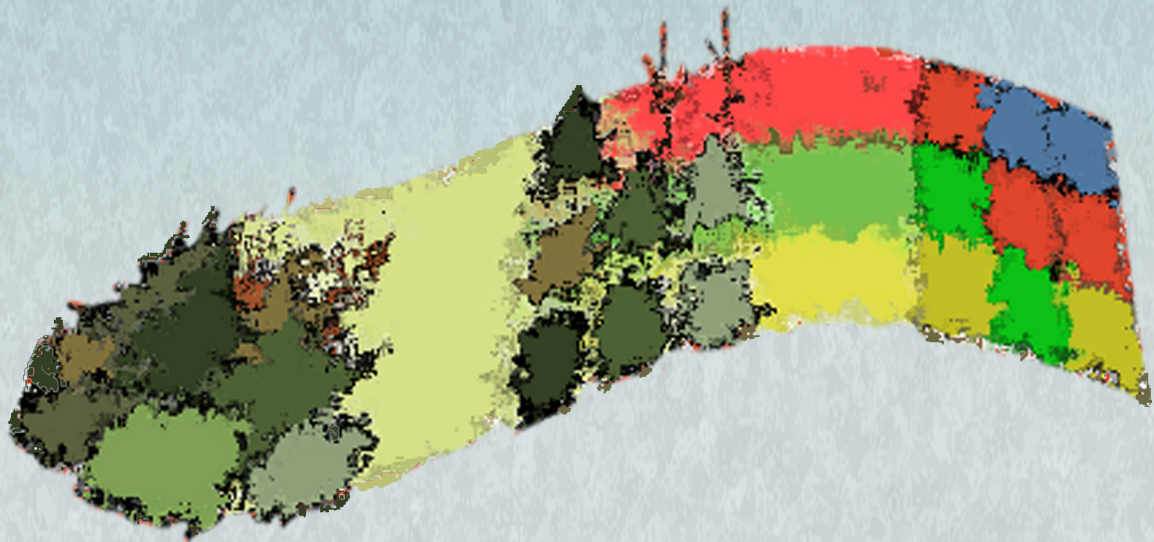UNIVERSITY OF IOANNINA

SCHOOL OF HEALTH SCIENCES

DEPARTMENT OF BIOLOGICAL APPLICATIONS AND TECHNOLOGY

Ph.D. Thesis

Predicting Rates of Biodiversity Decline Following Habitat Losses

Vasiliki Sgardeli

Ioannina
2015

Εικόνα εξωφύλλου: Βασιλική Σγαρδέλη και Λίτσα Μπουλάκη

Cover image: Sgardeli Vasiliki and Litsa Boulaki

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

ΤΜΗΜΑ ΒΙΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Πρόβλεψη Ρυθμών Μείωσης Βιοποικηλότητας που Οφείλονται σε Απώλεια Ενδιαιτήματος

Βασιλική Σγαρδέλη

Ιωάννινα
2015

Ορισμός Τριμελούς Συμβουλευτικής Επιτροπής και Θέματος από τη Γ.Σ.Ε.Σ.: 173/ 6-5-2011
(Members of supervising committee)

Επιβλέπων: John Maxwell Halley, Καθηγητής τμήματος Βιολογικών Εφαρμογών &
(supervisor)   Τεχνολογιών Πανεπιστημίου Ιωαννίνων, επιβλέπων

Μέλη: Χάρης Βάρβογλης, Καθηγητής τμήματος Φυσικής Αριστοτελείου
(members)   Πανεπιστημίου Θεσσαλονίκης

                Yoh Iwasa, Καθηγητής τμήματος Βιολογίας, Kyushu University Ιαπωνίας

Θέμα (Subject): "Predicting rates of biodiversity decline following habitat losses"
 «Πρόβλεψη ρυθμών μείωσης βιοποικιλότητας που οφείλονται σε απώλεια ενδιαιτήματος»

Ορισμός Επταμελούς Εξεταστικής Επιτροπής από τη Γ.Σ.Ε.Σ.: 227/05-10-2015
(Members of examination board)

1.   John Maxwell Halley, Καθηγητής τμήματος Βιολογικών Εφαρμογών &
     Τεχνολογιών Πανεπιστημίου Ιωαννίνων, επιβλέπων
2.   Yoh Iwasa, Καθηγητής τμήματος Βιολογίας, Kyushu University Ιαπωνίας,
     συνεπιβλέπων
3.   Χάρης Βάρβογλης, Καθηγητής τμήματος Φυσικής Αριστοτελείου Πανεπιστημίου
     Θεσσαλονίκης, συνεπιβλέπων
4.   Δέσποινα Βώκου, Καθηγήτρια τμήματος Βιολογίας Αριστοτελείου Πανεπιστημίου
     Θεσσαλονίκης, μέλος εξεταστικής επιτροπής
5.   Κωνσταντία Λύκα, Αναπληρώτρια Καθηγήτρια τμήματος Βιολογίας
     Πανεπιστημίου Κρήτης, μέλος εξεταστικής επιτροπής
6.   Αθανάσιος Καλλιμάνης, Επίκ. Καθηγητής τμήματος Διαχείρισης Περιβάλλοντος
     & Φυσικών Πόρων Πανεπιστημίου Πατρών, μέλος εξεταστικής επιτροπής
7.   Αναστάστιος Τρογκάνης, Αναπλ. Καθηγητής τμήματος Βιολογικών Εφαρμογών &
     Τεχνολογιών Πανεπιστημίου Ιωαννίνων, μέλος εξεταστικής επιτροπής

Έγκριση Διδακτορικής Διατριβής με βαθμό «Άριστα» στις 29-10-2015
Ph.D. thesis approval on 29-10-2015 with grade: "Distinction"

Ευχαριστίες

Acknowledgments

Στους γονείς μου,
Στέφανο και Λίτσα

Table of Contents

Περίληψη

Η απώλεια ενδιαιτήματος αποτελεί σήμερα την κυριότερη αιτία εξαφάνισης ειδών (Millennium Ecosystem Assessment, 2005). Ωστόσο, η πρόβλεψη των εξαφανίσεων και των γενικότερων συνεπειών της απώλειας ενδιαιτήματος για τη βιοποικιλότητα δεν είναι εύκολη, κυρίως επειδή αυτές δεν περιορίζονται μόνο στις περιοχές που έχουν πληγεί άμεσα αλλά επηρεάζουν μακροπρόθεσμα και το εναπομένον ενδιαίτημα. Συγκεκριμένα, η απώλεια ειδών που ακολουθεί την απώλεια ενδιαιτήματος γίνεται σε δύο φάσεις. Η πρώτη φάση αφορά στην απώλεια ειδών που ενδημούν στο κατεστραμμένο ενδιαίτημα (ενδημικές ή *άμεσες* εξαφανίσεις). Η δεύτερη φάση αφορά στην σταδιακή μείωση του αριθμού ειδών στο εναπομένον ενδιαίτημα (*έμμεσες* εξαφανίσεις). Υπάρχουν τουλάχιστον τρεις λόγοι για τις έμμεσες εξαφανίσεις: α) είδη που δεν εξαφανίζονται άμεσα μπορεί να υποστούν μεγάλη μείωση του πληθυσμού τους. Αν ο πληθυσμός που απομείνει είναι πολύ μικρός, αυτά τα είδη θα έχουν αυξημένη πιθανότητα εξαφάνισης. β) Κάποια είδη μπορεί να διαφύγουν στις περιοχές που δεν επηρεάστηκαν από την απώλεια ενδιαιτήματος. Αυτό όμως οδηγεί στην αύξηση του ανταγωνισμού σε αυτές τις περιοχές, που τελικά αναμένεται να οδηγήσει σε αύξηση του ρυθμού εξαφάνισης. Τέλος, γ) η απώλεια ενδιαιτήματος συνοδεύεται συνήθως από μεταβολές των περιβαλλοντικών συνθηκών που επηρεάζουν το εναπομένον ενδιαίτημα. Η πιο συνηθισμένη μεταβολή είναι η απομόνωση του ενδιαιτήματος (π.χ. λόγω κατακερματισμού). Ως αποτέλεσμα, το ενδιαίτημα δεν μπορεί να υποστηρίξει τον ίδιο αριθμό ειδών όπως πριν την απομόνωσή του, ώστε σε βάθος χρόνου κάποια από τα είδη θα εξαφανιστούν. Η σταδιακή απώλεια ειδών και επαναφοράς της κοινότητας στη νέα ισορροπία που ακολουθεί την απώλεια ενδιαιτήματος μπορεί να διαρκέσει ως και χιλιάδες χρόνια (Diamond 1972). Αυτή η διαδικασία των εξαφανίσεων ονομάζεται χαλάρωση (relaxation), ενώ ο αριθμός ειδών που αναμένεται να εξαφανιστεί αναφέρεται ως χρέος εξαφάνισης (extinction debt, Jackson and Sax 2010, Halley et al. 2014).

Την τελευταία δεκαετία έχει διαδοθεί η χρήση στοχαστικών μοντέλων για την περιγραφή της οργάνωσης των βιοκοινοτήτων. Η επιτυχία αυτών των μοντέλων οφείλεται στο

ότι μπορούν να προβλέψουν μακροσκοπικά χαρακτηριστικά των οικολογικών κοινοτήτων στην βάση διαδικασιών που συμβαίνουν σε επίπεδο ατόμων (όπως γεννήσεις, θάνατοι και διασπορά) αλλά και της διαδικασίας ειδογένεσης (speciation). Το πιο γνωστό και επιτυχημένο μοντέλο είναι η ουδέτερη θεωρία της βιοποικιλότητας (Hubbell 2001). Σημαντικό πλεονέκτημα του μοντέλου είναι ότι μπορεί να δώσει λύσεις κλειστής-μορφής για την κατάσταση μιας βιοκοινότητας σε ισορροπία, όπως κατανομές αφθονίας-ειδών και σχέσεις-επιφάνειας ειδών (McKane et al. 2000, Volkov et al. 2003, Vallande and Houchmandzadeh 2003, McKane et al. 2004, Etienne and Alonso 2007, Rosindell and Cornell 2007, O'Dwyer and Green 2010). Παράλληλα, δίνει τη δυνατότητα να μελετηθεί η δυναμική των βιοκοινοτήτων. Έτσι το ουδέτερο μοντέλο της βιοποικιλότητας φαίνεται κατάλληλο για τη μελέτη της απόκρισης των οικολογικών κοινοτήτων στην απώλεια ενδιαιτήματος. Να σημειωθεί εδώ ότι το ουδέτερο μοντέλο περιγράφει την οικολογική κοινότητα ως ένα σύνολο ειδών του ίδιου τροφικού επιπέδου (όπως για παράδειγμα τα δέντρα ενός δάσους) (Hubbell 2001). Στο εξής, όταν αναφερόμαστε στην οικολογική κοινότητα και τις εξαφανίσεις ειδών, θα εννοούμε την οικολογική κοινότητα με την παραπάνω έννοια.

Η παρούσα διατριβή κινείται στην κατεύθυνση της δημιουργίας ενός ενιαίου πλαισίου για την ερμηνεία και την πρόβλεψη της απώλειας βιοποικιλότητας που ακολουθεί την απώλεια ενδιαιτήματος. Για το σκοπό αυτό, χρησιμοποιούμε το ουδέτερο μοντέλο για να αντιμετωπίσουμε τρία μείζονα ζητήματα που σχετίζονται με την απώλεια βιοποικιλότητας. (1) Κατασκευάζουμε ένα εννοιολογικό μοντέλο που εξηγεί τη διαδικασία χαλάρωσης έπειτα από απώλεια ενδιαιτήματος. (2) Χρησιμοποιούμε το ουδέτερο μοντέλο ως μοντέλο-βάση για τον έλεγχο υποθέσεων που αφορούν τη μεταβολή της δομής των βιοκοινοτήτων λόγω κλιματικής αλλαγής. (3) Κάνουμε μια μαθηματική περιγραφή της διαδικασίας χαλάρωσης, που προβλέπει τη μεταβολή του αριθμού των ειδών με τον χρόνο σε μια κοινότητα εκτός ισορροπίας. Τα παραπάνω αποτελούν επίσης και τρεις βασικές εφαρμογές του ουδέτερου μοντέλου στην οικολογία, δηλαδή τη χρήση του μοντέλου ως εννοιολογικού μοντέλου (conceptual model), ως μοντέλου-βάσης (null model), και ως μιας πρώτης προσέγγισης της δυναμικής των βιοκοινοτήτων (first approximation to reality).

Στόχος της πρώτης εφαρμογής είναι να γίνει μια ποιοτική περιγραφή της απώλειας βιοποικιλότητας που ακολουθεί την απώλεια ενδιαιτήματος. Αναφερόμαστε στην περίπτωση

όπου ένα ενδιαίτημα που είναι τμήμα μιας ηπειρωτικής περιοχής απομονώνεται λόγω κατακερματισμού. Εξαιτίας της απομόνωσής του, το ενδιαίτημα μετατρέπεται σταδιακά σε νησίδα (isolate) (Preston 1962, Triantis et al. 2012). Στην συγκεκριμένη περίπτωση, εκτός από την μείωση της επιφάνειας, η μεταβολή της αφθονίας ειδών του ενδιαιτήματος οφείλεται και στην απομόνωσή του, δηλαδή στη μείωση του ρυθμού μετανάστευσης. Χρησιμοποιούμε το ουδέτερο μοντέλο για να εκτιμήσουμε το χρέος εξαφάνισης για το ενδιαίτημα. Ταυτόχρονα από το μοντέλο προκύπτουν και οι σχέσεις επιφάνειας-ειδών για το ενδιαίτημα πριν και μετά την απομόνωσή του. Όπως προκύπτει, το χρέος εξαφάνισης είναι η διαφορά μεταξύ των δύο σχέσεων επιφάνειας-ειδών. Έτσι το ουδέτερο μοντέλο παρέχει μια σύνδεση ανάμεσα στη σχέση επιφάνειας-ειδών και τη δυναμική της διαδικασίας χαλάρωσης. Επιπλέον δικαιολογεί την χρήση της σχέσης επιφάνειας-ειδών για την πρόβλεψη των εξαφανίσεων. Εκτιμώντας το μέγεθος του χρέους εξαφάνισης κάτω από διάφορα σενάρια καταλήγουμε σε ένα σημαντικό συμπέρασμα: το χρέος εξαφάνισης (έμμεσες εξαφανίσεις) μπορεί να είναι τάξεις μεγέθους μεγαλύτερο από τις άμεσες εξαφανίσεις.

Εκτός από τη μείωση της επιφάνειας ενός ενδιαιτήματος λόγω καταστροφής, η απώλεια ενδιαιτήματος μπορεί να είναι έμμεση και να οφείλεται και σε μεταβολή των περιβαλλοντικών συνθηκών. Για παράδειγμα, μια αλλαγή της θερμοκρασίας μπορεί να καταστήσει ένα ενδιαίτημα ακατάλληλο για κάποια ή όλα τα είδη που φιλοξενούνται σε αυτό, αναγκάζοντάς τα να μεταναστεύσουν σε άλλες περιοχές. Στο επίπεδο μιας κοινότητας, η επίδραση μιας κλιματικής αλλαγής εκδηλώνεται ως μια μεταβολή των αφθονιών των ειδών στον χρόνο (temporal turnover). Ωστόσο, η ανίχνευση αυτών των αλλαγών και η σύνδεσή τους με την κλιματική αλλαγή δεν είναι ένα απλό πρόβλημα. Η δυσκολία οφείλεται στο ότι οι αφθονίες των ειδών μιας βιοκοινότητας μεταβάλλονται συνεχώς, ακόμα και χωρίς την επίδραση συστηματικών εξωτερικών επιδράσεων. Αυτή η φυσική διακύμανση των αφθονιών οφείλεται στη δημογραφική στοχαστικότητα (ecological drift) αλλά και στη τυχαία διασπορά των ατόμων (dispersal) (Vellend 2010). Το ουδέτερο μοντέλο εμπεριέχει τις δύο παραπάνω πηγές αβεβαιότητας και έτσι είναι κατάλληλο για την περιγραφή της μεταβολής των αφθονιών στο χρόνο σε συνθήκες ισορροπίας, δηλαδή όταν δεν υπάρχει συστηματική μεταβολή περιβαλλοντικών συνθηκών.

Στη δεύτερη εφαρμογή χρησιμοποιούμε το ουδέτερο μοντέλο ως μοντέλο βάση (null model) για να ελέγξουμε την υπόθεση της επίδρασης της αύξησης της θερμοκρασίας στην κοινότητα πεταλούδων του δάσους της Δαδιάς. Η ανάλυση δείχνει ότι η φυσική διακύμανση εξηγεί μεγάλο μέρος της παρατηρούμενης μεταβολής των αφθονιών. Παράλληλα, υπάρχει και σημαντικός αριθμός ειδών των οποίων η μεταβολή αφθονιών δεν εξηγείται από τη φυσική διακύμανση. Η μεταβολή της αφθονίας των συγκεκριμένων ειδών συμφωνεί με τις αναμενόμενες μεταβολές λόγω αύξησης της θερμοκρασίας. Δεδομένου ότι η αύξηση της θερμοκρασίας είναι η μόνη γνωστή συστηματική μεταβολή συνθηκών στην περιοχή, καταλήγουμε στο συμπέρασμα ότι είναι και η πιο πιθανή αιτία για τις παρατηρούμενες μεταβολές. Η ανάλυση οδηγεί και σε ένα γενικότερο συμπέρασμα: η φυσική διακύμανση των αφθονιών μπορεί να είναι πολύ μεγάλη, και χωρίς τον κατάλληλο στατιστικό έλεγχο μπορεί εύκολα να αποδοθεί λανθασμένα σε εξωτερικά αίτια, όπως η κλιματική αλλαγή. Αυτό είναι ιδιαίτερα επίκαιρο λόγω του μεγάλου αριθμού των δημοσιεύσεων που αποδίδουν παρατηρούμενες μεταβολές σε κοινότητες στην κλιματική αλλαγή, χωρίς τον απαραίτητο στατιστικό έλεγχο (Meshinev et al. 2000, Walther et al. 2002, Poloczanska et al. 2013). Καταλήγουμε, ότι ένα μοντέλο για την εκτίμηση του μεγέθους της φυσικής διακύμανσης είναι απαραίτητο εργαλείο στην αξιολόγηση τέτοιων παρατηρήσεων. Συγκεκριμένα, το ουδέτερο μοντέλο είναι κατάλληλο για αυτό το σκοπό, επειδή, σε αντίθεση με προηγούμενα μοντέλα, εμπεριέχει και τη διακύμανση λόγω διασποράς η οποία επηρεάζει σημαντικά τη σύνθεση τοπικών κοινοτήτων.

Το ουδέτερο μοντέλο δίνει τη δυνατότητα της μαθηματικής περιγραφής της διαδικασίας χαλάρωσης. Παρόλο που η δυναμική του ουδέτερου μοντέλου έχει μελετηθεί διεξοδικά, υπάρχουν λίγες δημοσιεύσεις που να αφορούν τη διαδικασία χαλάρωσης. Συγκεκριμένα οι Halley και Iwasa (2011) έχουν διατυπώσει μια εξίσωση που περιγράφει τη μεταβολή του αριθμού ειδών με τον χρόνο σε απομονωμένα ενδιαιτήματα, δηλαδή ενδιαιτήματα που δεν επηρεάζονται από μετανάστευση ή ειδογένεση. Σε αυτήν την περίπτωση, το μοντέλο προβλέπει ότι στην ισορροπία επιβιώνει ένα μόνο είδος. Στην πραγματικότητα καμία κοινότητα δεν είναι εντελώς απομονωμένη και καινούργια είδη εγκαθίστανται μέσω μετανάστευσης ή δημιουργούνται μέσω ειδογένεσης. Στην τρίτη εφαρμογή, κάνουμε μια γενίκευση της σχέσης των Halley και Iwasa (2011) η οποία

περιλαμβάνει και εμφάνιση νέων ειδών με τη διαδικασία ειδογένεσης τυχαίας διαίρεσης (Haegeman and Etienne 2010). Η εξίσωση που προκύπτει έχει μια επιπλέον παράμετρο που είναι ο σταθερός ρυθμός ειδογένεσης. Η λύση της εξίσωσης οδηγεί σε μια αναλυτική σχέση για τη μεταβολή του αριθμού των ειδών με τον χρόνο. Αυτή έχει τρεις παραμέτρους: τον αρχικό αριθμό ειδών, τον αριθμό των ειδών στην ισορροπία και τον ρυθμό ειδογένεσης. Σε αντίθεση με την απομονωμένη κοινότητα, ο αριθμός ειδών στην ισορροπία προκύπτει ως συνάρτηση του μεγέθους της κοινότητας και του ρυθμού ειδογένεσης.

Η εξίσωση μπορεί να χρησιμοποιηθεί για την πρόβλεψη της απώλειας ειδών σε πραγματικές κοινότητες έπειτα από απώλεια ενδιαιτήματος ή άλλη διαταραχή. Χρησιμοποιήσαμε πραγματικά δεδομένα από εξαφανίσεις πτηνών σε νησιά και ηπειρωτικές νησίδες για να παραμετροποιήσουμε την εξίσωση και υπολογίσαμε χαρακτηριστικούς χρόνους χαλάρωσης. Βρίσκουμε ότι, σε αντίθεση με το μοντέλο για απομονωμένες κοινότητες, ο χρόνος χαλάρωσης αυξάνεται πιο αργά με την επιφάνεια του νησιού και συγκεκριμένα με την τετραγωνική ρίζα της επιφάνειας. Παραμετροποιώντας την εξίσωση από δεδομένα εξαφανίσεων πτηνών στο Barro Colorado Island (BCI), προβλέπουμε ότι η ορνιθοπανίδα του νησιού βρίσκεται ακόμα στην διαδικασία χαλάρωσης και ο αριθμός των ειδών πρόκειται να σταθεροποιηθεί γύρω στα 116 είδη πτηνών περίπου σε 100 χρόνια από σήμερα.

Το μοντέλο που χρησιμοποιήσαμε καλύπτει μόνο έναν από τους μηχανισμούς με τους οποίους γίνεται ανανέωση των ειδών μιας κοινότητας. Παραμένει ανοιχτό να βρεθούν αντίστοιχες καμπύλες χαλάρωσης για κοινότητες που επηρεάζονται από μετανάστευση ή άλλους μηχανισμούς ειδογένεσης πέραν της ειδογένεσης τυχαίας διαίρεσης. Επιπλέον ένα μοντέλο που περιγράφει πιο λεπτομερώς τον μηχανισμό ειδογένεσης θα δώσει μια πιο ρεαλιστική περιγραφή της διαδικασίας χαλάρωσης.

Ερευνήσαμε τη δυνατότητα του ουδέτερου μοντέλου της βιοποικιλότητας να περιγράψει τη διαδικασία χαλάρωσης έπειτα από απώλεια ενδιαιτήματος. Το μοντέλο απλοποιεί την οικολογική πραγματικότητα, αγνοώντας τις διαφορές μεταξύ των ειδών και τις μεταξύ τους αλληλεπιδράσεις και λαμβάνοντας υπ' όψιν μόνο στοχαστικές διαδικασίες σε επίπεδο ατόμων, δηλαδή τη δημογραφική στοχαστικότητα και τη διασπορά αλλά και την διαδικασία της ειδογένεσης. Αυτό σε καμία περίπτωση δεν αποτελεί ισχυρισμό ότι οι άλλες

οικολογικές διαδικασίες δεν υπάρχουν ή δεν παίζουν σημαντικό ρόλο, όπως συχνά αναφέρεται σε αρνητικές κριτικές εναντίον του μοντέλου. Η απλοποίηση αποσκοπεί στην περιγραφή μακροσκοπικών χαρακτηριστικών των οικολογικών κοινοτήτων, για την περιγραφή των οποίων ένα πιο λεπτομερές και ρεαλιστικό μοντέλο πιθανώς να μην είχε την ίδια προβλεπτική ικανότητα ενώ η εξαγωγή συμπερασμάτων θα ήταν δύσκολη. Παράλληλα, ένα μοντέλο που βασίζεται στις βασικές οικολογικές διεργασίες της δημογραφικής στοχαστικότητας και της διασποράς μπορεί να αποτελέσει τη βάση για την διερεύνηση επιπλέον διαδικασιών που επηρεάζουν σημαντικά την οργάνωση και την εξέλιξη των οικολογικών κοινοτήτων (Rosindell et al. 2012, McGill and Nekola 2010, Alonso et al. 2006).

Παρά τις απλοποιήσεις, το ουδέτερο μοντέλο αποτελεί ένα σημαντικό εργαλείο για την περιγραφή της διαδικασίας χαλάρωσης. Το μοντέλο εμπεριέχει τα στοιχεία για την ποιοτική περιγραφή της διαδικασίας, αλλά και για ποσοτικές εκτιμήσεις του χρέους εξαφάνισης. Επίσης, μπορεί να περιγράψει τη δυναμική της διαδικασίας χαλάρωσης και να προβλέψει τη μεταβολή του αριθμού των ειδών με τον χρόνο σε μια κοινότητα εκτός ισορροπίας. Τέλος, το ουδέτερο μοντέλο είναι κατάλληλο για την περιγραφή της φυσικής διακύμανσης των αφθονιών, λόγω δημογραφικής στοχαστικότητας και διασποράς. Ως τέτοιο μπορεί να χρησιμοποιηθεί για τον έλεγχο υποθέσεων σχετικά με τις μεταβολές αφθονίας λόγω κλιματικών μεταβολών ή άλλων συστηματικών εξωτερικών επιδράσεων. Καταλήγουμε ότι το ουδέτερο μοντέλο μπορεί να αποτελέσει βάση για την οικοδόμηση μιας ενοποιημένης περιγραφής της διαδικασίας χαλάρωσης. Η κατανόηση της διαδικασίας χαλάρωσης μέσα από απλά μοντέλα είναι η αρχή για μια πιο ρεαλιστική περιγραφή της απώλειας βιοποικιλότητας έπειτα από διαταραχές.

Abstract

Biodiversity loss is accelerating as a result of the destruction of species natural habitats (habitat loss, fragmentation etc.). A direct effect of habitat loss is species extinctions, however not all of the extinctions are imminent. On the contrary, most extinctions are spread out in a long period of time following the loss (delayed extinctions), a process that is called relaxation. Today there is no single framework for predicting the effects of habitat loss on biodiversity. However, the recent development of stochastic models of community assembly, in particular Hubbell's neutral theory of biodiversity, has given the opportunity to model community dynamics. The purpose of this thesis is to investigate the capacity of the neutral model to describe the process of biodiversity loss in general and to provide a mathematical description of the relaxation process in particular. We investigated the above matters in three applications of the neutral model.

In the first application the model is used to qualitatively describe the relaxation process in habitat fragments that became isolated due to habitat loss. It is found that the number of delayed extinctions on the fragments can be predicted as the difference between the sample species-area relationship (SAR), describing the area before the loss, and the isolate SAR (ISAR), characterizing the remaining fragments. Both types of SARs can be predicted from the neutral model. Interestingly, delayed extinctions can be up to two orders of magnitude more than imminent extinctions (at least for the set of parameters used). This shows that large errors can arise in extinction forecasts if delayed extinction are ignored, a fact that has been neglected in relevant studies.

The second application deals with the problem of attribution of community temporal turnover to external drivers, like climate change; an observed turnover can be the result of a systematic external driver or it can just be the result of natural drift, i.e. the natural fluctuation of species abundances due to stochastic demography and random dispersal. Using the model, we estimated the expected (due to natural drift) temporal turnover of a community of butterflies and compare it with the observed turnover. The comparison shows that the observed turnover is higher than expected by natural drift, hence there must be additional

causes for the observed turnover; possibly the systematic temperature rise recorded in the study area. Still, a considerable amount of the observed turnover is explained by natural drift. This highlights the need for careful statistical tests before attributing observed alterations in communities to climate change or other external drivers. We argue that the neutral model provides a good basis for such tests.

In the third application we present a mathematical description of the relaxation process. Solving the neutral model's master equation, we derive a closed-form expression for the variation of species richness with time in a community that is subject to speciation. This is an extension of already existing results applying to isolated habitats, namely habitats with no immigration or speciation. We used the equation to estimate relaxation times in islands and forest fragments that suffered habitat loss and found that the new formula improves the predictions. In particular, it predicts a scaling of the relaxation time with habitat area, which is in better agreement with observations.

In summary, we have assessed the use of the neutral model of biodiversity in extinction forecasts and conclude that this can provide a basis for building a unified framework to study the process of biodiversity loss.

Chapter 1

Introduction

1.1 Land use change, a major threat for biodiversity

As human population grows, the demand for area to support it increases rapidly. Habitat destruction or conversion to agricultural land, housing and associated infrastructure like transport and energy facilities, reduces the area and resources available to other species. Apart from habitat loss caused directly by human activities, this can also be the result of other environmental changes, as for example the melting of ice due to temperature rise, desertification, etc. Today, habitat loss is the number one cause of species extinctions and constitutes a major threat to biodiversity (Millennium Ecosystem Assessment 2005).

At the same time, the effects of habitat loss on ecological communities cannot be accurately predicted. There are several reasons for this. Firstly, the effects of habitat loss on species richness are not easily accessed by observation (Collen et al. 2010). For instance, it is very difficult to confirm the absence of a species (extinction) and many species that were thought extinct have reappeared (Priddel et al. 2003, Fisher and Blomberg 2011, Scheffers et al. 2011). Thus, theoretical predictions cannot be easily confirmed or falsified. Secondly, habitat loss has long-term effects on the affected communities, which can cause species extinctions long after the loss has been realized. However, the traditional method for estimating extinctions, namely the SAR method, predicts extinctions on the basis of area reduction, which is the final result of habitat loss, ignoring the dynamics of the extinction process. As a result, the predictions cannot be matched with observations at intermediate stages of biodiversity loss. A dynamical approach to the extinctions process comes from the theory of island biogeography (MacArthur and Wilson 1967, Diamond 1972). Still this model does not take into account the community composition, as all species are considered to have the same probability of going extinct. Thus currently, there is no single theory of biodiversity that can view all the matters regarding biodiversity loss under a unified framework.

Furthermore, as the effects of habitat loss are long-term, such a framework needs to take into account the assembly and dynamics of ecological communities. It should be noted at this point that we are interested in the ecological community in a restricted sense. In particular, we think of an ecological community as a group of species that belong to the same trophic level and live in the habitat at the time of the loss.

The recent development of stochastic models of community assembly has established a different view of the organization of ecological communities. In these, community assembly is explained on the basis of individual based processes (births, deaths and dispersal) as well as speciation. Hence, such models give the opportunity to explicitly model ecological community dynamics. The Neutral Theory of Biodiversity (NTB) (Hubbell 2001) is the most popular and successful among others and its equilibrium and dynamical features have been studied thoroughly (McKane et al. 2000, Volkov et al. 2003, Vallande and Houchmandzadeh 2003, McKane et al. 2004, Azaele et al. 2006, Etienne and Alonso 2007, Rosindell and Cornell 2007, O'Dwyer and Green 2010, Chisholm 2011). However, to date there haven't been many studies that examine the consequences of these theoretical considerations on the process of biodiversity loss following habitat loss. In this thesis we explore the prediction of the neutral theory of biodiversity with respect to the process of extinctions following habitat loss, which facilitates the aim of developing a unified framework for understanding and predicting the effects of habitat loss on ecological communities.

1.2 The effects of habitat loss and models of community assembly

Habitat loss can have both direct and indirect effects on species. Firstly, when a habitat area is destroyed, all individuals that cannot escape from the area are killed, which means that species lose part of their population. If all individuals of a species are found entirely in the lost habitat at the time of habitat loss, then this species becomes locally or globally extinct. These extinctions, observed immediately after habitat loss, are called **imminent** or **endemic** extinctions. Apart from these direct effects, there are also long-term effects on the remaining species. The extinctions that happen some time after habitat loss are called **delayed extinctions** and the relevant phenomenon is called **extinction debt**. There are at least three

reasons for delayed extinctions: 1) even if a species is not directly extinct, it could still lose a large part of its population, which can put it near an extinction threshold. This means that this species has a high probability of going extinct in the time following the loss. 2) Some species may have escaped and found refuge in the remaining habitat areas. However, as there is a limit to the number of individuals and consequently the number of species that the remaining habitat can support, some of these species are expected to go extinct at a later time due to increased competition. Finally, 3) habitat loss is in many cases accompanied by a change of conditions affecting the remaining habitats. A usual such change is isolation due to fragmentation. Due to isolation, the remaining fragments cannot host the same number of species as when they were parts of a continuous habitat and in time some of these species are bound to go extinct. From the above it is clear that species extinctions following habitat loss are spread out in a period of time after the destruction of the habitat. This period of extinctions, known as the **relaxation process**, can take up to thousands of years (Diamond 1972, Terborg 1974).

To understand the effects that habitat loss has on the ecological community it is first appropriate to understand the effects on single species. When considering the population of a single species, all the interactions that this has with its environment (abiotic) and with other species (biotic) collectively determine its reproduction and death rate. Under a deterministic perspective, if the birth rate is higher than the death rate, then the species population is expected to increase exponentially. If the death rate is higher, then the species population will fall exponentially to extinction. For big enough population sizes, the change in abundance of a single species is well approximated by a deterministic growth curve, which is a solution to the logistic equation (Renshaw, 1991). The same is not true if the population size falls below an extinction threshold. This is a critical population size, below which the species has high probability of becoming extinct even with a positive growth rate, simply because of demographic stochasticity or other random disturbances. For example, suppose that a species population is represented by one individual and the birth rate is twice as big as the death rate. This means that in every 3 events there are on average two births and one death. Deterministically, this species average population will increase exponentially and there is no risk of extinction. However, since births and deaths occur randomly (i.e. at random times),

there is still a probability of 1/3 that this one individual dies before giving birth and thus the species will go extinct (Renshaw, 1991).

When considering two co-existing species, the interaction between them is an important determinant of their population growth. The species could be independent (i.e. not affecting one another) or could interact (e.g. competition for the same resources). Again, the species population growth can be described by logistic growth equations. This system of coupled equations can be solved explicitly for some sets of parameters and successfully predicts the dynamics of the species populations (for examples see Renshaw 1991, chapter5). Still in this case, the stochastic drift due to demographic events can play an important role at low abundances. Adding more species, the number of interactions increases rapidly and the system of equations cannot be solved explicitly. Moreover, the solution that can be found through simulations becomes less informative of the overall state of the community and usually lacks prediction power. In this case, it is more effective to replace the model of inter-specific interactions with a model that considers the assembly of the community as a whole.

There are two opposing theories that offer an explanation to the assembly of ecological communities. The first is the niche-based community assembly. According to this view, every species in a community occupies its own niche, which refers to the set of conditions and resources that enable the species to persist (Hutchinson 1957). The co-existence of species in a community is possible if these occupy different or slightly overlapping niches. If two species have widely overlapping niches, then competitive exclusion will in time lead to the extinction of one of them. In the end, the abundances of species in a community reflect their success of adaptation to their biotic and abiotic environment. The second view is dispersal-assembly. As defined by Hubbell (2001), dispersal-assembly asserts that species abundances are governed by random speciation, demography and dispersal. That is to say, there is no need to incorporate species differences in order to explain the different relative species abundances in a community; these are the result of the random dispersal patterns of individuals and stochastic demography. The two theories seem contradictory; however they can both describe reality very well on different spatial and temporal scales. Niche theory offers an explanation of competitive exclusion between two similar species (species scale) (Gause 1934). On the other

hand, models based on dispersal-assembly can capture the collective effect of demographic stochasticity and dispersal on a macroscopic scale, where many species are involved (macroecological scale).

The neutral theory of biodiversity is currently the most popular dispersal-assembly model that can successfully capture macroecological community patterns. The theory views the ecological community as a number of trophically similar species that coexist in the same area and compete in equal terms for the same limited resources (Hubbell 2001). The limitation of resources requires that the sum of the abundances of all species is constant over time (community size constraint). Apart from this, species dynamics are independent and all inter-specific interactions are ignored. The neutral community evolves in time through demographic events (birth and deaths), dispersal and speciation, which are all modeled as stochastic events (chance events). The community size constraint is very relevant in the case of habitat loss, as the basic effect of habitat loss is the reduction of area, which sets a limitation on the maximum community size that can be supported. Thus, within the neutral theory, the response of a community to habitat loss is modeled as a rearrangement of species abundances that follows: a) the reduction of populations of single species and b) the reduction of available resources, which constrains the total community size.

1.3 Aim and structure of the thesis

Aim of the thesis

The ultimate goal of this thesis is to contribute towards developing a unified framework for studying the effects of habitat loss on biodiversity. More specifically we use the Neutral Theory of Biodiversity (NTB) of Hubbell (2001) to address three major issues:

1. To develop a conceptual framework for understanding the process of extinction following habitat loss.
2. To demonstrate how the neutral model can be used as a null model to estimate the magnitude of ecological drift and assess the observed turnover in real communities.
3. To build a mathematical description of the dynamics of extinctions following habitat loss.

Addressing the above questions, we also provide three major applications of NTB in ecology, namely the use of the NTB as a conceptual model, as a null model and as an approximation to the relaxation process.

First Application: a theoretical framework for understanding the process of extinction following habitat loss

There are several frameworks for understanding the process of relaxation following habitat loss. Currently, the most commonly used framework for extinction forecast is the SAR method, which, however, has been criticized for its assumptions (e.g. Connor and McCoy 2001) and even as fundamentally flawed (He and Hubbell 2011). In the theoretical introduction chapter (Chapter 2) we include a short review of the SAR method, which attempts to: 1) recognize the method's limitations and possible errors that arise at its implementation and 2) outline the correct implementation of the method that allows the correct interpretation of its predictions. The review concludes that the SAR method is justified. The neutral model provides yet another way of looking at the process of extinctions. In the first application, we use the neutral model to construct a conceptual framework that describes the relaxation process in habitats that become isolated due to fragmentation. The model predicts delayed extinction as the difference between the sample SAR describing the habitat before the loss and the isolate SAR describing the remaining fragments, while both SARs are provided by the model. Hence, the model provides a generalization to the SAR method that takes into account the delayed extinctions due to the isolation of the remaining habitat. Overall, the framework provides a unification of the different concepts and ideas that are used to describe the biodiversity loss following habitat loss.

Second Application: the use of NTB as a null model to estimate the magnitude of ecological drift and assess the observed turnover in real communities

Along with habitat loss, systematic changes of environmental conditions (e.g. climate change) can increase the local rate of species extinctions and alter a community's composition with time. Hence, it is important to be able to recognize the signs of climate change early enough in order to take appropriate measures. This proves not an easy task, as the effects of climate

change cannot be easily distinguished from the effects of chance, namely the natural drift of species abundances that is caused by demographic stochasticity and dispersal. In this application, we use the neutral model as a null model to simulate natural drift and use statistical hypothesis testing to assess the observed turnover in a community of butterflies. To parameterize the model, we follow the usual methodology of maximum likelihood parameter estimation described by Etienne 2007 and Jabot et al. 2008. However, we introduce a slight modification to the previous methodology by directly estimating one of the parameters from available data. Overall, the method is not new; nevertheless it is topical as there are many studies that claim to prove the effects of climate change on ecological communities by showing that observed changes are correlated with the expected changes under climate change. However, without testing if the magnitude of the observed turnover is within the limits of natural drift, such conclusions are not sound. Although this application is not directly related to habitat loss, the model is used to decide whether the community is at equilibrium or whether it is drifting out of equilibrium due to the effects of climate change. In the latter case, habitat loss is indirect as species are forced to migrate to other habitats due to the changing conditions.

Third Application: predicting the rate of extinctions following habitat loss

An ultimate goal in the study of the relaxation process is to be able to predict the decline of species richness with time. To this end, the neutral model of biodiversity provides a very promising starting point. Paradoxically, although the neutral model dynamics have been thoroughly studied, there haven't been many studies that apply to the process of relaxation. Among the few exceptions are the studies of Leigh et al. (1993), Gilbert et al. (2006) and Halley and Iwasa (2011). In particular, Halley and Iwasa (2011) have used the neutral model and found a relaxation equation that describes the variation of species richness with time in a fully isolated habitat (i.e. a habitat with no recruitment of new species). The aim of this last application is to expand the work of Halley and Iwasa (2011) to "open" habitats, i.e. habitats where new species are recruited through speciation or immigration. For this purpose we use the neutral model with random-fission speciation (Haegeman and Etienne 2010) and derive a closed-form equation for the variation of species richness with time. To demonstrate how the

equation can be used in extinction forecasts, we present two separate applications where the equation is parameterized from data of avifaunal extinctions. In the first application we estimate the relaxation times of the avifaunal communities in islands and forest fragments. In the second application we predict the avifaunal relaxation of the Barro Colorado Island.

Structure of thesis

The thesis is organized as follows. Chapter 2 is an introduction to the theoretical tools that were used in the research and are essential to comprehend the remainder of the thesis. This is divided in three parts. In the first part, we introduce terminology and conceptualize habitat loss and the relaxation process. The second part is a short review of the SAR method for extinction forecasts. In the third part we introduce the neutral theory of biodiversity, giving a detailed mathematical description. In Chapter 3, we use the neutral model to build a conceptual framework for understanding extinctions following habitat loss (first application). In Chapter 4, we use the neutral model as a null model to simulate ecological drift in a community of butterflies (second application). In Chapter 5, we use the neutral model to mathematically describe the relaxation process (third application). Finally, in Chapter 6, we present the main conclusions and assess the use of the neutral theory of biodiversity in extinction forecasts and in general.

Chapter 2

Theoretical Background[1]

In this chapter we introduce the theoretical tools that were used in this research and are necessary to comprehend the work presented in the remainder of this thesis. In section 2.1 we introduce terminology and define habitat loss and the relaxation process. Section 2.2 is a review of the SAR method for extinction forecasts. Finally, in section 2.3 we introduce the neutral theory of biodiversity and give a detailed mathematical description.

2.1 Habitat loss

Before defining other terms, it is appropriate to define habitat loss. The habitat loss paradigm that will be used throughout this thesis is depicted schematically in Figure 2.1.
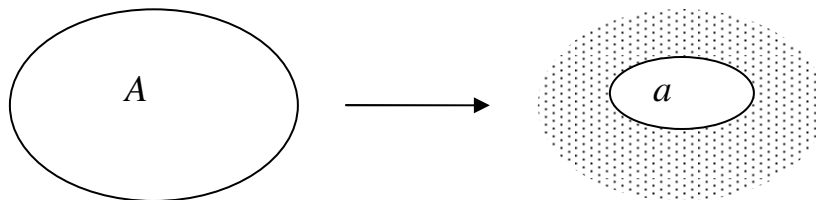
Figure 2.1 Schematic representation of a habitat loss scenario. An initial habitat of area *A* suffers habitat loss which causes the reduction of its area to *a*. Habitat loss happens rapidly, which means that all individuals found in the destroyed habitat (dotted area) are killed. The destroyed habitat is considered contiguous and inhospitable for the community under study, while no habitat regeneration takes place afterwards.

---

[1] Parts of this chapter have been published in Halley, Sgardeli and Triantis (2014) (parts of sections 2.1.1 and 2.1.2) and Halley, Sgardeli and Monokrousos (2013) (section 2.2).

