álvarez (Guerrero, México) y Manantlán (Colima y Jalisco, México)*. 94 pp. Tesis de Mestratge. UNAM. Mèxic D. F.
- Monteagudo-Sabaté, D., Luis-Martínez, A., Vargas-Fernández, I. & Llorente-Bousquets, J. 2001. Patrones altitudinales de diversidad de mariposas en la Sierra Madre del Sur (México) (Lepidoptera: Papilionoidea). *SHILAP Revta lepid.*, 29(115): 207-237.
- Murguía, M. & Rojas, F. 2001. Biogeografía cuantitativa. In: *Introducción a la Biogeografía en Latinoamérica: Teorías, Conceptos, Métodos y Aplicaciones* (Llorente-Bousquets, J. & Morrone, J. J. eds.): 39-47. Las Prensas de Ciencias, Facultat de Ciències, UNAM. Mèxic D. F.
- Nelson, G. & Platnick, N. 1981. *Systematics and Biogeography. Cladistics and Vicariance*. xiv + 567 pp. Columbia University Press, EUA.
- Peterson, A. T., Flores-Villela, O., León-Paniagua, L. S., Llorente-Bousquets, J. E., Luis-Martínez, M. A., Navarro-Sigüenza, A. G., Torres-Chávez, M. G. & Vargas-Fernández, I. 1993. Conservation priorities in Mexico: moving up to the world. *Biodiversity Letters*, 1: 33-38.
- Rzedowski, J. 1998. Diversidad y orígenes de la flora fanerogámica de México. In: *Diversidad biológica de México. Orígenes y distribución*. (Ramamoorthy, T. P., Bye, R., Lot, A. & Fa, J. comp.): 129-145. Instituto de Biología. Universidad Nacional Autónoma de México. México D. F.
- Salinas-Gutiérrez, J. L. 1999. *Análisis de la diversidad de los Papilionoidea (Lepidoptera: Rhopalocera) de los bosques tropicales de la vertiente atlántica de México*. 65 pp. Tesis de licenciatura. UNAM. Mèxic D. F.
- Sánchez, O. & López, G. 1988. A theoretical analysis of some indices of similarity as applied to biogeography. *Folia ent. mex.*, 75: 119-145.
- Vargas-Fernández, I., Llorente-Bousquets, J. & Luis-Martínez, A. 1992. Listado lepidoptero-faunístico de la Sierra de Atoyac de Álvarez en el Estado de Guerrero: Notas acerca de su distribución local y estacional (Rhopalocera: Papilionoidea). *Folia ent. mex.*, 86: 41-178.
- Vargas-Fernández, I., Llorente-Bousquets, J. & Luis-Martínez, A. 1999. Distribución de los Papilionoidea (Lepidoptera: Rhopalocera) de la sierra de Manantlán (250-1.650 m) en los estados de Colima y Jalisco. UNAM. *Publicaciones especiales del Museo de Zoología, Facultad de Ciencias, UNAM*, 11: 1-153.

Data de recepció: 13 de maig de 2003

Data d'acceptació: 7 de gener de 2004