

Aus der Schweizerischen Vogelwarte Sempach

## Grundlagen der Bestandserfassung am Beispiel von Vorkommen und Verbreitung

Marc Kéry



KÉRY, M. (2008): Foundations of bird surveying with examples from occurrence and distribution. *Ornithol. Beob.* 105: 353–386.

Occurrence and distribution are central topics in ornithology. However, every ornithologist knows very well that birds can be overlooked, that is, that detection probability ( $p$ ) is usually less than 1 ( $p < 1$ ). Typically therefore, our observations map the true population sizes and true occurrence of birds only in an incomplete and potentially distorted, that is, biased, way. Mathematically, all our bird observations are random variables. Whenever  $p < 1$ , our observations aren't fixed numbers anymore. Instead, by pure chance replicate observations will typically differ even under identical conditions and will be predictable only on average. The binomial distribution of statistics represents the theoretical basis for a bird survey where  $N$  units (such as occupied quadrats) all independently have the same chance  $p$  of being recognized as occupied by a species of interest; the resulting count of occupied quadrats is binomially distributed. The flip of a loaded coin is a perfect analogy to the observation process involved in any bird distribution study: all the birds in each of  $N$  occupied quadrats flip the coin, and only when any of the inhabitants of a quadrat tosses heads is that quadrat recognized as occupied and therefore appears in the count of occupied quadrats. Loaded means that a coin's probability of heads can be any number between 0 and 1 in principle, not only 0.5. The binomial model of bird surveys perfectly illustrates three key features of bird observations: (1) The observed number of occupied quadrats ( $n$ ) will almost always differ in repeated surveys, even under identical conditions, and varies between zero and the true number of occupied quadrats ( $N$ ). (2) We can't say much about any individual count  $n$ , but the mean count, i.e., averaged over many replicate counts, will be equal to  $N * p$ , the expected value of a binomial random variable. In addition, counts will automatically vary one from another and the magnitude of that variation is known to be  $\sqrt{N * p * (1 - p)}$ , the standard deviation of a binomial random variable. Furthermore, since  $p \leq 1$ , we will almost always underestimate true distributions. (3) Changes in counts of occupied quadrats over time or differences between sites or habitats in the number of occupied quadrats can be due to differences in the true number  $N$  of occupied quadrats, differences in the probability to detect an occupied quadrat ( $p$ ), a combination of both or even just due to (binomial) chance variation. Therefore, trends in detection probability  $p$  can mask genuine distributional trends or feign changes in a distribution that is in fact static. Whenever we need to know the absolute distribution ( $N$ ) of a species or when we fear the presence of «dangerous patterns» in  $p$  (by which term is meant that  $p$  and  $N$  depend on the same factor of interest) we must estimate

N directly. This involves estimation of  $p$  and therefore correction of our observations for imperfect detection ( $p < 1$ ). Key to this in the context of distribution studies are observations that are replicated in the short-term, so that the true occupancy state (occupied or not) does not change. Under this condition, the pattern of detection and nondetection of a species contains the information required to estimate  $p$  separately from the true distribution. Using such replicate observations, a new class of statistical models that goes by the peculiar name «site-occupancy model» (MacKenzie et al. 2002) enables one to estimate true distribution and its relationship with external (e.g., habitat or climatic) variables corrected for imperfect detection probability. Conceptionally, site-occupancy models are based directly on the binomial model of bird surveys and represent two coupled logistic regression models: the first describes the imperfectly observed true pattern of presence and absence, and the second describes the pattern of detection/nondetection, given the true pattern of occurrence. Covariates can be introduced into both regressions, that is, both distribution and detection probability can be made a function of the values of measured factors of interest. In this essay, I describe site-occupancy models and in a simulation study show how they are much better able to estimate the true distribution of a species than conventional ways of interpreting distributional data. In a second simulation study I show that, unlike conventional methods, estimates under a site-occupancy model can correct estimates of species distributions for the presence of «dangerous patterns» in detection probability, for instance, when elevation affects occurrence probability and detection probability of a species alike. It is argued that many species distribution studies conducted by professional or amateur ornithologists alike would benefit from more conceptual rigour. The binomial distribution of statistics represents the theoretical underpinning for the design, conduct, analysis and interpretation of all empirical studies on distribution and occurrence of species. I hope that this essay fosters a deeper understanding of the fundamentals of bird distribution surveys and that increased application of site-occupancy models for estimation of species distributions will enhance the quality and interpretability of such studies.

Marc Kéry, Schweizerische Vogelwarte, CH-6204 Sempach, E-Mail marc.kery@vogelwarte.ch

Das Zählen von Vögeln im weitesten Sinne ist die definierende Tätigkeit einer Vielzahl von professionellen und vor allem von Amateur-Ornithologen. In vielen Fällen geht es dabei nicht um die absoluten Bestände, sondern bloss um das Vorkommen an einem Ort oder in mehreren Gebieten, oder anders ausgedrückt, um die Verbreitung einer Art. Zumindest implizit werden dabei immer räumliche und zeitliche Einheiten gebildet, und es wird versucht herauszufinden, welche Arten darin vorkommen. Verbreitungsatlantanten wie jene von Schifferli et al. (1980) und Schmid et al. (1998) sind klassische Beispiele, wo versucht wurde, das Vorkommen al-

ler Brutvögel in räumlichen Einheiten von 10 bzw. 1- oder 10-km-Quadraten und über Zeiträume von 4–5 Jahren zu dokumentieren. In wieder anderen Fällen mag die Untersuchung der regionalen Verbreitung einer Art äquivalent zur Erfassung ihres Bestands sein, z.B. wenn die räumlichen Einheiten gleich den Revieren einer Art definiert werden (s. Uhu- und Wanderfalkenbeispiele, Kap. 9, Beispiel 1). Interessanterweise können also die Grenzen zwischen Bestand und Verbreitung verschwimmen. Ferner kann die Verbreitung natürlich nicht nur von Brutvögeln erhoben werden, sondern auch von Überwinterern und Durchzügler (vgl. z.B.