In this, a habitat of area *A*, called the **initial habitat**, suffers habitat loss that results in the destruction of part of its area (dotted area). The habitat that is left after the destruction, called the **remaining habitat** has reduced area $a < A$. Note that the initial habitat represents any habitat area of interest. This could be embedded in a larger habitat of the same or of different type, or could be surrounded by inhospitable matrix. Following Halley and Iwasa (2011) we make the following assumptions about habitat loss. We consider that the destroyed habitat is contiguous with no special edge effects and is rendered inhospitable for the community under consideration. For example, if the destroyed habitat is covered by water and the target species in the remaining habitat are trees, then the destroyed habitat cannot host the community any more. This kind of habitat loss is called complete. Furthermore, we assume that there is no habitat regeneration taking place after the loss, namely the area that is lost stays lost. Finally, habitat loss is considered to take place instantaneously or at least very rapidly. This means that all individuals that are found in the destroyed habitat at the time of habitat loss and cannot escape are killed. For a definition of the main terminology relevant to habitat loss see Box 2.1.

The community that is left after habitat loss will be in general out of equilibrium and species extinctions are expected to occur in the period following the loss, during a process that is called **relaxation**. In order to define the relaxation process and the non-equilibrium state, there should first be a notion of an equilibrium state for the community. In reality, ecological communities are affected by so many processes that equilibrium is not easy to define and it is even questionable whether a community is ever at equilibrium. However, for the purpose of this study, we define equilibrium as a state in which all the macroscopic characteristics of the community fluctuate around a constant average value with a fixed variance. Since we are interested in species extinctions, the macroscopic property of interest is the species richness. Hence, we say that the community is at equilibrium when its species richness is stabilized around a constant average value. The above definition of equilibrium is a loosely defined stationarity property. Strict stationarity, as defined in stochastic processes analysis, requires that the joint probability distribution of a given process is constant over time (Lindgren et al. 2013, section 2.3). As we will see in section 2.3, within the neutral model of biodiversity, which we use as a model of ecological communities, stationarity and thus the equilibrium is strictly defined.

Box 2.1 Glossary

Habitat loss: the loss of an area that was previously habitat to an ecological community, in a way that this area can no longer support the community as before. If the lost area is rendered completely inhospitable for the community in question, then habitat loss is complete.

Initial habitat: the habitat area of interest before habitat loss occurs.

Destroyed habitat: the part of the initial habitat that is destroyed by habitat loss.

Remaining habitat: the part of the initial habitat that is not destroyed by habitat loss.

Equilibrium: the state of an ecological community in which all its macroscopic properties are stabilized around a constant average value. In the case of extinction, the property of interest is species richness, which is stabilized when the rate of species extinctions is balanced by the rate of species origination through speciation and/or immigration. This kind of equilibrium can also be referred to as a stationarity state, a steady state or as dynamical equilibrium.

Relaxation process: the process by which an ecological community that has been disturbed returns to equilibrium.

Imminent extinctions: species extinctions (local or global) that are a direct result of habitat loss or other disturbances (e.g. spatially correlated disturbance on endemic species with a narrow geographic range (Kallimanis et al. 2005)).

Delayed extinctions: species extinctions that happen during the relaxation process. These could be considered as indirect extinctions due to habitat loss.

Extinction debt: the number of species committed to extinction at some time, t, of the relaxation process. According to this definition, extinction debt equals the delayed extinctions minus the delayed extinctions already realized by time t.

Total extinctions: the sum of imminent and delayed extinctions.

Temporal turnover: the change of a community's composition with time. This includes the variation of species abundances, but also species extinctions and introduction of new species.

2.1.1 Two phases of species loss: imminent and delayed extinctions

The decline of species richness following habitat loss happens in two phases. In the first phase species whose range is restricted to the lost habitat become immediately extinct. Such extinctions, being the direct result of habitat loss, are called **imminent** or **endemic** extinctions (Jackson and Sax 2010, Halley et al. 2014). After habitat loss is complete, the community that is left in the reduced habitat is in general out of equilibrium, as many species may have reduced populations and the loss of habitat is usually accompanied by changes in habitat structure (e.g. increased isolation). As a result, the number of species of the remaining habitat

soon after habitat loss is usually greater than that the affected area can support. The second phase regards the gradual relaxation of the community to a new equilibrium during which more species are expected to go extinct. These are called **delayed** extinctions. The number of species that are bound to go extinct at a given time after habitat loss is called **extinction debt** (Kuussaari et al. 2009, Jackson and Sax 2010, Halley et al. 2014). Figure 2.2 shows schematically the two phases of extinctions and the corresponding reduction of species richness. The total number of extinctions (S_{tot}) that result from habitat loss is the sum of imminent and delayed extinctions. In practice, total extinctions can be estimated using Species-Area Relationships (SARs) as the difference in species richness between the initial habitat and the remaining habitat at equilibrium (SAR method) (Halley et al. 2013). If the initial habitat area is A and the remaining habitat area is a we can write,

$$S_{tot} = S_A - S_a, \tag{2.1}$$

where S_A is the number of species of the initial habitat and S_a the number of species of the remaining habitat at equilibrium, both predicted from the appropriate SARs. Note that, due to changes induced by habitat loss (e.g. isolation of the remaining habitat), the SAR describing the initial and remaining habitats might be different. The total number of extinctions also equals the sum of delayed and imminent extinctions, hence we can write,

$$S_{tot} = S_A - S_a = E_{A-a} + D_a, \tag{2.2}$$

where E_{A-a} are the imminent extinctions in the destroyed habitat and D_a the delayed extinctions in the remaining habitat. Taking into account the time dependence (Halley et al. 2014) equation (2.2) reads,

$$S_{tot}(t) = S_A - S_a(t) = E_{A-a} + D_a(t). \tag{2.3}$$

In equation (2.3), $S_a(0)=S_0$ is the initial number of species of the remaining habitat soon after habitat loss, $D_a(t)$ represents the number of delayed extinctions already realized by time t and $D_a(0)=0$ are the delayed extinctions at time $t = 0$ (Figure 2.2).
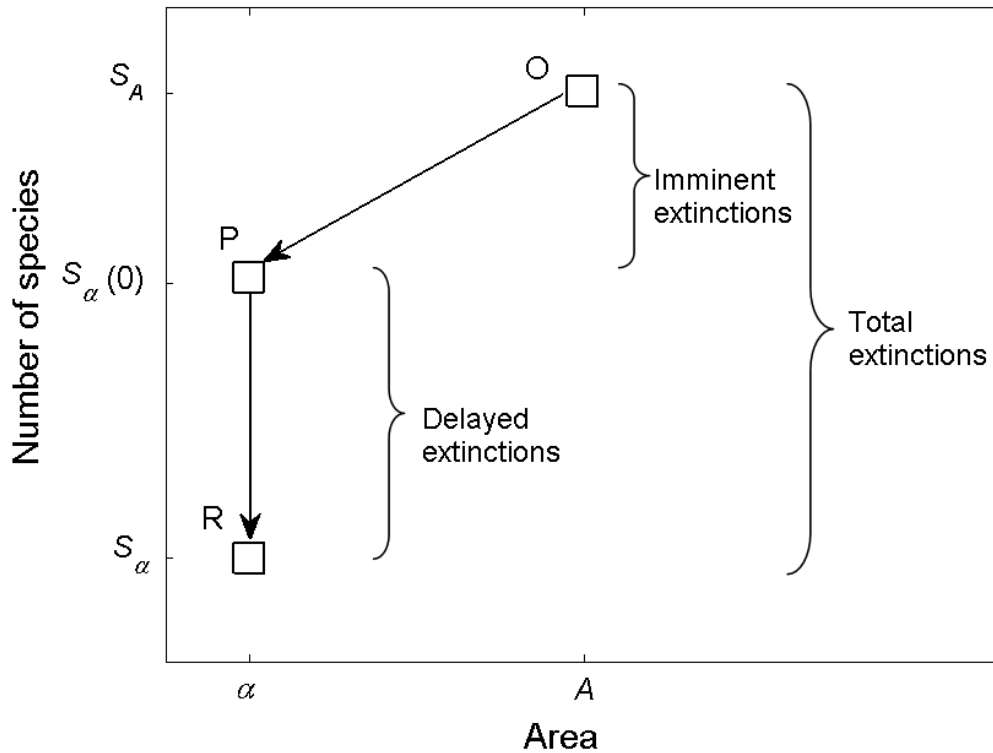
Figure 2.2 The phases of extinction following habitat loss. Initially the community of S_A species is in equilibrium (point O) in a habitat of area A. When habitat is lost, some extinctions happen immediately because there are species found entirely in the lost area (OP trajectory). More extinctions follow later due to increased isolation and smaller populations in the remaining area a. Eventually the community arrives at the new equilibrium (R) where local extinction and colonization are in balance (c.f. Rosenzweig, 2001). It should be noted that in particular cases (high aggregation, uneven species-abundance distribution or high immigration), imminent extinctions can be so many that the species left in the remaining area after habitat loss are fewer than the area can support. In such a case, the species richness will increase to equilibrium (immigration credit) (Jackson and Sax 2010). (Figure taken from Halley et al. 2014).

The magnitude of imminent extinctions can vary based on the degree of aggregation, the dominance pattern of the community and the size of the destroyed habitat (Green and Ostling 2003, Halley et al. 2014, Kitzes and Harte 2015). If we focus on the degree of aggregation, we can describe two extreme scenarios shown in Figure 2.3 This shows two

communities both having four species but with very different spatial distributions; in the first, the distributions of species are not overlapping, while in the second the distributions of species are overlapping. In the case of extreme aggregation, where the species spatial distributions are not overlapping, a possible area loss can deplete the populations of two of the species but leave the other two unaffected (Figure 2.3, left). In this case, imminent extinctions deplete half of the species (50% of initial species richness). However, the rest of the species stay unaffected, so there are no delayed extinctions. In the other extreme, where the spatial distributions of species are completely overlapping, there will be no imminent extinctions and the habitat reduction will equally affect all species populations (Figure 2.3, right). However, all four species are now forced to live in half the initial area, and in time it is expected that some of these species will go extinct. If the remaining area can only support two species at equilibrium (based on a species-area relationship), then two of the species are expected to go extinct. In this case, imminent extinctions are zero, but delayed extinctions correspond to 50% of the initial species richness. At equilibrium, both habitats (Figure 2.3, left and right) are left with two species.

Apart from the spatial configuration, the size of imminent extinctions also depends on the Species-Abundance Distribution (SAD) of the given community (see Box 2.2). As explained in Green and Ostling (2003), imminent extinctions are more for more uneven SADs. A SAD is even when the individuals are evenly distributed over species. An extreme case of an even distribution is one where all species have the same abundance. In the other extreme, the most uneven distribution is one where there is one species with many individuals and the rest of the species have one individual each. For an uneven distribution, the depletion of a habitat area is more likely to completely deplete a species, since many species have very few individuals. For example, if the spatial distribution of individuals is random, the probability that 1 individual of a given species is found in the destroyed area x is x/A. If this species has n individuals then the probability that all of them are found in x is $(x/A)^n$. Hence, the fewer individuals a species has, the more likely it is that these are found within area x, which makes it more likely to go directly extinct in an event of habitat loss.
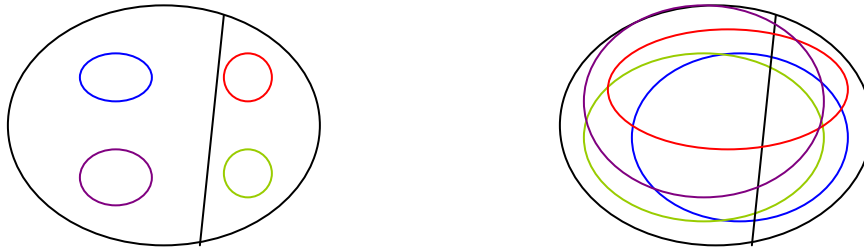
Figure 2.3 Two extreme scenarios of spatial distribution of species and the effects of habitat loss. Species spatial distributions are defined by circles of different colors. On the left, the species distributions are non-overlapping and the loss of the area results in the extinctions of two of the species but leaves the other two unaffected. In this case imminent extinctions correspond to 50% of the initial species richness and delayed extinctions correspond to 0% since the remaining habitat can support the two remaining species as before. On the right, the species spatial distributions are almost completely overlapping. The loss of area affects all species equally but does not immediately lead to their extinction. However, the four species are now restricted in a smaller habitat and this can lead to the extinction of some of them. In this case imminent extinctions are 0%. If the remaining area can only support two species at equilibrium (based on species-area relationship) then delayed extinctions comprise 50% of the initial species richness. In the end both habitats (left and right) will end up having two species.

Box 2.2 Species-abundance and rank-abundance distributions

In a sample of abundance data taken from an ecological community, different species are expected to be represented by different numbers of individuals. The distribution of individuals to species can be summarized by a vector, $N=\{n_1, n_2,..., n_S\}$, where n_i represents the abundance of the ith species and S is the total number of species. The species-abundance distribution (SAD) and the rank-abundance distribution (RAD) are two different ways of representing the distribution of individuals to species. There are several theoretical models for this distribution; four of them are shown in Figure I using (a) the RAD and (b) the SAD representation.

RAD: in the RAD representation, species are ranked in descending order according to their abundance. This gives a vector $rad=\{n_1, n_2,..., n_S\}$, in which n_1 is the abundance of the most abundant species, n_2 the abundance of the second most abundant species and so on. The RAD can be represented by a graph, in which the abundances of all species are plotted against their rank (rank-abundance plot) (Figure Ia). The RAD curve is flatter if the individuals are more evenly distributed to species. The extreme case is the even distribution in which all species have the same abundance.

SAD: in the SAD representation, the species are first grouped according to their abundance and then the number of species in each abundance class is counted. The SAD can be represented by a vector $sad=\{S_1, S_2,..., S_J\}$, in which S_1 is the number of species that have abundance 1, S_2 the number of species with abundance 2 and so on (Figure Ib). Note that some authors use the term species-abundance distribution (SAD) to refer to both RAD and SAD. Here we keep the distinction between the two terms as defined above.
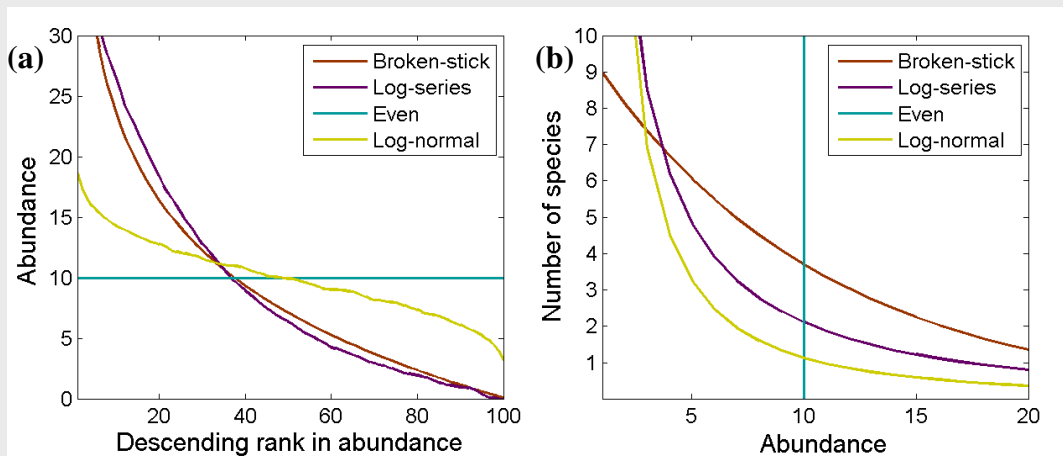
Figure I. Four theoretical models used to describe the distribution of individuals to species (log-series, broken-stick, log-normal and even distribution) represented in two different ways: (a) using the rank-abundance distribution (RAD) and (b) using the species-abundance distribution (SAD).

2.1.2 Estimation of imminent and delayed extinctions

The size of imminent extinctions depends on the degree of aggregation and on the species-abundance distribution (SAD) of the community. Green and Ostling (2003) provide formulas for the estimation of imminent extinctions for random placement of individuals across the habitat and for aggregated configurations described by the negative binomial distribution. Consider that N_i, $i=1-S_A$ represents the abundance of species i in a community. If habitat loss results in the removal of area A-a, then the expected number of imminent extinction in the case of random placement is given by,

$$E_{A-a} = \sum_{i=1}^{S(A)} \left(1 - \frac{a}{A}\right)^{N_i} , \tag{2.4}$$

where a is the remaining habitat area. For a spatial distribution following the negative binomial distribution, the expected number of imminent extinctions is,

$$E_{A-a} = \left(1 - \frac{a}{A}\right) \sum_{i=1}^{S(A)} \left[1 + \frac{N_i a}{A k_i}\right]^{k_i} , \tag{2.5}$$

where k_i is a parameter controlling the aggregation, which can vary between ($-\infty$, $-N_i\,a/A$) and ($0,+ \infty$). For positive k, the distribution of species is aggregated with smaller values corresponding to greater aggregation. When k is negative the distribution of species is regular (He and Gaston 2000). An important property of the negative binomial distribution to keep in mind is that k increases proportionally with scale (see Box 2.3).

Box 2.3 The story of the negative binomial distribution and imminent extinctions

The negative binomial distribution (NBD) is frequently used to model the spatial distribution of individuals of organisms (Eberhardt 1967, Perry and Taylor 1985, He and Gaston 2000, He and Legendre 2002, Green and Ostling 2003, He and Gaston 2007, Conlisk et al. 2007). In 2011 it was used by He and Hubbell, in a very cited and controversial paper, which showed that Species-Area Relationships (SARs) always overestimate extinctions from habitat loss due to an unnoticed difference between the regular SAR and the endemics SAR (He and Hubbell 2011). In the core of it, this article defied the laws of probability, stating that the probability of finding a species entirely in the lost area (*A-a*) and the probability of not finding it in the remaining area (*a*) are not equal (given that the species exists in the initial area *A*). After a year of confusion and strong opinions for and against the article, the solution to the mystery was given by Axelsen et al. (2013), who discovered that He and Hubbell forgot to scale the aggregation parameter (*k*) of the negative binomial distribution! As He and Gaston (2007) explain, "the assumption of a constant NBD *k* across scales is a simple violation of a theoretical premise of the NBD. It is well established in statistics that *k* increases proportionally with scale; that is, if *x*1 and *x*2 are from an NBD with aggregation parameter *k*, then *y* = *x*1 + *x*2 follows a NBD with 2*k*.".

If the imminent extinctions can be estimated using the above formulas, then the delayed extinctions can be estimated from equation (2.2) as the difference between total and imminent extinctions. However, there is no way to estimate the delayed extinctions at any given time after habitat loss without considering the dynamics of the relaxation process. In recent years, the development of stochastic models of community assembly has given the opportunity to model the dynamics of ecological communities. In particular, the neutral model of biodiversity has proven very successful in predicting both the equilibrium and dynamical features of communities.

For a community whose initial SAD is close to broken-stick there is a solution to the neutral model (Halley and Iwasa 2011) applicable for systems without immigration or speciation. According to this solution, the number of delayed extinctions at time *t* is,

$$D_a(t) = S_0 - \frac{S_0}{1 + t/t_{50}},$$
(2.6)

with $t_{50}=\tau J/S_0$ the time to lose half of the species, J the community size, τ one generation time and $S_0 = S_A\text{-}E_a$ the number of species in the remaining habitat immediately after habitat loss. Equation (2.6) was used by Halley and Iwasa (2011) to derive the rate of biodiversity decline on isolated islands. A time-dependent neutral formula was also derived by Gilbert et al. (2006) assuming a log-series SAD, which is applicable at short timescales ($t << J$),

$$D_a(t) = \alpha \ln \frac{J + \alpha}{J/(t+1) + \alpha}, \tag{2.7}$$

with J the community size and α Fisher's alpha-diversity.

There have been other studies of the dynamical aspects of the neutral model and analytical results regarding species turnover distributions and the time evolution of probability distributions (McKane et al. 2000, Azaele et al. 2006, Chisholm 2011). Nevertheless, these authors don't report any closed-form expressions for the variation of species richness with time and this area has been left unexplored. In Chapter 5 we derive such an expression for a neutral community that is subject to speciation, which expands the result of Halley and Iwasa (2011) to communities where there is recruitment of new species. Still, even without explicit formulas, the dynamics of relaxation under the neutral model can be efficiently explored using simulations.

2.2 SAR method for extinctions

2.2.1 Introduction

Species-Area Relationships (SARs) relate the number of species found in a habitat with the habitat's area. Many functions have been used to describe this relationship, but the most commonly used one is the Arrhenius curve,

$$S=cA^z. \tag{2.8}$$

In equation (2.8), S is the number of species, A is the area of the habitat, c is a constant representing the number of species per unit area (i.e. the alpha diversity) and the exponent z is a constant that controls how quickly the species richness increases with area. SARs are commonly used to predict extinctions following habitat loss (Wilson 1988, Pimm and Askins 1995, Brooks et al. 1997, Triantis et al. 2010). Let us consider again the scenario of habitat loss depicted in Figure 2.1. An initial habitat of area A suffers habitat loss as a result of which its area reduces to a. The number of extinctions can be defined as the difference between the species richness of the initial and the remaining habitat at equilibrium, namely $\Delta S = S_A - S_a$, which according to equation (2.8) is,

$$\Delta S = S_A - S_a = cA^z - ca^z. \tag{2.9}$$

The reduction of species richness with area is shown in Figure 2.4. Graphically, the number of extinctions can be found by tracing the SAR curve backwards until the remaining area is reached. For this reason, this method of estimating extinctions has been called the backward SAR method. Note that the SAR method as defined by equation (2.9) and Figure 2.4 implies that the SAR describing the initial and the remaining habitat is the same. However, this is not always the case. As we will see in Chapter 3, the SAR of the remaining habitat may have a different slope, z, as a result of a change in the habitat structure (e.g. increased isolation due to fragmentation). From now on, when referring to the SAR method we refer to the general method in which the species richness of the initial and the remaining areas can be estimated independently, namely using different SARs.
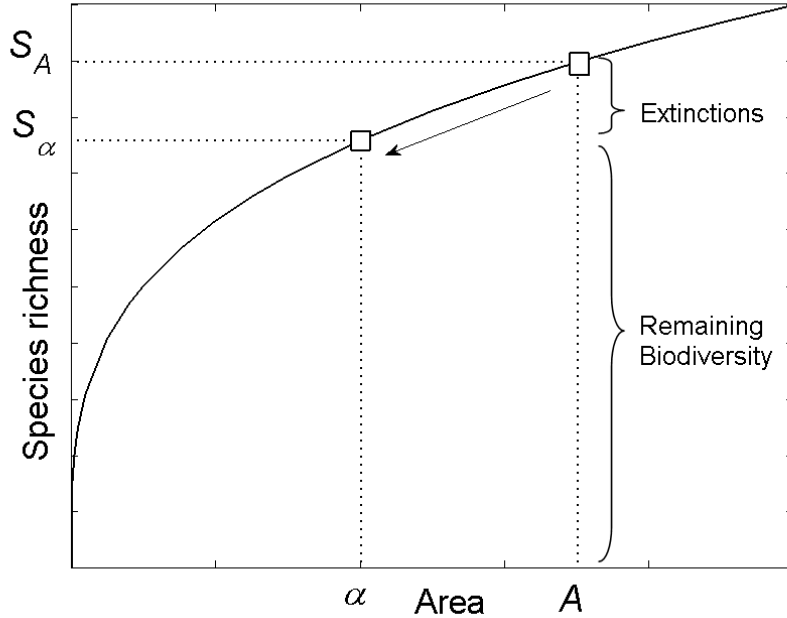
Figure 2.4 The backward SAR method for estimating extinctions following habitat loss. If a habitat's area is reduced from A to a, the species richness falls from S_A to S_a.

It should be emphasized that the SAR method described above predicts the total number of extinctions (imminent plus delayed extinctions). However, this fact is often misjudged. He and Hubbell (2011) claimed that equation (2.9) gives the number of imminent extinctions, in the lost habitat, A-a. According to this claim, the use of equation (2.9) to estimate total extinctions is flawed, while its failure to predict total extinctions is erroneously attributed to extinction debt. We note, however, that He and Hubbell interpret the species richness of the remaining habitat, S_a, in a different way than it is meant by equation (2.9). If S_a represents the number of species found in area a at the time of habitat loss, then surely $\Delta S = S_A - S_a$ represents the imminent species extinctions in the lost habitat area, A-a. However, if this is the case, this habitat should be considered as a sample from the larger habitat, A, in which case its species richness is not given by the Arrhenius curve of equation (2.9), but by the appropriate sample SAR for this habitat. We want to make clear however that S_a in equation (2.9) refers to the species richness of the remaining habitat long after this has become isolated, namely it refers to the species richness of the habitat at equilibrium. This is no more a sample from the initial area, and an island type SAR should be used to estimate its species

richness. As we explain in the next section, ignoring the above fact, namely that the SAR method predicts total extinctions, can lead to large errors in the SAR extinction forecasts.

2.2.2 Uncertainty of SAR extinctions forecasts

Given the species richness of the initial habitat, the size of the remaining area and the parameters of the Arrhenius curve, the backward SAR estimation is straightforward. However, as the estimation of the parameters, but also the choice of the correct functional form for the SAR, is not obvious, there is a great uncertainty to this prediction. The method itself has been criticized for the assumptions it bears (Connor and McCoy 2001). Nevertheless, as explained in Halley et al. (2013), the mismatch between SAR predictions and observation is in many cases due to the omission of secondary phenomena (like habitat regeneration or extinction debt). In studies where these phenomena were taken into account, SAR estimates held up rather well (Halley et al. 2013). Another objection to the SAR method came from He and Hubbell (2011), who claimed to prove that the method was fundamentally flawed, but as it proved later there was a mistake in their derivation (see Box 2.3).

Halley et al. (2013) reviewed the possible sources of error arising when the SAR method is used to forecast extinctions. These errors can be grouped into four categories. The first type of error and one that cannot be avoided relates to the natural scatter of the SAR. The second type is related to the choice of the SAR functional form and the choice of parameters. The third type regards the erroneous interpretation of the SAR predictions. The fourth type of error has to do with secondary phenomena that the SAR method ignores, like partial habitat conversion instead of complete loss and habitat regeneration. Below we present a summary of the more extended review of Halley et al. (2013).

Natural scatter

When measurements of species richness are plotted against area, it is unlikely that the points will fall on a monotonically increasing curve; in other words the SAR is an approximate law. Hence, an unavoidable source of uncertainty of the SAR method is due to the natural scatter of

SARs. Triantis et al. (2012) found that when fitting Arrhenius models, the coefficient of determination (goodness of fit on a scale 0–1) was on average R^2=0.64 (±0.20 standard deviation). As reported in Halley et al. (2013) a R^2 =0.68 is associated with a root mean square deviation of 0.32, which means that the prediction can be 2.07 times larger or smaller than it should be (based on one-sigma level).

Different types of SAR, choice of functional form and parameters

According to Preston (1962), SARs can be categorized into two types depending on the habitat units that are used to construct them, namely '**sample areas**' or '**isolates**' (Scheiner 2003, Tjørve and Turner 2009, Triantis et al. 2012). Isolates are self-contained geographical units, such as islands, mountain tops, lakes and habitat islands surrounded by inhospitable matrix. Sample areas are fractions of isolates or of large continental areas, e.g. a plot within a forest (Halley et al. 2013). The above distinction leads to the two types of SARs, namely the isolate SAR (ISAR) and the sample or continental SAR (CSAR). An example of each one can be seen in Figure 2.5. The ISAR is steeper and lies lower than the CSAR. If an Arrhenius curve is fitted to species-area data, typical values of the *z* exponent lie between 0.25 and 0.35 for ISARs and less than 0.15 for the CSAR. Despite the variety of values of the *z* exponent, only a limited range has been used in SAR forecasts (Halley et al. 2013).
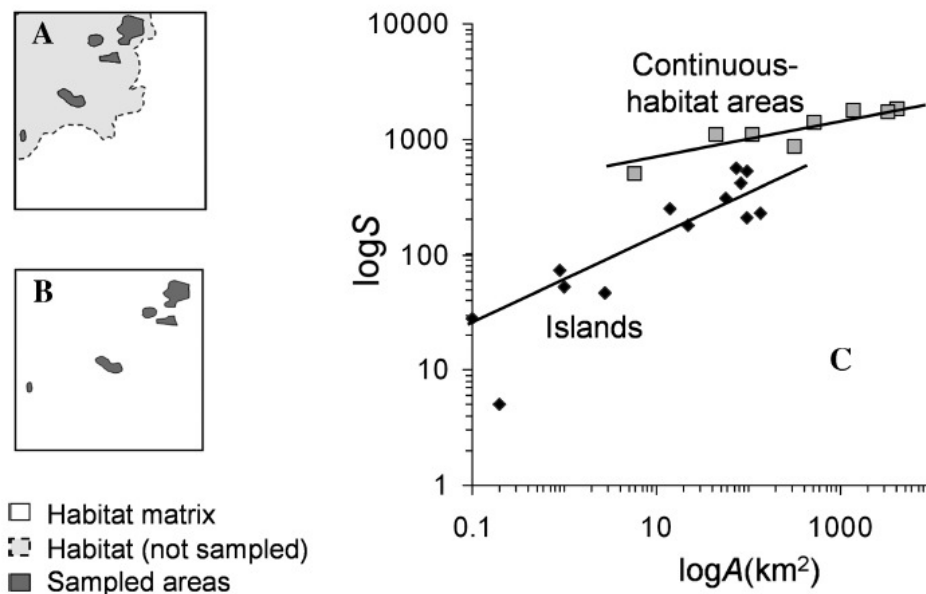
Figure 2.5 Habitat patches as sample areas and as isolates. (A) A number of sample plots in a continuous habitat, e.g. non-nested plots of various sizes within a tropical rain forest. (B) A set of true geographical islands or isolates surrounded by inhospitable matrix. Note that the sizes of sample areas and isolates in A and B are the same, so A indicates the architecture and initial state of what remains after habitat loss, while B represents the final state. (C) Data for sample areas and for isolates for Californian plants (Data from Johnson et al. 1968). Species–area relationships for isolates tend to be lower and steeper; the slopes are 0.37 (for islands) and 0.16 (for continental areas). (Figure taken from Halley et al. 2013).

Depending on the scale but also on the taxonomic group sampled, the scatter plot of species richness versus area may appear to follow a linear, convex, concave or sigmoidal function (Connor and McCoy 2001). With respect to scale, it is often observed that the SAR follows a triphasic pattern, which is described by a rapid increase of species richness at local scales, a lower slope at intermediate scales, often described by an Arrhenius curve, and an accelerating increase at continental scales (O'Dwyer and Green 2010, He and Legendre 1996). Triantis et al. (2012) lists 20 different functions that have been used to model SARs. Most of them can be categorized in three basic families, each associated with a different ecological theory. The first family includes variants of the Arrhenius curve. Preston (1962) showed that in a community following a canonical lognormal species-abundance distribution, the SAR has an Arrhenius form with exponent approximately $z = 0.26$. The second family of SARs is of a log-linear form, such as the Gleason and Kobayashi equations, which often appear in the

neutral theory of biodiversity (Hubbell 2001). Finally, the third family includes asymptotic forms, namely SARs in which there is an upper bound on species richness. The three families comply with the following general functional forms,

$$S = \begin{cases} cA^z \\ \log(cA^z) \\ S_{max} \dfrac{cA^z}{cA^z + 1} \end{cases} .$$

(2.10)

In the last function, S_{max} is the upper bound for the species richness. Halley et al. 2013 fitted the three SAR models to the data of Figure 2.5 (Figure 2.6). As it is shown in Figure 2.6, the three models can fit well to the observed data, but when extrapolated outside the range of values for which the SAR was fitted, the predictions can differ by several orders of magnitude (Figure 2.6 B).
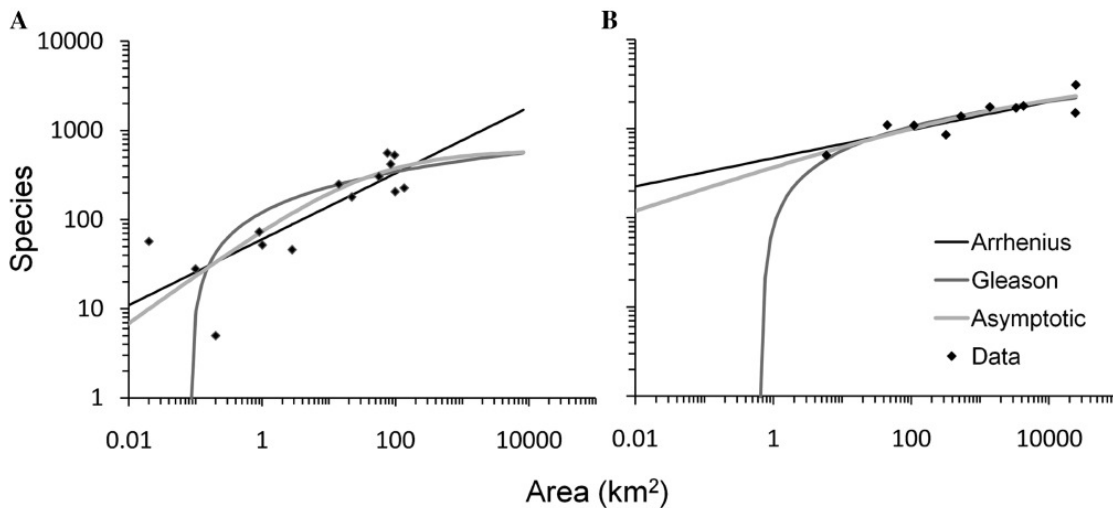
Figure 2.6 Fitting Arrhenius, Gleason, and asymptotic forms to predict extinctions for (A) islands and (B) sample areas, for the data appearing in Figure 2.5. In both cases, the Arrhenius model was fitted by the least-squares method to the line $\log S = z\log A + b$, the Gleason was fit as $S = z\log A + b$, and the asymptotic was fit using $\log[S/(S_0-S)] = z\log A + b$, having first assumed that $S_0 = 4000$ is the limiting value. (Figure taken from Halley et al. 2013).

Extinction debt

The effects of habitat loss are rarely limited to the loss of endemic species in the cleared habitat. Most likely, habitat loss 'kicks' the remaining habitat out of equilibrium. Depending on the pattern of habitat loss, the remaining community might have an excess or a deficit of species (i.e. more or fewer species than it can support). Even more, the community will have to adjust to a likely change of condition (e.g. isolation). In all of the above cases, the relaxation of the remaining habitat to the new equilibrium takes a considerable time. A well known example of this phenomenon are the New Guinea islands that were isolated at the end of the Pleistocene, approximately 12,000 years ago, due to the rising sea levels. Diamond (1972) constructed an ISAR for these islands and noticed that there is a group of big islands which contain much more species than the ISAR predicts. By contrast, the species richness of smaller islands of the same region agrees with the ISAR predictions. Based on the theory of island biogeography, he found that the relaxation time for the big islands is around 16,700 years, which explains their higher biodiversity, as these have not yet reached equilibrium. Hence, the fact that SAR estimates do not always agree with observations does not mean that the SAR forecast is flawed. This shows that if extinction debt is ignored, the SAR method is liable to overestimate the extinctions observed before the relaxation process is complete. But how long does the relaxation process take to complete? Halley and Iwasa (2011) used data of bird extinctions and found that the half-life time to relaxation inferred from the data scales with area as $T_{50} = 4.35A^{0.652}$. As we show in Chapter 5, the neutral model with random-fission speciation predicts a scaling exponent of 0.5 for the same data set, while the isolated-fragment model of Halley and Iwasa (2011) predicts an exponent of 0.8.

Incomplete habitat loss and habitat regeneration

The SAR method assumes that habitat loss is complete, so that the species under consideration cannot survive in the destroyed or converted habitat. This is true in the case the lost area is replaced by inhospitable matrix. In cases of habitat conversion like deforestation, there is always the possibility that biodiversity can be retained outside these boundaries, in the converted habitat. This effect can be taken into account if one measures habitat loss in terms

of the net primary productivity (NPP) of the converted habitat. For example, if a former forest habitat of area *A* is transformed to agricultural land, humans use part of the NPP e.g. 58%, while the remaining 42% is left for other organisms (e.g. birds could eat wheat or vegetables, so that the loss is not complete for these species). In such a case, habitat loss corresponds to 0.58*A* (Halley et al. 2013). There are studies where the effects of a partial conversion of habitat have been taken into account (Koh et al. 2010, Tjørve 2009, Pereira et al. 2012). Koh et al. 2010 compared the predictions of these modified methods with the conventional SAR method, in habitats where the total extinctions were known. They showed that modified SAR methods had a root mean squared error of 35%, while the use of a conventional method had an error of 51% when using a CSAR or 98% when using an ISAR. Another reason for a possible error of the SAR calculation is habitat regeneration. If habitat regeneration happens before the relaxation is complete, then there is a chance of full recovery of the habitat to its previous state.

2.2.3 Summary and conclusions

Table 2.1 summarizes the possible sources of error that can arise in SAR extinction forecasts and the associated maximum percentage error (corresponding to the logarithm of species richness). The natural scatter of the SAR alone can lead to an error of 100%. As the natural scatter is unavoidable, errors of less than 100% due to other sources cannot be considered significant. Large errors can arise from a bad choice of the SAR functional form or a bad choice of parameters, but also from ignoring extinction debt. On the one hand, if the SAR estimate (which refers to total extinctions) is compared with observations of species richness soon after habitat loss, then the SAR method will apparently largely overestimate extinctions; however this is a result of ignoring the fact that the relaxation process is not yet complete. On the other hand, in many cases of habitat loss, the remaining habitat becomes isolated and its equilibrium species richness is more appropriately described by an island SAR (i.e. an SAR with a larger *z* exponent). If this fact is ignored and the same type of SAR is used to describe both the initial and the remaining habitat, then the total extinctions will be apparently

underestimated. This comes from ignoring the extinction debt caused by the habitat's isolation.

Table 2.1 Possible sources of error arising when using the SAR method to forecast extinction and the corresponding maximum percentage error on the logarithm of species richness. + sign denotes overestimation of extinctions, - sign denotes underestimation, ± either underestimation or overestimation and *** denotes a very large number.

Description of limitation or problem	Maximum Error
1 SAR natural scatter	±100%
2 Error in exponent of Arrhenius SAR	±200%
3 Wrong functional form (prediction inside range of fitted data)	±200%
4 Wrong functional form (prediction outside range of fitted data)	±***
5 Habitat loss not complete. Diversity retained in the "matrix"	+16 to 63%
6 Regeneration of habitat	±***
7 Extinction debt (prediction timescale << relaxation time)	+***
8 Extinction debt (prediction timescale >> relaxation time)	-200%

Based on the above review we arrive at three main conclusions. Firstly, there is no fundamental flaw in the SAR method. The concerns about the validity of the method are actually related to its misuse or the misinterpretation of its predictions (i.e. ignoring extinction debt, wrong use of SAR functional form, etc.). Whether the requirements for the appropriate use of the SAR method can be met in reality is a different question. Secondly, the SAR method predicts extinctions based only on area, which means that other factors affecting extinctions are ignored (e.g. habitat regeneration, incomplete habitat loss). In some cases, the estimate can be corrected by taking these other factors into account. Finally, the SAR method has a great uncertainty that limits its prediction power. Nevertheless there is, currently, no alternative method that gives more accurate predictions. As noted by Pereira et al. (2010), in current biodiversity assessments using multiple methods "the range of uncertainty across models and scenarios is close to three orders of magnitude." In this light, most of the uncertainties in Table 1 are negligible and do not undermine the credibility of the SAR method (Halley et al. 2013).

2.3 Hubbell's neutral theory of biodiversity

The neutral theory of biodiversity was introduced in ecology by S. Hubbell and popularized with his book "The Unified Neutral Theory of Biodiversity and Biogeography" in 2001 (Hubbell 2001). The development of the theory and the corresponding model were based on the neutral theory of molecular evolution, which already existed in population genetics since 1964 (Kimura and Crow 1964). The aim of the theory is to give a dispersal-assembly explanation to the organization of ecological communities, namely to explain community assembly as the result of the combined effects of random demography and dispersal. It should be emphasized, however, that its application is limited to the diversity within a given trophic level. For the purposes of neutral theory, Hubbell defines an ecological community as "a group of trophically similar, sympatric species that compete in a local area for the same or similar resources, as for example might be trees in a forest".

In adapting the model to ecology, Hubbell was largely inspired by the Island Biogeography Theory (ITB) (MacArthur and Wilson 1967). In the IBT, the equilibrium species richness of islands results from a balance between species extinctions and species introductions through immigration from a regional community. The neutral theory of biodiversity (NBT) is based on the same principle but differs from the IBT in a major aspect: in IBT, the neutral units are species having the same extinction and immigration rates and thus the same abundance, while in NBT the neutral units are individuals all having the same demographic rates and colonization ability. The species are formed as sets of individuals, whose abundances at a given time reflects their different times of "origination" through either dispersal or speciation and their subsequent evolution through stochastic demographic events. Due to this distinction, species have different abundances and lifespans, which come as predictions of the theory rather than being assumed (Hubbell 2001). Because of modeling at the individual level, the NTB can predict macroscopic features of communities, like species-abundance distributions and species-area relationships (McKane et al. 2000, Volkov et al. 2003, Vallande and Houchmandzadeh 2003, McKane et al. 2004, Etienne and Alonso 2007, Rosindell and Cornell 2007, O'Dwyer and Green 2010). Furthermore, the neutral model is a dynamical model and can be used to predict the time evolution of communities (McKane

et al. 2000, McKane et al. 2004, Gilbert et al. 2006, Azaele et al. 2006, Chisholm 2011, Halley and Iwasa 2011).

2.3.1 The theory

Neutral theory makes the following simplifying view of an ecological community. A community is a collection of individuals, belonging to different species that coexist in the same area. Each individual occupies a space associated to resources (i.e. a site). As resources are limited, so are the available sites, which sets a constraint on the total number of individual that can exist in the community. If furthermore the area is saturated with individuals at all times (i.e. there is no empty site), then a new individual can arise only if a site becomes available following the death of another individual. This is the so called **zero-sum rule**. The **neutrality** of the theory lies in the fact that all individuals independent of species are equal competitors, namely have equal chances of reproducing, dying, mutating and dispersing (i.e. colonizing an empty site). As Etienne and Alonso (2007) have noted, neutrality does not require that the species are functionally equivalent, but that they have the same demographic rates and colonization ability, which can happen even if they follow completely different strategies.
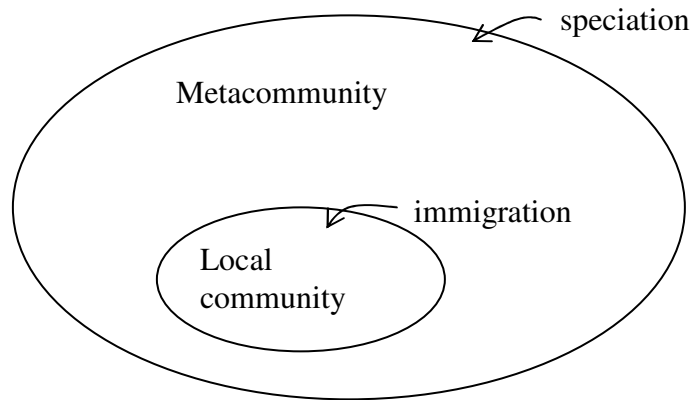
Figure 2.7 The implicit-space neutral theory of biodiversity assumes two spatial scales; the metacommunity and the local community. The metacommunity's biodiversity is sustained through a balance between extinctions and introduction of new species through speciation. In the local community, the species richness is sustained by a balance between local extinctions and immigration of individual from the metacommunity. The local community can be thought of as being part of the metacommunity but it can also be a geographically separate habitat (i.e. an island). In both cases, the metacommunity is much bigger than the local community, so that its composition is not affected by the local community dynamics.

The basic form of the model is called the **implicit-space** neutral model. This assumes two spatial scales; the metacommunity and the local community (Figure 2.7). The metacommunity, which represents a regional species pool, evolves in time through demographic events (births and deaths of individuals) and speciation. At the steady state, the extinction of species is balanced by the introduction of new species through speciation. The local community represents a sample from the metacommunity. The local scale is introduced in order to model the effects of dispersal limitation on community composition. Dispersal limitation is the restriction of individuals to move and colonize places far from their origination. Due to this limitation, the compositional similarity of samples decreases with distance. Furthermore, the composition of a sample does not represent all the species found at a regional level. For example, if a sample is taken from a locality within a forest, the trees that are found in the sample will mostly represent trees that exist in the neighborhood of the sample and not all the trees found in the forest. In Hubbell's local community model, dispersal limitation is modeled as a restriction on the ability of species from the metacommunity to enter the local community. In particular, in every death event in the local community, local

species compete with the metacommunity species for the empty site, with local species having a greater chance of colonizing the site. At the steady state, the local species richness results from a balance between local extinctions and introduction of new species through immigration.

Although the metacommunity species cannot freely enter the local community, the metacommunity itself has no spatial structure, so that its individuals can be found anywhere within its area (except the local community). This means that in an immigration event, all the metacommunity individuals have the same chance of colonizing an empty site in the local community. Finally, the metacommunity is considered much bigger than the local community, so that its composition is not affected by the local community dynamics.

2.3.2 Metacommunity dynamics
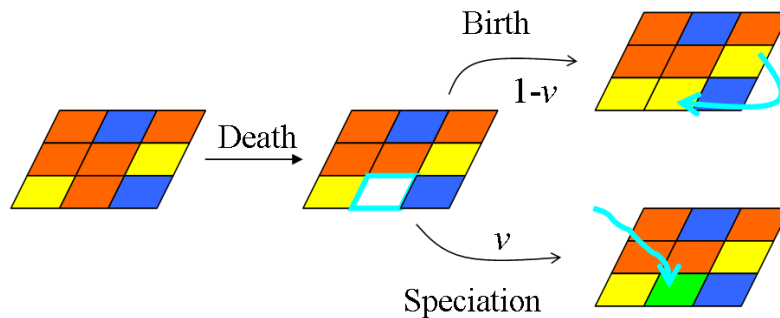
Figure 2.8 One step in a neutral metacommunity with point mutation speciation. The community is represented by a grid in which each site is occupied by an individual of a given species (color). At each step, a random individual is killed (here an orange) and a site becomes available (white site). At the same step, another random individual is picked (here a yellow) to give the descendant that will occupy the empty site. However, with probability v, this descendant suffers a point mutation that converts it to a different species (here a green). If these rules are applied many times, it is proved mathematically that the community reaches a steady state, namely a state where all its statistical measures fluctuate around a fixed value.

Hubbell described the evolution of the metacommunity as a discrete time stochastic process (Figure 2.8). The process proceeds according to the following rules: at each step a random individual is selected to die (death). The death leaves an empty site, which in the same step is replaced by a descendant of another random individual from the metacommunity (birth).

However, with a small probability v, the individual that will occupy the site mutates to become a single member of a new species (point mutation speciation). The above rules define an irreducible Markov chain, where the state of the community at each step can be determined by its state at the previous step. It can be shown that after enough time, the process reaches a steady state characterized by a unique stationary distribution (Haegeman and Etienne 2010, Van Kampen 1992).

The maximum number of species that can exist in the metacommunity at any given time is constrained by the total number of individuals, J_M. In an extreme case, the metacommunity can have up to J_M species with one individual each. However, because speciation constantly creates new species, the actual number of species that have lived in the community can be much bigger than J_M. This means that if one wants to describe the state space of the metacommunity as a vector with the abundances of all species, this should be infinite, as given infinite time there is an infinite number of species that can be created. We note, however, that species identities do not matter, as ultimately one seeks to find the species-abundance distribution of the community. To slide over the problem of infinite state space, we can keep the state space fixed by assuming that there are at most J_M species, which can reappear by speciation after they have gone extinct. In this case, the state of the metacommunity is described by a vector $\mathbf{n_M}=(n_1, n_2,\ldots, n_{JM})$, where n_i is the abundance of species i. The metacommunity can be found in all the states that comply with the community size constraint, namely the total number of individuals in the community is fixed to $n_1+n_2+\ldots+n_{JM}=J_M$. A more elegant way to overcome the problem of the infinite state space is to define the state of the community using the species-abundance distribution, namely a vector $\mathbf{S_M}=(S_1, S_2,\ldots, S_{JM})$, with S_i representing the number of species with abundance i (unlabelled species description) (see Etienne and Alonso 2007, Haegeman and Etienne 2010). In the following we present the model using the former approach.

In a metacommunity with a size constraint, the dynamics of species are coupled; a species can increase its abundance (birth) only following the death of an individual of another species (zero-sum rule). Nevertheless, because all individual, independent of species, have the same probability of death, birth, dispersal and speciation (neutrality), the dynamics of a particular species with abundance n can be considered independently, with all other species

considered as one species with abundance J_M-n (Haegeman and Etienne 2010). A way to represent the dynamics of the process is by constructing the master equation. This is a differential equation for the probability that the community is in a given state. If we focus on one species, we can write a master equation for the probability that this species is at a state with abundance n. This has the following general form,

$$\frac{dp_{n,i}(t)}{dt} = \sum_{k \neq n}\left[R(k,n)_i\, p_{k,i}(t) - R(n,k)_i\, p_{n,i}(t)\right], \qquad (2.11)$$

where $p_{n,i}(t)$ is the probability that species i has abundance n at time t and $R(k,n)_i$ is the rate at which species i moves from a state with abundance k to a state with abundance n. Equation (2.11) has a simple interpretation. It states that the probability of being at state n changes with time because of moving into state n from other states k and because of moving out of state n to other states k. All possible transitions and the corresponding rates are described in Box 2.4. There are three events that can cause a change in abundance for a species; birth, death or speciation. The birth and death events result in an increase or decrease of its abundance by one individual. Point mutation speciation also leads to an increase or decrease by one individual. Hence a species with abundance n can only reach or be reached by states $n+1$ and $n-1$ in one step. Note that because all species have the same transition rates (see Box 2.4) we can discard the index i referring to species. Equation (2.11) takes the following form,

$$\frac{dp_n(t)}{dt} = g_{n-1}p_{n-1}(t) + r_{n+1}p_{n+1}(t) - (r_n + g_n)p_n(t), \qquad (2.12)$$

where g_n and r_n are the rates of increase and decrease in abundance respectively. To find the steady state probabilities, we set $dp_n(t)/dt = 0$. This leads to the metacommunity equilibrium species-abundance distribution (Etienne and Alonso 2005),

$$E(S_n \mid \theta, J_M) = \frac{\theta}{n}\frac{(J_M + 1 - n)_n}{(J_M + \theta - n)_n}, \qquad (2.13)$$

where $(x)_n$ denotes the rising factorial (Pochhammer symbol). The detailed derivation of equation (2.13) is presented in Appendix A. In equation (2.13), $E(S_n|\theta,J_M)$ denotes the expected number of species with n individuals and $\theta = (J_M-1)/(1-v)$ is a constant parameter called the **fundamental biodiversity parameter**. This controls the total species richness of the community at equilibrium. In the limit of large J_M, equation (2.13) tends to a probability density distribution:

$$p(x)dx = \frac{\theta(1-x)^{\theta-1}}{x}dx,$$ (2.14)

which represents the number of species with relative abundance between $(x, x+dx)$ (Vallade and Houchmandzadeh 2003, Alonso and McKane 2004). Summing equation (2.14) over all n from $n=1$ to $n=J_M$ gives the expected number of species at equilibrium (Etienne and Alonso 2005, Hubbell 2001),

$$S_M \equiv E(S \mid \theta, J_M) = \sum_{i=1}^{J} \frac{\theta}{\theta+i-1},$$ (2.15)

which practically, for large enough J_M (>100) and θ (>1), is well approximated by (Hubbell 2001),

$$S_M = 1 + \theta \ln\left(1 + \frac{J}{\theta}\right).$$ (2.16)

Box 2.4 Transition rates for a species with abundance n in a <u>metacommunity</u> with point-mutation speciation

The following diagram shows the possible transitions *in* and *out* of a state with abundance n, for a species in a community with point-mutation speciation. There are two types of transitions *out* and two transitions *into* state n, due to birth, death and speciation events.

Transition	*Rate*	*Description*
$n \to n+1$	$g_n = (1-v)\dfrac{n}{J_M}\dfrac{J_M - n}{J_M - 1}$, $n>0$	An individual of another species dies with prob. $(J_M - n)/J_M$ and an individual of the target species gives birth with prob. $(1-v)n/(J_M-1)$
$n \to n-1$	$r_n = \dfrac{n}{J_M}\left((1-v)\dfrac{J_M - n}{J_M - 1} + v\right)$	An individual of the target species dies with prob. n/J_M. Then either an individual of another species gives birth with prob. $(1-v)(J_M-n)/(J_M-1)$ or speciation happens with prob. v.
$n-1 \to n$	$g_{n-1} = (1-v)\dfrac{n-1}{J_M}\dfrac{J_M - n + 1}{J_M - 1}$, $n>1$	An individual of another species dies and an individual of the target species gives birth
$n+1 \to n$	$r_{n+1} = \dfrac{n+1}{J_M}\left((1-v)\dfrac{J_M - n - 1}{J_M - 1} + v\right)$	An individual of the target species dies. Then either an individual of another species gives birth or speciation happens

Note: A special case is the transition from a state with abundance zero to a state with abundance one, due to speciation. In such a case, the species can be reintroduced in the community when a speciation event happens. A speciation event happens with probability v per step, however as there exist S_0 species with abundance 0, the probability of reappearance for the target species is $g_0 = v/S_0$.

2.3.3 Local community dynamics
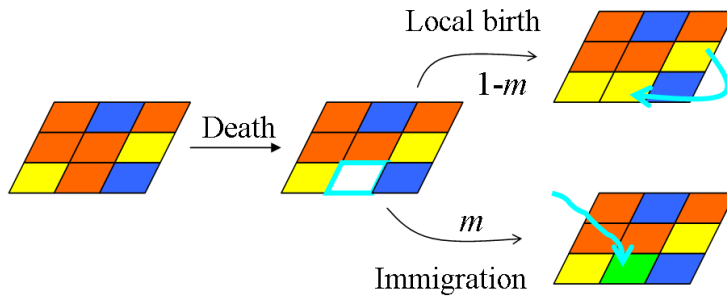
Figure 2.9 One step in a neutral local community. The community is represented by a grid in which each site is occupied by an individual of a given species (color). At each step, a random individual is killed (here an orange) and a site becomes available (white site). Then with probability 1-*m*, another random individual is picked (here a yellow) to give the descendant that will occupy the empty site. With probability *m* the site is occupied by an immigrant from the metacommunity (here a green), which is a descendant of a random individual from the metacommunity. Note that the species to immigrate could already exist in the local community (i.e. could be one of the blue, yellow or orange species).

Due to its bigger size, the turnover time of the metacommunity is much slower than that of the local community. Hence, during the time it takes the local community to reach a steady state, the metacommunity species-abundance distribution (SAD) does not change substantially. A common practice followed to derive the local community dynamics is to consider that the metacommunity SAD is fixed and follows equation (2.14) (Hubbell 2001, Volkov et al. 2003, Etienne and Alonso 2007). Fixing the metacommunity SAD defines a finite state space for the local community, as the species that can exist in the local community are limited to the species that exist in the metacommunity. The state space can be described by a vector $\boldsymbol{n_l}=(n_1, n_2, \ldots, n_{SM})$, with n_k the abundance of the kth species and S_M the number of species in the metacommunity. In the local community, species abundances change due to birth and death of individuals and colonization from the metacommunity (see Figure 2.9). This process can be described in discrete time: at each step a random individual is selected to die (death). Then the empty site is either taken by the descendant of another individual from the local community

(birth) or with probability[2] *m* it is taken by a descendant of an individual from the metacommunity (immigration). In an immigration event, the individual to immigrate is selected randomly from the metacommunity, which means that this species is selected with probability proportional to its metacommunity abundance. In contrast to the metacommunity, every species in the local community has different rates of increasing and decreasing its abundance per step, due to its different abundance in the metacommunity. The process described above defines an irreducible Markov chain with a unique stationary distribution (Haegeman and Etienne 2010, Van Kampen 1992).

As before, the dynamics of each species can be considered separately. Denoting by $p_{n,k}(t)$ the probability that species *k* has abundance *n* at time *t*, one can write the following master equation,

$$\frac{dp_{n,k}(t)}{dt} = g_{n-1,k}\, p_{n-1,k}(t) + r_{n+1,k}\, p_{n+1,k}(t) - (r_{n,k} + g_{n,k})\, p_{n,k}(t)\,. \tag{2.17}$$

The rates $g_{n,k}$ and $r_{n,k}$ are listed and explained in Box 2.5. To find the steady state probabilities we set $dp_{k,n}(t)/dt=0$ and solve the resulting equations (Appendix B). The solution gives the expected number of species with *n* individuals in the local community (i.e. the expected species-abundance distribution), which Hubbell called **zero-sum multinomial distribution** (Hubbell 2001),

$$E(S_n \mid m, \theta, J) = \theta \binom{J}{n} \int_0^1 \frac{(Ix)_n \left(I(1-x)\right)_{J-n}}{(I)_J} \frac{(1-x)^{\theta-1}}{x}\, dx\,, \tag{2.18}$$

where *I=m(J-1)/(1-m)* is called the **fundamental dispersal number**, $(I)_n$ denotes the rising factorial (Pochhammer symbol) and $\binom{J}{n}$ the combinatorial (Etienne and Alonso 2005). Summing equation (2.18) over all *n* from *n=1* to *n=J* gives the expected number of species in

---

[2] The parameter *m* has been called immigration probability by Hubbell (2001). However, this is essentially the colonization probability, i.e. the probability that the individual will immigrate and successfully establish in the local community. Here, we use the original term for *m*, as this has been established in the relevant literature.

the local community. Since the summation does not yield a closed form expression, the expected species richness is usually found computationally with simulations.

Box 2.5 Transition rates for a species with abundance *n* in the <u>local</u> community

The following diagram shows the possible transitions *in* and *out* of a state with abundance *n*, for a species in the local community. There are two types of transitions *out* and two transitions *into* state *n* due to birth, death and immigration events.

Transition	Rate	Description
$n \rightarrow n+1$	$g_{n,k} = \dfrac{J-n}{J}\left(\dfrac{(1-m)n}{J-1} + mx_k\right)$	An individual of another species dies with prob. $(J\text{-}n)/J$ and then, either an individual of the target species gives birth with prob. $(1\text{-}m)n/(J\text{-}1)$ or an individual of the target species immigrates from the metacommunity with prob. mx_k, with x_k the relative abundance of the species in the metacommunity.
$n \rightarrow n\text{-}1$	$r_{n,k} = \dfrac{n}{J}\left(\dfrac{(1-m)(J-n)}{J-1} + m(1-x_k)\right)$	An individual of the target species dies with prob. n/J and then, either an individual of another species gives birth with prob. $(1\text{-}m)(J\text{-}n)/(J\text{-}1)$ or immigration happens and an individual of another species immigrates $m(1\text{-}x_k)$.
$n\text{-}1 \rightarrow n$	$g_{n-1,k} = \dfrac{J-n+1}{J}\left(\dfrac{(1-m)(n-1)}{J-1} + mx_k\right)$	An individual of another species dies and then, either an individual of the target species gives birth or immigrates from the metacommunity
$n+1 \rightarrow n$	$r_{n+1,k} = \dfrac{n+1}{J}\left(\dfrac{(1-m)(J-n-1)}{J-1} + m(1-x_k)\right)$	An individual of the target species dies. Then either an individual of another species gives birth or immigrates from the metacommunity

Note: A species that became extinct in the local community can be reintroduced through immigration from the metacommunity. Hence, a transition from a state with abundance 0 to a state with abundance 1 can happen due to immigration. This happens with rate $g_{0,k}$

2.3.4 Sampling nature of neutral theory

When a sample is taken from a community, the composition of the sample will be in general different from that of the whole community. Suppose that the species-abundance distribution (SAD) of a community of size J is $P(n)$. Namely, $P(n)$ represents the probability that a species has n individuals. Consider that a sample of size $J_1 < J$ is taken from this community and that $p(n)$ is the relative number of species with n individuals in the sample. $P(n)$ and $p(n)$ are connected through the following relation,

$$p(n) = \sum_{x=1}^{J} p(n \mid x, J, J_1) P(x). \tag{2.19}$$

In equation (2.19), $p(n|x,J,J_1)$ is the probability to find n individuals of a species in a sample of size J_1, given that the abundance of this species in the community of size J is x. This is called the **sampling distribution** and depends on the way the sampling is performed. If the sampling is random without replacement, then $p(n|x,J,J_1)$ is the hypergeometric distribution,

$$p_{\mathrm{hyp}}(n \mid x, J, J_1) = \frac{\binom{x}{n}\binom{J-x}{J_1-n}}{\binom{J}{J_1}}. \tag{2.20}$$

Etienne and Alonso (2005) showed that the SAD of the metacommunity (equation (2.13) or (2.14)) is invariant under hypergeometric sampling. This means that if a hypergeometric sample (i.e. a random sample without replacement) is taken from the metacommunity, then its SAD will be the same as the SAD of the whole metacommunity. In other words, in equation (2.19), $p(n)$ and $P(x)$ represent the same distribution (i.e. $p(n) = P(x)$).

  As already mentioned, the local community is also a sample from the metacommunity. Due to dispersal limitation, however, this is not a random hypergeometric sample, but a dispersal-limited sample. Etienne and Alonso (2005) found the sampling distribution for this kind of sampling, which they call the **dispersal-limited hypergeometric distribution**. Using

this distribution, the local community SAD can be connected to that of the metacommunity through equation (2.19), namely,

$$E(S_n \mid m,\theta,J) = \sum_{x=1}^{J_{\mathrm{M}}} p_{\mathrm{hyp}}^{\mathrm{DL}}(n \mid m,x,J_{\mathrm{M}},J)\, E(S_x \mid \theta,J_{\mathrm{M}}),$$

(2.21)

where $E(S_n \mid m,\theta,J)$ and $E(S_x \mid \theta,J_{\mathrm{M}})$ represent the local community's and metacommunity's species-abundance distribution respectively and the dispersal-limited hypergeometric distribution is expressed as,

$$p_{\mathrm{hyp}}^{\mathrm{DL}}(n \mid m,x,J,J_1) = \binom{J}{n}\sum_{A=1}^{J}\sum_{a=1}^{n}\bar{s}(n,a)\bar{s}(J-n,A-a)\frac{I^{\mathrm{A}}}{(I)_{\mathrm{J}}}\frac{1}{\binom{A}{a}}p_{\mathrm{hyp}}(a \mid x,J_{\mathrm{M}},A),$$

(2.22)

where $\bar{s}(n,a)$ denotes the unsigned Stirling number of the first kind. The sampling framework described above shows how dispersal, which is a spatio-temporal process, manifests as a sampling effect when local samples are drawn from a community. This is because the composition of local samples is affected by the dispersal pattern of individuals in the community. Also note that equation (2.19) provides yet another way of estimating the composition of a local sample (or a dispersal-limited sample), when the metacommunity species-abundance distribution is known.

2.3.5 Neutral theory success and criticism

Undoubtedly, the success of the neutral theory is that it derives realistic macroscopic community features, e.g. species-abundance distributions and species-area relationships, from fundamental ecological processes (e.g. see Hubbell 2001, Volkov et al. 2003, Etienne and Alonso 2005). By contrast, the commonly used models of species-abundance distributions (geometric, broken-stick, log-series and log-normal) are descriptive, namely they have been developed to fit observations based on assumptions that are not directly linked to ecological processes. Surprisingly, the neutral model provides justification for two of the above

distribution models. As Hubbell (2001) first noted, the regional community with point mutation speciation follows approximately the log-series distribution (see section 2.3.2). This allows interpreting the log-series parameters; Fisher's alpha is identified to the biodiversity parameter, θ, while the parameter x represents the probability of a local birth (i.e. local birth rate/death rate). As it was later shown by Haegeman and Etienne (2010), the broken-stick distribution also emerges as a prediction of the neutral model. This is followed by the regional community that is subject to random fission speciation (see Chapter 5).

Apart from the above, there are other aspects of the neutral theory that constitute it a useful model in ecology. Firstly, the neutral model is a **dynamical model**. As such it gives the opportunity to predict the evolution of communities, e.g. estimate species temporal turnover, times to extinction, test a community's response to disturbances etc. (e.g. see McKane et al. 2004, Azaele et al. 2006, Halley and Iwasa 2011, Chisholm 2011). Furthermore, the dynamical nature of the model enables the explicit description of the relaxation process following habitat loss. Secondly, the neutral model is a **stochastic model**. Stochasticity is an important element of community evolution, not only because unknown and uncontrollable forces affect communities, but mainly because the fundamental ecological processes are stochastic in nature. In other words, demography, dispersal and speciation are inherently stochastic. By explicitly modeling these processes (i.e. by including stochasticity), we are able to predict the range of possible states that a community can be found in. This provides the basis for statistical hypothesis testing whereby observations are compared with model predictions, and where apart from an average state of a community one needs to predict the expected deviation. As we will see in Chapter 4, the above two properties (in one word: stochastic dynamics) constitute the neutral model an appropriate null model of temporal community turnover.

Finally, a great advantage of the neutral model compared to previous models of stochastic community dynamics (e.g. birth-death models) is that it includes stochastic dispersal. Dispersal crucially affects the composition of local communities/samples (see section 2.3.4). As explained by Etienne and Alonso (2007), when a local sample is drawn from a community, the information that one takes for the whole community is filtered by the limited ability of its individuals to disperse to the local community. Hence, by modeling dispersal, the

neutral model provides a link between the composition of local samples and the composition of the whole community, namely the neutral theory is also a **sampling theory** (section 2.3.4, Etienne and Alonso 2005).

Overall, the neutral theory includes three of the four processes that according to Vellend (2010) influence the composition of ecological communities, namely speciation, ecological drift (i.e. stochastic demography) and dispersal. The fourth process that is not incorporated in the model is selection. By involving these basic processes, the model provides a baseline for detecting the relative importance of any additional processes affecting the community assembly and evolution. Finally, the neutral model has an important conceptual and explanatory value; that is the model's predictions can be traced back to the relevant underlying processes, which facilitates the understanding of the reasons leading to observed macroecological patterns. For example, the form of the species-abundance distribution at the local scale arises from an interplay between local extinctions and dispersal, with the immigration probability defining its shape; an explicit-space neutral model predicts a triphasic species-area relationship, which shape was found to depend strongly on the speciation rate (Rosindell and Cornell 2007); the process of extinctions following habitat loss can be understood to a satisfactory level under the implicit-space neutral model (Halley and Iwasa 2011, Halley et al. 2014, Chapter 3, Chapter 5).

The neutral model has been used in several different applications. Many such examples can been found in Alonso et al. (2006), Rosindell et al. (2011) and Rosindell et al. (2012). These fall into two categories. In the first type of applications, the model is used as an approximation, i.e. the model is proven or assumed to be a good approximation for the community in study and is used to make predictions for the community. In the second type of applications, the neutral model is used as a null model, i.e. a model that by its failure to fit observations can reveal other factors affecting the community in study or disprove one of the models assumptions (e.g. neutrality) (Rosindell et al. 2012).

The neutral model has been criticized for making unrealistic assumptions about ecological communities (e.g. Ricklefs 2003, Clark 2009, Purves and Turnbull 2010, Clark 2012). Currently it has become clear that from the four assumptions of the implicit-space neutral model, namely the zero-sum rule, the neutrality, the implicit treatment of space and the

point mutation speciation, the first two are the least unjustified ones (e.g. see Rosindell et al. 2011 for a review of these assumptions). Firstly, as it was proven by Etienne et al. (2007a) (see also Haegeman and Etienne 2008), the analytical results regarding the equilibrium state of the community are the same even if the zero-sum rule is discarded. Secondly, the neutrality assumption is a good starting point if one wants to have a fairly simple model with some predictive power. But even as a first approximation to reality, neutrality is not such an unreasonable assumption. As is has been shown, the species-abundance distribution of the neutral model is very robust to the breaking of neutrality (Rosindell et al. 2012, Allouche and Kadmon 2009, Chisholm and Pacala 2010, Etienne and Haegeman 2011). The consequences of breaking the neutrality have been studied by He et al. (2012). These authors showed that, compared to a neutral model, in a birth-death trade-off model species coexist for longer times.

The implicit-space assumption and the assumed speciation mechanism are more restrictive to the application of the model. In the implicit-space model there is no sense of dispersal distance, as all metacommunity individuals have the same probability of colonizing empty sites in the local community. This can be problematic when the model is used to parameterize real data, where many local samples are taken from the same regional area (Etienne and Rosindell 2011, Etienne 2007, Jabot et al. 2008). This creates the following contradiction: the local samples are subject to dispersal-limitation, but at the same time, the union of these local samples forms the regional community, which in the implicit-space model is considered to be well-mixed (i.e. has no spatial structure). To overcome these problems, explicit space models have been developed (Etienne and Rosindell 2011, Rosindell and Cornell 2007). These have been successful in predicting species-area relationships (Rosindell and Cornell 2009, O'Dwyer and Green 2010), but perform worse than the implicit-space model in predicting species-abundance distributions (Etienne and Rosindell 2011).

The final assumption refers to the mechanism of speciation and its implications on the mean lifetime of species and the number of species at equilibrium. As noted by Ricklefs (2003), the **point mutation** mode of speciation predicts many species with very short lifetimes while an alternative mode (the **random fission** speciation mode) predicts very long lifetimes resulting in unrealistically high equilibrium species richness, for realistic speciation rates. To correct for this, Rosindell et al. (2010) proposed the mechanism of protracted speciation. In

this, the establishment of speciation events is delayed by τ generations. This means that a new species appears with a higher abundance, given that it survives for τ generations, which results in more realistic speciation rates per individual. A related issue has to do with the interpretation of species richness and speciation rate. While in the model every speciation event results to a new species, in reality there are many species that are created and go extinct before they can be observed. Hence, there is a mismatch between the species richness and speciation rates estimated from observation (in which many species are unobserved) and the model predictions where all species are counted (see Ricklefs 2003, Hubbell 2003, Chave 2004 for further discussion).

There has also been a different kind of criticism against the neutral model, which focuses on more general matters. For example, the role of stochasticity and demographic stochasticity (ecological drift) in particular is questioned by some ecologists (e.g. see Clark 2009, Clark 2012). As explained in Rosindell et al. (2012), this kind of criticism refers to semantic issues that relate to the use of models in general. These semantic issues regard questions like: should we use complex or simple models, stochastic or deterministic models? Does pattern reveal process? Should a model's assumptions be strictly accurate? A discussion of these matters can be found in Rosindell et al. (2012), McGill and Nekola (2010) and Alonso et al. (2006). We also catch up on this discussion in the concluding chapter of the thesis (Chapter 6).

Chapter 3

Neutral theory as a conceptual model of community relaxation[3]

3.1 Introduction

In this chapter we use the neutral model to develop a framework for understanding the mechanisms of species extinctions following habitat loss. This is based on the conceptual model of extinction debt built by Kuussaari et al. (2009) and provides a justification of this model in terms of the dynamics of the relaxation process. Although the developed framework can also be used for predictions, here we emphasize more its conceptual value that conveys the fundamental principles and basic functionality of the system to be described (Strickland 2011). Thus, the focus is to evaluate the model's ability to describe the extinction process, including only the necessary and sufficient variables and mechanisms needed to explain the phenomenon.

There are various frameworks aiming to understand and interpret the effects of habitat loss on biodiversity. Traditionally, the total extinctions following habitat loss are estimated as the difference between the species richness of the initial habitat and the species richness of the remaining habitat at equilibrium, using Species-Area Relationships (SARs). Under this perspective, the total loss of species (imminent plus delayed extinctions) is explained by one variable, namely area. However, habitat loss is usually accompanied by the formation of isolates (e.g. islands of forest created by fragmentation). In such cases, the loss of species in the remaining fragments can be explained in terms of sink and source species. According to Rosenzweig (1995), sink populations of a species in a given habitat are characterized by more deaths than births and their persistence is due to immigration from source populations inhabiting a habitat where births are more than deaths. Areas that are part of a continuous

---

habitat have more sink species (i.e. species that are not resident, but are occasional immigrants). On the other hand, isolates have mostly source species. So, the number of species of an area that was previously part of a continuous habitat and becomes isolated will decline due to the loss of sink species that cannot sustain their population on the isolate without support from the regional area. Under this view, extinction debt equals the number of sink species in a sample plot before habitat loss occurs. Another framework for extinctions due to habitat loss is provided by the theory of island biogeography (MacArthur and Wilson 1967). This predicts that the species richness of an island or a local habitat results through a balance between local extinctions and immigration from the mainland. In an event of habitat loss, the rates of extinction and immigration are out of balance so that the species richness increases or decreases until a new balance is met.

The neutral theory of biodiversity (NTB) provides yet another way of looking at extinctions. As we demonstrate, the neutral theory can link many of the concepts related to habitat loss (i.e. extinction debt, imminent and delayed extinctions, island and sample SARs) in one unified framework. It predicts that under fragmentation, extinction debt is precisely the number of species lost in the process by which a sample area transforms to an isolate and is equal to the difference between the sample SAR of the initial habitat and the island SAR of the remaining habitat, which can both be derived from the model.

3.2 Habitat loss under the NTB paradigm

As a case of habitat loss we consider a fragmentation scenario, where a sample area becomes isolated by clearing part of the habitat surrounding it (Figure 3.1). Due to the clearing, the former sample area losses the support of the surrounding habitat and gradually transforms to an isolate, that is a habitat area surrounded by inhospitable matrix, as for example an island, a mountain peak or a plot of forest surrounded by agriculture land (see Preston 1962 or Chapter 2).
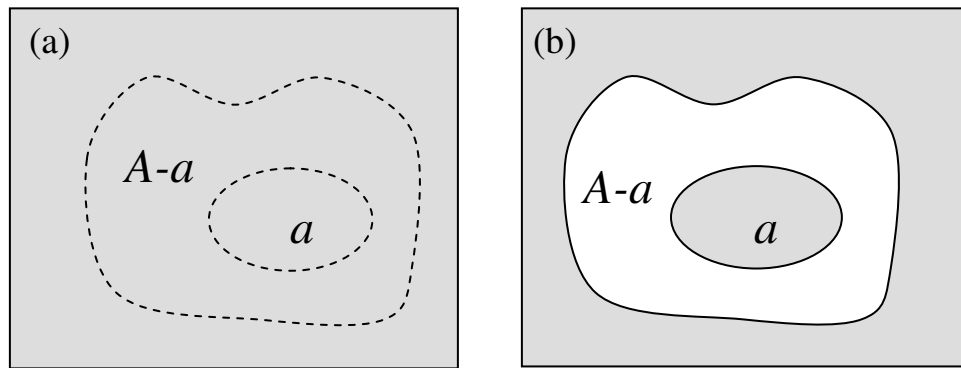
Figure 3.1 Definition of sample areas (a) and isolates (b). Under fragmentation, a habitat area (*A-a*) surrounding a sample plot (*a*) is destroyed, leaving the sample plot isolated, surrounded by inhospitable matrix (white). The percentage of habitat loss is *f*=100*(*A-a*)/*A*.

During the isolation procedure, the sample area will lose its sink species and relax to a lower equilibrium. The main difference between an isolate and a sample area (e.g. a plot within a continuous habitat) is the degree of isolation. While the sample area is embedded within a habitat and species can easily enter and exit its area, an isolate has a higher degree of isolation and is connected to other areas only through rare immigration events. For this reason, sample areas can accommodate more species than isolates of the same size. Furthermore, the species-area relationship for isolates (ISAR) is steeper than the species-area relationship for sample areas (sample SAR or CSAR) (see section 2.2.2). To make a connection between extinction debt and the process by which a sample area transforms to an isolate, Halley et al. (2014) used the neutral model of biodiversity to compute the extinction debt and the SARs of sample areas and the corresponding isolates.

In terms of the neutral model of biodiversity, the initial sample plot can be considered as a local community, namely a dispersal-limited sample from the metacommunity (e.g. the surrounding forest) (see Chapter 2, Etienne and Alonso 2007). The Species-Abundance Distribution (SAD) of this sample plot follows Hubbell's zero-sum multinomial distribution (equation 2.18, Chapter 2). This provides an estimate for the expected number $E(S_n|\theta,I,J)$ of species with *n* individuals, given the fundamental biodiversity parameter θ of the metacommunity, the total number of individuals in the sample, *J*, and the fundamental dispersal parameter, *I*,

$$E(S_n \mid \theta, I, J) = \frac{\theta}{(I)_J} \binom{J}{n} \int_0^1 (xI)_n \left((1-x)I\right)_{J-n} \frac{(1-x)^{\theta-1}}{x} dx, \tag{3.1}$$

where $(I)_J$ and $\binom{J}{n}$ denote the rising factorial (Pochhammer symbol) and the combinatorial respectively. The fundamental dispersal parameter is defined as $I=m(J\text{-}1)/(1\text{-}m)$, where m is the per birth probability of immigration from the surrounding area in the sample plot, which can take values from $m=0$ (total isolation) to $m=1$ (no dispersal limitation). To find the number of species within the sample plot of area a we can simply sum the SAD of equation (3.1) from $n=1$ to $n=J_a$, where $J_a=\rho*a$ is the number of individuals in area a, that is:

$$S(a) = \sum_{n=1}^{J_a} E(S_n \mid \theta, I, J_a). \tag{3.2}$$

After the sample plot is isolated, this is still connected to the metacommunity through immigration; however the immigration probability has now decreased. The isolate is still a dispersal-limited sample from the metacommunity, its SAD is described by equation (3.1) and its equilibrium species richness can be computed from equation (3.2) with dispersal parameter $I'=m'(J_a\ \text{-}1)/(1\text{-}m')$, where $m'<m$ is the reduced immigration probability. Based on the above, the difference in species richness between the initial sample area and the isolate area of the same size is given by,

$$D_a = S_{\text{sample}}(a) - S_{\text{isolate}}(a) = \sum_{n=1}^{J_a} \left[E(S_n \mid \theta, I, J_a) - E(S_n \mid \theta, I', J_a)\right]. \tag{3.3}$$

Equation (3.3) gives the number of delayed extinctions that are expected to happen when a sample area transforms to an isolate. Imminent extinctions can also be estimated from equation (3.2). The number of imminent extinctions is just the number of species found in the destroyed habitat area $(A\text{-}a)$, hence,

$$E_{A-a} = S_{\text{sample}}(A-a) = \sum_{n=1}^{J_{A-a}} E(S_n \mid \theta, I, J_{A-a}) \,. \tag{3.4}$$

Note that the above estimates of imminent and delayed extinctions are consistent with random placement of individuals across the initial habitat (an assumption of the implicit-space neutral model). That is equations (3.3) and (3.4) cannot predict imminent and delayed extinctions in aggregated habitats. This fact however, does not affect the equilibrium species richness of the remaining isolate, which is independent of the initial condition (i.e. independent of the state of the remaining community at the time of habitat loss). As we show in section 3.3, one can estimate imminent extinctions independently, on the basis of an aggregated distribution of species, and then find the equilibrium species richness under the neutral model by evaluating equation (3.2). The equilibrium species richness can also be found from direct simulations of the local community dynamics, or more efficiently using Conlinsk's colonization method (Conlisk et al. 2010) (see Appendix C for a Fortran 95 source code).

3.2.1 Immigration rate in isolates

From equation (3.3) it is clear that extinction debt is largely due to the isolation of the habitat fragments (isolates) remaining after habitat loss. In this section we explore what determines the immigration rate in isolates. It is reasonable to assume that the number of immigrants arriving on an isolate per unit time depends on both its area and its distance from the nearest regional community or mainland, that is $\mu = \mu(a, d)$. For example, consider the case depicted in Figure (3.2) left. Two islands of different area lie in the same distance from a mainland. Suppose that the immigrants are seeds and that the density of seeds decreases with the distance from the mainland, as illustrated by the color gradient (i.e. darker color corresponds to higher density of seeds). For a given distance, the number of seeds that will land on the island per unit time is proportional to the island's area. In a different scenario where the immigrants arrive by sea, the immigration rate could instead be proportional to the island's perimeter. In this case, the number of immigrants arriving per unit time increases with the square root of

area. The immigration rate also depends on the island's distance, d, from the mainland. For an island of given area, the number of immigrants it receives per unit time will decrease with distance (see Figure 3.2 right). Hence the immigration rate for isolates can be expressed by the general relationship $\mu=ka^x d^y$, with $x\geq0$, $y\leq0$ and k a constant.

The immigrants that arrive on the island compete with the local population for available sites. The probability that an immigrant will take up an empty site instead of a local, i.e. the immigration probability m', depends on the relative abundances of locals and immigrants but also on their establishment capability. It could be the case that immigrant species have a disadvantage of establishment compared to the local species, a phenomenon that is termed biotic resistance. Studies of real communities show that biotic resistance is greater in richer communities (see Levine et al. 2004); hence the establishment capability of immigrants is expected to decrease with area, as bigger areas support richer communities. If we denote by c the establishment capability of immigrants (defined here as the probability that an immigrant that is selected to immigrate will establish in the community), then this is expected to scale with area as $c(a)\sim a^{-w}$, $w\geq0$. Because the number of local individuals is proportional to the island's area, the relative abundance of immigrant to local individuals will be proportional to $a^x d^y/a$. Taking the above effects into account, the immigration probability is expected to scale as $m'\sim a^{-w}a^{x-1}d^y$. Hence in the general case, the immigration probability for isolates depends on both its area and its distance from the mainland. Note, that although both the immigration rate (μ) and the immigration probability (m') scale with the isolate's area, there is one importance distinction: while the immigration rate increases with area, the immigration probability (which is in practice the probability of immigration and establishment (see section 2.3.3)) might as well decrease with the area of the isolate.

There is however a situation where the expected distance of an island from the mainland depends on its area. This is the fragmentation scenario described in Figure 3.1, where an island (or fragment) of area, a, is created by clearing the habitat area, $A-a$, surrounding it. If the lost area is small, the fragment will be on average closer to the mainland and the immigration probability will tend to be equal to that of the corresponding sample plot. At the other extreme, when the percentage of habitat loss increases (i.e. for smaller fragments), the average distance of the fragment from the mainland also increases, which means that the

immigration probability will decrease until it becomes zero when there is no remaining fragment. This apparent dependence of the expected immigration probability on the area of the fragment is derived in Appendix D and is,

$$m' = m_A \left(\frac{a}{A} \right)^{0.5} , \qquad (3.5)$$

where m_A is the immigration probability for $a=A$ (no habitat loss) (see figure 3.1).

Adopting different assumptions about the immigration pattern, leads to different dependencies of the immigration probability on area. To incorporate these, equation (3.5) can be generalized to,

$$m' = m_A \left(\frac{a}{A} \right)^{-\omega} , \qquad (3.6)$$

where the exponent ω can be positive, negative or zero and the equation is made to comply with the fragmentation scenario of Figure 3.1, so that when there is no area loss, the immigration probability is equal to that of a sample area of size A, i.e. $m'(a=A)=m_A$. For $\omega > 0$ the immigration probability decreases with the fragment's area and assumes its maximum value for the smallest possible fragment (i.e. for a fragment accommodating only one individual). For $\omega < 0$ the immigration probability increases with the area of the fragment and assumes its maximum value, m_A, for the largest possible fragment, namely that of area A. Finally, for $\omega=0$ the immigration probability is independent of area.
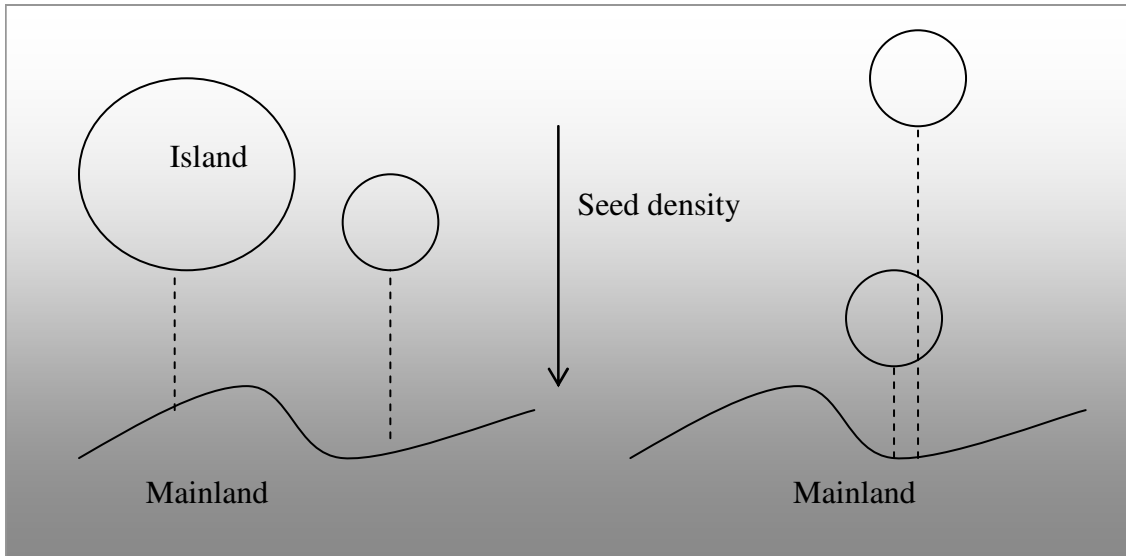
Figure 3.2 Schematic presentation showing how the number of immigrants arriving on an island (isolate) depends on the island's area (left) and its distance from the mainland (right). It is assumed that the density of immigrants (e.g. seeds) that disperse from the mainland to the island decreases with the distance from the mainland (as represented by the color gradient). The number of immigrants landing on the island can be approximated by the product of the island's area and the average density of seeds over its area. In general, the number of immigrants increases with the island's area and decreases with its distance from the mainland.

3.2.2 Extinction debt and Species-Area Relationships

We now turn to examining what the neutral framework for extinctions (equations 3.1-3.4) predicts in terms of extinction debt and SARs. We consider three habitat loss scenarios implying different types of dependence of the immigration probability, m', on isolate's area a. In each case, we use equation (3.2) to construct the sample SAR of the habitats before habitat loss and the ISAR of the isolates remaining after the loss. Analytically, the three cases are:

1. *Immigration probability constant ($\omega=0$).* This is based on the fragmentation scenario of Figure 3.1 for which we assume that the immigration probability for the isolates, m', is independent of their area and their distance from the mainland. Hence the ISAR is constructed with m'=constant.

2. *Immigration probability scales with distance.* This is again based on the fragmentation scenario of Figure 3.1, but where the immigration probability for the isolates scales

only with their distance from the mainland. This leads to an apparent dependence of m' on the isolate's area: $m'(a) = m_A(a/A)^{0.5}$ (equation 3.5).

3. *Immigration probability scales with area.* In this we consider a scenario of habitat loss in an island as described by Halley et al. (2014). The immigration probability is assumed to decrease with area according to $m'(a)=m_A(a/A)^{-0.5}$. This could be interpreted as the result of biotic resistance of the isolate's community, which is stronger in bigger (and thus more species-rich) habitats.

In all three cases, the sample SAR is constructed using a constant dispersal number, I, which means that the immigration probability for sample areas scales as $m(a)=I/(I+J_a-1)$, where $J_a=\rho a$ and ρ is the constant density of individuals (number of individuals per unit area).

Case A: Immigration probability constant

Consider the scenario of habitat loss of Figure 3.1 also shown in Figure 3.3 (a), (b). A continuous habitat of area A suffers habitat loss, as a result of which a former sample plot of area a becomes isolated. Due to its increased isolation, the newly formed fragment has a lower immigration probability than the initial sample area ($m'<m$). Although the immigration probability for the isolate, m', is expected to depend on its area (as explained in section 3.2.1), in this first example we ignore this dependence, i.e. we consider the case $\omega=0$, which leads to $m'=$constant. Figure 3.3 (c) shows the SAR for sample areas (upper curve) and the corresponding isolates (lower curve), both produced using equation (3.2). The sample SAR is produced using a constant dispersal number $I=m'(J_A-1)/(1-m')$. The ISAR is produced using $m'=$constant. The arrows in Figure 3.3 (c) represent the decline of species richness due to imminent and delayed extinctions. The different colors represent the magnitude of imminent (light grey) and delayed (grey) extinctions and the species richness of the remaining habitat at equilibrium (dark grey). Note that the two curves intersect at $a=A$, at which point there is no habitat loss and the habitat in question is just a sample plot of area A within the surrounding continuous habitat.

Figure 3.3 A scenario of habitat loss and the corresponding SARs derived from the neutral model of biodiversity. (a) A sample plot of area a within a continuous habitat. The plot is dispersal-limited with immigration probability m. (b) The isolate remaining after clearing the area surrounding the sample plot. The isolate has lower immigration probability than the sample plot ($m'<m$). (c) SARs of the two habitats. The first describes the species richness of sample areas within the initial habitat (sample SAR, upper curve). This is computed using equation (3.2) with parameters $\theta=10.0$, $I=103.0$. The second describes the species richness of the isolates remaining after habitat loss (ISAR, lower curve). This is computed using equation (3.2) with parameters $\theta=10.0$, $m'=0.1$. Different colors represent imminent extinctions (light grey), delayed extinctions (grey) and remaining species richness (dark grey). (Parameters: $\rho=928$ ind./km^2 $J_A=928$).

Case B: Immigration probability scaling with distance

In this second example, we consider the case in which the immigration probability declines proportionally with the distance from the mainland. Within the fragmentation scenario we consider, this leads to a scaling of immigration probability with area described by equation (3.5) and corresponds to $\omega=-0.5$. The resulting SARs are shown in Figure 3.4 (c). As in case A, we construct the sample SAR using a constant dispersal number $I=m_A(J_A-1)/(1-m_A)$. The ISAR is produced using $m'(a) = m_A(a/A)^{0.5}$. Although the sample SAR is the same as in Figure 3.3, the ISAR has a different form and species richness increases more rapidly with area. Compared to case A, for a given percentage of habitat loss (e.g. 80%), delayed extinctions (grey) are increased at the expense of imminent extinctions (light grey) and remaining biodiversity (dark grey).
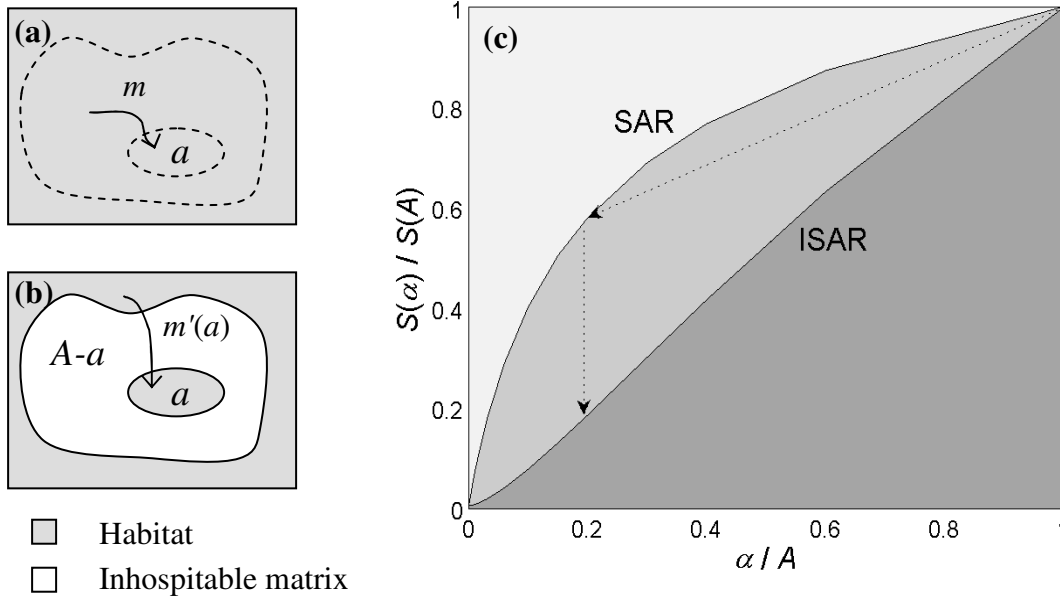
Figure 3.4 A scenario of habitat loss and the corresponding SARs derived from the neutral model of biodiversity. (a) A sample plot of area *a* within a continuous habitat. The plot is dispersal limited with immigration probability *m*. (b) The isolate remaining after clearing the area surrounding the sample plot. The immigration probability for the isolate scales with the square root of its area according to equation (3.5). (c) SARs of the two habitats. The first describes the species richness of sample areas within the initial habitat (sample-SAR, upper curve). This is computed using equation (3.2) with parameters $\theta=10.0$, $I=103.0$. The second describes the species richness of the isolates remaining after habitat loss (ISAR, lower curve). This is computed using equation (3.2) with parameters $\theta=10.0$, $m'(a) = m_A(a/A)^{0.5}$ and $m_A=0.1$. Different colors represent imminent extinctions (light grey), delayed extinctions (grey) and remaining species richness (light grey). (Parameters: $\rho=928$ ind./km^2 $J_A=928$).

Case C: Immigration probability scaling with area

Halley et al. (2014) described the situation of habitat loss where an island (already an isolate) losses part of its area (e.g. due to a volcano eruption or a rising sea level). This scenario is shown in Figure 3.5 (a), (b). Halley et al. (2014) considered an immigration pattern, where the immigration probability of the isolate decreases with the isolate's area according to equation (3.6) with ω=0.5. The SARs for sample plots within the initial island and for isolates remaining after habitat loss are estimated using equation (3.2) (Figure 3.5(c)). The ISAR curve is constructed using $m'(a)=m_A(a/A)^{-0.5}$. The sample SAR is constructed using a constant dispersal number $I=m_A(J_A-1)/(1-m_A)$. In addition to the SAR curves, the species richness of the sample areas and isolates were estimated from simulations performed using Conlisk's colonization method (points) (Conlisk et al. 2010). Both the sample SAR and the ISAR initially rise more rapidly compared to cases A and B. Although the parameters used in this case are different, the examples show how the different immigration scenarios result in different relative magnitudes of imminent and delayed extinctions.
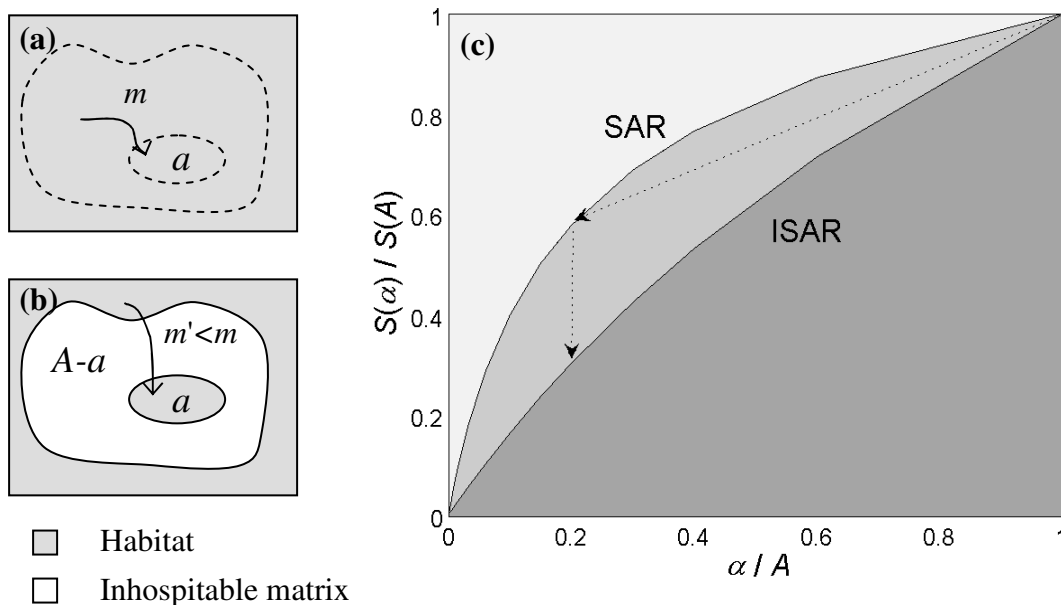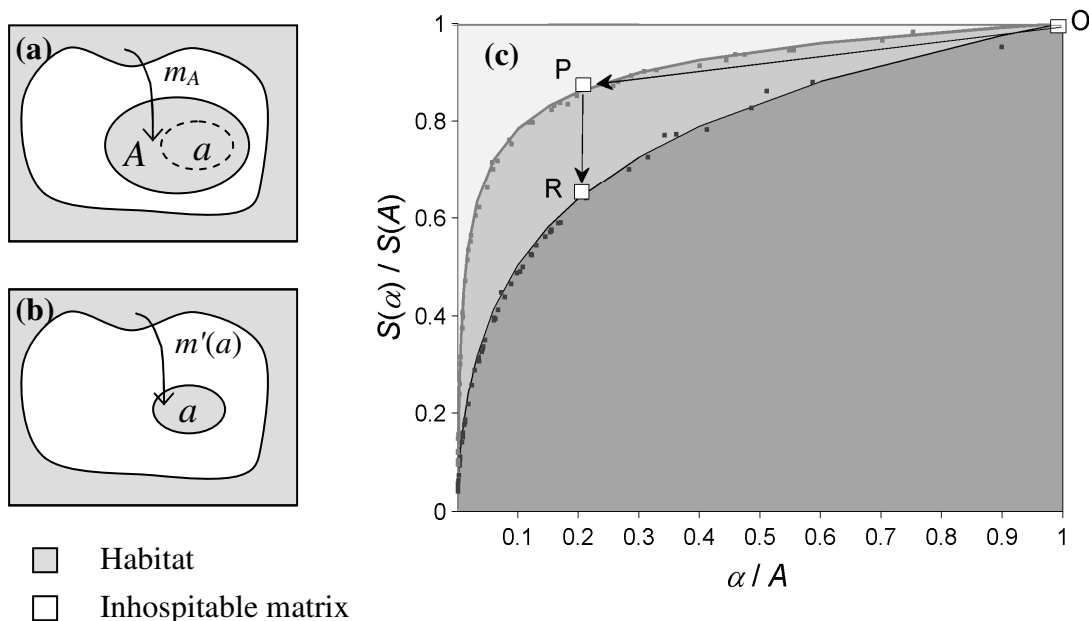
Figure 3.5 A scenario of habitat loss in an island and the corresponding SARs derived from the neutral model of biodiversity. (a) A sample plot of area a within the island. The immigration probability for the island is m_A. The sample plot is considered a random sample from the island (no dispersal-limitation within the island). (b) The island remaining after the partly submersion of the initial island. The new immigration rate is $m'(a)$. (c) SARs of the two habitats. The first describes the species richness of sample areas within the initial island (sample-SAR, upper curve), computed using equation (3.2) with parameters $\theta=10.0$, $I=1.18$, $m_A=0.0013$. The second describes the species richness of the islands (isolates) remaining after habitat loss (ISAR, lower curve). This is computed using equation (3.2) with parameters $\theta=10.0$, $m(a)=0.3a^{-0.5}$. Different colors represent imminent extinctions (light grey), delayed extinctions (grey) and remaining species richness (light grey). The black and grey points represent the species richness computed using Conlisk's colonization rule (Conlisk et al. 2010) (see Appendix C). (Parameters: $\rho=928$ ind./km^2 $J_A=928$).

3.3 Relative magnitude of imminent and delayed extinctions

The different colors in Figures 3.3, 3.4 and 3.5 express the message of equation (2.1) that, following habitat loss, species from the original habitat can meet three fates: some go extinct immediately (light grey), others are lost in the subsequent relaxation process (grey) and the remainder form the new equilibrium community (dark grey). The relative magnitude of imminent and delayed extinctions depends on the size of habitat loss and the extent of isolation of the remaining habitat, as it is shown in the above figures. Note that these results are specific to the implicit-space neutral model of biodiversity, which is consistent with a

random spatial distribution of individuals and a zero-sum multinomial species-abundance distribution (SAD). In reality, the SAD and the spatial distribution of individuals differ from that assumed by the neutral model and this difference is expected to affect the size of imminent extinctions. To explore the influence of the initial community configuration, Halley et al. (2014) considered two different SADs, the broken-stick and the Zero-Sum Multinomial (ZSM) and two initial spatial distributions, random placement (RP) of individuals and an aggregated distribution (AGG) based on a negative binomial distribution.

To estimate imminent and delayed extinctions the following procedure was followed. The first step was to construct the RAD or the SAD of the initial community. For the ZSM distribution this was done by running simulations based on Conlisk's colonization method (Conlisk et al. 2010, see source code in Appendix C). For the broken-stick distribution the SAD can be found given the size and the species richness of the community (see Table 1 in Green and Ostling 2003). The second step was to estimate imminent extinctions, E_{A-a}, for a given spatial configuration (either RP of AGG) using equations (2.4) and (2.5) (Chapter 2, section 2.1.2). Due to imminent extinctions (loss of endemic species), the remaining community is left with $S_0=S_A-E_{A-a}$ species and $J_a=(a/A)J_A$ individuals. The third step was to determine the equilibrium species richness, S_{eq}, of this remaining community. This was done using simulations and letting the community to gradually relax to its new equilibrium state. Note that in estimating S_{eq}, the initial abundance vector of the community after habitat loss is irrelevant, since the community will always relax to the same equilibrium independent of the magnitude of imminent extinctions. In other words, in all four cases (broken-stick and RP, broken-stick and AGG, ZSM and RP, ZSM and AGG) the equilibrium species richness is the same. The delayed extinctions were then estimated as the difference between the initial species richness after habitat loss and the species richness at equilibrium, namely $D_a=S_0-S_{eq}$. The percentage of delayed extinctions for the four initial community configurations is reported in Table 3.1, for initial habitats of various sizes and different percentages of habitat loss. The results correspond to the scenario of habitat loss in an island described in case C of section 3.2.2.

Table 3.1 Percentage of delayed extinctions (Delayed / (Delayed+Imminent))*100 as a function of the percentage of habitat loss, in three habitats of different initial sizes. Each column corresponds to a different initial community configuration: broken-stick or zero-sum multinomial (ZSM) SAD and aggregation (AGG) or random placement (RP) of individuals across the habitat. Parameters: constant density of individuals: ρ=928 (ind./km^2), initial habitat area: A, initial species richness: S_A initial community size: $J_A=\rho A$, immigration probability: $m'=1/(\rho \cdot a)^{0.5}$ (equation 3.6). Metacommunity: individuals: J_M=10,000,000, species: S_M=1,400, biodiversity parameter: θ=123.7. The aggregation model used to estimate imminent losses is based on a finite negative binomial distribution with aggregation parameter k=0.5 for all species (see equation (2.5), Chapter 2).

Parameters of initial community	Area Lost (%)	Broken-stick		Zero-sum multinomial	
		Random	Aggregated	Random	Aggregated
$A =$ 10 km^2 $N_A =$ 9,280 $S_A =$ 232 $J_A/S_A \approx$ 40	10	90.4	38.5	34.7	9.8
	30	90.3	45.2	58.5	20.7
	50	89.4	49.6	65.0	28.7
	70	85.4	46.6	57.7	27.3
	90	70.9	37.6	47.2	22.8
	99	22.4	15.0	17.1	10.1
$A =$ 100 km^2 $N_A =$ 92,800 $S_A =$ 397 $J_A/S_A \approx$ 233	10	96.4	40.2	58.4	14.3
	30	97.3	60.3	72.7	28.3
	50	96.5	56.7	69.3	26.7
	70	95.5	57.2	68.3	30.2
	90	91.1	52.5	63.2	30.2
	99	61.7	33.7	44.1	23.2
$A =$ 1,000 km^2 $N_A =$ 928,000 $S_A =$ 562 $J_A/S_A \approx$ 1,651	10	99.5	75.3	82.4	35.3
	30	99.3	72.5	78.7	28.9
	50	99.2	73.0	77.1	31.6
	70	99.0	74.4	76.1	37.8
	90	98.2	72.3	71.5	40.1
	99	90.3	56.8	59.5	33.3

The results reported in Table 3.1 can be summarized in three main conclusions:

- For a relatively even SAD (broken-stick) and with random placement of individuals, delayed extinctions dominate. This no longer prevails when the SAD is more asymmetrical (zero-sum multinomial) or if there is aggregation or both. Thus, in communities with a large percentage of rare species and a strong degree of localization, (that is low dispersal abilities), the relative numbers of imminent losses is much larger than that of delayed extinctions (see also Green and Ostling 2003).

- The percentage of delayed extinctions does not change monotonically with the percentage of habitat loss. Delayed losses initially increase with lost area but then

decrease again and become zero when the total area is lost, in which case all extinctions are imminent. This pattern is more apparent for the aggregated spatial distribution and the more uneven SADs (i.e. the zero-sum multinomial). In conclusion, the proportion of delayed losses is largest for low and intermediate levels of habitat loss.

- Delayed extinctions are more for larger initial habitats. We conjecture, however, that this is a secondary effect. The decisive variable is the ratio J_A/S_A, namely the average number of individuals per species, which with our choice of parameters happens to be higher in the larger habitats. When this ratio is small, there are a lot of species with few individuals, which are more prone to imminent extinction and thus delayed extinctions are less. As this ratio becomes bigger, species are less prone to imminent extinction and thus delayed extinctions increase.

The above results are indicative of the scaling of the percentage of delayed extinctions with area and community structure. However, the values are also expected to depend on the particular immigration pattern. In addition, it should be emphasized that imminent extinctions were estimated independently of the neutral framework described by equations (3.3) and (3.4). According to the implicit-space neutral model, the configuration of the initial community is zero-sum multinomial with random placement of individuals. Thus, there is a mismatch between the configuration used to estimate imminent extinctions (e.g. aggregation and broken-stick) and the configuration used to estimate the equilibrium species richness and thus delayed extinctions (i.e. zero-sum multinomial with random placement). For a fully consistent estimation of delayed extinctions for aggregated habitats, one should start from an explicit spatial model.

3.4 Summary and conclusions

The neutral model was used to construct a conceptual framework for understanding the mechanism of species extinctions following habitat loss. This links the dynamics of the relaxation process with the SAR method for extinction forecasts. The main result is

schematically presented in Figure 3.6, which shows two cases. In the first case (Figure 3.6a) habitat loss causes the isolation of the remaining habitats. In this case, imminent extinctions can be estimated from the sample SAR of the initial habitat used in a backward fashion. Then, delayed extinctions are species lost from the remaining habitat due to increased isolation. These can be found from the difference between the sample SAR and the ISAR describing the habitat before and after its isolation. In the second case (Figure 3.6b) the remaining habitat is not isolated due to habitat loss. This can describe a situation where the initial habitat is already an isolate. In this case, one can estimate the total number of species extinctions (Imminent + Delayed) from the ISAR used in a backward fashion, a method that is usually used to predict extinctions in islands (for example see Triantis et al. 2010).

The described framework predicts that delayed extinctions are the sole result of the reduction of the immigration probability for the remaining habitat. Hence, if the remaining habitat is no further isolated there are no delayed extinctions. In reality, however, there might be extinction debt even without further isolation of the remaining habitat. For example, if habitat loss results in the reduction of species populations without causing any imminent extinctions (as in the example of Figure 2.3, Chapter 2), there are no species lost as a direct result of habitat loss. However, the species are now forced to reside in a smaller habitat, which will lead to delayed extinctions due to increased competition. The limitation of the model to describe this situation originates from the built-in assumption that the individuals are randomly placed across the habitat, i.e. the model cannot describe the spatial distribution of species. To describe this situation an explicit space model is needed.

Nevertheless, as the equilibrium species richness of the habitat does not depend on the initial configuration, imminent and delayed extinctions can be estimated independently. Following this approach, we estimated imminent extinctions based on sampling formulas on aggregated distributions developed by Green and Ostling (2003), while we estimated the equilibrium species richness using the implicit-space neutral model of biodiversity. The most important outcome of the analysis is that the number of delayed extinction can be up to two orders of magnitude greater than imminent extinctions (for the range of parameters used). In particular, delayed extinctions are expected to be more for less aggregated spatial distributions

and more even species-abundance distributions. This shows the importance of incorporating extinction debt in extinction forecasts.
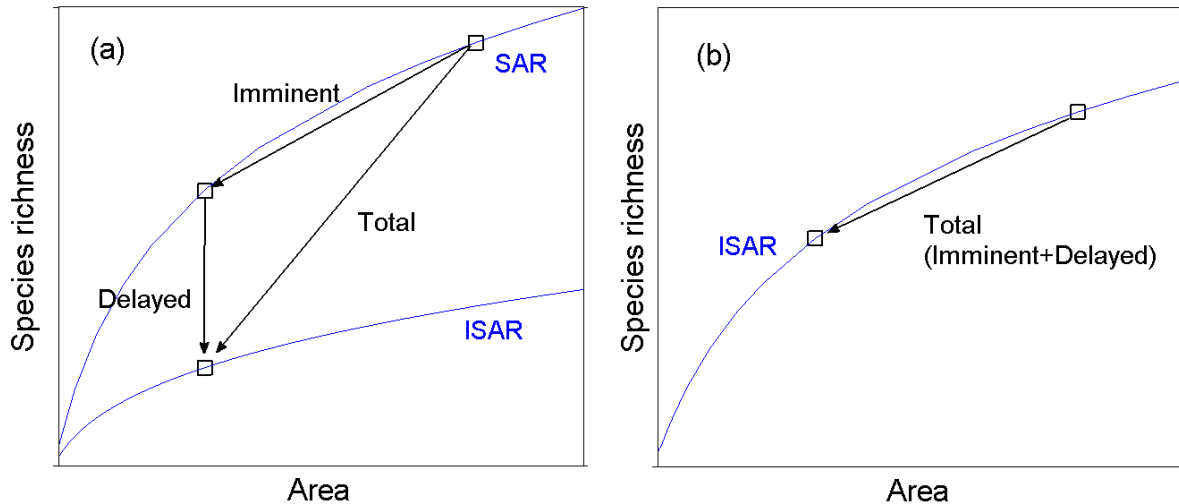
Figure 3.6 A schematic framework for extinctions following habitat loss. (a) The general case: habitat loss causes the isolation of the remaining habitat (e.g. due to fragmentation). Imminent extinctions can be predicted from the sample SAR of the initial habitat used in a backward fashion. Delayed extinctions can be estimated as the difference between the sample SAR describing the reduced habitat before its isolation and the ISAR describing the same habitat after its isolation. (b) The initial habitat and the remaining isolate are described by the same ISAR. In this case, the initial habitat, which is already an isolate, loses part of its area (e.g. the submersion of part of an island due to a volcanic eruption) without this affecting its degree of isolation (i.e. the immigration rate stays the same). In this case, the total number of extinctions (imminent + delayed) can be estimated from the ISAR used in a backward fashion.

Chapter 4

Neutral theory as a null model of temporal community turnover[4]

4.1 Introduction

In this chapter we demonstrate how the neutral model can be used as a null model to test hypothesis regarding the effects of climate change on real communities. Climate change can increase the rate of species extinctions, as it may disfavor some species and render their habitat unable to support them. At a local scale, the effects of climate change manifest as a temporal community turnover, namely a variation of species abundances with time including possible local extinctions. However, an observed community turnover cannot be readily attributed to climate change or to other external forces, because communities are always changing even in the absence of external drivers. The main cause of natural community drift is demographic stochasticity, namely the random births and deaths of individuals (also called ecological drift). Additionally, dispersal increases this drift by affecting the composition of local samples (i.e. sampling effects, section 2.3.4, Chapter 2). Hence, a model of natural community drift that includes demographic stochasticity and dispersal is appropriate to assess the significance of an observed turnover in local communities. The purpose of the model is to capture the extent of variation expected due to natural drift and thus reveal any additional drift, which can then be linked to external forces (see Box 4.1), i.e. the model is used as a null model for temporal community turnover. In this study, we demonstrate how the neutral model can be used as a null model to asses the observed temporal turnover of a community of butterflies for which there is indication of a response to temperature rise.

---

[4] The work presented in this chapter was done in collaboration with Konstantina Zografou and John Halley. This has been submitted for publication in an international peer-reviewed journal and is currently under review.

4.1.1 Natural community drift

An inevitable element of ecological community dynamics is stochasticity, namely the effect of processes that involve chance. In population ecology, a main source of stochasticity is demographic stochasticity, which refers to the random events of births and deaths of individuals, also called ecological drift. Other sources of stochasticity include environmental variability (e.g. changes in temperature), but also the stochasticity due to dispersal (i.e. the random movements of individuals). In this thesis, we will refer to the combined action of demographic stochasticity and dispersal as natural drift. Because of natural drift, communities are never static. Nevertheless a community can be in a state of dynamical equilibrium if the forces affecting it are stationary, namely if they are on average constant and have a fixed variance (e.g. the temperature fluctuates around a mean value with a fixed variance). At the state of dynamical equilibrium, the macroscopic characteristics of the community are also stationary (e.g. the number of species fluctuates around a given average value) (see Chapter 2). Hence, a community that is otherwise at equilibrium can be found in a multitude of different states that are all consistent with the same average conditions.

Even at equilibrium conditions as defined above, natural drift causes a fluctuation of species abundances which results in a gradual change of a community's composition with time (i.e. species go extinct and are substituted by other species). Thus, natural drift is associated to a temporal community turnover. Apart from natural drift, a community's turnover is also affected by other forces, i.e. biotic factors (interactions between species) and abiotic factors (change of environmental conditions). In a set of ecological data, natural drift and other stochastic components can usually be recognized as noise, which can be removed to reveal an overall pattern or trend. However, depending on the size of a community, the magnitude of natural drift can be comparable to a systematic drift caused by non-stationary ecological forces (e.g. temperature rise), making it difficult to distinguish between the two. If the available data are not appropriate to distinguish between the two processes (natural drift and systematic drift due to external forces), then the observed turnover can only be significant if it is greater than the expected turnover under natural drift. Thus, in order to test the significance of an observed turnover, it is necessary to first estimate the magnitude of natural drift in a community.

Box 4.1 Statistical hypothesis testing and null models

Stochasticity is inevitable in ecological data and *statistical hypothesis testing* is frequently used to make sense of or extract information from data. In such cases, a null model can be used in order to capture stochasticity and thus reveal any underlying ecological mechanism or relation between the measured variables or to test hypotheses. Gotelli and McGill (2006) make the distinction between two types of null models used in ecology. The first type is that of *statistical null models*. Statistical null models are not actual models, but refer to randomization techniques on observed data. In these, the data are randomized to produce an ensemble of possible states of a system by stochastically varying according to some degrees of freedom, while keeping some constraints implied from the observed data. The second type of null models is that of dynamical or *mechanistic null models*. These are actual models as they assume some mechanisms by which the observed system works. In these, the observed data are only involved in estimating the parameters of the model. Then, the parameterized model is used to produce a number of possible states of the system that are consistent with these parameters.

Example of a mechanistic null model. Suppose that a coin is tossed 10 times and the sequence of heads and tails is recorded e.g. [T, T, H, T, H, H, T, T, T, T]. Then it is asked whether the coin is fair. To be able to answer this question one needs to compare the observed sequence with the sequence produced by a fair coin, namely a coin which gives heads or tails with probability 0.5. As there is stochasticity (one cannot predict the result of a given coin toss), every sequence of 10 successive tosses of the fair coin will be different. For this reason, the fair coin has to be tossed many times to produce a fair amount of sequences of 10 successive tosses. Then, to decide whether the initial coin was fair is a matter of counting how many times the fair coin produced a result like the one observed. This is the basis for a statistical hypothesis testing, where the null model is the fair coin and the corresponding null hypothesis is that "the coin that produced the observed sequence is fair". Luckily, one need not have a fair coin, but it is sufficient to construct a conceptual model of a fair coin. In this particular example, the probability of obtaining 3 heads and 7 tails when tossing a fair coin 10 times is given by the binomial distribution and equals $p = 0.117$ (where in order to keep things simple, we ignored the particular order of heads and tails). This probability is large enough to say that the observed sequence could be produced by a fair coin. Hence, the null hypothesis cannot be rejected.

The neutral model in one of several possible models that can be used to estimate the magnitude of natural drift. Compared to other drift models, the neutral model pioneers in one important aspect; that is it takes into account the sampling effects that arise when local samples are drawn from a dispersal-limited community (see section 2.3.4, chapter 2 or Etienne and Alonso 2007). The model has been used before as a null model to assess species temporal turnover in real communities (Leigh et al. 1993, McGill et al. 2005, Gilbert et al. 2006,

Ricklefs 2006). Using the neutral model as a null model, one assumes that the natural drift of the community in study is well approximated by neutral drift.

4.1.2 Introduction to the application

In recent years, there is a growing literature that cites the effects of climate change and temperature rise on ecological communities. A typical sign of temperature rise is the expansion of population's ranges to higher altitudes or latitudes as they track the movement of temperature isotherms (Walther et al. 2002). On a local scale, what is usually observed is the invasion of species from lower altitudes (or latitudes) and the decline or extinction of local species populations that are on the lower temperature edge of their distribution. The invasion from lower altitudes is a much faster procedure compared to the extinction of local populations. The latter is a slow relaxation process that can take years to complete (extinction debt, Jackson and Sax 2010, Halley et al. 2014). This difference in the rates of species introduction (through invasion) and species extinctions, causes an apparent increase of species diversity at a local level, which if conditions remain constant will be restored by the eventual loss of species that are not favored by the changed conditions.

Butterflies and other insects provide an excellent example for examining the effects of global change on populations, as they react faster than other groups to the changes of temperature (Bale et al. 2002, Devictor et al. 2012). Furthermore, due to their short life cycle, changes on their distribution, abundances and community composition can become visible over a short time period (Robinson et al. 2012). The best documented effect on butterfly populations as climate warms is geographic range expansions to cooler areas, towards higher latitudes or altitudes (Parmesan et al. 1999, Parmesan and Yohe 2003). As a result, species adapted to warm conditions begin to invade ecological communities (Barry et al. 1995, Dapporto and Dennis 2013, Wilson et al. 2007).

There have been many studies proving the northward expansion of species ranges (Walther et al. 2002, Parmesan 1996). To make the connection with climate change, a common approach is to show that the observed changes in a community over time are consistent with the expected changes under a projected scenario (Southward et al. 1995,

Holbrook et al. 1997, Brown et al. 1997, Sagarin et al. 1999, Meshinev et al. 2000, Parmesan 2006, Poloczanska et al. 2013). In studies where the whole species range is observed, a possible range expansion may be easy to show and provides a strong evidence of the effects of climate change (Parmesan 1996). However, proving the effects of climate change from the change of community composition on a local scale is tricky, because a turnover caused by natural drift can easily be mistaken for an effect of climate change. Given the possible bias towards publishing studies that prove the effects of climate change (and possibly not publishing studies that prove opposite effects), a statistical test for natural drift is even more important.

In this study, we used the neutral model as a null model to asses the observed turnover of the butterfly community of Dadia National Park, Greece, for which there is an indication of turnover due to temperature rise. Dadia NP is located in northeastern Greece and is part of the NATURA 2000 network. Due to its conservation status, the reserve has remained relatively unmodified by humans for the last 30 years, which excludes land use changes as an explanation for the observed turnover. The butterfly community of Dadia NP has been surveyed by Grill and Cleary in 1998 and by Zografou in 2011 and 2012 following the same sampling scheme (Grill and Cleary 2003, Zografou et al. 2014). A comparative analysis of the data between the two samplings (1998 and 2011), done by Zografou et al. (2014), showed an increase of warm-adapted species and a decrease of cold-adapted species. This was found to be significant based on randomization tests. As reported in Zografou et al. (2014), the observed turnover was consistent with the temperature rise recorded in the area between 1990 and 2012. In all habitats except one, the community temperature index increased significantly, an indication that the community is being dominated by warm-adapted species.

The aim is to see whether the observed turnover of the Dadia butterfly community in a period of 13 years can be explained by natural drift alone. If this is the case, then there is one more mechanism that explains the observed turnover as well as climate change and thus the hypothesis of climate induced changes is not sound. If on the contrary natural drift is not sufficient to explain the observed turnover, then the hypothesis of climate induced changes can be further assessed.

4.2 Methods

4.2.1 Model Parameterization

The parameterization of a model consists of two steps. In the first step one needs to make a connection between the model's concepts (or entities) and the real system in study. The second step is the estimation of the model parameters. The main elements of the implicit-space neutral model (described in Chapter 2) are the metacommunity and the local community. We make the connection between the model and our study system based on the following assumptions: a) we identify the metacommunity to Evros region, which is the regional unit of Greece containing Dadia NP, b) we assume that the seven sampled areas are independent local samples from this metacommunity and c) we assume that the metacommunity has a fixed rank-abundance distribution. Considering a fixed metacommunity distribution is a common practice, justified by the fact that the metacommunity turnover times are much larger than those of the local communities, hence its distribution does not change significantly during the time scale of examination (Volkov et al. 2003, Etienne and Alonso 2007).

Given the above, the parameters that need to be determined are the size of the metacommunity, J_M, the fundamental biodiversity parameter, θ (Hubbell 2001), and the immigration probabilities for each local sample, m_j. Ideally, these parameters should be measured directly or estimated independently of the data used to test the model (Gotelli and McGill 2006). However, as this is very difficult to do, the usual methodology followed is to estimate the parameters that maximize the likelihood of the observed data set (Etienne 2007, Jabot et al. 2008). In this study we followed the usual methodology of maximum likelihood parameter estimation that is described in Etienne (2005), Jabot et al. (2008) and Etienne (2007). However, in contrast to previous studies, we introduced a parameterization of the metacommunity that is partly independent of the test data set. In particular, we assumed that the metacommunity species-abundance distribution has a log-series form and specified its species richness from estimates of regional diversity that exist for our study area.

We first estimated the biodiversity parameter of the metacommunity, θ, using Ewen's sampling formula (Ewens 1972). As explained in Hankin (2007), to estimate θ it is sufficient to maximize the following function,

$$L(\theta \mid J, S) = \frac{\theta^{S}}{\prod_{i=1}^{J} \theta + i - 1}, \tag{4.1}$$

where S and J are the number of species and individuals of the pooled sample (i.e. the sample that results from the merge of the data of all the samples). Given θ, we estimate the immigration probability (m_j) for each local sample independently, using Etienne's sampling formula (Etienne 2005) as applied by Jabot et al. (2008). For one sample with abundance vector $N_j = \{n_1, \ldots, n_{SM}\}$, the likelihood of immigration probability m and metacommunity vector $X = \{x_1, \ldots, x_{SM}\}$ is,

$$L(m, X \mid N_j) = \frac{J!}{\prod_{i=1}^{S_M} n_i} \frac{\prod_{i=1}^{S_M} (I\, x_i)_{n_i}}{(I)_J}, \tag{4.2}$$

where $I = m(J-1)/(1-m)$ is the fundamental dispersal number of the sample, $(I)_J$ denotes the rising factorial (Pochhammer symbol), J the number of individuals in the sample and S_M the number of species in the metacommunity (Jabot et al. 2008). Following Jabot et al. (2008), we assumed that the metacommunity abundance vector, X, is fixed and the same for all samples. However, instead of identifying the metacommunity to the pooled sample as proposed by Jabot et al. 2008, we assumed that this follows a log-series rank-abundance distribution. To create the metacommunity abundance vector, we used the estimated θ and the recorded number of species of the regional area (i.e. the species richness of Evros region (S_M=128) reported in Pamperis (2009)). We then found the metacommunity size using the log-series species individual curve, $S_M = \theta \cdot \ln[1 + J_M/\theta]$. We produced 100 random metacommunity vectors with these parameters and took the average. To enforce a correspondence between the species in the metacommunity and the species in the samples, we sorted the metacommunity vector according to the observed pooled abundance vector averaged in both years. The parameters estimated using our approach and Jabot's approach are reported in Table E1 (Appendix E). Because the samples taken from the same habitat in 1998 and 2011 differ in size, we estimated

the parameters again after re-sampling the bigger sample (either 1998 or 2011) to make the two equal in size (Table E2, Appendix E). The effect of using a larger metacommunity than the pooled sample used by Jabot et al. (2008) is that the estimated immigration probabilities for all samples are smaller than those estimated using Jabot's method (Table E1, Appendix E).

4.2.2 Simulations

We used direct simulations of the stochastic process to produce an ensemble of equilibrium neutral samples corresponding to each of the seven parameterized communities, using the 1998 parameterization. For the simulation of the local community dynamics, a source code was developed in Fortran 95 programming language. This can be found in Appendix F. To produce one sample, we run the simulation for 10,000 generations (years) to allow it to reach stationarity (dynamical equilibrium) and recorded the species-abundance vector of the sample at the end of the simulation. The produced samples correspond to possible states of the communities when these are at dynamical equilibrium. We also produced random samples from the log-series metacommunity. These are produced by running the simulations with immigration probability $m=1$, which corresponds to a random sampling of the metacommunity (Etienne and Alonso 2007).

4.2.3 Estimation of temporal turnover

In order to quantify the turnover of the observed and the simulated communities in a period of 13 years, we used two measures of turnover. The first measure is the Root Mean Squared Distance (RMSD) between an abundance vector sampled at time t and the same vector sampled at time $t + \Delta t$,

$$\text{RMSD} = \sqrt{\sum_{i=1}^{S} \left(n_{t,i} - n_{t+\Delta t,i}\right)^2 / S} \ , \tag{4.3}$$

with $n_{t,i}$ and $n_{t+\Delta t,i}$ denoting the abundances of species i at the two instances, S the total number of species found in both samples (the union) and $\Delta t = 13$ years. As a second measure of turnover we used the Bray-Curtis dissimilarity index defined as,

$$BC = \frac{\sum_{i=1}^{S}|n_{t,i} - n_{t+\Delta t,i}|}{\sum_{i=1}^{S} n_{t,i} + n_{t+\Delta t,i}}. \tag{4.1}$$

The two measures of turnover differ in that the RMSD measures the absolute change in abundance, while the BC index measures the relative change. Because the samples taken from the same habitat in 1998 and 2011 differ in size, before applying equations (4.3) and (4.4) we re-sampled the bigger sample (either 1998 or 2011) and re-estimated the dispersal probabilities, m_j.

To estimate the turnover of the corresponding neutral samples we performed simulations using the 1998 parameterization. For each sample we performed the simulation for 13 years at stationarity, recording the species-abundance vector at the beginning and the end of the simulation. We repeated each simulation 1,000 times to get 1,000 sets of vectors 13 years apart and applied equation (1). Apart from the community turnover, we are also interested in the population drift of separate species. For this we performed the simulation for 10,000 years at stationarity and kept record of the abundances of all species at the end of each year. We then selected an initial abundance value n_0, found all species that have this abundance at some point in time as well as their abundances after 13 years. This way we constructed a distribution of abundances after 13 years, for each initial abundance, n_0. We assumed that a species has significant change in abundance if its observed abundance after 13 years lies outside the 95% confidence intervals of the distribution of neutral abundances.

4.3 Results

The first step is to asses the goodness of the parameterization. This can be done by comparing the rank-abundance distributions (RADs) of the observed samples with the corresponding distributions of the simulated samples. Note that there isn't a one to one correspondence

between the species in the real samples and the species in the simulated samples. As far as the neutral model is concerned, all species have the same a-priory probability of reproduction and extinction. Furthermore, the number of species predicted by simulations is not the same in every simulation and does not coincide with the observed number of species in the corresponding samples. The RAD plots are shown Figures 4.1 and 4.2. There are two main conclusions from Figures 4.1 and 4.2. (a) The observed distributions fall within the bulk of the neutral distributions for all seven habitats and for both years and (b) the observed pooled-sample RADs do not fall within the bulk of the simulated metacommunity distributions. Overall the parameterization is good. The mismatch between the observed pooled-sample and the simulated metacommunity samples means that the merge of local samples cannot be considered a random sample from the assumed log-series metacommunity.
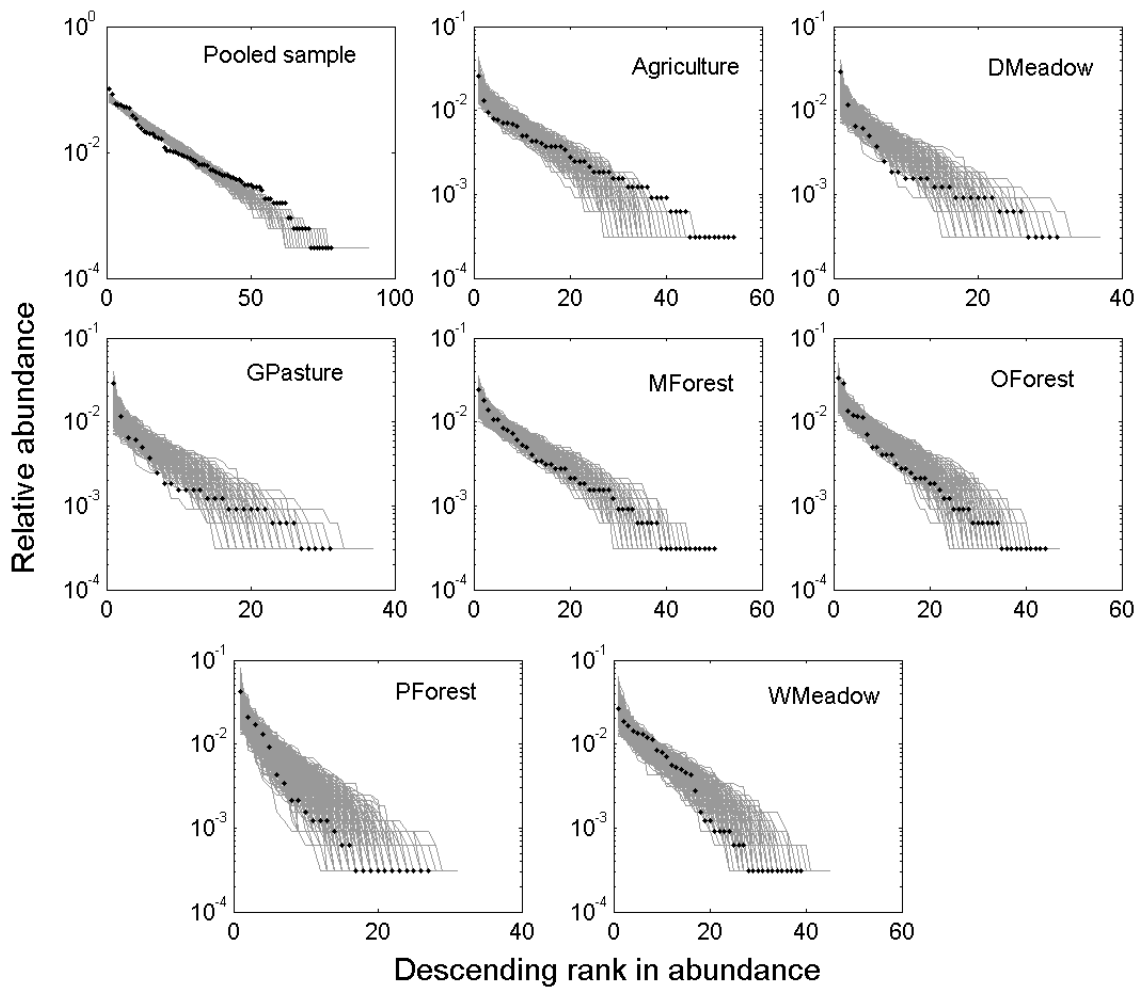
Figure 4.1 Rank-abundance distributions of the seven local samples and of the pooled-sample in 2011. Black dots represent the observed relative abundances ranked in descending order. Grey lines represent 1,000 rank-abundance distributions of the simulated neutral communities (parameterized using the maximum likelihood parameters of Table E1, Appendix E). The pooled-sample simulated distributions are produced using immigration parameter m=1 (no dispersal limitation), which stands for random sampling of the metacommunity.
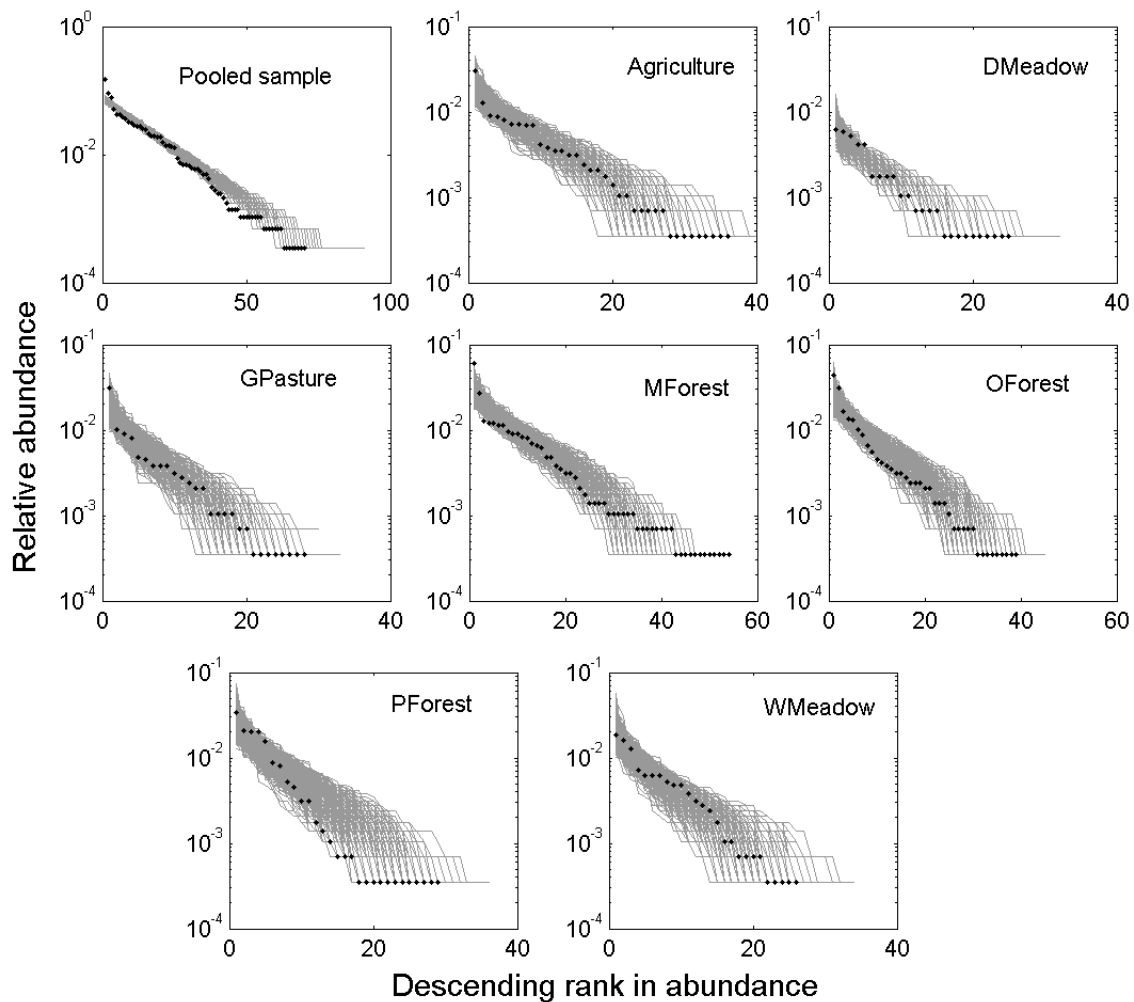
Figure 4.2 Rank-abundance distributions of the seven local samples and of the pooled-sample in 1998. Black dots represent the observed relative abundances ranked in descending order. Grey lines represent 1,000 rank-abundance distributions of the simulated neutral communities (parameterized using the maximum likelihood parameters of Table E1, Appendix E). The pooled-sample simulated distributions are produced using immigration parameter m=1 (no dispersal limitation), which stands for random sampling of the metacommunity.

Figure 4.3 shows the turnover of the real samples (grey dots) and the turnover distribution of the simulated samples (boxes). Using both measures of turnover (RMSD or Bray-Curtis index), we find that in all habitats (except one) the observed turnover is higher than the median neutral turnover. The exception is the Wet Meadow habitat, for which the RMSD measure predicts lower turnover. The higher than neutral turnover found in 6 of the 7 habitats is a result that could have occurred by chance. However, using a Wilcoxon

signed-rank test (Wilcoxon 1992), we find that this has a probability of less than 2.5% to have occurred by chance. Finally, in the three forest habitats, the observed turnover is significantly greater than expected by neutral drift.
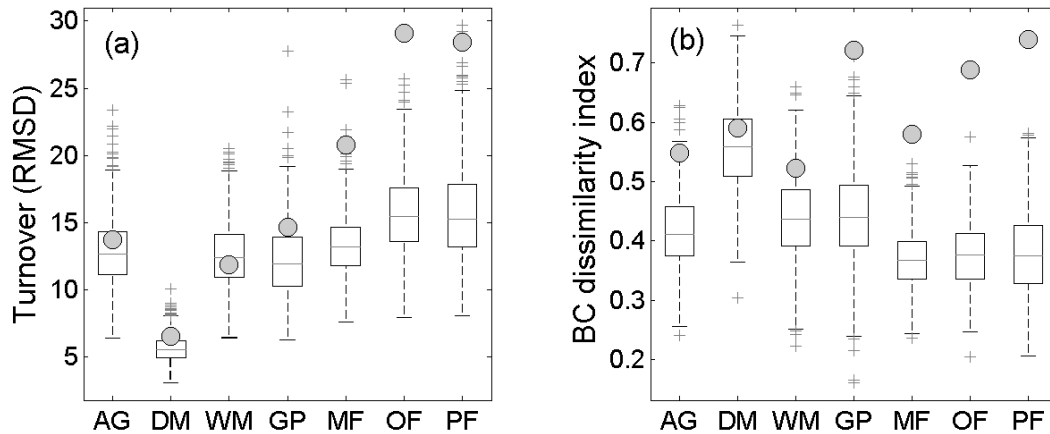
Figure 4.3 Observed turnover between 1998 and 2011 (grey circles) and turnover of 1,000 simulated neutral samples (boxes) for each of the seven habitats (AG=Agriculture, DM=Dry meadow, WM=Wet meadow, GP=Grazed pasture, MF=Mixed forest, OF=Oak forest, PF=Pine forest). (a) Turnover defined as the Root Mean Squared Distance (RMSD). (b) Turnover defined as the Bray-Curtis dissimilarity index. Boxes define the upper and lower quartiles of the distribution, i.e. they contain 50% of the values around the median, which is represented by the horizontal grey line. Whiskers extend to 1.5 times the boxes range. The real turnover is computed after re-sampling the bigger vector (either 1998 or 2011) (see section 4.2.1 and Table E2, Appendix E). Based on the RMSD the turnover is significant for the MF, OF and PF habitats (p=0.004, <0.001 and 0.002). Based on the BC dissimilarity index the turnover is significant for the MF, OF, PF (p<0.001) and the GP habitat (p=0.002).

To visualize the change in abundance of separate species, we plot the abundance of every species in 2011 against its abundance in 1998, for each of the seven habitats separately (Figure 4.4). On the same plots, we draw the corresponding 95% confidence intervals of the distribution of simulated abundances, which are computed by the method described in section 4.2.3. Figure 4.4 shows that in two of the habitat (Agriculture and Wet Meadow) the abundances of all species lie within 95% confidence intervals of the neutral distributions, while in the rest of the habitats there are a few species which lie outside the 95% confidence intervals. We identified 16 such cases, which correspond to 11 species (out of the 88 species studied). Almost all of these species have increased their abundance except from *Aporia*

crataegi and *Argynnis paphia*, while three of the species (*Aricia agestis*, *A. crataegi* and *Kirinia roxelana*) show significant increase or decrease in more than one habitat.

These species are listed in Table 4.1, along with the 19 species that were reported by Zografou et al. (2014) to contribute more to turnover. We see that species that contribute more to turnover do not all have significant change in abundance according to the drift model. In particular, from the 19 species reported by Zografou et al. (2014), only 7 have an observed trend in abundance that cannot be attributed to drift. Furthermore, we found 4 other species whose change in abundance is greater than expected by drift. Table 4.1 also reports the Species Temperature Index (STI); that is the mean temperature value per species across its range (data taken from Schweiger et al. 2014). Viewing our results in conjunction with the STI values shows that species with high STI (warm-adapted) have increased abundance; while species with low STI (cold-adapted) have either increased or reduced abundance. The above is a sign of increased domination of warm-adapted species in the communities.
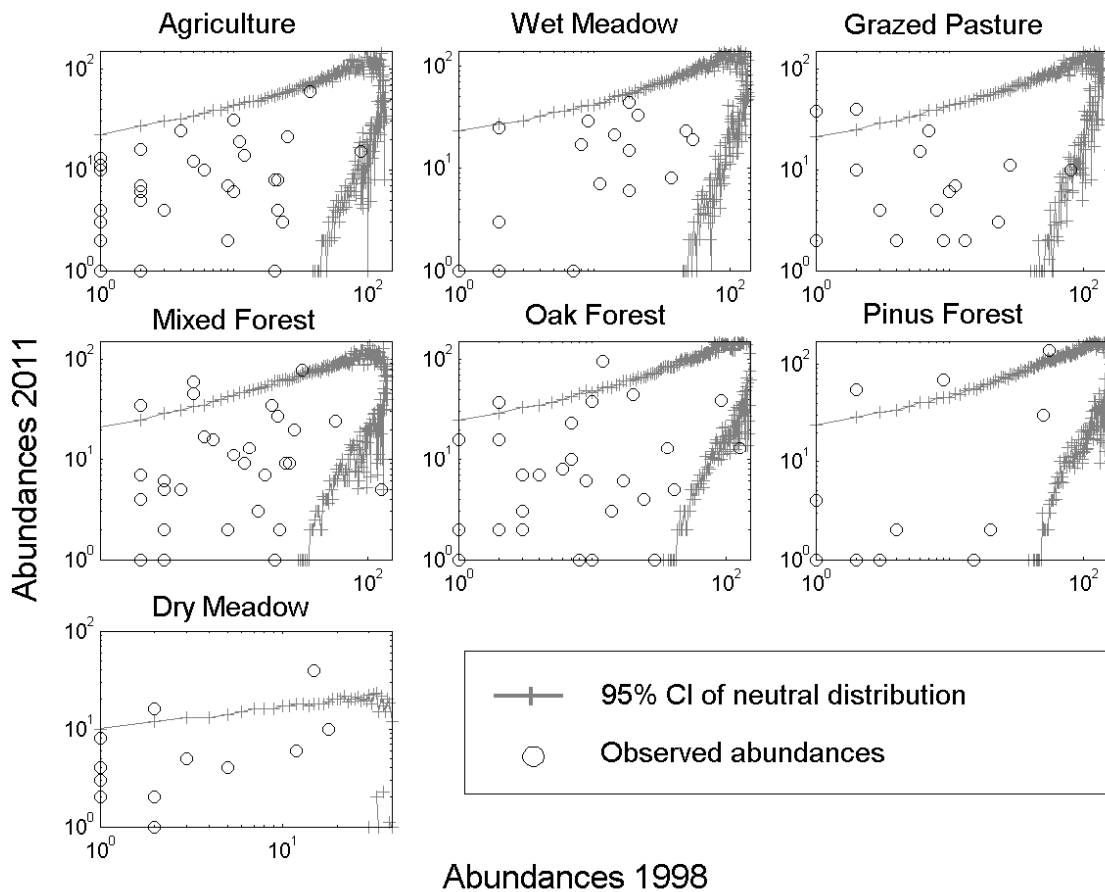
Figure 4.4 Observed species abundances in 2011 versus their abundances in 1998 for the seven local samples (open circles) and 95% confidence intervals (CI) of the corresponding distribution of simulated neutral abundances (grey lines). Simulated data are drawn from stationarity by picking random species with a given abundance and finding the distribution of abundances after 13 years. Species with abundances that lies outside the grey lines have higher or lower abundance in 2011 than expected by the model. The real data are plotted after re-sampling the bigger sample (either 1998 or 2011) to make the 1998 and 2011 samples comparable in size.

Table 4.1 Butterfly species with significant increase (+) or decrease (-) in abundance between 1998 and 2011 (i.e. species that lie outside the 95% confidence intervals of the distribution of simulated abundances of Figure 4.2) and species that contributed more to turnover according to Zografou et al. (2014). STI ± SD: Species Temperature Index with ± 1 Standard Deviation (Schweiger et al. 2014, supplementary material 1). Note that the average STI value of all 88 species in our samples is 10.5. Habitats: the habitats where significant changes occurred in this study (AG=Agriculture, DM=Dry meadow, WM=Wet meadow, GP=Grazed pasture, MF=Mixed forest, OF=Oak forest, PF=Pine forest). Nomenclature follows the updated taxonomy of European Butterflies, made for the Red List of 2010 (Van Swaay et al. 2010).

Species	Zografou et al. (2014)	This study	STI ± SD	Habitats
Coenonympha pamphilus	-		8.96 ± 3.89	
Argynnis paphia	-	-	9.02 ± 2.91	MF
Vanessa cardui	-		9.04 ± 4.12	
Polyommatus icarus	+	+	9.07 ± 4.11	GP
Aporia crataegi	-	-	9.14 ± 3.53	GP, OF
Issoria lathonia	-		9.33 ± 3.08	
Favonius quercus		+	9.49 ± 2.79	MF
Maniola jurtina	+	+	9.85 ± 3.29	MF
Thymelicus sylvestris	+		9.87 ± 2.96	
Aricia agestis	+	+	10.16 ± 2.89	MF, OF, DM
Satyrium ilicis	+		10.21 ± 2.91	
Melitaea didyma	+		10.42 ± 3.10	
Hipparchia fagi	+	+	10.53 ± 2.62	PF
Brenthis daphne	-		10.60 ± 2.90	
Arethusana arethusa	-		10.63 ± 2.22	
Colias crocea	+		10.69 ± 3.33	
Melitaea trivia	-		10.97 ± 2.71	
Brintesia circe	-		11.07 ± 2.74	
Pieris mannii	-		11.46 ± 3.13	
Argynnis Pandora		+	11.92 ± 3.06	GP
Kirinia roxelana	+	+	13.10 ± 2.34	MF, OF, PF
Hipparchia syriaca		+	13.13 ± 2.61	PF
Hipparchia fatua		+	14.03 ± 2.28	DM

4.4 Discussion

Ecological drift and dispersal are two inevitable sources of stochasticity affecting populations. This natural drift can cause a community temporal turnover that can easily be mistaken for an environmental trend. For this reason, statistical inference and hypothesis testing are necessary to interpret observations. The proposed framework is to use the neutral model of biodiversity

as a null model for temporal community turnover. That is to use the model to simulate natural drift and be able to asses the observed turnover using standard methods of statistical hypothesis testing. There are two main assumptions behind this framework: 1) the assumption that the community in study is affected by natural drift (ecological drift and dispersal) and 2) the assumption that the natural drift of the observed community is well approximated by neutral drift.

The framework was used to asses the turnover of a butterfly community (in Dadia-Leukimi-Soufli National Park, Greece), for which there is an indication of shift due to temperature rise over a period of 13 years. We found that neutral drift explains most of the variation of species abundances. At the same time, there is a considerable number of species that don't behave as expected by neutral drift (i.e. have significant change in abundance) and hence their change in abundance could be more reliably associated with changing conditions and in particular climate change. Thus, we reject the null hypothesis of natural drift as the only explanation of the observed turnover. Although there might be other hypothesis explaining the observations, the climate change hypothesis is very plausible and can be further assessed.

The drift model reveals a significant increase in abundance of warm-adapted species (i.e. species with high Species Temperature Index), consistent with the temperature rise recorded in Dadia NP since 1990 (Zografou et al. 2014). The above is in agreement with the study of Zografou et al. (2014), which reported an increased alpha-diversity of warm-adapted species and an increase of the Community Temperature Index. At the level of particular species, we find that the observed trends of 11 species cannot be attributed to drift, 8 of which showed significant turnover in forest habitats, where the observed community turnover is also higher than expected by drift. Zografou et al. (2014) also report species that contribute more to community turnover. However, we found that many of these species trends are actually explained by neutral drift, while we found a few more species, not reported by Zografou et al. (2014), whose change in abundance is greater than expected by drift.

The increase of populations of species associated with warm conditions is a frequently observed pattern (Barry et al. 1995, Wilson et al. 2007, Dapporto and Dennis 2013) explained by the northward expansion of species ranges (Parmesan and Yohe 2003). A typical example of a northward range expansion is that of *A. agestis* in Britain (Pateman et al. 2012). The same

species abundance increased significantly in Dadia, unlike in some other Mediterranean areas where its population declined (Stefanescu et al. 2011). Similarly, the increased abundance of *Favonius quercus* in our system contrasts with the stable and declining population trends in Catalonia and Andorra (Stefanescu et al. 2011). However, our findings for *A. crataegi* and *A. paphia* (decreased abundance) are in accordance with corresponding observations from central Spain (Merrill et al. 2008), North Europe (Parmesan et al. 1999) and Catalonia and Andorra (Stefanescu et al. 2011). Finally, the significant increase in abundance of *Maniola jurtina* is in agreement with the prediction of Isaac et al. (2011) for the UK.

The present study is not the first to use a drift model to assess real communities' turnover. Leigh et al. (1993) found that the turnover (over 9 years) of the tree community on the islands of Gatun Lake is not explained by demographic stochasticity, attributing this result to a violent change in environmental conditions. McGill et al. (2005) examined fossil records of mammalian communities across 1 million years and found that they changed less than the neutral drift predicts, concluding that there are also deterministic forces at work in structuring communities. More recently, Dornelas et al. (2014) carried out a meta-analysis of time series from different taxa and geographical regions (with most series concentrated in the last 40 years). They found that the observed community turnover is higher than expected by a neutral model, attributing this to changes of environmental conditions like habitat loss, climate change, species range shifts and invasion. Like Dornelas et al. (2014) we find the observed community turnover to be higher than expected by the neutral model but the only known systematic change in conditions in our study area is temperature rise, which we consider as the most likely cause of the large turnover.

Although the model's applicability can be limited by its assumptions, it has one great advantage compared to previous models of community drift. That is it takes into account the sampling effects that arise when local samples are drawn from a dispersal limited community. This property (i.e. the sampling property) extends out of the model's limits (Alonso and Etienne 2005) and should be incorporated in statistical inference based on samples.

A large number of papers report changes in ecological communities and attribute them to climate change. Our results show that natural drift can also lead to substantial rates of species turnover, which could be mistaken for a response to climate change. Furthermore,

autocorrelations of environmental variability will substantially increase these rates of turnover. This underlines the need for careful statistical analysis when attributing ecological community changes to climate change or other large-scale processes. In this respect, a null model that simulates natural drift can be used to exclude the expected natural turnover and thus investigate the causes of any additional turnover. To this end, neutral model offers a reasonable starting point.

Chapter 5

Neutral theory with random fission speciation as a model of the relaxation process[5]

5.1 Introduction

In this chapter we attempt the mathematical description of the relaxation process within the neutral theory of biodiversity. The aim is to describe the variation of species richness with time in a community that is out of equilibrium as a result of a given disturbance, as for example habitat loss, change of climatic conditions, diseases, natural destructions or other. Today habitat loss is the main disturbance on ecological communities and the leading cause of species extinctions. As we discussed in Chapter 2, the main effects of habitat loss are the reduction of the available habitat area (i.e. the reduction of available resources) and the change of conditions affecting the remaining habitats (e.g. isolation). Relaxation after habitat loss affects every species to a different extent, while the interactions between species and the interaction of species with the changed environment are important (e.g. extinctions of co-adapted species, environmental changes favor some species but not other etc.). Nevertheless, in the core of it, relaxation is the competition of species to take hold of the remaining available resources and the resulting change of species abundances that this causes. On this basis, the neutral model seems like the appropriate model to describe the community dynamics following habitat loss or other disturbances. In such a case, the aim is to estimate the

---

average number of species going extinct with time, ignoring the particular interaction between species, which provides a first approximation to the more complex process of relaxation.

Earlier attempts to describe the relaxation process have been based on the theory of island biogeography of MacArthur and Wilson (1967). According to this theory, the equilibrium species richness of an island (or a local community) is determined by a balance between immigration from the mainland (or the regional community) and local extinctions. By this view, a community that is out of equilibrium is characterized by an imbalance between local extinctions and immigration. For example, in a habitat area that became recently isolated, the immigration rate is reduced and there are more extinctions per unit time than introductions of new species through immigration. So the species richness will decrease until a new balance is met. The simplest version of the model assumes that the extinction rate on the island is proportional to the number of existing species and the immigration rate is proportional to the number of species in the mainland that are absent from the island. The theory of island biogeography was used by Diamond (1972) to predict the decline of species richness in islands following an event of habitat loss. Diamond (1972) considered a constant rate of extinction per species and derived an exponential decline of the number of species with time,

$$S(t) = S_{\text{eq}} + (S_0 - S_{\text{eq}})e^{-t/t_r} ,$$

(5.1)

with S_0 the initial species richness, S_{eq} the equilibrium species richness and t_r the relaxation time. Along similar lines but including competition between species, Terborg (1974) considered an isolated habitat (no immigration) and found that the decline of species richness with time is described by,

$$S(t) = \frac{S_0}{1 + kS_0 t} .$$

(5.2)

The neutral theory of biodiversity was build on the basis of the island biogeography theory, with the difference that the dynamics of the community are modeled by stochastic processes at the individual level (see Chapter 2). This difference, apart from leading to a much

richer behaviour, gives the opportunity to directly link the macroscopic features of communities to the fundamental processes of birth, death and dispersal, thus providing a tool for discovering the relative role of these processes in shaping macroecological patterns. The dynamical behaviour of the neutral model has already been studied. A time-dependent species-abundance distribution for a neutral community subject to immigration has been independently derived by McKane et al. (2000, 2004), Azaele et al. (2006) and Chisholm (2011)). At the same time, there hasn't been much effort in deriving relaxation curves analogous to those derived from the theory of island biogeography. The only exceptions are the works of Gilbert et al. (2006) and Halley and Iwasa (2011) who derived relaxation curves for isolated communities (with no immigration or speciation). The solution of Gilbert is based on a branching process and is applicable to short timescales ($t \ll J$). Halley and Iwasa (2011) gave a solution that also applies to long timescales, by solving the neutral model equations assuming a broken-stick species-abundance distribution. This is,

$$S(t) = \frac{S_0}{1 + t/t_{50}}, \tag{5.3}$$

in which t_{50} denotes the half life time to equilibrium. Note that in an isolated community, like the one described by Halley and Iwasa (2011), the equilibrium state is fixation. Namely in the end, one species dominates the community ($S_{eq}=1$), provided that the species are bound by a zero-sum rule. If the species dynamics are independent (no zero-sum rule), then eventually all species drift to extinction ($S_{eq}=0$).

In this study, we extend the result of Halley and Iwasa (2011) to communities where species diversity is sustained through speciation. In such a case, the equilibrium species richness is non-trivial and results from a balance between extinctions and introduction of new species. To introduce speciation we use the neutral model with random fission speciation, which has been studied before by Haegeman and Etienne (Haegeman and Etienne 2010, Etienne and Haegeman 2011). Based on this model, we derive a closed-form equation for the variation of species richness with time. This has three parameters: the initial species richness, the speciation rate and the species richness at equilibrium and can be parameterized to predict

the relaxation curve of real communities or to estimate characteristic times to relaxation and the species richness at equilibrium. We present two possible applications of the equation. In the first, we estimate extinctions times for a number of avifaunal communities that have suffered habitat loss. In the second, we predict the relaxation curve of the avifaunal community of Barro Colorado Island. Since all of the communities considered are also affected by immigration, we clarify the conditions under which the model can be used to account for immigration as well as speciation.

5.1.1 Speciation modes in neutral model

In his original model, Hubbell uses the mechanism of point mutation speciation as the mechanism that sustains the biodiversity of the metacommunity. However, he also briefly discusses the random fission mode of speciation, which he introduces as a mechanistic description of allopatric speciation. Unlike the point mutation mode where new species arise by random mutations of individuals at birth, in the random fission mode, new species arise by randomly splitting the population of existing species. From the two population fragments resulting from the split, one remains to the mother species while the other forms the initial population of the new species (Haegeman and Etienne 2010). A crucial difference between the point mutation and the random fission mechanism is the initial abundance of the newly created species. While in point mutation new species arise as singletons (i.e. with one individual), in the random fission mode the abundance of the new species is a random proportion of the abundance of the mother species. Because species with high abundance are more likely to speciate, newly created species are likely to have a high initial abundance and thus a rather high survivorship. This leads to higher species richness at equilibrium than in the point mutation mode. In general, the random fission speciation mode predicts more realistic speciation rates and mean lifetimes of species and is considered by some authors more realistic than the point mutation mode (Etienne and Haegeman 2011, Rosindell et al. 2011).

An important parameter of the neutral speciation models is the community-level speciation rate, namely the number of speciation events happening in the community per unit time. Hubbell described both point mutation and random fission speciation as individual based

processes. Namely, individuals are picked at a constant rate to initiate a speciation event (either point mutation or random fission) (see Chapter 2). This means that the community-level speciation rate is proportional to the number of individuals in the community. Hence, in a community of a given size, speciation events happen at a constant rate. The above description facilitates the simplicity of describing and implementing the model, however it might not be realistic at least for some speciation modes (Etienne et al. 2007b). For example, the random fission mode of speciation, involves dividing the population of species, so that the speciation rate is expected to depend on the existing number of species. Such a model, in which the speciation rate depends on both the number of individuals and the number of species, has been solved by Etienne et al. (2007b). By contrast, in point mutation speciation, the linear dependence of community-level speciation rate on community size seems reasonable. Hubbell (2001) supports that the above dependence offers an explanation to the mechanism of radiation bursts. He suggests that these bursts are a result of the population expansion of species after a mass extinction. As species expand their populations, they have more births than deaths. Because speciation events happen at birth, more births lead to more speciation events, which explains the rapid appearance of new species.

Still in reality, speciation is a complex spatio-temporal process that requires the reproductive isolation of population. In particular, allopatric speciation requires the geographic isolation of species populations. A hint on how allopatric speciation works in a group of isolated islands was given by Rosenzweig (1995), who suggested that if the geographic barriers are occasionally crossed by individuals, but after enough time for a speciation to be established, then "this system acts like a speciation machine, rapidly cranking out new species". This mechanism was implemented by Yamaguchi and Iwasa (2013) in a model of neutral accumulation of genetic differences, allowing a small (smaller than speciation) but recurrent immigration between islands. It was shown that there is a value of immigration rate for which the rate of species formation is maximized. In a situation like this, the number of speciation events happening per unit time increases with the number of existing species. However, the number of species cannot increase indefinitely. If available resources put a constraint on community size, then as species numbers increase, every species will have less individuals, which will reduce its probability of speciating.

From the above, it is clear that the dependence of speciation rate on population size and species richness is not trivial, as it largely depends on spatial processes affecting the gene flow within or between populations. Studies of real speciation events suggest that the speciation rate increases with habitat area and decreases with increasing levels of gene flow, while the level of gene flow determines the minimum area for which speciation can occur (Kisel and Barraclough, 2010). Since the random fission speciation model we use here is spatially implicit, these dependences cannot arise as predictions of the model. To avoid a complicated description, we assume that the community-level speciation rate is constant throughout the relaxation process and thus it does not depend on the number of existing species. The dependence on the number of individuals does not matter in our case, as the community size is kept constant (zero-sum rule). This description is consistent with the original description of the random fission speciation model by Hubbell (2001) and Haegeman and Etienne (2010).

5.2 Derivation of the differential equation

The model of a neutral community subject to random fission speciation has been studied by Haegeman and Etienne, who found closed-form solutions for the equilibrium and the time evolution of the community (Haegeman and Etienne 2010, Etienne and Haegeman 2011). Haegeman and Etienne described the evolution of the community in continuous time, so as time passes, birth-death events and speciation events happen at a constant rate and independently of each other (decoupled). In a birth-death event which happens with rate μ, a random individual is selected to die (death) and is immediately replaced by the descendant of another random individual from the community (birth). In a speciation event, which happens at rate v, a species is selected with probability proportional to its abundance to undergo speciation. The population of this species splits into two fragments, one corresponding to a newly formed species and the other to the old species. If the initial abundance of the old species was k, then after the split, the old species can have from 1 up to k-1 individuals, with all possible ways of splitting happening with the same probability. Note that both birth-death

and speciation events conserve the total number of individuals in the community, so that the community size is constant at all times.

Like in the local community or metacommunity model described in Chapter 2, the evolution of the random fission speciation metacommunity can be described with a master equation. Below we derive the master equation for the average number of species with abundance n (this is equation (34) in Haegeman and Etienne (2010)). We will refer to a species with abundance n as being at state n. If we denote by S_n the expected number of species being at state n and by $R(n,k)$ the transition rate from state n to another state k, we can write the following general master equation,

$$\frac{dS_n}{dt} = \sum_{k \neq n} \left[R(k,n)S_k - R(n,k)S_n \right].$$

(5.4)

The possible transitions in the random fission model happen due to birth and death of individuals and random fission speciation events. All possible transitions and their corresponding rates for a species at state n are listed in Box 5.1. Based on these, the master equation (5.4) takes the following form,

$$\frac{dS_n}{dt} = r_{n-1}S_{n-1} + r_{n+1}S_{n+1} - (2r_n + s_n)S_n + \sum_{m>n}^{J} s_m 2 f^{(m)}(n) S_m ,$$

(5.5)

where $r_n = \mu(n/J)((J-n)/(J-1))$ is the rate related to birth-death events, $s_n = vn/J$ for $n>1$ and $s_1=0$ is a rate related to speciation events and $f^{(m)}(n)=1/(m-1)$ is the probability that a species of abundance m that undergoes speciation will break into fragments $(n, m-n)$. Note that s_1, the rate of speciation for a species with one individual, should be set to zero. Even if we consider that the individual of this species is replaced by an individual of a new species when speciation happens, this has no net effect on the total number of species. In order to proceed, we need to introduce the total number of species, S, into equation (5.5). We note that S can be expressed by the sum: $S = \sum_{n=1}^{J} S_n$. Taking the time derivative of this equation gives

$\frac{dS}{dt} = \frac{d}{dt}\left(\sum_{n=1}^{J} S_n\right) = \sum_{n=1}^{J} \frac{dS_n}{dt}$, i.e. the time derivative of the total number of species can be found by summing equation (5.5) from $n=1$ to $n=J$. Doing so, leads to the following equation,

$$\frac{dS}{dt} = \sum_{n=1}^{J} r_{n-1} S_{n-1} + \sum_{n=1}^{J} r_{n+1} S_{n+1} - \sum_{n=1}^{J} (2r_n + s_n) S_n + \sum_{n=1}^{J} \sum_{m>n}^{J} s_m 2 f^{(m)}(n) S_m \, . \tag{5.6}$$

Substituting the rates and evaluating the sums of the right hand side (see Appendix G), we obtain the following differential equation,

$$\frac{dS}{dt} = v - \frac{\mu + v}{J} S_1 .$$ (5.7)

Note that equation (5.7) can also be derived from equation (58) of Etienne and Haegeman (2011).

Equation (5.7) describes the change of species richness with time as a function of S_1, which is the number of species with one individual at time t (singleton species). Since S_1 is unknown, in order to proceed we have to express S_1 as a function of S. For a community that is out of equilibrium, for example after an event of habitat loss, S_1 will be an unknown function of time. However, in a community at or close to stationarity, S_1 is a given fraction of the total species richness, S, which can be determined by the species-abundance distribution of the community. Here, in order to be able to express S_1 as function of S, we assume that the species-abundance distribution of the community is fixed throughout the relaxation process. In the random fission model, the equilibrium species-abundance distribution is very close to the broken-stick distribution (see Etienne and Haegeman (2011) equation (74) of Appendix 4). The expected number of species with n individuals in the broken-stick distribution is: $S_n = \frac{S(S-1)}{J}\left(1-\frac{n}{J}\right)^{S-2}$. So, the expected number of species with one individual is: $S_1 = \frac{S(S-1)}{J}\left(1-\frac{1}{J}\right)^{S-2}$. For $J\!>\!>\!S\!>\!>\!1$ this formula is well approximated by $S_1 \approx \frac{S^2}{J}$ (see Appendix H). Substituting into equation (5.7) leads to a differential equation for S that is of a "proper" form,

$$\frac{dS}{dt} = v - \frac{\mu + v}{J^2} S^2 .$$ (5.8)

By setting dS/dt=0 in equation (5.8) we find the equilibrium species richness,

$$S_{eq} = J\sqrt{\frac{v}{v+\mu}} .$$ (5.9)

In equation (5.9), $v+\mu$ is the total rate at which events (either births-deaths or speciation) happen in the community, so that the fraction $v'=v/(v+\mu)$ represents the probability of speciation, given that an event happens. Using v', the equilibrium species richness can be also be expressed as $S_{eq}=\sqrt{v'}\, J$.

5.3 Solution of differential equation (5.8)

Equation (5.8) can be solved by separation of variables. The step by step derivation can be found in Appendix I, which leads to our main result,

$$S = S_{eq} + \frac{2S_{eq}}{\dfrac{S_0 + S_{eq}}{S_0 - S_{eq}}e^{\gamma t} - 1}, \tag{5.10}$$

with S_{eq} the equilibrium species richness of equation (5.9), S_0 the species richness at time $t = 0$ and $\gamma = 2v/S_{eq}$ a constant determining the rate of increase or decrease of species richness, which we will be calling the **relaxation rate**.

5.4 Characteristic times of relaxation

Solving equation (5.10) for t we can find characteristic times of relaxation (see Appendix J):

$$t = \frac{1}{\gamma}\ln\left[\frac{(S + S_{eq})(S_0 - S_{eq})}{(S - S_{eq})(S_0 + S_{eq})}\right]. \tag{5.11}$$

We define as t_p the time needed for a fraction p of the total extinctions to be realized. As the total number of species that will eventually go extinct is S_0-S_{eq} and the remaining species richness when a fraction p of these extinctions happen is $S_p = S_0 - p\,(S_0 - S_{eq})$, substituting $S = S_p$ into equation (5.11) gives,

$$t_p = t(S_0 - p(S_0 - S_{eq})) = \frac{1}{\gamma} \ln \left[\frac{(S_0 + S_{eq}) - p(S_0 - S_{eq})}{(1 - p)(S_0 + S_{eq})} \right].$$ (5.12)

For $p=1/2$, equation (5.12) gives the time for half of the extinctions to happen (half-life time), which is,

$$t_{50} = \frac{1}{\gamma} \ln \left[\frac{S_0 + 3S_{eq}}{S_0 + S_{eq}} \right].$$ (5.13)

5.5 Dimensionless equations

The isolated-fragment and random fission speciation (RFS) equations (5.3) and (5.8) can be transformed to dimensionless equation by applying appropriate transformation of variables. The isolated-fragment equation can be transformed to a dimensionless equation by defining the relative species richness $s_I = S/S_0$ and relative time $t_I = t/t_{50}$. This leads to,

$$s_I = 1/(1 + t_I).$$ (5.14)

This equation is similar to that found by Terborg (1974) (equation (5.2)). To create a parameter independent version of the RFS equation, we begin from equation (5.11). The definition of new variables $s_{RF} = \left[(S - S_{eq})(S_0 + S_{eq}) \right] / \left[(S + S_{eq})(S_0 - S_{eq}) \right]$ and $t_{RF} = \gamma t$ leads to the following equation,

$$s_{RF} = \exp(-t_{RF}) \ .$$ (5.15)

This has an exponential form similar to the equation derived by Diamond (equation (5.1)). The dimensionless forms of the equations could be used to compare data from different studies. To do so, the real data have to be normalized according to the above transformations. Plotting the

data along with the dimensionless curves is a way to test the model and can also provide an estimate of the expected uncertainty when the model is used to fit a new data set.

5.6 Comparing with simulation results

The derivation of equation (5.10) is based on the assumption that the species-abundance distribution (SAD) of the community is broken-stick throughout the relaxation process. This is not generally true. For example, soon after a disturbance, the SAD can be widely different from the equilibrium broken-stick distribution. For this reason, the relaxation curve of equation (5.10) is expected to initially deviate from the full solution (i.e. the solution where the initial state is taken into account). To quantify this deviation, we compare the predictions of the equation with direct simulations of the stochastic process. A detailed description of our algorithm and the Fortran 95 source code is provided in Appendix K. We simulate three initial community configurations that are characterized by different initial SADs: a) a broken-stick SAD, b) a log-series SAD and c) an even SAD (where all species have the same number of individuals).

Figure 5.1 shows the theoretical relaxation curve of equation (5.10) along with the average relaxation curve from simulations, for the three cases. For a broken-stick initial SAD, the theoretical relaxation curve coincides with the simulation average (Figure 5.1a). For an even initial SAD, the theoretical curve initially declines more rapidly than the simulation average (21.3% maximum divergence) and converges after 80 generations (Figure 5.1b). Finally, for a log-series initial SAD, the theoretical curve initially declines more slowly than the simulation average (10.2% maximum divergence) and converges after 50 generations (Figure 5.1c). We conjecture that this difference in the initial relaxation rate is explained by the proportion of rare species in the initial community. The log-series community has more rare species than the theoretically assumed broken-stick community. As rare species quickly drift to extinction, the initial relaxation rate is faster in the log-series community. Accordingly, in a community with an even initial SAD there are no rare species and the initial relaxation rate is slower than the theoretical curve predicts, because species need on average more time to go extinct.
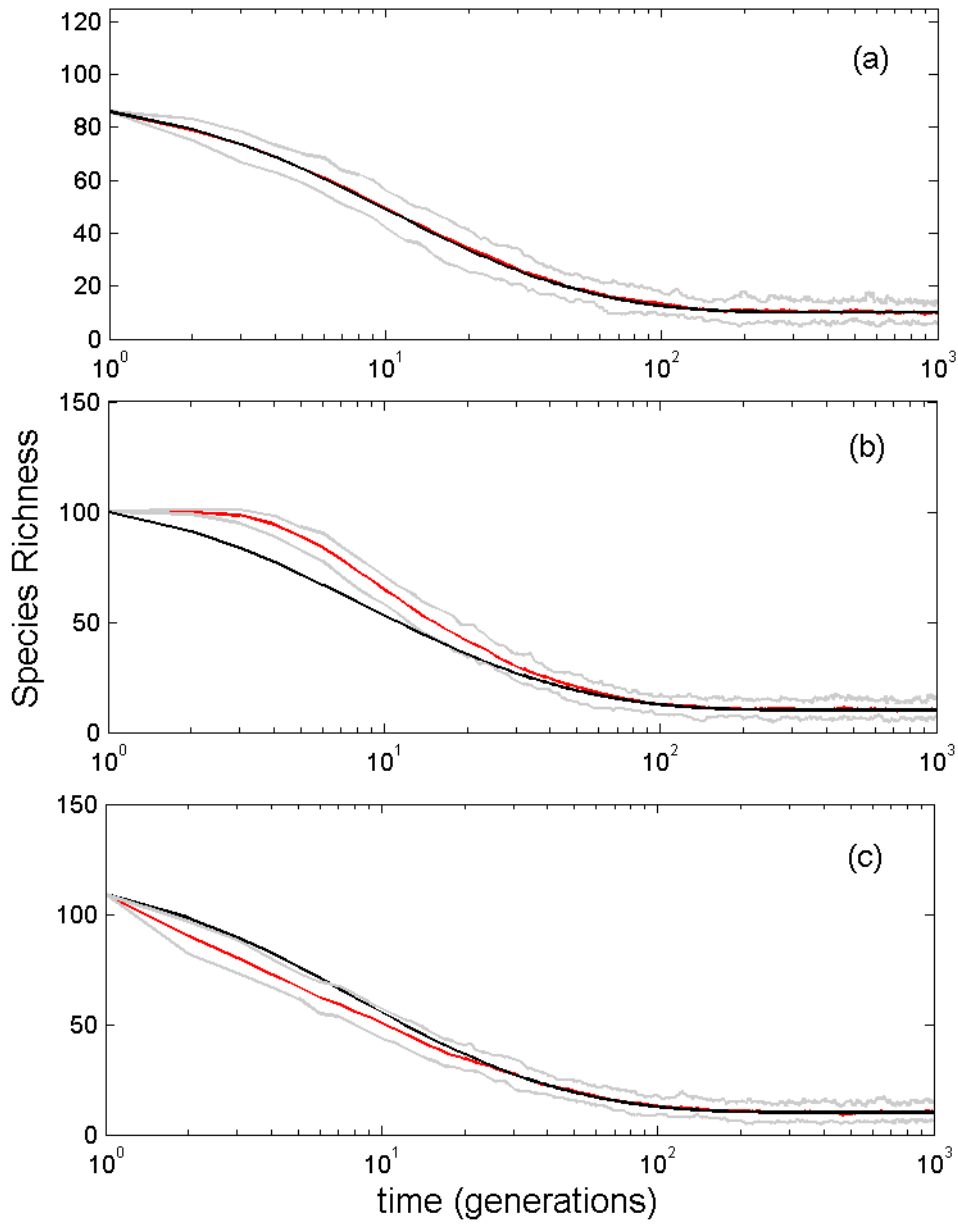
Figure 5.1 Comparison between the theoretical relaxation curve of equation (5.10) (black line) and the average relaxation curve from simulations (red line: average and grey lines: 95% confidence intervals), for three different initial conditions: a) a community with a broken-stick species-abundance distribution (SAD), (b) a community with a log-series SAD and (c) a community with an even SAD. Average and confidence intervals are calculated from 50 simulations. Parameters: Community size J=1000, speciation probability v'=0.0001.
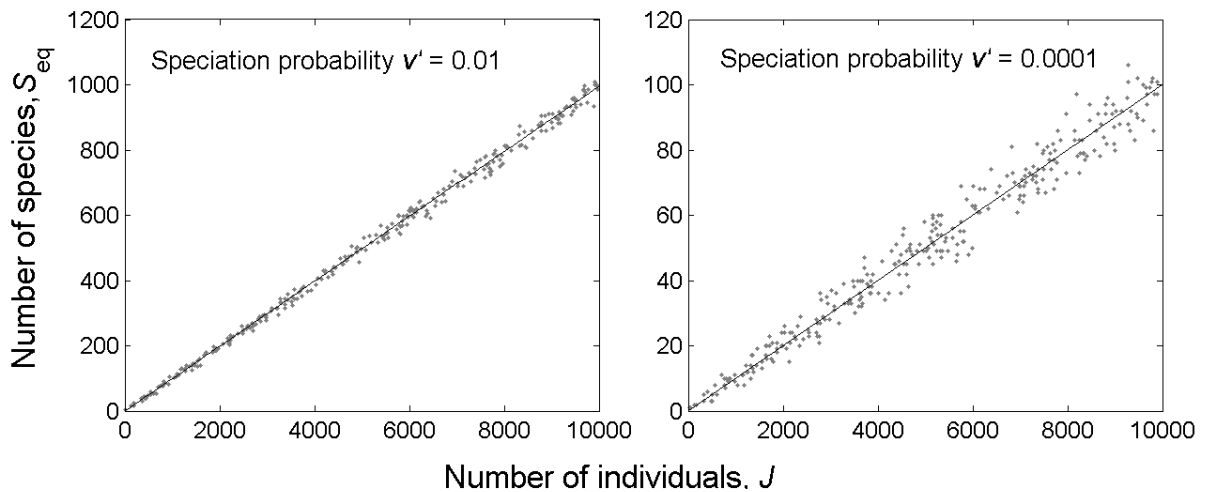
Figure 5.2 Equilibrium species richness, S_{eq}, as a function of community size, J, for two different speciation probabilities (v'=0.01 and v'=0.0001); comparison between the theoretical curve of equation (5.9) (line) and simulation results (points). Each simulation point represents the number of species at the end of a 10,000 generation run of the stochastic process. At this point the largest community (J=10,000 individuals) has reached stationarity.

Another prediction of the model is the equilibrium species richness of the community (equation (5.9)). We performed simulations for different community sizes and speciation probabilities and recorded the equilibrium species richness at the end of each simulation (after 10,000 generations). Figure 5.2 shows the equilibrium species richness as function of community size as predicted by simulations and by equation (5.9) for two different values of the speciation probability. The simulation results seem scattered symmetrically around the line hence the predictions of equation (5.9) are fairly accurate. By further investigating the parameter space we find that equation (5.9) is accurate for big enough communities ($J >$ 1,000) and low enough speciation probabilities ($v' < 0.1$) (see Appendix L). For high speciation probabilities or small communities the equilibrium species richness is systematically underestimated.

5.7 Species-Area Relationships

Equation (5.9) gives the species richness, S, of the community as a function of its size, J and the rates of birth-death, μ and speciation, v,

$$S = \sqrt{\frac{v}{v+\mu}} J \, .$$ (5.16)

Etienne and Haegeman (2011) derived a similar equation for the same model (equation (31) on page 91). This is,

$$S = \frac{{}_1F_1(1-J,1,-(v/\mu)J)}{{}_1F_1(1-J,2,-(v/\mu)J)},$$ (5.17)

with ${}_1F_1(a,b,c)$ the confluent hypergeometric function. Equations (5.16) and (5.17) can be thought of as species-individual curves, namely curves that predict the species richness of communities of different sizes. For a community saturated with individuals, the number of individuals is approximately proportional to the area of the habitat accommodating the community, so that equations (5.16) and (5.17) can be transformed to species-area relationships by substituting $J = \rho A$, where A is the habitat area and ρ is the average density. However, in order to derive the species-individual curve or species-area relationship one also has to take into account the scaling of the community level speciation and birth rates (v and μ) with the community size (J) and the number of species (S). For the community birth rate it is reasonable to assume that it is proportional to community size, namely $\mu = mJ$, where m is a constant birth rate per individual. However, it is not generally know how the speciation rate scales with J and S. As we mentioned in section 5.1.1, the speciation process is a complex spatio-temporal process, which depends greatly on the geometry of the habitat (e.g. see Yamaguchi and Iwasa 2013). Hence, the dependence of the speciation rate on J and S is not straightforward to predict. Here, we consider a general power law dependence of the speciation rate on community size and explore the predicted SAR forms. In particular, we assume that the speciation rate scales as a power of community size, namely $v=kJ^x$ with k a constant. The possible dependence on the species richness, S, is ignored as this was already not considered in deriving equation (5.16). Substituting v into equation (5.16) we get,

$$S = \sqrt{\frac{kJ^x}{kJ^x + mJ}} J = \frac{J}{\sqrt{1+(m/k)J^{1-x}}} = \frac{J}{\sqrt{1+cJ^{1-x}}} , \qquad (5.18)$$

with $c = m/k$ a constant. Below we consider two specific values of x, namely $x=0$ and $x=-1$. In the first case ($x = 0$), the community level speciation rate is constant ($v = k$ = constant, i.e. independent of community size) and equation (5.18) takes the form,

$$S = \frac{J}{\sqrt{1+(m/k)J}} = \frac{J}{\sqrt{1+J/\theta_{pm}}} \approx \sqrt{\theta_{pm}} J^{0.5} , \qquad (5.19)$$

where $\theta_{pm} = J(v/\mu) = k/m$ = const. is the fundamental biodiversity parameter as defined in the point mutation speciation model (Etienne and Alonso 2007, Haegeman and Etienne 2010 eq. (33)) and the last step results by considering that $J \gg \theta_{pm}$. In the second case ($x = -1$) the community level speciation rate is inversely proportional to community size $v = k/J$ and equation (5.18) takes the form,

$$S = \frac{J}{\sqrt{1+(m/k)J^2}} = \frac{J}{\sqrt{1+J^2/\theta_{rf}^2}} , \qquad (5.20)$$

where $\theta_{rf} = J\sqrt{(v/\mu)} = \sqrt{(k/m)} = const.$ is the fundamental biodiversity parameter as defined in the random fission speciation model (Etienne and Haegeman 2011). In Figure 5.3 we compare the species-individual curves of equations (5.19) and (5.20) with the corresponding curves resulting from equation (5.17). The curves we derived are quite good approximation of the more complex curves of Haegeman and Etienne. For values of x in the range $-1 < x \leq 0$, the resulting species-individual curves have an approximate Arrhenius form with exponents between 0 and 0.5. For $x=-1$ the curve is asymptotic, meaning that species richness is bounded below a maximum value.
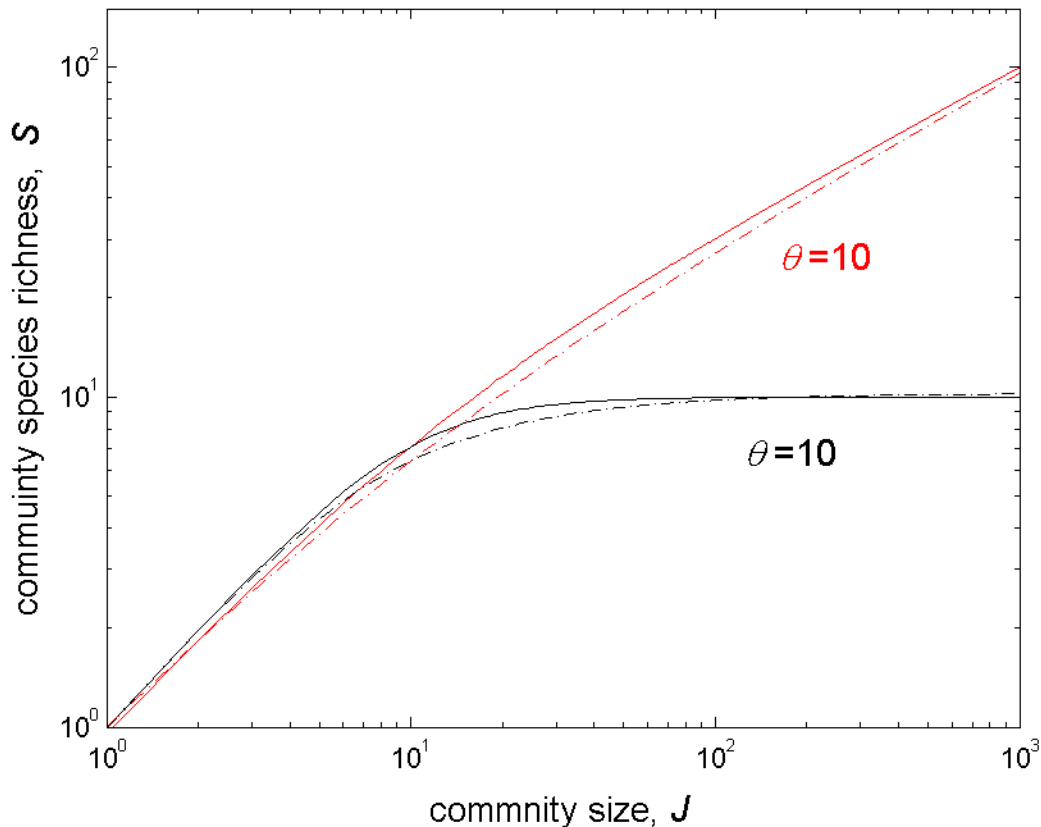
Figure 5.3 Two types of Species-Individual Curves resulting from the random fission speciation model (equation (5.16)) (solid lines). The red solid line (equation 5.19) is produced by considering that the community level speciation rate is constant (i.e. does not depend on community size, J). The resulting curve has no saturation point, namely the number of species increases indefinitely with the number of individuals. For large community sizes the curve converges to an Arrhenius curve with exponent 0.5. The black solid line (equation 5.20) is produced by scaling the speciation rate inversely proportional to community size. This has the property of keeping the species richness bounded as the number of individuals, J, increases. The dotted red and black lines are the corresponding curves produced from equation (31) of Etienne and Haegeman (2011) (equation (5.17)) (see Muller (2001) for methods of computation of the confluent hypergeometric function). All curves correspond to a constant biodiversity parameter θ=10.

5.8 Real Data Applications

Observational data on speciation under habitat loss are hard to find. In the following we parameterize the RFS relaxation curve from data of avian extinctions from islands and forest fragments that are also affected by immigration. We use the RFS mechanism to account for

the compound effects of immigration and speciation assuming that, under certain conditions, the RFS mechanism would have similar effects on species richness as immigration. Haegeman and Etienne (2010) approximated the random fission speciation mechanism using an immigration model with multiple immigrants colonizing the community at each immigration event. If immigrants replace a number of individuals of an existing species to keep the community size fixed, then immigration has the same effect for the target community as random fission speciation. There are however two basic differences between the two mechanisms: 1) while in immigration, the species to immigrate and the number of immigrants depend on the composition of the regional species pool (i.e. are determined externally); in speciation, both the species to speciate and the abundance of the new species depend on the current community composition and the fission mechanism. 2) All species created by speciation are new to the community. In immigration this is true only in the limit of an infinite species pool. For a finite species pool, the rate of inflow of new species in the community depends on the relative number of species that are common in the community and the regional species pool. Based on the above, immigration can be modeled as random fission speciation if: a) in immigration events the immigrating individuals belong to the same species and replace a number of individuals of a species of the target community. b) the broken‐stick distribution is retained at least approximately (this means that immigrants replace a random portion of individuals of an existing species) and c) the regional species pool is big and the immigration rate is low, so that most of the immigrating species are new to the community.

5.8.1 Relaxation rates in avifaunal communities

In this first example, we parameterize equation (5.10) using the same data set that was used by Halley and Iwasa (2011) to parameterize the isolation equation (5.3). The data come from different studies of avian extinctions from around the world. These are summarized in Appendix M, Table M1 (columns 1-5). All studies provide two values of species richness, the initial species richness, S_0, and a record of species richness at a later time, $S(t)$. Halley and Iwasa also provide the equilibrium species richness, S_{eq}, which is estimated based on a species-area relationship. Given the two records of species richness at times t_0 and t, we can

compute the proportion of species gone extinct in the time interval $\Delta t = t - t_0$, between the two observations. This is $p = (S_0 - S(t))/(S_0 - S_{eq})$. Knowing p and Δt we can find the relaxation rate, γ, from equation (5.12) and the time for half of the extinction to happen, t_{50}, from equation (5.13). Note that because the islands and forest fragments considered here are also largely affected by immigration, the relaxation rate will reflect the compound effects of extinctions, speciation and immigration in these communities. The estimated relaxation rates and half-life times are given in columns 6-8 of Table M1, Appendix M.

In Figure 5.4 (a) the half-lives estimated from the two models (isolated-fragment model and the RFS model) are plotted against fragment area. As illustrated by the power law fits to the data, the dependence of relaxation time on area is weaker in the RFS speciation model. For the RFS model, the half-life time increases with the square root of area ($t_{50} \sim A^{0.5}$), while for the isolated-fragment model the increase with area is much sharper ($t_{50} \sim A^{0.8}$). Furthermore, the two curves cross at around 50ha of fragment area, so that the half-life time predicted by the RFS model is longer than that predicted by the isolated-fragment model for small habitats (less than 50 ha), but is shorter for larger habitats (more than 50 ha). This result can be understood in the two limiting cases. First note that the extinctions predicted by the FRS model are always less or equal to those predicted by the isolated-fragment model, where the equilibrium is always one species. Hence, everything else being equal, the time needed to reach the equilibrium will be shorter in the RFS model. This effect is dominant in large fragments, for which the two models predict very different number of extinctions. By contrast, in smaller fragments (less than 50 hectares), the number of extinctions predicted by the two models are comparable. In this case, a slower relaxation rate in the RFS model leads to longer relaxation times. Note that the above results characterize the particular data set and don't necessarily describe a general case. In general, the relaxation time in the random fission model depends on the equilibrium species richness, the initial species richness and the speciation probability and hence its dependence on area is not straightforward.

5.8.2 Avifaunal relaxation in Barro Colorado Island

In this application we use data of avian extinction from Barro Colorado Island (BCI) to parameterize equation (5.10). Barro Colorado is a land-bridge island which was isolated in 1914 after the flooding of Gatun Lake to create the Panama Canal (Robinson 2001). Since 1920 there has been regular monitoring of the island's avifauna. The surveys conducted in early 1930s by F.M. Chapman reported 208 bird species including water birds (according to Willis and Eisenmann 1979). We use data from 4 subsequent surveys conducted in 1970, 1977, 1995 and 2006, whose findings are reported in Willis and Eisenmann (1979), Robinson (2001) and Watson (2010) (Table I). Since Willis and Eisenmann (1979) and Robinson (2001) provide lists of the species recorded by name, we were able to derive presence/absence of these species in 1970, 1977 and 1995 by comparing these lists. Following Robinson (2001), we consider only resident species including water birds but excluding migrant species. However, because the classification of species in Robinson (2001) and Willis and Eisenmann (1979) do not always coincide, we adopt the classification of Robinson (2001) and include only the species that are mentioned in both studies. The resulting list is included in Appendix M, Table M2. The last record of species richness in 2006 is derived from Table 2 of Watson 2010. In order to include the water birds, we added the 24 water birds (i.e. all 186 species minus the 162 forest species) to the 122 observed island resident species (see page 282 of Watson 2010 for the species classifications). Table 5.1 shows the species richness derived from the 5 surveys and the parameters estimated from a least-squares fit to equation (5.10). The observed data and the fitted model are plotted in Figure 5.4 (b). This predicts an equilibrium species richness of 116 ± 23 (95% confidence intervals). It is also predicted that 90% of the total extinctions will have been realized by 2079.

Table 5.1 Numbers of bird species recorded in Barro Colorado Island in 5 surveys between 1930 and 2006, after the islands isolation in 1914. The species richness reported here refers to all resident bird species, including water birds but excluding migrant species (Robinson 2001). Parameters estimated by a least-square fit of equation (5.10) to the data. Relaxation rate (γ), initial species richness (S_0) (corresponding to year 1914 or t=0), equilibrium species richness (S_{eq}) and the corresponding 95% confidence intervals in parenthesis. The half-life time (t_{50}) (the time needed for half of the extinction to happen) and the time needed for 90% of extinctions to happen estimated from equations (5.12) and (5.13) respectively.

Data source	Year of survey	t, Years since isolation	$S(t)$, Species richness
Willis and Eisenmann 1979	1930	16	208
Robinson 2001	1970	56	166
Willis and Eisenmann 1979	1977	63	161
Robinson 2001	1995	81	152
Watson 2010	2006	92	146
Non linear least square fit parameters			
γ	S_0	S_{eq}	t_{50}	t_{90}
0.0012 (0.0055 - 0.0179)	237 (228 - 245)	116 (93 – 139)	43	165
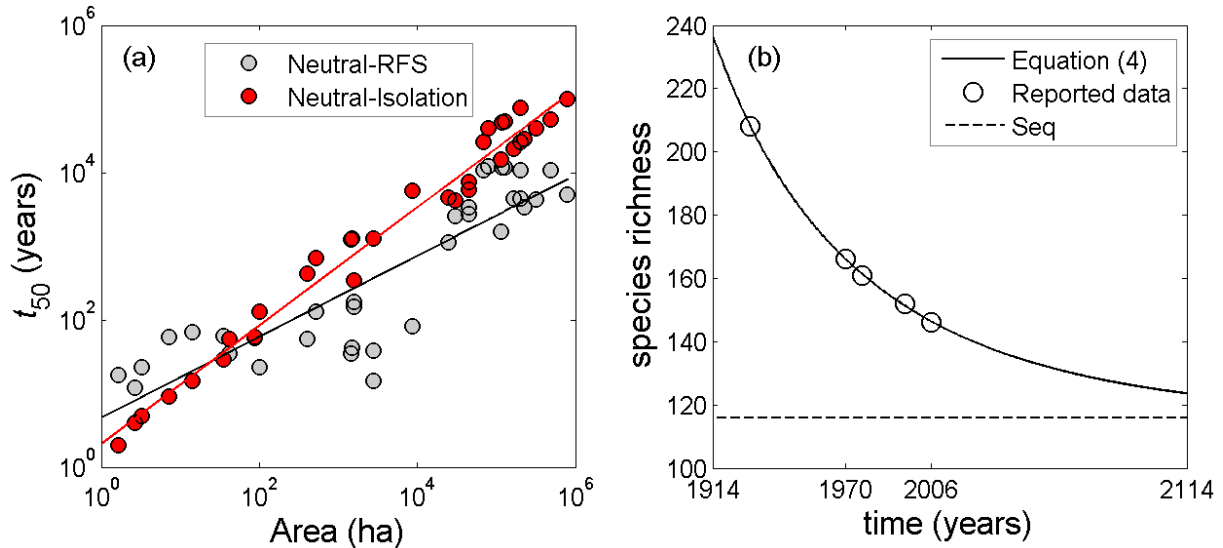
Figure 5.4 (a) Time for half of the extinctions to be realized (t_{50}) as function of the habitats' area for the avifaunal communities of Table M1, Appendix M. Half-life time estimated from the isolated-fragment model of Halley and Iwasa (2011) (black points) and from the random fission speciation model (grey points) and fitted power models $t_{50\_isol}=2A^{0.8}$ (black line) and $t_{50\_RFS}=4.7A^{0.5}$ (grey line). (b) Relaxation of the avifauna of Barro Colorado Island. Observed species richness from 5 surveys (open circles) and least-squares fit to equation (5.10) (solid line). The observed records of species richness refer to island resident species reported in Willis and Eisenmann (1979), Robinson (2001) and Watson (2010). The fit predicts 116 ± 23 (95% confidence intervals) species at equilibrium (dotted-line) and 237 ± 9 (95% confidence intervals) species in 1914, the year of the island's isolation. The time for 90% of the total extinctions to be realized is estimated to be around 65 years from now.

5.9 Discussion

Many of the results for extinction relaxation have been based on isolated communities (Richman et al. 1988, Gilbert et al. 2006, Halley and Iwasa 2011). In this study, we have generalized earlier results by deriving a closed-form expression for the variation of species richness with time in a community where new species can enter through speciation. For this we used the neutral model of biodiversity with random fission speciation (RFS). Our equation has three parameters (one more than the isolated-community model of Halley and Iwasa (2011)): the speciation rate, the initial species richness and the species richness at equilibrium. In contrast to an isolated community where the equilibrium state is fixation (one species survives), the RFS equation predicts a non-trivial equilibrium resulting from a balance between speciation and extinction. Mathematical formulas for the variation of species richness

with time (i.e. the relaxation process) have been derived using MacArthur's and Wilson's theory of island biogeography (Diamond 1972) and competition theory (Terborg 1974). However, being able to describe the relaxation process with the neutral model of biodiversity is important, as this predicts the dynamics of communities in terms of fundamental ecological processes taking place at the level of individuals. With neutral biodiversity models nowadays used as baseline models for processes that produce empirical biodiversity patterns, the model we use here allows us to explore a wider variety of factors affecting biodiversity loss.

To derive the equation we assumed that the species-abundance distribution (SAD) of the community is broken-stick throughout the relaxation process. Within the neutral model we used, this is true for a community that is at or close to a steady state (equilibrium). For a community that is far from equilibrium (e.g. just after a disturbance) the SAD can differ. To investigate the validity of our assumption, we compared the predictions of the equation with direct simulations of the stochastic process for different initial states (SADs) of the community. We find that, the assumption provides a very accurate approximation to the actual relaxation curve (i.e. the one found using simulations) for communities whose initial configuration is broken-stick. For initial species-abundance distributions (SADs) that are more even/uneven than the broken-stick, the initial relaxation rate is overestimated/underestimated. Still, it can be concluded that for commonly observed SADs (ranging from broken-stick to log-series) the initial deviation from the actual relaxation curve is less than 9%.

The equation also provides a simple expression for the species richness of the community at equilibrium, which is $S_{eq} = \sqrt{v'}\, J$, with v' the speciation probability per birth and J the community size. The above is a simplification to the more accurate formula derived by Haegeman and Etienne (2010) for the same model. By exploring the parameter space using simulations, we find that for realistic values of speciation probabilities ($v'<0.1$) and big enough communities ($J>1000$), the formula agrees very well with the simulation average. For small communities or high speciation probabilities the equilibrium species richness is systematically underestimated. Given the correct scaling of speciation rate with community size, the same formula can also provide species-individual curves or species-area relationships. Here we explored the possible species-individual curves resulting from the model by considering that the speciation rate scales as a power of community size. This results in two different types of

curves, which have either an Arrhenius or an asymptotic form. As we discuss below, these are not the only possible species-individual curve that can result from the model. Possibly a different curve will arise if a scaling of speciation rate with the number of species is included, but as this requires a completely different approach we didn't consider such a case.

The RFS equation can be used to predict the decline of species richness in real communities after catastrophic events (e.g. habitat loss). To demonstrate how this works, we parameterized the equation using data from avian extinctions. In the first application we used the data that were used by Halley and Iwasa (2011) to parameterize the isolated-fragment model. We find that the RFS model predicts shorter relaxation times than the isolated-fragment model for fragments of more than 50 hectares, but longer relaxation times for smaller fragments. By fitting a power model to the data of half-life times versus area we find that the half-life time predicted by the RFS speciation model depends less strongly on fragment area than predicted by the isolated-fragment model. In particular, this increases with the square root of area in the RFS model and with the 0.8 power of area in the isolated fragment-model. In the second example, we fit the equation to observations of avian species richness from Barro Colorado Island. This predicts that the equilibrium species richness is around 116 species and that 90% of the total extinctions will have been realized by 2079. In both the above examples immigration is likely to be playing an important role, for some of the data at least. Thus, we have used a random fission speciation mechanism to account for the compound effects of immigration and speciation. This will be a reasonable assumption a) if the regional species pool is big and immigration rate low, so that most of the immigrating species are new to the community and b) if immigration events are such that multiple individuals immigrate at once and replace a random portion of individual of a species in the community.

Although both speciation and immigration can be considered as just an inflow of new species in the community, the two processes are fundamentally different. While immigration depends on the abundances of species of the regional community (i.e. it is externally determined), the abundance distribution of the "new" species in speciation depends on the species-abundance distribution of the community itself (Haegeman and Etienne 2010). Hence, a possible extension of the model will be to include an explicit description of immigration

events. An important parameter of the random fission speciation model is the speciation rate. Here, in agreement with the original model by Hubbell (2001) and Haegeman and Etienne (2010), we considered a constant speciation rate per individual so that the speciation rate of the community is also constant given that the community size is fixed. The above is a fair assumption for an equilibrium community, in which all macroscopic characteristics are on average constant over time. However this might not be the case for a community far from equilibrium. In particular, as speciation is a species-level process, the speciation rate is also expected to depend on the number of species (Etienne et al. 2007b). Such a case has been described by Yamaguchi and Iwasa (2013) who studied a model of allopatric speciation in a group of isolated islands where there are rare immigration events. In this case, the rate of species formation is at least initially increasing with the existing number of species. In general, the speciation process is a complex spatio-temporal process that requires some kind of reproductive isolation of populations (e.g. geographic isolation). Hence, a natural extension of the model will be to combine neutral drift with an explicit modeling of the speciation mechanism. This can reveal the dependence of speciation rate on species richness and community size and give more accurate description of the relaxation process in general. Furthermore it can provide the correct form of species-area relationships.

The model we used here combines random fission speciation and neutral extinction to predict the dynamics of species richness. Although there are many reasons for which such a simple model deviates from reality, it is very important to have a simple and analytically tractable model in which all the situations can be analyzed. Such models can help crystallize our ideas on what are the leading factors affecting biodiversity loss and formulate a paradigm within which matters of biodiversity loss can be discussed in the future.

Chapter 6

Summary and Conclusions

Habitat loss is today the main causes of species extinctions and a major threat to biodiversity as a consequence of the increase of human population and activities (Millennium Ecosystem Assessment, 2005). Predicting its effects is thus an important matter in conservation biology. Currently, there is no complete theory of biodiversity and no single framework to describe biodiversity loss. The traditional SAR method for extinction forecasts, although correct in principle, gives only an equilibrium estimate of biodiversity capacity based on area and does not take into account community composition or dynamics (Halley et al. 2013). To this end, dynamical models of community organization are a helpful tool for predicting the response of ecological communities to habitat loss or other global disturbances. The purpose of this thesis was to investigate how a dynamical model, and in particular the neutral theory of biodiversity, can improve our understanding of the process of extinctions following habitat loss and whether this can provide the basis for building a unified description of the process of biodiversity loss. The model was used in three separate applications: 1) as a conceptual model to construct a framework for understanding the relaxation process after habitat loss, 2) as a null model for temporal community turnover to reveal possible effects of climate change on a given community and finally 3) to develop a simple mathematical description of the relaxation process in a habitat where new species are recruited through speciation.

 To date, the commonly-used method to predict extinctions following habitat loss is the SAR method. This predicts extinctions as the difference in the carrying capacity of the initial habitat (before the loss) and the habitats remaining after the loss. Nevertheless, the method's assumptions and predictions are questioned (Connor and McCoy 2001, He and Hubbell 2011). As we argue in Halley et al. (2013) (Chapter2, section 2.2), although much of the criticism is justified, there is nothing fundamentally wrong with the SAR method itself, while the errors arising in SAR predictions are due to unavoidable sources of uncertainty (i.e. the natural

scatter of SARs), but also due to either the misuse of the method or misinterpretation of its predictions. A common misuse of the SAR method is due to ignoring extinction debt, i.e. the fact that some of the extinctions following habitat loss are delayed in time. This can lead to two types of error. The first arises in cases where habitat loss is accompanied by environmental changes that alter the character of the remaining habitats (e.g. isolation of the remaining habitat). Ignoring this alteration will result in using a wrong functional form for the SAR of the remaining habitat (e.g. use of continental SAR instead of isolate SAR in cases of isolation), which can lead to underestimation of the total extinctions (Halley et al. 2013). The second error arises when the SAR predictions are compared with observations of species richness before delayed extinctions are realized. Then, the apparent mismatch between the observation and the prediction of the method can be falsely interpreted as an overestimation of extinctions, although this mismatch is due to ignoring extinction debt (Halley et al. 2013).

In the first application, we used the neutral model to construct a conceptual framework for understanding the process of extinctions following habitat loss. The value of this framework is that it connects the SAR method for extinctions with the dynamics of the relaxation process, thus providing a justification of the SAR method from the point of view of the community dynamics. The framework also clarifies some ideas on the origin of extinction debt. Firstly, extinction debt following habitat loss can be interpreted as the result of two phenomena. The first is the reduction of resources (i.e. the loss of area). Although the reduction of resources is immediate, few species will go directly extinct, even though many of them will have reduced populations. Thus, soon after habitat loss, the number of remaining species usually exceeds the biodiversity capacity of the remaining habitat(s). The second reason is the possible isolation of the remaining habitat(s). This further reduces their capacity, so that these cannot support the same number of species as before their isolation. In this latter case (i.e. when there is isolation), extinction debt is precisely the difference between the sample SAR describing the habitat in question and the island SAR describing the same habitat after its isolation. Finally, for common species-abundance distributions and spatial configurations, delayed extinctions (extinction debt) can be orders of magnitude greater than imminent extinctions. This result highlights the important consequences of neglecting extinction debt in extinction forecasts (Halley et al. 2014).

Together with habitat loss, the change of environmental conditions is another major driver of species extinctions. Such changes (e.g. temperature rise) can render a habitat inadequate to support some species, thus having an effect on these species that is similar to the effects of habitat loss. In the second application, we used the neutral model as a null model to reveal the possible effects of temperature rise in a community of butterflies. Approximating the natural drift of the community by neutral drift, we estimated the magnitude of temporal turnover expected in the community under equilibrium conditions. We found that neutral drift is not sufficient to explain all the observed variation of species abundances in the community. The additional turnover was mainly due to the increased abundance of species associated to warm conditions, which is consistent with the expected effects of the temperature rise that was recorded in the area. Nevertheless, the results show that natural drift can clearly lead to large rates of turnover, rates which could easily be mistaken for a response to climate change. We conclude that without an appropriate statistical test, one cannot derive safe conclusions on the origin of turnover. To this end, the neutral theory of biodiversity provides a good starting point.

Models of stochastic community drift have been used before to assess temporal turnover in real communities (Leigh et al. 1993, Gilbert et al. 2006, Ricklefs 2006). Nevertheless, the neutral theory adds an important component to these tests. Apart from accounting for stochastic demography (i.e. births and deaths of individuals), it also includes the stochasticity due to sampling effects that arises when localized samples are drawn from a community. In such a case, the composition of the sample is not the same as that of the community, as this is affected by the dispersal pattern of the individuals in the community. The implicit-space neutral model accounts for this effect by introducing dispersal-limitation between the community and the local samples, which is modeled as a restriction of the individuals of the community to enter the local communities. This is achieved by introducing one more parameter, which is the immigration rate. Although including dispersal-limitation provides a better explanation of the composition of local sample, the additional parameter can rarely be estimated directly from data. The usual methodology for estimating the model parameters is maximum likelihood parameter estimation based on observed species-abundance distributions (Etienne 2007, Jabot et al. 2008), which however is a fit rather than a

parameterization from observed data (Etienne 2007). In the present thesis we have slightly altered the maximum likelihood methodology of Jabot et al. 2008 and estimated the regional-community species richness directly from data, thus reducing the effect of the test data set on the parameterization.

In the last application, we used the neutral model to mathematically describe the relaxation process. In this, we derived a closed-form equation for the variation of species richness with time in a community that is subject to speciation, thus expanding a previous result by Halley and Iwasa (2011) that applied to fully-isolated habitats. The equation can be used to forecast the species richness at any time after habitat loss or any other disturbance. In contrast to the isolated-fragment model for which the equilibrium state is fixation to one species, here the equilibrium species richness is non-trivial and can be predicted as a function of the community size and the speciation rate. To model the recruitment of new species in the community, we used the mechanism of random fission speciation (RFS) (Haegeman and Etienne 2010). Although this restricts the use of the equation, we clarified the conditions under which the random fission mechanism can also be used to account for immigration. This allowed parameterizing the equation using real data of avian extinctions from islands and forest fragments that are also subject to immigration. Estimating the relaxation time, we found that compared to the isolated-fragment equation, the RFS equation predicts a less steep increase of the relaxation time with the area of the island or fragment, which grows as the square root of area. Parameterizing the equation from data of avian extinctions in Barro Colorado Island we found that the island is still in the relaxation process and that it should settle to around 116 bird species approximately 100 years from now.

Although an important step towards mathematically describing the relaxation process, the solution presented here refers only to the mechanism of random fission speciation and there is the need to describe the relaxation process under different mechanisms of species recruitment like immigration, point mutation speciation and other. Furthermore, an explicit description of the speciation process that models the process at the level of species and takes into account its spatial dimensions will allow a more realistic description of the relaxation dynamics (Etienne et al. 2007b, Yamaguchi and Iwasa 2013).

We have explored the use of the neutral model of biodiversity in predicting biodiversity loss following habitat loss. The neutral model assumes a simplified view of the ecological communities' assembly, which ignores interactions between species and introduces only a weak coupling through competition for limited resources. The model itself has been heavily criticized for making unrealistic assumptions. However, as noted by Rosindell et al. (2012), a model's assumptions do not need to be strictly accurate for it to provide a successful description of reality. Regarding the neutrality assumption, nobody seriously believes that species behave neutrally (Alonso et al. 2006, Rosindell et al. 2012). Namely, neutrality is not offered as a law of ecology to replace niche differentiation and adaptation. On the contrary, this is a simplifying assumption, used to predict macroecological patterns in communities of many species, for which the inclusion of species differences and interactions would create an extremely complicated model with no predictive power. Ultimately, the model's success should be judged on the basis of how well it predicts these patterns (McGill and Nekola 2010).

Unfortunately, most of the criticism against the use of the neutral model regards semantic issues that relate to the use of models in general (Rosindell et al. 2012). A common criticism questions the use of stochastic models in ecology (Clark 2009). Clark believes that stochastic processes do not describe mechanisms but the lack of mechanisms and that ecologists should always seek to find deterministic rules. We feel the need to clarify that stochastic models used in ecology and other fields of science are based on sets of rules that relate to particular mechanisms. It is precisely because of these mechanisms, that the system (here the ecological community) is more likely to be found in some states compared to others. In the end, the state of the system, in a probabilistic sense, reflects the underlying mechanisms. Hence, stochasticity is not the equivalent of a mechanism-free process and should not be perceived as pure randomness. Whether nature is inherently stochastic or deterministic is a philosophical question, but in terms of understanding nature by the use of models and theories, stochasticity and determinism can be used without any need to answer the above question. In fact, as Rosindell et al. (2012) point out, the same phenomenon can be described as stochastic or deterministic depending on the level of description that one is willing to make.

We conclude that despite the fact that some of its assumptions are unrealistic, the neutral model of biodiversity is a valuable tool for investigating the process of biodiversity

loss, as it includes the basic ingredients of community organization, namely stochastic demography, random dispersal and speciation. The neutral theory can improve our understanding of the relaxation process, by providing a qualitative description of the process (conceptualization), but also by making quantitative predictions (i.e. estimation of extinction debt). Most importantly, it gives the opportunity to model explicitly the dynamics of the relaxation process and can provide closed-form equations for the relaxation curve (i.e. the variation of species richness with time). Finally, the neutral model is an appropriate model for natural drift, as it also takes into account the sampling effects due to dispersal-limitation. Despite its simplicity, the neutral theory can provide a basis for building a unified framework for describing the effects of biodiversity loss under many conditions. Understanding the relaxation process through simple models is a first step towards creating more realistic descriptions of biodiversity loss.

Bibliography

1. Allouche, O. and Kadmon, R. (2009) A general framework for neutral models of community dynamics. *Ecology letters*, **12**:1287–1297.

2. Alonso, D. and McKane, A.J. (2004) Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters*, **7**:901–910.

3. Alonso, D., Etienne, R.S., and McKane, A.J. (2006) The merits of neutral theory. *Trends in Ecology and Evolution*, **21**:451–457.

4. Axelsen, J.B., Roll, U., Stone, L., and Solow, A. (2013) Species-area relationships always overestimate extinction rates from habitat loss: comment. *Ecology*, **94**:761–763.

5. Azaele, S., Pigolotti, S., Banavar, J. R., and Maritan, A. (2006) Dynamical evolution of ecosystems. *Nature*, **444**:926–928.

6. Bale, J.S., Masters, G.J., Hodkinson, I.D., Awmack, C., Bezemer, T.M., Brown, V.K., Butterfield, J., Buse, A., Coulson, J.C., Farrar, J., Good, J.E.G., Harrington, R., Hartley, S., Jones, T.H., Lindroth, R.L., Press, M.C., Symrnioudis, I., Watt, A.D. and Whittaker, J.B. (2002) Herbivory in global climate change research: Direct effects of rising temperature on insect herbivores. *Global Change Biology*, **8**:1−16.

7. Barry, J.P., Baxter, C.H., Sagarin, R.D. and Gilman, S.E. (1995) Climate-related, long-term faunal changes in a California rocky intertidal community. *Science*, **267**:672−675.

8. Brooks T.M., Pimm S.L., and Oyugi J.O. (1999) Time lag between deforestation and bird extinction in tropical forest fragments. *Conservation Biology*, **13**:1140–1150.

9. Brooks, T.M., Pimm S.L. and Collar N.J. (1997) Deforestation predicts the number of threatened birds in insular Southeast Asia. *Conservation Biology*, **11**:382–394.

10. Brown, J.H., Valone, T.J. and Curtin, C.G. (1997) Reorganization of an arid ecosystem in response to recent climate change. *Proceedings of the National Academy of Sciences*, **99**:9729–9733.

11. Chave, J. (2004) Neutral theory and community ecology. *Ecology letters*, **7**:241–253.

12. Chisholm, R. (2011) Time-dependent solutions of the spatially implicit neutral model of biodiversity. *Theoretical Population Biology*, **80**:71–79.

13. Chisholm, R. A. and Pacala, S. W. (2010) Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proceedings of the National Academy of Sciences*, **107**:15821–15825.

14. Clark, J.S. (2009) Beyond neutral science. *Trends in Ecology & Evolution*, **24**:8–15.

15. Clark, J.S. (2012) The coherence problem with the unified neutral theory of biodiversity. *Trends in ecology & evolution,* **27**:198–202.

16. Collen, B., Purvis, A., and Mace, G.M. (2010) When is a species really extinct? Testing extinction inference from a sighting record to inform conservation assessment. *Diversity and Distributions*, **16**:755–764.

17. Conlisk, E., Bloxham, M., Conlisk, J., Enquist, B., and Harte, J. (2007) A new class of models of spatial distribution. *Ecological Monographs*, **77**:269–284.

18. Conlisk, J., Conlisk, E. and Harte, J. (2010) Hubbell's local abundance distribution: insights from a simple colonization rule. *Oikos*, **119**:379–383.

19. Connor, E.F. and McCoy, E.D. (2001) Species–area relationships. *Encyclopedia of biodiversity*, **5**:397–411.

20. Dapporto, L. and Dennis, R.L.H. (2013) The generalist-specialist continuum: Testing predictions for distribution and trends in British butterflies. *Biological Conservation*, **157**:229–236.

21. Devictor, V., Van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölö, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, A., Reif, J., Roy, D.B., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., Wynhoff, I. and Jiguet, F. (2012) Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*, **2**:121–124.

22. Diamond, J. M. (1972) Biogeographic kinetics: estimation of relaxation times for avifaunas of southwest pacific islands. *Proceedings of the National Academy of Sciences*, **69**:3199–3203.

23. Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C. and Magurran, A.E. (2014) Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. *Science*, **344**:296–299.

24. Eberhardt, L. (1967) Some developments in 'distance sampling'. *Biometrics*, **23**:207–216.

25. Etienne, R.S. (2005) A new sampling formula for neutral biodiversity. *Ecology Letters*, **8**:253–260.

26. Etienne, R. S. (2007) A neutral sampling formula for multiple samples and an exact test of neutrality. *Ecology Letters*, **10**:608–618.

27. Etienne, R. S., Alonso, D., and McKane, A. J. (2007a) The zero-sum assumption in neutral biodiversity theory. *Journal of Theoretical Biology*, **248**:522–536.

28. Etienne, R. S., Apol, M. E. F., Olff, H. and Weissing, F. J. (2007b) Modes of speciation and the neutral theory of biodiversity. *Oikos*, **116**:241–258.

29. Etienne, R. S. and Alonso, D. (2005) A dispersal-limited sampling theory for species and alleles. *Ecology Letters*, **8**:1147–1156.

30. Etienne, R. S. and Alonso, D. (2007) Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. *Journal of Statistical Physics*, **128**:485–510.

31. Etienne, R. S. and Haegeman, B. (2011) The neutral theory of biodiversity with random fission speciation. *Theoretical Ecology*, **4**:87–109.

32. Etienne, R. S. and Rosindell, J. (2011) The spatial limitations of current neutral models of biodiversity. *PloS one*, **6**:1–8.

33. Ewens, W.J. (1972) The sampling theory of 1 selectively neutral alleles. *Theoretical Population Biology*, **3**: 87–112.

34. Fisher, D.O. and Blomberg, S.P. (2011) Correlates of rediscovery and the detectability of extinction in mammals. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**:1090–1097.

35. Gause, G.F. (1934) *The struggle for existence.* Williams & Wilkins Co. Baltimore.

36. Gilbert, B., Laurance, W. F., Leigh, E. G., and Nascimento, H. E. M. (2006) Can neutral theory predict the responses of amazonian tree communities to forest fragmentation? *The American Naturalist*, **168**:304–317.

37. Green, J. L. and Ostling, A. (2003) Endemics-area relationships: the influence of species dominance and spatial aggregation. *Ecology*, **84**:3090–3097.

38. Grill, A. and Cleary, D.F.R. (2003) Diversity patterns in butterfly communities of the Greek nature reserve Dadia. *Biological Conservation*, **114**:427–436.

39. Gotelli, N.J. and McGill, M.J. (2006) Null versus neutral models: what's the difference? *Ecography*, **29**:793–800.

40. Haegeman, B. and Etienne, R. S. (2008) Relaxing the zero-sum assumption in neutral biodiversity theory. *Journal of Theoretical Biology*, **252**:288–294.

41. Haegeman, B. and Etienne, R. S. (2010) Self-consistent approach for neutral community models with speciation. *Physical Review E*, **81**:1–13.

42. Halley, J. M. and Iwasa, Y. (2011) Neutral theory as a predictor of avifaunal extinctions after habitat loss. *Proceedings of the National Academy of Sciences*, **108**:2316–2321.

43. Halley, J. M., Sgardeli, V., and Monokrousos, N. (2013) Species–area relationships and extinction forecasts. *Annals of the New York Academy of Sciences*, **1286**:50–61.

44. Halley, J. M., Sgardeli, V., and Triantis, K. A. (2014) Extinction debt and the species–area relationship: a neutral perspective. *Global Ecology and Biogeography*, **23**:113–123.

45. Hankin, R.K.S. (2007) Introducing untb, an R package for simulating ecological drift under the unified neutral theory of biodiversity. *Journal of Statistical Software*, **22**:1–15.

46. He, F. and Gaston, K.J. (2000) Estimating species abundance from occurrence. *The American Naturalist*, **156**:553–559.

47. He, F. and Gaston, K.J. (2007) Estimating abundance from occurrence: an underdetermined problem. *The American Naturalist*, **170**:655–659.

48. He, F. and Hubbell, S.P. (2011) Species-area relationships always overestimate extinction rates from habitat loss. *Nature*, **473**:368–371.

49. He, F. and Legendre, P. (1996) On species-area relations. American Naturalist, **148**:719–737.

50. He, F. and Legendre, P. (2002) Species diversity patterns derived from species-area models. *Ecology*, **83**:1185–1198.

51. He, F., Zhang, D.-Y. and Lin, K. (2012) Coexistence of nearly neutral species. *Journal of plant ecology*, **5**:72–81.

52. Holbrook, S.J., Russell, J.S. and Stephens Jr., J.S. (1997) Changes in an Assemblage of Temperate Reef Fishes Associated with a Climate Shift. *Ecological Applications*, **7**:1299–1310.

53. Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press. Princeton, New Jersey.

54. Hubbell, S.P. (2003) Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs. *Oikos*, **100**:193–199.

55. Hutchinson, G.E. (1957) Concluding remarks. *Population Studies: Animal Ecology and Demography. Cold Spring Harbor Symposium on Quantitative Biology*, **22**:415–457.

56. Isaac, N.J.B., Girardello, M., Brereton, T.M. and D.B. Roy. (2011) Butterfly abundance in a warming climate: Patterns in space and time are not congruent. *Journal of Insect Conservation*, **15**:233–240.

57. Jabot, F., Etienne, R. S., and Chave, J. (2008) Reconciling neutral community models and environmental filtering: theory and an empirical test. *Oikos*, **117**:1308–1320.

58. Jackson, S. T. and Sax, D. F. (2010) Balancing biodiversity in a changing environment: extinction debt, immigration credit and species turnover. *Trends in Ecology and Evolution*, **25**:153–160.

59. Johnson, M. P., Mason, L. G., and Raven, P. H. (1968) Ecological parameters and plant species diversity. *The American Naturalist*, **102**:297–306.

60. Kallimanis, A.S., Kunin, W.E., Halley J.M. and Sgardelis, S.P. (2005) Metapopulation extinction risk under spatially autocorrelated disturbance. *Conservation Biology*, **19**:534-546.

61. Kimura, M. and Crow, J. F. (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**:725–738.

62. Kisel, Y. and Barraclough T.G. (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, **175**:316–334.

63. Kitzes, J. and Harte, J. (2015) Predicting extinction debt from community patterns. *Ecology*, **96**:2127–2136.

64. Koh, L. P., Lee, T. M., Sodhi, N. S., and Ghazoul, J. (2010) An overhaul of the species–area approach for predicting biodiversity loss: incorporating matrix and edge effects. *Journal of applied ecology*, **47**:1063–1070.

65. Kuussaari, M., Bommarco, R., Heikkinen, R. K., Helm, A., Krauss, J., Lindborg, R., Öckinger, E., Pärtel, M., Pino, J., Roda, F., Stefanescu, C., Teder, T., Zobel, M. and Steffan-Dewenter, I. (2009) Extinction debt: a challenge for biodiversity conservation. *Trends in ecology & evolution*, **24**:564–571.

66. Leigh Jr., E.G., Wright, S.J., Herre, E.A. and Putz, F.E. (1993) The decline of tree diversity on newly isolated tropical islands: A test of a null hypothesis and some implications. *Evolutionary Ecology*, **7**:76−102.

67. Levine, J. M., Adler, P. B. and Yelenik, S. G. (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters*, **7**:975–989.

68. Lindgren, G., Rootzén, H. and Sandsten, M. (2013) *Stationary Stochastic Processes for Scientists and Engineers*. CRC Press.

69. MacArthur, R. H. and Wilson, E. (1967) *The theory of island biogeography*. Princeton University Press. Princeton, New Jersey.

70. MacHunter, J., Wright, W., Loyn, R., and Rayment, P. (2006) Bird declines over 22 years in forest remnants in southeastern Australia: Evidence of faunal relaxation? *Canadian Journal of Forest Research*, **36**:2756–2768.

71. McGill, B. J. and Nekola, J. C. (2010) Mechanisms in macroecology: Awol or purloined letter? towards a pragmatic view of mechanism. *Oikos*, **119**:591–603.

72. McGill, B.J., Hadly, E.J. and Maurer, B.A. (2005) Community inertia of Quaternary small mammal assemblages in North America. *Proceedings of the National Academy of Sciences*. **102**: 16701−16706.

73. McKane, A., Alonso, D., and Solé, R.V. (2000) Mean-field stochastic theory for species-rich assembled communities. *Physical Review E*, **62**:8466–8484.

74. McKane, A., Alonso, D. and Solé, R.V. (2004) Analytic solution of Hubbell's model of local community dynamics. *Theoretical population biology*, **65**:67–73.

75. Millennium Ecosystem Assessment (2005) *Ecosystems and human well-being: biodiversity synthesis*. World Resources Institute, Washington, DC.

76. Merrill, R.M., Gutiérrez, D., Lewis, O.T., Gutiérrez, J., Díez, S.B. and Wilson, R.J. (2008) Combined effects of climate and biotic interactions on the elevational range of a phytophagous insect. *Journal of Animal Ecology*, **77**:145−155.

77. Meshinev, T., Apostolova, I. and Koleva, E. (2000) Influence of warming on timberline rising: a case study on Pinus peuce Griseb. in Bulgaria. *Phytocoenologia*, **30**:431−438.

78. Muller. K.E. (2001) Computing the confluent hypergeometric function $M(a,b,x)$. *Numerical Mathematics*, **90**:179-196.

79. O'Dwyer, J. P. and Green, J. L. (2010) Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. *Ecology letters*, **13**:87−95.

80. Pamperis, L.N. (2009) *The butterflies of Greece*. Pamperis Editions.

81. Parmesan, C. (1996) Climate and species' range. *Nature*, **382**:765−766.

82. Parmesan, C. (2006) Ecological and evolutionary responces to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**: 637−669.

83. Parmesan, C. and Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**:37−42.

84. Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon, H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W.J., Thomas, J.A. and Warren, M. (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, **399**:579−583.

85. Pateman, R.M., Hill, J.K., Roy, D.B., Fox, R. and Thomas, C.D. (2012) Temperature-dependent alterations in host use drive rapid range expansion in a butterfly. *Science*, **336**: 1028−1030.

86. Pereira, H. M., Borda-de Água, L., and Martins, I. S. (2012) Geometry and scale in species area relationships. *Nature*, **482**:368−371.

87. Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. L., Chini, L. Cooper, H.D., Gilman, E.L., Guénette, S., Hurtt, G.C., Huntington, H.P., Mace, G.M., Oberdorff, T., Revenga, C., Rodrigues, P., Scholes, R.J., Sumaila, U.R. and Walpole, M. (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**:1496−1501.

88. Perry, J. and Taylor, L. (1985) Ades: new ecological families of species-specific frequency distributions that describe repeated spatial samples with an intrinsic power-law variance-mean property. *The Journal of Animal Ecology*, **54**:931–953.

89. Pimm, S.L. and Askins, R.A. (1995) Forest losses predict bird extinctions in eastern North America. *Proceedings of the National Academy of Sciences USA*, **92**:9343–9347.

90. Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K. , Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., O'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F., Thompson, S.A. and Richardson, A.J. (2013) Global imprint of climate change on marine life. *Nature Climate Change*, **3**:919−925.

91. Preston, F. W. (1962) The canonical distribution of commonness and rarity: Part II. *Ecology*, **43**:410–432.

92. Priddel, D., Carlile, N., Humphrey, M., Fellenberg, S. and Hiscox, D. (2003) Rediscovery of the 'extinct' Lord Howe Island stick-insect (*Dryococelus australis* (Montrouzier))(Phasmatodea) and recommendations for its conservation. *Biodiversity and Conservation*, **12**:1391−1403.

93. Purves, D. W. and Turnbull, L. A. (2010) Different but equal: the implausible assumption at the heart of neutral theory. *Journal of Animal Ecology*, **79**:1215–1225.

94. Renshaw, E. (1991) *Modelling biological populations in space and time*. Cambridge University Press.

95. Richman, A.D., Case, T.J., and Schwaner, T.D. (1988) Natural and unnatural extinction rates of reptiles on islands. *American Naturalist*, **131**:611–630.

96. Ricklefs, R. E. (2003) A comment on Hubbell's zero-sum ecological drift model. Oikos, **100**:185–192.

97. Ricklefs, R.E. (2006) The Unified Neutral Theory of Biodiversity: Do the Numbers Add Up? *Ecology*, **87**:1424−1431.

98. Robinson, N., Armstead, S. and Bowers, M.D. (2012) Butterfly community ecology: the influences of habitat type, weather patterns, and dominant species in a temperate ecosystem. *Entomologia Experimentalis et Applicata*, **145**:50−61.

99. Robinson, W.D. (2001) Changes in abundance of birds in a Neotropical forest fragment over 25 years: a review. *Animal Biodiversity and Conservation*, **24**:51–65.

100. Rosenzweig, M. L. (2001) Loss of speciation rate will impoverish future diversity. Proceedings of the National Academy of Sciences, **98**:5404–5410.

101. Rosenzweig, M.L. (1995) *Species Diversity in Space and Time*. Cambridge University Press. Cambridge.

102. Rosindell, J. and Cornell, S. J. (2007) Species-area relationships from a spatially explicit neutral model in an infinite landscape. *Ecology letters*, **10**:586–95.

103. Rosindell, J. and Cornell, S. J. (2009) Species-area curves, neutral models, and long-distance dispersal. *Ecology*, **90**:1743–1750.

104. Rosindell, J., Cornell, S. J., Hubbell, S. P., and Etienne, R. S. (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology letters*, **13**:716–27.

105. Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., and Etienne, R. S. (2012) The case for ecological neutral theory. *Trends in ecology & evolution*, **27**:203–208.

106. Rosindell, J., Hubbell, S.P., and Etienne, R.S. (2011) The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology and Evolution*, **26**:340–348.

107. Sagarin, R.D., Barry, J.P., Gilman, S.E. and Baxter, C.H. (1999) Climate-related change in an intertidal community over short and long time scales. *Ecological Monographs*, **69**:65–490.

108. Scheffers, B.R., Yong, D.L., Harris, J.B.C., Giam, X. and Sodhi, N.S. (2011) The world's rediscovered species: back from the brink. *PloS ONE*, **6**: e22531. doi: 10.1371/journal.pone.0022531

109. Scheiner, S.M. (2003) Six types of species-area curves. *Global Ecology and Biogeography*, **12**:441–447.

110. Schweiger, O., Harpke, A., Wiemers, M. and Settele, J. (2014). CLIMBER: Climatic niche characteristics of the butterflies in Europe. *Zookeys*, **367**:65–84.

111. Southward, A.J., Hawkins, S.J. and Burrows, M.T. (1995) Seventy years' observations of changes in distribution and abundance of zooplankton and intertidal organisms in the

western English Channel in relation to rising sea temperature. *Journal of Thermal Biology*, **20**:127−155.

112. Stefanescu, C., Torre, I., Jubany, J. and Páramo, F. (2011) Recent trends in butterfly populations from north-east Spain and Andorra in the light of habitat and climate change. *Journal of Insect Conservation*, **15**:83−93.

113. Strickland, J. (2011) *Simulation Conceptual Modeling*. Lulu.com.

114. Terborgh, J. (1974) Preservation of natural diversity: The problem of extinction prone species. *Bioscience*, **24**:715−722.

115. Tjørve, E. (2009) Shapes and functions of species–area curves (ii): a review of new models and parameterizations. *Journal of Biogeography*, **36**:1435–1445.

116. Tjørve, E. and Turner, W. R. (2009) The importance of samples and isolates for species–area relationships. *Ecography*, **32**:391–400.

117. Triantis, K.A., Borges, P.A.V., Ladle, R.J., Hortal, J., Cardoso, P., Gaspar, C., Dinis, F., Mendonça, E., Silveira, L.M.A., Gabriel, R., Melo, C., Santos, A.M.C., Amorim, I.R., Ribeiro, S.P., Serrano, A.R.M., Quartau, J.A. and Whittaker, R.J. (2010) Extinction debt on oceanic islands. *Ecography*, **33**:285–294.

118. Triantis, K.A., Guilhaumon, F., and Whittaker, R.J. (2012) The island species-area relationship: biology and statistics. *Journal of Biogeography*, **39**:215–231.

119. Vallade, M. and Houchmandzadeh, B. (2003) Analytical solution of a neutral model of biodiversity. *Physical review E*, **68**:1–5.

120. Van Kampen, N.G. (1992) *Stochastic Processes in Physics and Chemistry*. Elsevier. Amsterdam.

121. Van Swaay, C., Cuttelod, A., Collins, S., Maes, D., López Munguira, M., Šašić, M., Settele, J., Verovnik, R., Verstrael, T., Warren, M., Wiemers, M. and Wynhof, I. (2010) *European Red List of Butterflies*. Publications Office of the European Union.

122. Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**:183–206.

123. Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003) Neutral theory and relative species abundance in ecology. *Nature*, **424**:1035–1037.

124. Walther, G.-R. Post, E. Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O. and Bairlein, F. (2002) Ecological responses to recent climate change. *Nature*, **416**:389−395.

125. Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B., Telfer, M.G., Jeffcoate, S., Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N., Moss, D. and Thomas, C.D. (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature*, **414**:65−69.

126. Watson, D.M. (2010) Optimizing inventories of diverse sites: insights from Barro Colorado Island birds. *Trends in Ecology and Evolution*, **1**:280-291.

127. Wilcoxon, F. (1992) *Individual comparison by ranking methods. Breakthroughs in Statistics*. Springer, New York.

128. Willis, E.O. and E. Eisenmann (1979) *A Revised List of Birds of Barro Colorado Island, Panama. Smithsonian Contribution to Zoology*, no. **291**, Smithsonian Institution Press, Washington.

129. Wilson, E.O. (1988) The current state of biological diversity. *Biodiversity* (ed. by E.O. Wilson and F.M. Peter), pp. 3–18. National Academy Press, Washington, DC.

130. Wilson, R.J., Gutiérrez, D., Gutiérrez, J. and Monserrat V.J. (2007) An elevational shift in butterfly species richness and composition accompanying recent climate change. *Global Change Biology*, **13**:1873−1887.

131. Yamaguchi, R. and Iwasa, Y. (2013) First passage time to allopatric speciation. *Interface Focus*: **3**:20130026. doi: 10.1098/rsfs.2013.0026.

132. Zografou, K., Kati, V., Grill, A., Wilson, R.J., Tzirkalli, E., Pamperis, L.N. and Halley, J.M. (2014) Signals of Climate Change in Butterfly Communities in a Mediterranean Protected Area. *PloS ONE*, **9**: e87245. doi:10.1371/journal.pone.0087245.

Appendices

Appendix A: Derivation of the metacommunity species-abundance distribution (equation 2.13)

To find the stationary (equilibrium) distribution for p_n we set $\mathrm{d}p_n(t)/\mathrm{d}t = 0$ in equation (2.12). This leads to $g_{n-1}p_{n-1} + r_{n+1}p_{n+1} - (r_n+g_n)p_n = 0$. Writing down this equation for all n gives:

$$n=0 \qquad\qquad r_1 p_1 - g_0 p_0 = 0 \Rightarrow \qquad\qquad p_1 = \frac{g_0}{r_1}p_0$$

$$n=1 \qquad\qquad g_0 p_0 + r_2 p_2 - (r_1 + g_1)p_1 = 0 \Rightarrow \qquad\qquad p_2 = \frac{g_1}{r_2}p_1 = \frac{g_0 g_1}{r_1 r_2}p_0$$

$$\dots$$

$$\text{for } n \qquad\qquad g_n p_n + r_{n+1}p_{n+1} - (r_n + g_n)p_n = 0 \Rightarrow \qquad\qquad p_n = \frac{g_0 \cdot g_1 \cdot \dots \cdot g_{n-1}}{r_1 \cdot r_2 \cdot \dots \cdot r_n}p_0$$

So, the probability that a species has abundance n at equilibrium is,

$$p_n = \prod_{i=1}^{n} \frac{g_{i-1}}{r_i}p_0, \qquad\qquad (A1)$$

with $g_0 = v/S_0$, $g_n = (1-v)\frac{n}{J_M}\frac{J_M - n}{J_M - 1}$ for $n>0$ and $r_n = \frac{n}{J_M}\left((1-v)\frac{J_M - n}{J_M - 1} + v\right)$. Before substituting the rates g_n and r_n into equation (A1), we can first sum the equation over all species in order to get the average number of species with abundance n, which we denote by $S_n \equiv E(S_n|\theta, J_M)$. Note, that since all species have the same probability p_n to have abundance n, we can simply write $S_n = J_M p_n$ for each n, where J_M is the total number of species including species with zero abundance. This gives,

$$S_n = \sum_{i=1}^{J_M} p_n = \sum_{i=1}^{J_M} \frac{g_0 \cdot g_1 \cdot \ldots \cdot g_{n-1}}{r_1 \cdot r_2 \cdot \ldots \cdot r_n} p_0 = S_0 \prod_{i=0}^{n-1} \frac{g_i}{r_{i+1}}, \tag{A2}$$

where S_0 is the average number of species with zero abundance. Next we substitute the rates g_n and r_n into equation (A2), which gives,

$$S_n = v\,J_M \prod_{i=1}^{n-1} \frac{\dfrac{i(1-v)(J_M-i)}{J_M(J_M-1)}}{\dfrac{(i+1)\big((1-v)(J_M-i-1)+v(J_M-1)\big)}{J_M(J_M-1)}} \Rightarrow$$

$$S_n = v\,J_M \prod_{i=1}^{n-1} \frac{i(1-v)(J_M-i)}{(i+1)\big(J_M-1-i(1-v)\big)} \Rightarrow \quad S_n = \frac{v\,J_M}{n} \prod_{i=1}^{n-1} \frac{J_M-i}{\dfrac{J_M-1}{1-v}-i}. \tag{A3}$$

Introducing the fundamental biodiversity parameter $\theta = \dfrac{(J_M-1)}{1-v}v$ and changing the product index as $i \to i+1$ gives,

$$S_n = \frac{v\,J_M}{n} \prod_{i=2}^{n} \frac{J_M-i+1}{\theta+J_M-i}. \tag{A4}$$

Finally, multiplying and dividing by the factor corresponding to $i=1$, we derive equation (2.13):

$$S_n = \frac{v\,J_M}{n} \frac{J_M+\theta-1}{J_M} \prod_{i=1}^{n} \frac{J_M+1-i}{J_M+\theta-i} \Rightarrow \quad S_n = \frac{v}{n} \frac{J_M-1}{1-v} \prod_{i=1}^{n} \frac{J_M+1-i}{J_M+\theta-i} \Rightarrow$$

$$E(S_n \mid \theta, J_M) \equiv S_n = \frac{\theta}{n} \frac{\left(J_M+1-n\right)_n}{\left(J_M+\theta-n\right)_n}, \tag{A5}$$

140

where $(x)_N = \prod_{i=1}^{N}(x+i-1)$ denotes the rising factorial.

Appendix B: Derivation of the local community species-abundance distribution (equation 2.18)

In order to find the stationary (equilibrium) distribution for $p_{n,k}$ we set $dp_{n,k}(t)/dt = 0$ in equation (2.17). This leads to $g_{n-1,k}p_{n-1,k} + r_{n+1,k}p_{n+1,k} - (r_{n,k}+g_{n,k})p_{n,k} = 0$. Note that in this case, each species has different metacommunity abundance and thus a different probability of having abundance n in the local community. For this reason, we retain the species index, k. Writing down the resulting equation for all n gives,

$n=0$ $r_1 p_{1,k} - g_{0,k} p_{0,k} = 0 \Rightarrow$ $p_{1,k} = \dfrac{g_{0,k}}{r_{1,k}} p_{0,k}$

$n=1$ $g_{0,k} p_{0,k} + r_{2,k} p_{2,k} - (r_{1,k} + g_{1,k}) p_{1,k} = 0 \Rightarrow$ $p_{2,k} = \dfrac{g_{1,k}}{r_{2,k}} p_{1,k} = \dfrac{g_{0,k} g_{1,k}}{r_{1,k} r_{2,k}} p_{0,k}$

$\dots$

for n $g_{n,k} p_{n,k} + r_{n+1,k} p_{n+1,k} - (r_{n,k} + g_{n,k}) p_{n,k} = 0 \Rightarrow$ $p_{n,k} = \dfrac{g_{0,k} \cdot g_{1,k} \cdot \dots \cdot g_{n-1,k}}{r_{1,k} \cdot r_{2,k} \cdot \dots \cdot r_{n,k}} p_{0,k}$

So, the probability that the equilibrium abundance of species k is n is,

$$ p_{n,k} = \frac{g_{0,k} p_{0,k}}{r_{1,k}} \prod_{i=2}^{n} \frac{g_{i-1}}{r_i} . \tag{B1} $$

Substituting the rates $g_{n,k} = \dfrac{J-n}{J}\left((1-m)\dfrac{n}{J-1} + mx_k \right)$ and $r_{n,k} = \dfrac{n}{J}\left((1-m)\dfrac{J-n}{J-1} + m(1-x_k) \right)$ into equation (B1) gives,

$$p_{n,k} = \frac{J\,m\,x_k\,p_{0,k}}{1-m\,xk}\prod_{i=2}^{n}\frac{\dfrac{J-(i-1)}{J}\left[\dfrac{i-1}{J-1}(1-m)+mx_k\right]}{\dfrac{i}{J}\left[\dfrac{J-i}{J-1}(1-m)+m(1-x_k)\right]} \Rightarrow$$

$$p_{n,k} = \frac{J\,m\,x_k\,p_{0,k}}{1-m\,x_k}\prod_{i=2}^{n}\frac{J-(i-1)}{i}\,\frac{(i-1)(1-m)+(J-1)mx_k}{(J-i)(1-m)+(J-1)m(1-x_k)} \Rightarrow$$

$$p_{n,k} = \frac{m\,x_k\,p_{0,k}}{1-m\,x_k}\binom{J}{n}\prod_{i=2}^{n}\frac{(i-1)(1-m)+(J-1)mx_k}{(J-i)(1-m)+(J-1)m(1-x_k)},$$

where we used the combinatorial $\binom{J}{n}=\dfrac{J!}{n!(J-n)!}$. Multiplying and dividing by the factor corresponding to $\underline{i}$=1 leads to,

$$p_{n,k} = \frac{m\,x_k\,p_{0,k}}{1-m\,x_k}\binom{J}{n}\frac{(J-1)(1-m)+(J-1)m(1-x_k)}{(J-1)mx_k}\prod_{i=1}^{n}\frac{(i-1)(1-m)+(J-1)mx_k}{(J-i)(1-m)+(J-1)m(1-x_k)} \Rightarrow$$

$$p_{n,k} = p_{0,k}\binom{J}{n}\prod_{i=1}^{n}\frac{(i-1)(1-m)+(J-1)mx_k}{(J-i)(1-m)+(J-1)m(1-x_k)}.$$

Introducing the fundamental dispersal number $I = \dfrac{(J-1)}{1-m}m$,

$$p_{n,k} = p_{0,k}\binom{J}{n}\prod_{i=1}^{n}\frac{i-1+Ix_k}{J-i+I(1-x_k)} \Rightarrow$$

$$p_{n,k} = p_{0,k}\binom{J}{n}(Ix_k)_n\prod_{i=1}^{n}\frac{1}{J-i+I(1-x_k)} \Rightarrow$$

$$p_{n,k} = p_{0,k} \binom{J}{n} (Ix_k)_n \prod_{i=1}^{n} \frac{1}{J - i + I(1 - x_k)} \cdot \prod_{i=1}^{J} \frac{J - i + I(1 - x_k)}{J - i + I(1 - x_k)} \Rightarrow$$

$$p_{n,k} = p_{0,k} \binom{J}{n} (Ix_k)_n \frac{\prod_{i=n+1}^{J} J - i + I(1 - x_k)}{\prod_{i=1}^{J} J - i + I(1 - x_k)} .$$

Next we apply the index transformation $i \rightarrow J-i+1$ and use the definition of the rising factorial

$$(x)_N = \prod_{i=1}^{N} (x + i - 1),$$

$$p_{n,k} = p_{0,k} \binom{J}{n} (Ix_k)_n \frac{\prod_{i=1}^{J-n} i - 1 + I(1 - x_k)}{\prod_{i=1}^{J} i - 1 + I(1 - x_k)} \Rightarrow$$

$$p_{n,k} = p_{0,k} \binom{J}{n} \frac{(Ix_k)_n (I(1 - x_k))_{J-n}}{(I(1 - x_k))_J} . \tag{B2}$$

As we are interested in the average number of species having abundance n, namely $S_n \equiv E(S_n|m,\theta,J)$, we have to sum equation (B2) over all species, namely:

$$S_n = \sum_{k=1}^{S_M} p_{n,k} . \tag{B3}$$

Note that in equation (B3), $p_{n,k}$ is not the same for all species, because every species has different relative metacommunity abundance, x_k. Although x_k is not known for every species, the species-abundance distribution of the metacommunity is given by equations (2.13) and (2.14). Hence, we can estimate S_n as follows,

$$S_n = \sum_{x=1}^{J_M} E(S_x \mid \theta, J_M) \, p_L(n \mid x),$$

<div align="right">(B4)</div>

where $E(S_x \mid \theta, J_M)$ is the expected number of species with abundance x in the metacommunity (equation (2.13)) and $p_L(n \mid x)$ is the probability that a local species has abundance n given that its metacommunity abundance is x. Note that $p_L(n \mid x)$ is provided by equation (B2). Accordingly, we can write a similar equation using the probability density, $p(x)$, of equation (2.14) for the metacommunity species-abundance distribution. This gives:

$$S_n = \int_0^1 p(x) p_L(n \mid x) dx = \int_0^1 p_L(0 \mid x) \frac{\theta(1-x)^{\theta-1}}{x} \binom{J}{n} \frac{(Ix)_n \left(I(1-x)\right)_{J-n}}{\left(I(1-x)\right)_J} dx,$$

<div align="right">(B5)</div>

where $p_L(n \mid x)$ is the same as above. Summing equation (B5) over n, from $n=1$ to $n=J$ gives the total number of species in the local community and provides an expression for $p_L(0 \mid x)$ (see Etienne and Alonso 2007). Finally, this leads to equation (2.18),

$$E(S_n \mid m, \theta, J) \equiv S_n = \theta \binom{J}{n} \int_0^1 \frac{(Ix)_n \left((1-x)\right)_{J-n}}{(I)_J} \frac{(1-x)^{\theta-1}}{x} dx.$$

<div align="right">(B6)</div>

Cited Literature

Etienne, R. S. and Alonso, D. (2007) Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. *Journal of Statistical Physics*, **128**:485–510.

Appendix C: Fortran95 source code applying Conlisk's colonization rule

The following subroutine finds the species-abundance vector of Hubbell's local community by applying the colonization rule described in Conlisk et al. (2010). The subroutine receives the relative abundances of species in the metacommunity (**p**), the number of species in the metacommunity (Sm), the immigration probability per birth (m) and the size of the local community (Jr) and computes the species-abundance vector of the local community (**N**) and the number of species (S).

```fortran
subroutine Conlisk_colonisation (N,S,Jr,p,m,Sm)

!N      integer vector dim(1,Sm), abundances of species in the local
!       community
!S      integer, number of species in the local community
!Jr     integer, number of individuals in the local community
!Sm     integer, number of species in the metacommunity
!p      real vector dim (1,Sm), relative abundances of species in the
!       metacommunity (fixed)
!m      real, immigration probability per birth
!Pr     real vector dim(1,Sm), probability of species to colonize the
!       community
!cumPr  real vector dim(1,Sm), cumulative sum of Pr

IMPLICIT double precision (A-H,O-Z)
integer i,j,N,S,Jr,Sm,ind,index
real m,p,Pr,A,h,cumPr
dimension N(Sm),Pr(Sm),cumPr(Sm),p(Sm)

A=m*(Jr-1)    !! a constant parameter
N=0           !! initialize the abundance vector
N(1)=1        !! begin with a random species having one individual in the
              !! local community

do j=2,Jr
  call random_number(r)    !! random number to select species to colonize
  ind=sum(N)               !! current number of individuals
  h=A/(A+(1-m)*ind)        !! a parameter
  Pr(1)=h*p(1)+((1-h)*N(1))/ind  !! probability that species 1 will
                                 !! colonize
   cumPr(1)=Pr(1)          !! first element of cumulative sum
   do i=1,Sm               !! repeat until species to colonize is found
     if (cumPr(i)>=r) then
        index=i            !! species i was selected to colonize
        exit               !! exit the loop
```

```
      end if
      Pr(i+1)=h*p(i+1)+((1-h)*N(i+1))/ind  !!no species was selected
      cumPr(i+1)=cumPr(i)+Pr(i+1)          !! compute the next element of cumPr
    end do
    N(index)=N(index)+1      !! add one individual to species i=index
  end do                     !! repeat until all Jr individuals colonize
  S=count(N/=0)              !! find the number of species in local community
end subroutine Conlisk_colonisation
```

Cited Literature

Conlisk, J., Conlisk, E. and Harte, J. (2010) Hubbell's local abundance distribution: insights from a simple colonization rule. *Oikos*, **119**: 379–383.

Appendix D: Immigration probability as a function of habitat's area (derivation of equation (3.6))
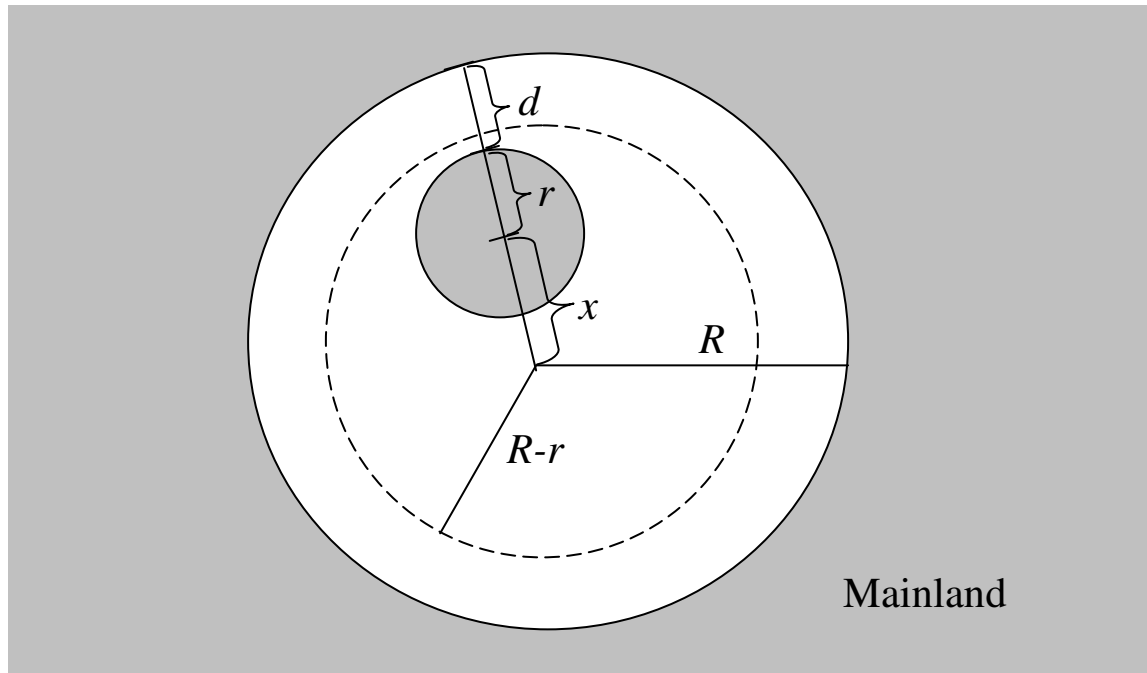
Figure D1. A simplified scenario of fragmentation. The white circular area of radius R represents a former habitat area that has been destroyed, e.g. a habitat basin that is flooded with water. After the loss of habitat, a circular fragment of radius r, remains within the basin. This is separated by distance d from the mainland, while its center lies in distance x from the center of the basin. The dotted line defines the area within which the center of the fragment of radius r should lie so as this is not connected to the mainland.

Figure D1 describes an idealized case of fragmentation in which a circular plot of radius R is destroyed (e.g. a basin is flooded with water), but leaves a circular fragment of radius r intact within the destroyed habitat. The remaining fragment is surrounded by inhospitable matrix (e.g. water, denoted by the white color) and lies in distance d from the mainland. For the remaining fragment of radius r to be considered an island, its center should lie within a circle of radius R-r from the center of the plot (i.e. the area delimited by the dotted line in Figure D1). If the fragmentation happens in a non systematic way, we can ask whether the expected distance of the island from the mainland is related to the fragment's area. We assume that the

fragmentation happens in such a way that the center of the fragment could be anywhere within the destroyed habitat, i.e. the fragment's center could fall anywhere within the circle of radius R-r with the same probability. The probability that the center falls within area A_1 from the center of the plot is,

$$P(A_1) = \frac{A_1}{\pi(R-r)^2} ,$$ (D1)

where A_1 is a circular plot of radius x in the center of the plot, namely $A_1 = \pi x^2$. Thus, the probability that the center of the fragment falls within distance x from the center of the plot is,

$$P(x) = P(A_1) = \frac{A_1}{\pi(R-r)^2} = \frac{\pi x^2}{\pi(R-r)^2} = \frac{x^2}{(R-r)^2} ,$$ (D2)

for which the associated probability density function is,

$$f(x) = \frac{dP(x)}{dx} = \frac{2x}{(R-r)^2} ,$$ (D3)

namely $f(x)dx$ is the probability that the center of the fragment is found within distance x and x+dx. The expected value of x, or in other words the average distance of the fragment's center from the center of plot is:

$$< x >= \int_0^{R-r} xf(x)dx = \int_0^{R-r} \frac{2x^2}{(R-r)^2} dx = \frac{2x^3}{3(R-r)^2}\bigg|_0^{R-r} = \frac{2}{3}(R-r).$$ (D4)

From Figure D1 we can see that x is related to d like,

$$R = d + r + x \quad \text{or} \quad d = R - r - x ,$$ (D5)

from where we can find the expected d as,

$$< d >= R - r - < x >= R - r - \frac{2}{3}(R - r) = \frac{1}{3}(R - r) \,.$$

(D6)

Hence, the average distance of an island's edge from the mainland decreases proportionally with the island's radius. Now we assume that the immigration probability decreases proportionally with the distance d from the mainland. Namely,

$$m' = m_0 - k \frac{d}{R} \,,$$

(D7)

where m_0 is the immigration probability when $d=0$, in which case the fragment is connected to the mainland and $k \leq m_0$ is a constant. Substituting the average d from equation (D6) into equation (D7) leads to,

$$m' = m_0 - \frac{k}{3} \frac{R - r}{R} \Rightarrow m' = m_0 - \frac{k}{3} \left(1 - \frac{r}{R}\right) \Rightarrow m' = m_0 - \frac{k}{3} \left(1 - \sqrt{\frac{a}{A}}\right) \,.$$

(D8)

This has to satisfy the boundary conditions $m'(a=0)=0$ and $m'(a=A)=m_A$. The first condition suggests that when there is no fragment area left (i.e. $a=0$) there is no immigration (i.e. $m'=0$). For this to be satisfied the constant k should be equal to $3m_0$. The second condition ensures that when the remaining fragment's area is $a=A$ (i.e. there is no habitat loss) the immigration probability should equal that of a sample area of size A. This second condition is satisfied if $m_0=m_A$, so that equation (D8) reads,

$$m' = m_A \left(\frac{a}{A}\right)^{0.5} \,.$$

(D9)

Appendix E: Maximum likelihood estimates of model parameters

Table E1. Parameter values and Maximum Likelihood Estimates (MLE) of model parameters for all habitats (local samples) in 1998 and 2011. Metacommunity: number of species, S_M, number of individuals, J_M and fundamental biodiversity parameter, θ. Local samples: sample size, J, MLE of immigration probability, $\hat{m}$, MLE of fundamental dispersal number, $\hat{I}$, maximum log likelihood, Loglike and the corresponding values using Jabot's method, $\hat{m}_J$, $\hat{I}_J$ and Loglike$_J$.

Metacommunity		S_M			J_M		θ
		128			172,260		13.544

Local communities	J	$\hat{m}$	$\hat{I}$	Loglike	$\hat{m}_J$	$\hat{I}_J$	Loglike$_J$
Local samples (1998)							
Agriculture	395	0.064	27.13	-153.88	0.090	38.87	-145.71
Dry Meadow	118	0.187	26.87	-83.44	0.201	29.35	-86.29
Grazed Pasture	288	0.049	14.69	-128.19	0.076	23.52	-117.76
Mixed Forest	754	0.071	57.69	-233.35	0.086	70.81	-225.15
Oak Forest	570	0.058	34.96	-158.98	0.107	67.98	-141.55
Pine Forest	445	0.034	15.85	-119.61	0.040	18.73	-118.88
Wet Meadow	314	0.048	15.79	-115.72	0.06	18.33	-112.58
Local samples (2011)							
Agriculture	557	0.095	58.42	-231.14	0.091	55.93	-227.63
Dry Meadow	283	0.068	20.43	-134.19	0.093	28.96	-117.10
Grazed Pasture	259	0.127	37.47	-163.25	0.128	38.03	-156.78
Mixed Forest	553	0.125	79.10	-201.42	0.220	155.88	-172.83
Oak Forest	569	0.071	43.23	-176.32	0.097	61.21	-161.95
Pine Forest	408	0.023	9.74	-141.20	0.028	11.73	-131.18
Wet Meadow	619	0.054	35.32	-164.53	0.066	43.77	-156.25

Table E2. Maximum Likelihood Estimates (MLE) of immigration probabilities estimated after re-sampling the bigger sample (either 1998 or 2011) to make the samples from both years equal in size (see section 4.2.1). Sample size, J, MLE of immigration probability and fundamental dispersal number for 1998 and 2011 samples, $\hat{m}_{98}$, $\hat{I}_{98}$ and $\hat{m}_{11}$, $\hat{I}_{11}$.

Local samples	J	$\hat{m}_{98}$	$\hat{I}_{98}$	$\hat{m}_{11}$	$\hat{I}_{11}$
Agriculture	395	0.064	27.13	0.118	52.68
Dry Meadow	118	0.187	26.87	0.139	18.90
Grazed Pasture	259	0.044	11.88	0.127	37.47
Mixed Forest	553	0.079	47.48	0.125	79.10
Oak Forest	569	0.048	28.76	0.071	43.23
Pine Forest	408	0.029	12.30	0.023	9.74
Wet Meadow	314	0.048	15.79	0.080	27.28

Appendix F: Fortran95 source to simulate the local community dynamics

The following subroutine iterates the local community dynamics for one generation (Jn steps). The subroutine receives the local community size (Jn), the immigration probability (m), the metacommunity species richness (Sm), the metacommunity species-abundance vector (**Nm**) and the local community species-abundance vector (**n**) and number of species (St). It updates and outputs the new local species-abundance vector (**n**) and species richness (St). The iteration proceeds in discrete time steps. At each step one of two events can happen: 1) with probability m, a random individual dies and a random individual from the metacommunity takes its place. In this case, a species is selected with probability proportional to its abundance to loss one individual and another species is selected with probability proportional to its metacommunity abundance to gain one individual, 2) with probability 1-m, a random individual dies and another random individual gives birth. In this case, two species are selected with probabilities proportional to their local abundances to loss one individual and gain one individual respectively. The abundances of the two species are updated together (after these have been picked), so that the individual that dies can still contribute to the pool of offsprings competing for the vacant site. This corresponds to the Moran version of the neutral model (see Etienne and Alonso 2007).

```fortran
subroutine iterate(Jn,n,m,St,Sm,Nm)

!Jn      integer parameter, number of individuals in local community
!Sm      integer parameter, number of species in metacommunity
!n       integer vector dim(1,Sm), number of individuals of each species
!        in the local community
!Nm      real vector dim(1,Sm), number of individuals of each species in
!        the metacommunity (fixed)
!St      integer, number of species in the local community
!m       real parameter, immigration probability per step
!cum     integer vector dim(1,Sm), cumulative sum of n
!cumm    integer vector dim(1,Sm) cumulative sum of Nm

IMPLICIT double precision (A-H,O-Z)
integer Jn,n,St,spec1,spec2,cum,Sm,cumm,q,p,s1
real m,r1,r2,mm,Nm,s2
dimension n(Sm),cum(Sm),Nm(Sm),cumm(Sm)
```

```fortran
logical flag1,flag2

s1=sum(n)    !! total number of individuals in local community (integer)
s2=sum(Nm)   !! total number of individuals in metacommunity (real)

do q = 1,Jn                  !! repeat for Jn steps (one generation)
   cum(1) = n(1)             !! first element of cumulative sum of n
   cumm(1) = Nm(1)           !! first element of cumulative sum of Nm
   call random_number(r1)    !! random number to select species 1
   call random_number(r2)    !! random number to select species 2
   call random_number(mm)    !! random number to decide immigration from
                             !! the metacommunity (event 1) or birth of a
                             !! local species (event 2)
!--------IMMIGRATION------------------------------------------------

 if (mm<=m) then         !! with probability m perform immigration (event 1)
     do p = 1,Sm         !! loop to select the species to loss an individual
        if (cum(p)>=r1*s1)then    !! check if the pth species is selected
            spec1=p              !! the pth species is selected
          exit                   !! exit the do loop
        endif                    !! a species is not yet selected
        cum(p+1) = cum(p) + n(p+1)!! find the next element of the
     enddo                       !! cumulative sum and repeat until
                                 !! the species is selected
     do p = 1,Sm         !! loop to select the species to immigrate
       if (cumm(p)>=r2*s2)then   !! check if the pth species is selected
           spec2=p              !! pth species is selected to immigrate
           exit                 !! exit the loop
       endif                    !! a species is not yet selected
        cumm(p+1) = cumm(p) + Nm(p+1)!! find the next cumulative sum
     enddo                       !! entry and repeat until a species is
                                 !! selected
!-----Update abundances--------------------------------------------
     n(spec1)= n(spec1)-1  !! spec1 losses one individual (death)
     n(spec2)= n(spec2)+1  !! spec2 gains one individual (immigration)


!------LOCAL BIRTH-------------------------------------------------
   else if (mm>m) then  !! with probability 1-m a local individual
                        !! gives birth (event 2)
     flag1=.true.       !! becomes false if the first species is selected
     flag2=.true.       !! becomes false if the second species is selected
     do p = 1,Sm        !! repeat until both species are selected (death
                        !! and birth)
       if (cum(p)>r1*s1.and. flag1)then !! check if the pth species is
          spec1=p                       !! selected (for death)
          flag1=.false.                 !! the pth species was selected
       endif                            !! update flag
       if (cum(p)>r2*s1.and. flag2)then !! check if the pth species is
          spec2=p                       !! selected (for birth)
          flag2=.false.                 !! pth species was selected
       endif                            !! update flag
       if (.not. flag1 .and. .not. flag2) exit  !! if both species have
                                                !! been selected exit
       cum(p+1) = cum(p) + n(p+1)       !! if not, find next cumulative
```

154

```
    enddo                                !! sum entry and repeat until
                                         !! both species are selected
!----Update abundances----------------------------------------------
    n(spec1)= n(spec1)-1                 !! spec1 losses an individual
    n(spec2)= n(spec2)+1                 !! spec2 gains an individual
   end if
end do
St = count(n/=0)                   !! update the total number of species
end subroutine iterate
```

Cited Literature

Etienne, R. S. and Alonso, D. (2007) Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. *Journal of Statistical Physics*, **128**:485–510.

Appendix G: Derivation of differential equation (5.7)

To simplify the right hand side of equation (5.6) we change the summation indices in the first and second sums, by applying the transformations $n \rightarrow n\text{-}1$ and $n \rightarrow n\text{+}1$ respectively:

$$\frac{dS}{dt} = \sum_{n=0}^{J-1} r_n S_n + \sum_{n=2}^{J+1} r_n S_n - \sum_{n=1}^{J} (2r_n + s_n) S_n + \sum_{n=1}^{J} \sum_{m>n}^{J} s_m 2 f^{(m)}(n) S_m \, . \tag{G1}$$

We can further simplify the first sum by noting that its first element, $r_0 S_0$, equals to zero, since $s_0=0$ and that the missing element $r_J S_J$ is also zero, since $s_J=0$. Hence, the first sum is equal to $\sum_{n=1}^{J} r_n S_n$. Accordingly for the second sum, we note that the last element, $r_{J+1} S_{J+1}$, is zero because $S_{J+1}=0$. By adding and subtracted $-r_1 S_1$ this reads $-r_1 S_1 + \sum_{n=1}^{J} r_n S_n$. Substituting into equation (G1) gives:

$$\frac{dS}{dt} = \sum_{n=1}^{J} r_n S_n - r_1 S_1 + \sum_{n=1}^{J} r_n S_n - \sum_{n=1}^{J} (2r_n + s_n) S_n + \sum_{n=1}^{J} \sum_{m>n}^{J} s_m 2 f^{(m)}(n) S_m \, . \tag{G2}$$

Now, the first three sums have the same range and can be combined to give:

$$\frac{dS}{dt} = -r_1 S_1 - \sum_{n=1}^{J} s_n S_n + \sum_{n=1}^{J} \sum_{m>n}^{J} s_m \frac{2}{m-1} S_m \, . \tag{G3}$$

Substituting the rates of r_n and s_n and noting that $s_1=0$ leads to the following equation:

$$\frac{dS}{dt} = -\frac{\mu}{J} S_1 - \frac{\nu}{J} \sum_{n=2}^{J} n S_n + \frac{2\nu}{J} \sum_{n=1}^{J} \sum_{m>n}^{J} \frac{m}{m-1} S_m \, . \tag{G4}$$

In order to evaluate the second sum we note that $\sum_{n=1}^{J} nS_n$ gives the total number of individual in the community, J. Thus we can write $\sum_{n=2}^{J} nS_n = \sum_{n=1}^{J} nS_n - S_1 = J - S_1$, which leads to:

$$\frac{dS}{dt} = -\frac{\mu}{J}S_1 - \frac{v}{J}(J - S_1) + \frac{2v}{J}\sum_{n=1}^{J}\sum_{m>n}^{J}\frac{m}{m-1}S_m \,.$$ (G5)

The outer sum in equation (G5) has the effect of repeating each terms m-1 times, so that this can simplify to:

$$\frac{dS}{dt} = -\frac{\mu}{J}S_1 - \frac{v}{J}(J - S_1) + \frac{2v}{J}\sum_{m=2}^{J}\frac{(m-1)m}{m-1}S_m \Rightarrow$$

$$\frac{dS}{dt} = -\frac{\mu}{J}S_1 - \frac{v}{J}(J - S_1) + \frac{2v}{J}\sum_{m=2}^{J}mS_m \,.$$ (G6)

As before, the sum on the right hand side equals to J-S_1. Substituting, we obtain equation (5.7):

$$\frac{dS}{dt} = -\frac{\mu}{J}S_1 - \frac{v}{J}(J - S_1) + \frac{2v}{J}(J - S_1) \Rightarrow$$

$$\frac{dS}{dt} = v - \frac{\mu + v}{J}S_1 \,.$$ (G7)

Appendix H: Number of species with one individual in the broken-stick distribution

The number of species with one individual for the broken-stick species-abundance distribution is:

$$S_1 = \frac{S(S-1)}{J}\left(1-\frac{1}{J}\right)^{S-2}. \tag{H1}$$

For $J \gg 1$ we can Taylor expand $\left(1-\frac{1}{J}\right)^{S-2}$ near $1/J = 0$. Setting $x=1/J$ and $f(x)=(1-x)^{S-2}$ we

have $\quad f(x) = f(0) + f'(0)x + O(x^2) = 1 - (S-2)x + O(x^2) \approx 1 - \frac{S-2}{J} \approx 1 - \frac{S}{J}, \quad$ for $\quad S \gg 1.$

Substituting back to equation (H1): $S_1 \approx \frac{S^2}{J}\left(1-\frac{S}{J}\right) = \frac{S^2}{J} - \frac{S^3}{J^2}$. If also $S/J \ll 1$, we can keep

only the first term, so that $S_1 \approx \frac{S^2}{J}$.

Appendix I: Solution of differential equation (5.8)

Equation (5.8) $\dfrac{dS}{dt} = v - \dfrac{\mu + v}{J^2} S^2$ can be solved by separation of variables. By setting $c=v$ and $a = -(\mu+v)/J^2$ this takes the form:

$$\frac{dS}{aS^2 + c} = dt \Rightarrow \qquad \int \frac{dS}{aS^2 + c} = \int dt \; . \tag{I1}$$

A general solution to the integral of the left hand side can be found in Abramowitz (1964), in "integrals of rational algebraic functions" (page 12). This is:

$$S = \frac{1}{\sqrt{-4ac}} \ln \left| \frac{2aS - \sqrt{-4ac}}{2aS + \sqrt{-4ac}} \right| . \tag{I2}$$

Substituting in equation (I1) and replacing the constants a and c we get:

$$\frac{1}{\sqrt{4v(\mu+v)/J^2}} \ln \left| \frac{-2\dfrac{(\mu+v)}{J^2} S - \sqrt{\dfrac{4v(\mu+v)}{J^2}}}{-2\dfrac{(\mu+v)}{J^2} S + \sqrt{\dfrac{4v(\mu+v)}{J^2}}} \right| + C = t \Rightarrow$$

$$\frac{J}{2\sqrt{v(\mu+v)}} \ln \left| \frac{-(\mu+v)S - J\sqrt{v(\mu+v)}}{-(\mu+v)S + J\sqrt{v(\mu+v)}} \right| + C = t$$

Dividing by $(\mu+v)$ and replacing the equilibrium species richness $S_{eq} = J\sqrt{\dfrac{v}{v+\mu}}$:

$$\frac{J}{2(\mu+v)\sqrt{v/(v+\mu)}}\ln\left|\frac{-S-J\sqrt{v/(v+\mu)}}{-S+J\sqrt{v/(v+\mu)}}\right|+C=t\Rightarrow\frac{J^2}{2(\mu+v)S_{eq}}\ln\left|\frac{-S-S_{eq}}{-S+S_{eq}}\right|+C=t\Rightarrow$$

$$\frac{J^2}{2(\mu+v)S_{eq}}\ln\left|\frac{S+S_{eq}}{S-S_{eq}}\right|=C+t.$$

Setting $\gamma=\dfrac{2(\mu+v)S_{eq}}{J^2}=2\dfrac{v}{S_{eq}^2}S_{eq}=\dfrac{2v}{S_{eq}}$ gives:

$$\ln\left|\frac{S+S_{eq}}{S-S_{eq}}\right|=\gamma C+\gamma t\Rightarrow\left|\frac{S+S_{eq}}{S-S_{eq}}\right|=C'e^{\gamma t}. \tag{I3}$$

Considering that $S>S_{eq}$, which is true when the initial species richness is above equilibrium (namely $S_0>S_{eq}$), we have:

$$\frac{S+S_{eq}}{S-S_{eq}}=C'e^{\gamma t}\Rightarrow\frac{2S_{eq}}{S-S_{eq}}=C'e^{\gamma t}-1\Rightarrow S=S_{eq}+\frac{2S_{eq}}{C'e^{\gamma t}-1}. \tag{I4}$$

Substituting the initial condition $S(t=0)=S_0$ gives:

$$S_0=S_{eq}+\frac{2S_{eq}}{C'-1}\Rightarrow S_0-S_{eq}=\frac{2S_{eq}}{C'-1}\Rightarrow C'=\frac{S_0+S_{eq}}{S_0-S_{eq}}. \tag{I5}$$

Finally, substituting back to equation (I4) we derive equation (5.8). It can be shown that for $S<S_{eq}$ and $S_0<S_{eq}$ the same equation (I5) is derived.

Cited Literature

Abramowitz, M. (1964) *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*. United States Department of Commerce, National Bureau of Standards. United States of America.

Appendix J: The time for a fraction p of the extinctions to be realized, equations (5.11) and (5.12)

To find times to extinction we solve equation (5.10) for t:

$$S = S_{eq} + \frac{2S_{eq}}{\dfrac{S_0 + S_{eq}}{S_0 - S_{eq}}e^{\gamma t} - 1} \Rightarrow \quad S - S_{eq} = \frac{2S_{eq}(S_0 - S_{eq})}{(S_0 + S_{eq})e^{\gamma t} - (S_0 - S_{eq})} \Rightarrow$$

$$e^{\gamma t} = \frac{2S_{eq}(S_0 - S_{eq})}{(S - S_{eq})(S_0 + S_{eq})} + \frac{(S_0 - S_{eq})}{(S_0 + S_{eq})} \Rightarrow \quad e^{\gamma t} = \frac{S_{eq}S_0 - S_{eq}^2 + S_0 S - S_{eq}S}{(S - S_{eq})(S_0 + S_{eq})} \Rightarrow$$

$$e^{\gamma t} = \frac{(S_0 - S_{eq})(S + S_{eq})}{(S - S_{eq})(S_0 + S_{eq})} \Rightarrow \quad t = \frac{1}{\gamma}\ln\frac{(S_0 - S_{eq})(S + S_{eq})}{(S - S_{eq})(S_0 + S_{eq})}. \tag{J1}$$

We can find the time, t_p, it takes for a fraction p of the total extinctions to be realized (equation (5.12)). The remaining species richness at this time will be $S = S_0 - p\,(S_0 - S_{eq})$, so:

$$t_p = t(S_0 - p(S_0 - S_{eq})) = \frac{1}{\gamma}\ln\left[\frac{(S_0 - S_{eq})(S_0 - p(S_0 - S_{eq}) + S_{eq})}{(S_0 - p(S_0 - S_{eq}) - S_{eq})(S_0 + S_{eq})}\right] \Rightarrow$$

$$t_p = \frac{1}{\gamma}\ln\left[\frac{(S_0 + S_{eq}) - p(S_0 - S_{eq})}{(1 - p)(S_0 + S_{eq})}\right]. \tag{J2}$$

Substituting $p=1/2$ we find the time needed for half of the extinctions to happen (equation (5.13)).

Appendix K: Fortran95 source code to simulate the dynamics of a metacommunity with random fission speciation

The following subroutine is to iterate the dynamics of a metacommunity subject to random fission speciation for one generation (Jr steps). As the total number of species in the metacommunity is not fixed, the process can be more easily described using the unlabeled species description (see Haegeman and Etienne 2010). In this case, the state of the metacommunity is described by the species-abundance distribution, namely the vector $\mathbf{Sp}=(S_{p1},S_{p2},S_{p3},\ldots,S_{pJr})$, with S_{pk} the number of species having abundance k and Jr the maximum abundance that a species can have. The subroutine receives the metacommunity size (Jr), the speciation probability (v), the metacommunity species-abundance distribution ($\mathbf{Sp}$), the number of individuals per abundance class ($\mathbf{N}$) and the total number of species in the metacommunity (S1) and updates and outputs the new species-abundance distribution ($\mathbf{Sp}$), the new total species richness (S1) and the new individuals per abundance class ($\mathbf{N}$). The iteration proceeds in discrete time steps. At each step one of two events can happen: 1) with probability 1- v there is a death event followed by a birth event. In this case a random species is selected (with probability proportional to its abundance) to lose one individual (death) and another species is selected (with probability proportional to its abundance) to gain one individual (birth), 2) with probability v there is a speciation event. In this case a species is selected with probability proportional to its abundance to undergo speciation. The selected species divides into two fragments in a random way. One of the fragments will correspond to the new species and the other remains to the old species. If a singleton species is selected to speciate, then the individual of this species is replaced by an individual of the new species. As this has no net effect on the species-abundance distribution ($\mathbf{Sp}$) or the species richness (S1), there is no need to update the vectors $\mathbf{Sp}$ and $\mathbf{N}$ in this case.

```
subroutine iterate1(Jr,Sp,v,S1,N)

!Jr    integer parameter, number of individuals in the metacommunity
!S1    integer variable, number of species in the metacommunity
!Sp    integer vector dim(1,Jr), number of species with abundance k
!N     integer vector dim(1,Jr), number of individuals of each abundance
!      class. N(k) is the sum of individual of all species with
```

```
!       abundance k (i.e. N(k)=Sp(k)*k).
!cum   integer vector dim(1,Jr), cumulative sum of N
!v     real parameter, speciation probability per step


IMPLICIT LOGICAL (A-Z)
integer Sp,S1,q,p,l,b,k,cum,Jr,N
real v,vv,r1,r2,r3,r4,r5,x
logical flag1,flag2
dimension Sp(Jr),N(Jr),cum(Jr)

do q = 1,Jr          !! Iterate for Jr steps
  do p = 1,Jr        !! Create vector N=(N1,...,NJr) with the abundance
     N(p)=Sp(p)*p    !! of each class. N(p) is the total number of
  end do             !! individuals of class p (i.e. the sum of
                     !! individuals of all species with p individuals)
  cum(1) = N(1)      !! first element of cumulative sum of N
  call random_number(r1)  !! to select species to lose 1 individual
  call random_number(r2)  !! to select species to gain 1 individual
  call random_number(vv)  !! to decide speciation or birth-death
  call random_number(r3)  !! to select abundance (l) of new species
  call random_number(r4)  !! to check if the same species is selected
                          !! for birth and death
  call random_number(r5)  !! to decide which species speciates

! ------BIRTH-DEATH-------------------------------------------------

  if (vv>v) then     !! with probability 1-v perform birth-death
    flag1=.true.     !! to check if species to die is selected
    flag2=.true.     !! to check if species to give birth is selected
    do p = 1,Jr
      if (cum(p)>=r1*Jr .and.flag1) then !! pick the class that losses
        k=p                              !! 1 individual (death)
        flag1=.false.              !! class k was selected, update flag
      endif
      if (cum(p)>=r2*Jr .and.flag2) then !! pick the class that gains
        b=p                              !! 1 individual (birth)
        flag2=.false.              !! class b is selected, update flag
      endif
      if (.not. flag1 .and. .not. flag2) exit  !! if both classes are
                                               !! selected, exit
          cum(p+1) = cum(p) + N(p+1)           !! if not, compute the
                                               !! next element of
      enddo                                    !! the cumulative sum


!--------Update the SAD vector---------------------------------------

      if (b==k) then    !! if the same class was selected to gain and
                        !! lose an individual
         x=1.0/Sp(k)    !! with probability x, the same species was
                        !! selected
        if (r4>=x) then  !! if the same species was not selected update
          Sp(k)=Sp(k)-1  !! a species from class k lost an individual
          Sp(b)=Sp(b)-1  !! a species from class b gained an individual
```

163

```fortran
        if (k/=1) then !! if the species that lost an individual had
          Sp(k-1)=Sp(k-1)+1  !! more than 1 individuals, then one
        end if              !! more species of class k-1 appears
        Sp(b+1)=Sp(b+1)+1    !! a species of class b+1 appears due to
      end if                !! birth at class b

    else                !! if b≠k (the same class was not selected)
      Sp(k)=Sp(k)-1   !! a species from class k losses an individual
      Sp(b)=Sp(b)-1   !! a species from class b gained an individual
      if (k/=1) then  !! if the species that lost an individual had
          Sp(k-1)=Sp(k-1)+1  !! more than 1 individuals, then one more
      end if              !! species of class k-1 appears
      Sp(b+1)=Sp(b+1)+1    !! a species of class b+1 appears due to
    end if                !! birth at class b

!-------SPECIATION-------------------------------------------------

  else if (vv<=v) then   !! with probability v perform speciation
    do p = 1,Jr
      if (cum(p)>=r5*Jr) then  !! pick the class that undergoes
        k=p                  !! speciation
        exit                 !! class k was picked
      endif
      cum(p+1)=cum(p)+ N(p+1) !! if no class was picked compute the
    end do                    !! next element of the cumulative sum

!-----------Update the SAD vector-----------------------------------
    if (k==1) then            !! if a singleton species is selected
      continue                !! continue without updating
    else                      !! else perform the split and update
      l=int(r3*(k-1))+1   !! l is the abundance of the newly created
                          !! species which is picked from a uniform
                          !! distribution between 1 and k-1 (where k
                          !! is the abundance of the species that
                          !! speciated). int() gives the integer part
      Sp(k)=Sp(k)-1       !! a species from class k divided(speciated)
      Sp(l)=Sp(l)+1       !! a species of class l appeared
      Sp(k-l)=Sp(k-l)+1   !! a species of class k-l appeared
    end if
  end if
end do
S1=sum(Sp)                 !! update number of species
end subroutine iterate1
```

Cited Literature

Haegeman, B. and Etienne, R.S. (2010) Self-consistent approach for neutral community models with speciation. *Physical Review E*, **81**:1–13.

Appendix L: Simulations

We performed direct simulations of the stochastic process emanating from the random fission speciation model described in Haegeman and Etienne (2010). The function used to iterate the community dynamics for one generation, written in Fortran 95 programming language, can be found in Appendix F. For each of the numerical experiments described below, we carried out the simulation for 10,000 generations, which is enough time for the largest community of $J = 10,000$ individuals to reach stationarity. In order to get the average behavior, each numerical experiment is repeated 1,000 times. Each time the seed of the random number generator is set in order to produce a different sequence of random numbers.

In the following we compare the equilibrium species richness of equation (5.9) ($S_{eq\_model} = J \sqrt{v'}$) with the equilibrium species richness from simulations (S_{eq_sim}), for several values of the speciation probability, v', and the community size, J. For each pair (v', J) we record the species richness at the end of a 10,000 generations simulation and repeat this numerical experiment 1,000 times to get the average behaviour. Figure L1 shows the relative difference between the theoretical value of species richness (equation 5.9) and the one resulting from simulations, computed as: $RDiff = (S_{eq\_model} - S_{eq\_sim}) / S_{eq\_model}$. For high speciation probabilities ($v'=0.1$), the model systematically underestimates the average equilibrium from simulations (relative error less than 5%). For speciation probabilities below $v'=0.01$ the relative difference fluctuates around 0, so there is no systematic under or overestimation except for very low community sizes, $J<1000$.
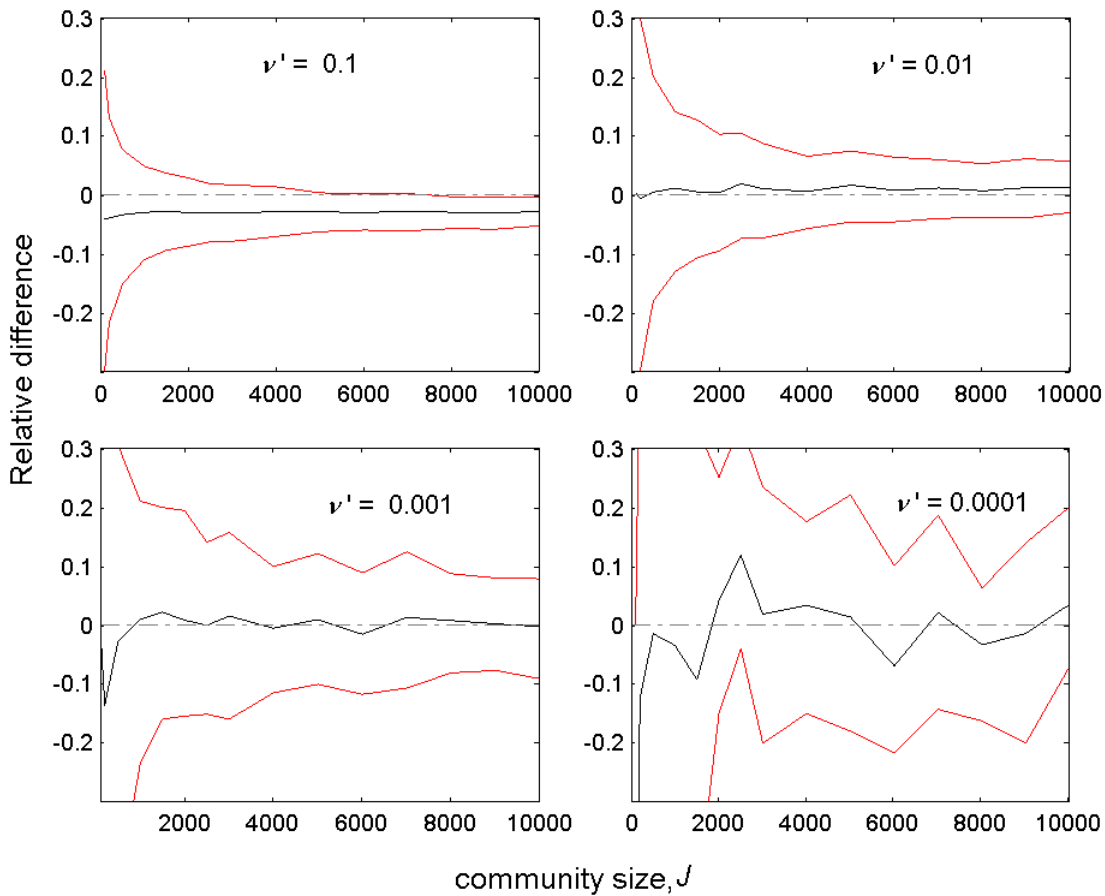
Figure L1. Relative difference between the equilibrium species richness of equation (5.9) (S_{eq_model}) and the equilibrium species richness estimated from direct simulations of the stochastic process (S_{eq_sim}), as a function of the community size, *J*. Black and red lines represent the average relative difference and the corresponding 95% confidence intervals from 1,000 simulations respectively. The grey dotted line is the line *y*=0, which represents a relative difference of zero. The increased fluctuation observed at low speciation probabilities is due to the very low species richness (<10 species) for this particular parameter combination.

Cited Literature

Haegeman, B. and Etienne, R.S. (2010) Self-consistent approach for neutral community models with speciation. *Physical Review E*, **81**:1–13.

Appendix M: Real Data Applications

Table M1 contains data of avian extinctions from different parts of the world that were gathered by Halley and Iwasa (2011) and were used to estimate half-life times to extinction. Part of the data refer to islands that contracted due to submersion or were cut off from the mainland due to submersion of land bridges at the end of the Pleistocene approximately 10,000 years ago. It is thought that these islands have not yet relaxed to equilibrium, as their biodiversity is considerably higher than that of oceanic islands of the same size (Diamond 1972, Terborg 1974). The rest of the data apply to fragments of forests formed after deforestation (Brooks et al. 1999, Newmark 1991, MacHunter et al. 2006, Castelletta et al. 2000, Diamond et al. 1987) and the island of Barro Colorado that was created after the flooding of the surrounding area to form Gatun Lake (Terborg 1974, Robinson 1999). Table M2 lists species recorded in Barro Colorado Island and their presence (1) or absence (0) in years 1970, 1977 and 1995, extracted from the studies of Willis and Eisenmann (1979) and Robinson (2001).

Table M1. Data of avian extinctions from islands and fragments around the world (columns 1-6 and 9 from Halley and Iwasa 2011, Table S1). Time since isolation or habitat loss in years (t), initial species richness (S_0), subsequent count of species richness ($S(t)$), equilibrium species richness (S_{eq}) estimated using a power-law species-area relationship with exponent $z = 0.22$, which was used by Diamond to estimate the initial species richness in Fergusson Island (Halley and Iwasa 2011, Diamond 1972), relaxation rate γ estimated from equation (5.12), time for half of the extinctions to happen in the RFS model (t_{50}) estimated from equation (5.13) and time for half of the extinctions to happen in the isolated-fragment model (t_{50}) (Halley and Iwasa 2011).

Region (Data source)	Fragment or island name	t Years since isolation	S_0 Initial species richness	$S(t)$ Subsequent observation	S_{eq} Equilibrium species richness	γ Relaxation rate (y^{-1})	t_{50} Random Fission Speciat. (y)	t_{50} Isolation (Halley-Iwasa) (y)
Southwest Pacific (Diamond 1972)	Fergusson	10000	108	86	59	$4.57*10^{-5}$	11719	49105
	Goodenough	10000	108	81	51	$4.63*10^{-5}$	10743	26143
	Aru	10000	325	158	88	$7.00*10^{-5}$	5068	99097
	Waigeu	10000	325	137	72	$7.23*10^{-5}$	4285	40233
	Japen	10000	325	119	67	$8.55*10^{-5}$	3430	28408
	Misol	10000	325	135	66	$6.50*10^{-5}$	4452	26029
	Salawati	10000	325	134	62	$6.19*10^{-5}$	4506	20810
	Batanta	10000	325	106	47	$6.62*10^{-5}$	3405	5780
	Batjan	10000	112	90	65	$5.15*10^{-5}$	10712	74574
	N. Hanover	10000	105	84	58	$4.54*10^{-5}$	11842	46735
	Amboina	10000	83	70	53	$4.67*10^{-5}$	12323	39237
West Indies (Terborg 1974)	Trinidad	10000	380	236	77	$2.67*10^{-5}$	10887	52663
	Margarita	10000	320	74	56	0.0002	1594	14896
	Coiba	10000	250	78	46	$9.80*10^{-5}$	2759	7511
	Tobago	10000	300	82	42	$8.48*10^{-5}$	2587	4145
	Rey	10000	225	46	40	0.0002	1130	4587
Kakamega forest, Kenya (Brooks et al. 1999)	Malava	101	32	19	18	0.0231	23	130
	Kisere	63	39	32	26	0.0105	56	425
	Ikuywa	20	48	44	36	0.0178	35	1252
	Yala	24	48	44	36	0.0149	42	1295
	Main	82	62	59	56	0.0081	82	5750
Usambasa mount., Tanzania (Newmark 1991)	Usambara/10	81	31	26	17	0.0042	128	697
	Usambara/4-9	81	31	15	10	0.0113	35	55
	Usambara/1	81	31	7	5	0.0209	12	4
	Usambara/2	81	31	8	5	0.0134	18	2
Southeastern Australia (MacHunter et al. 2006)	F21	22	28	17	6	0.0128	23	5
	F20	22	33	26	7	0.0050	58	9
	F24	22	41	33	8	0.0041	68	15
	T18	22	51	40	9	0.0045	61	29
Castelletta et al. 2000	Singapore	26	91	65	25	0.0094	38	1262
		75	91	30	25	0.0240	15	1262
Terborg 1974	BCI	60	208	163	22	0.0010	177	339
Robinson 1999	BCI	85	208	143	22	0.0012	152	339
Diamond et al. 1987	BBG	40	62	42	12	0.0047	57	58

Table M2. Presence (1) or absence (0) of bird species in Barro Colorado Island in 1970, 1977 and 1995. The data are extracted from the lists of species reported in Willis and Eisenmann (1979) and Robinson (2001). As Robinson (2001) reports only resident species but the classifications of species (resident, migrant, vagrant) in Willis and Eisenmann (1979) and Robinson (2001) do no completely coincide, we include only species that appear in both studies.

Scientific name	Common name	1970	1977	1995
Accipiter superciliosus	Tiny Hawk	1	1	0
Amazilia amabilis	Blue-chested Hummingbird	1	1	1
Amazilia tzacatl	Rufous-tailed Hummingbird	1	0	1
Amazona autumnalis	Red-lored parrot	1	1	1
Amazona farinosa	Mealy Parrot	1	1	1
Anthracothorax nigricollis	Black-throated Mango	1	1	0
Aramides cajanea	Gray-necked Wood-Rail	1	1	1
Attila spadiceus	Bright-rumped Attila	1	1	1
Baryphtengus martii	Rufous motmot	1	1	1
Brotogeris jugularis	Orange-chinnes Parakeet	1	1	1
Buteo albonotatus	Zone-tailed Hawk	1	1	0
Buteo brachyurus	Short-tailed Hwak	1	1	0
Cacicus cela	Yelow-rumped Cacique	1	1	1
Campephilus malenoleucos	Crimson-crested Woodpecker	1	1	1
Camptostoma obsoletum	Southern Beardless Tyrannulet	1	1	1
Capsiempis flavogaster	Yellow-bellied Elaenia	1	1	0
Cathartes aura	Turkey Vulture	1	1	1
Celeus loricatus	Cinnamon Woodpecker	0	0	1
Cercomacra tyrannina	Dusky antbird	1	1	1
Ceryle torquata	Ringed Kingfisher	1	1	1
Chaetura brachyura	Short-tailed Swift	1	1	1
Chaetura spinicauda	Band-rumped Swift	1	1	1
Chalybura buffonii	White-vented Plumeleteer	0	0	1
Chloroceryle aenea	Pygmy Kingfisher	1	1	1
Chloroceryle amazona	Amazon Kingfisher	1	0	0
Chloroceryle americana	Green Kingfisher	1	0	0
Chlorophanes spiza	Green Honeycreeper	1	1	1
Chlorostilbon canivetii	Fork-tailed Emerald	1	1	0
Chondrohierax uncinatus	Hook-billed Kite	1	1	1
Ciccaba nigrolineata	Black-and-White Owl	1	1	1
Ciccaba virgata	Mottled Owl	1	1	0
Claravis pretiosa	Blue Ground-Dove	0	0	1
Coereba flaveola	Bananaquit	1	1	1
Columba cayennensis	Pale-vented Pigeon	1	1	1
Columba nigrirostris	Shirt-billed Pigeon	1	1	1
Columba speciosa	Scaled Pigeon	1	1	1
Coragyps atratus	Black Vulture	1	1	1
Cotinga nattererri	Blue Contiga	1	1	1
Crax rubra	Great Curassow	0	0	1
Crotophaga ani	Smooth-billed Ani	1	1	0
Crotophaga major	Greater Ani	1	1	1

Crypturellus soui	Little Tinamou	0	1	1
Cyanerpes cyaneus	Red-legged Honeycreeper	1	1	1
Cyanerpes lucidus	Shining Honeycreeper	1	1	1
Cyanocompsa cyanoides	Blue-black Grosbeak	1	1	1
Cyanocorax affinis	Black-chested Jay	1	0	0
Cymbilaimus lineatus	Fasciated Antshrike	1	0	0
Dacnis venusta	Scarlet-thighed Dacnis	1	1	1
Damophila julie	Violet-bellied Hummingbird	1	1	1
Dendrocincla fuliginosa	Plain-brown Woodcreeper	1	1	1
Dromococcyx phasianellus	Pheasant Cuckoo	1	1	0
Dryocopus lineatus	Lineated Woodpecker	1	1	1
Dysithamnus puncticeps	Spot-crowned Antvireo	1	1	1
Electron platyrinchum	Broad-billed Motmot	1	1	1
Eucometis pencillata	Gray-headed Tanager	1	1	1
Euphonia fulvicrissa	Fulbus-vented Euphonia	1	1	1
Euphonia laniirostris	Thick-billed Euphonia	0	0	1
Euphonia minuta	White-vented Euphonia	1	1	0
Eurypyga helias	Sunbittern	1	0	0
Falco rufiguris	Bat Falcon	0	0	1
Florisuga mellivora	White-necked Jacobin	1	1	1
Geotrygon montana	Ruddy Quail-Dove	1	1	1
Geotrygon violacea	Violaceous Quail-Dove	1	1	1
Geranospiza caerulescens	Crane Hawk	1	1	1
Glyphorynchus spirurus	Wedged-billed Woodcreeper	1	1	1
Gymnopithys bicolor	Bicolored Antbird	1	1	1
Habia fuscicauda	Red-throated Ant-tanager	1	1	1
Harpagus bidentatus	Double-toothed Kite	1	1	1
Heliomaster longirostris	Long-billed Starthroat	1	1	0
Heliothryx barroti	Purple-crowned Fairy	1	1	1
Hylopezus perspicillatus	Streak-chested Antpitta	1	0	0
Hylophilus decurtatus	Lesser Greenlet	1	1	1
Hylophilus flavipes	Scrub Greenlet	1	1	0
Hylophylax naevioides	Spotted Antbird	1	1	1
Icterus chrysater	Yellow-backed Oriole	1	1	1
Laniocera rufescens	Specled Mourner	1	1	1
Legatus leucophaius	Piratic Flycatcher	1	1	1
Lepidopyga coeruleogularis	Sapphire-troated humminbird	0	1	1
Leptodon cayanensis	Gray-headed Kite	1	1	1
Leptotila cassinii	Gray-chested Dove	1	1	1
Leptotila verreauxi	White-tipped Dove	1	1	1
Leucopternis albicollis	White Hawk	1	1	1
Leucopternis semiplumbea	Semiplumbeous Hawk	1	1	1
Lipaugus unirufus	Rufous Pila	1	1	1
Lophornis delattrei	Rufous-crested Coquette	1	1	0
Lophostrix cristata	Crested Owl	1	1	0
Lurocalis semitorquatus	Shirt-tailed nighthawk	1	1	1
Malacoptila panamensis	White-whiskered Puffbird	1	1	1
Manacus vitelinus	Golden-collared Manakin	1	1	1
Megarhynchus pitagua	Boat-billed Flycatcher	1	1	1

Melanerpes pucherani	Black-ccheeked Woodpecker	1	1	1
Micrastur semitorquatus	Collared Forest-Falcon	1	1	1
Microrhpias quixensis	Dot-winged Antwren	1	1	1
Momotus momota	Blue-crowned motmot	0	0	1
Morphnus guianensis	Crested Eagle	1	1	0
Myiarchus panamensis	Panama Flycatcher	0	0	1
Myiarchus tuberculifer	Dusky-capped Flycatcher	1	1	1
Myiodynastes maculatus	Streaked Flycatcher	1	1	1
Myiopagis gaimardii	Forest Elaenia	1	1	1
Myiornis atricapillus	Black-capped Pygmy-Tyrant	1	1	1
Myiozetetes cayanensis	Rusty-margined Flycatcher	1	1	1
Myiozetetes similis	Social Flycatcher	1	1	1
Myrmeciza exsul	Chestnut-acked Antbird	1	1	1
Myrmeciza longipes	White-bellied Antbird	1	0	0
Myrmotherula axillaris	White-flanked Antwren	1	1	1
Myrmotherula fulviventhris	Checker-throated Antwren	1	1	1
Notharchus pectoralis	Black-breasted Puffbird	1	1	1
Notharchus tectus	Pied Puffbird	1	1	1
Nyctibius grandis	Great Potoo	1	1	0
Nyctidromus albicollis	Pauraque	1	1	1
Oncostoma olivaceum	Southern Bentbill	1	1	1
Ornithion Brunneicapillum	Brown-capped Tyrannulet	1	1	1
Ortalis cinereiceps	Gray-headed Chachalaca	1	1	0
Otus guatemalae vermiculatus	Vermiculated Schreech-Owl	1	1	1
Pachyramphus polychopterus	White-winged Becard	1	0	0
Panyptila cayennensis	Lesser Swallow-tailed Swift	1	1	1
Penelope purpurascens	Crested Guan	1	1	1
Phaenostictus mcleannani	Ocellated Antbird	1	1	0
Phaethornis longuemareus	Little Hermit	1	1	1
Phaethornis superciliosus	Long-tailed Hermit	1	1	1
Piaya cayana	Squirrel Cuckoo	1	1	1
Pionus menstrutus	Blue-headed Parrot	1	1	1
Pipra mentalis	Red-capped Manakin	1	1	1
Pipromorpha oleaginea	Ocre-bellied Flycatcher	1	1	1
Pitangus lictor	Lesser Kiskadee	1	1	1
Pitangus sulphuratus	Great Kiskadee	0	1	1
Pitylus grossus	Slate-colored Grosbeak	1	1	1
Platyrinchus coronatus	Golden-crowned Spadebill	1	1	1
Polioptila plumbea	Tropical Gnatcatcher	1	1	1
Progne chalybea	Gray-breasted Martin	1	1	1
Pteroglossus torquatus	Collared Aracari	1	1	1
Pulsatrix perspicillata	Spectacled Owl	1	1	1
Querula purpurata	Purple-throated Fruitcrow	1	1	1
Ramphastos sulphuratus	Keel-billed Toucan	1	1	1
Ramphastos swainsonii	Chestnut-mandibled toucan	1	1	1
Ramphocaenus rufiventris	Long-billed Gnatwren	1	1	1
Ramphocelus dimidiatus	Crimson backed Tanager	0	0	1
Rhinoptynx clamator	Striped Owl	1	1	0
Rhynchocyclus olivaceus	Olivaceous Flatbill	1	1	1

Rhytipterna holerythra	Rufous Mourner	1	1	1
Sarcoramphus papa	King Vulture	1	1	1
Scaphidura oryzivora	Giant-tailed Grackle	0	1	1
Schiffornis turdinus	Thrushlike Manakin	0	1	1
Sclerurus guatemalensis	Scaly-throated Leaftosser	1	1	1
Smaragdolanius pulchellus	Green Shrike-Vireo	1	1	1
Spizaetus ornatus	Crested Hawk-Eagle	1	1	0
Spizaetus tyrannus	Black Hawk-Eagle	1	1	1
Sporophila aurita	Variable Seedeater	1	1	1
Sporophila nigricollis	Yellow-bellied Seedeater	1	1	1
Stelgidopteryx rufipennis	Rough-winged Swallow	1	1	1
Streptprocze zonaris	White-collared Swift	1	1	0
Tachycineta albilinea	Mangrove Swallow	1	1	1
Tachyphonus luctuosus	White-shouldered tanager	1	1	1
Tangara gyrola	Bay-headed Tanager	0	1	1
Tangara inornata	Plain-colored tanager	1	1	1
Tangara larvata	Golden-hooded Tanager	1	1	1
Terenotriccus erythrurus	Ruddy-tailed Flycatcher	1	1	1
Thalurania colombica	Crowned Woodnymph	1	1	1
Thamnophilus punctatus	Slaty Antshrike	1	1	1
Thraupis episcopus	Blue-gray tanager	1	1	1
Thraupis palmarum	Palm tanager	1	1	1
Thryothorus modestus	Plain Wren	1	0	1
Tigrisoma lineatum	Rufescent Tiger-Heron	1	1	1
Tinamus major	Great Tinamou	1	1	1
Tityra inquisitor	Black-crowned Tityra	1	1	1
Tityra semifasciata	Masked Tityra	1	1	1
Todirostrum cinereum	Common Tody-Flycatcher	1	1	1
Tolmomyias assimilis	Yellow-margined Flycatcher	1	1	1
Trogon massena	Slaty-tailed Trogon	1	1	1
Trogon melanurus	Black-tailed Trogon	1	1	1
Trogon rufus	Black-throated Trogon	1	1	1
Trogon violaceus	Violaceous trogon	1	1	1
Trogon viridis	White-tailed Trogon	1	1	1
Turdus grayi	Clay-colored Robin	0	0	1
Tyranniscus vilissimus	Paltry Tyrannulet	1	1	1
Tyrannulus elatus	Yellow-crowned Tyrannulet	1	1	1
Tyrannus melancholicus	Tropical Kingbird	1	1	1
Vireo flavoviridis	Yellow-green Vireo	1	0	0
Xenops minutus	Plain Xenops	1	1	1
Xiphorhynchus guttatus	Buff-throated Woodcreeper	1	1	1
Xiphorhynchus lachrymosus	Black-striped Woodcreeper	1	1	1
Zarhynchus wagleri	Chestnut-headed Oropendola	1	1	1

Cited Literature

Brooks T.M., Pimm S.L., and Oyugi J.O. (1999) Time lag between deforestation and bird extinction in tropical forest fragments. *Conservation Biology*, **13**:1140–1150.

Castelletta, M., Sodhi, N.S. and Subaraj, R. (2000) Heavy Extinctions of Forest Avifauna in Singapore: Lessons for Biodiversity Conservation in Southeast Asia. *Conservation Biology*, **14**:1870–1880.

Diamond J.M. (1972) Biogeographic kinetics: Estimation of relaxation times for avifaunas of southwest Pacific islands. *Proceedings of the National Academy of Sciences*, **69**:3199–3203.

Diamond, J.M., Bishop, K.D. and Balen, S.V. (1987) Bird Survival in an Isolated Javan Woodland: Island or Mirror? *Conservation biology*, **1**:132–142.

Halley, J.M. and Iwasa, Y. (2011) Neutral theory as a predictor of avifaunal extinctions after habitat loss. *Proceedings of the National Academy of Sciences*, **108**:2316–2321.

MacHunter, J., Wright, W., Loyn, R., and Rayment, P. (2006) Bird declines over 22 years in forest remnants in southeastern Australia: Evidence of faunal relaxation? *Canadian Journal of Forest Research*, **36**:2756–2768.

Newmark, W.D. (1991) Tropical forest fragmentation and the local extinction of understory birds in the eastern Usambara Mountains, Tanzania. *Conservation Biology*, **5**:67–78.

Robinson, W.D. (1999) Long-Term Changes in the Avifauna of Barro Colorado Island, Panama, a Tropical Forest Isolate. *Conservation Biology*, **13**:85–97.

Robinson, W.D. (2001) Changes in abundance of birds in a Neotropical forest fragment over 25 years: a review. *Animal Biodiversity and Conservation*, **24**:51–65.

Terborgh, J. (1974) Preservation of natural diversity: The problem of extinction prone species. *BioScience*, **24**:715–722.

Willis, E.O. and Eisenmann, E. (1979) *A Revised List of Birds of Barro Colorado Island, Panama.* Smithonian Contribution to Zoology, Smithonian Institution Press, Washington.