die prominenten Verbreitungskarten, die aufgrund von ID-Daten, also Information in einer Datenbank für opportunistische Beobachtungen, für das neue umfassende Werk «Die Vögel der Schweiz» von Maumary et al. (2007) hergestellt wurden). Vorkommen und Verbreitung sind zentrale Grössen sowohl für die Ornithologie als auch für die gesamte Ökologie (Krebs 2000).

Nun weiss natürlich jeder Feldornithologe, dass nicht immer jeder anwesende Vogel auch entdeckt wird, und dass Vögel wie andere Tiere oder selbst Pflanzen (Kéry et al. 2006) oftmals übersehen werden. Das bedeutet, dass die Antreffwahrscheinlichkeit bei den meisten Arten, für die meisten Beobachter und unter den meisten Bedingungen kleiner als 1 ist (Schmid 2005). In gewisser Weise ist das sowohl ein Fluch als auch ein Segen. Denn zumindest für den Feldornithologen stammt ein grosser Teil der Spannung beim Beobachten gerade daher, dass er nicht a priori weiss, welche Arten da sind, und dass auch nicht alle Arten sofort und für jeden sichtbar sind. Viele Feldornithologen müssen wohl zugeben, dass in ihnen auch ein wenig ein Jäger schlummert. Und für einen richtigen Jäger ist nur jenes Wild interessant, das eine Herausforderung darstellt. Die Tatsache, dass nicht jeder Vogel gleich auf den ersten Blick entdeckt wird und dass der bessere Beobachter eben mehr sieht als der weniger talentierte, stellt somit eine enorme Herausforderung dar, die das Beobachten für viele erst interessant macht.

Die Kehrseite der Medaille ist allerdings, dass die Erfassung von Beständen und der Verbreitung meist ein sehr anspruchsvolles Unterfangen ist, weil die Beobachtungen die wirklichen Bestände und Vorkommen in aller Regel nur unvollständig abbilden. So weiss man bei Mustern (z.B. Zu- oder Abnahmen) in vielen Bestandszahlen oder Vorkommensbeobachtungen kaum, ob es sich dabei um die Realität oder einfach um Beobachtungseffekte handelt. Sehr viele Verbreitungsstudien sind somit nur sehr bedingt interpretierbar. Wenn beispielsweise im ersten Schweizer Brutvogelatlas (Schifferli et al. 1980; Feldarbeit 1972–1976) eine Art in viel weniger Quadraten von 10 km Seitenlänge nachgewiesen wurde als im zweiten (Schmid et

al. 1998; Feldarbeit 1993–1996), so kann ohne zusätzliche Information kaum entschieden werden, ob sich die Verbreitung der Art wirklich verändert hat oder ob im zweiten Atlas einfach viel mehr Leute intensivere Feldarbeit betrieben haben. Diese und sehr viele andere auch bloss regionale Verbreitungsstudien sind somit nur bedingt miteinander vergleichbar, und das ist schade.

Nun haben Ornithologen im Verlaufe von buchstäblich Jahrhunderten natürlich bis zu einem gewissen Grad intuitiv gelernt, mit dem Problem der unvollständigen Beobachtbarkeit umzugehen: Gute Ausbildung der Leute, lange und gewissenhafte Beobachtung unter guten Bedingungen (z.B. zur Jahreszeit der höchsten Balzaktivität sowie am frühen Morgen) und standardisierte Beobachtungsbedingungen tragen dazu bei, Vorkommensbeobachtungen vergleichbarer zu machen. Auch ist jedermann klar, dass das kombinierte Resultat mehrerer Aufnahmen näher an die effektive Verbreitung einer Art kommt als jenes einer einzigen Aufnahme, und so werden in grossen Inventaren wie den Brutvogelatlas (Schifferli et al. 1980, Schmid et al. 1998) oder Monitoringprogrammen wie dem Monitoring Häufige Brutvögel (MHB; Schmid et al. 2004) in jeder Brutsaison mehrere Aufnahmen durchgeführt. Viele dieser auch artspezifischen Erkenntnisse wurden zusammengetragen in Standardwerken zur Bestandserfassung wie Bibby et al. (2000) oder Südbeck et al. (2005), die eine Vielzahl von akribischen Empfehlungen für scheinbar jede mögliche Situation im Feld enthalten.

Und dennoch sind alle diese Optimierungen oft noch ungenügend. Beispielsweise kann es sein, dass der Bestand einer Art zunimmt, und bei einer grösseren Zahl von Individuen in einem Quadrat wird zwangsläufig eher mindestens eines gesehen als bei einem kleineren Bestand: Somit wird z.B. die Verbreitung von häufigen Arten weniger stark unterschätzt als jene von seltenen. Es wäre auch denkbar, dass sich im Rahmen der Klimaerwärmung das Verhalten einer Art verändert, so dass sie besser oder schlechter beobachtbar wird, und das wird sich selbst bei starker Standardisierung der Feldarbeit direkt verfälschend auf die wahrgenommene Verbreitung auswirken.

Mit meinem Essay verfolge ich zwei Ziele. Erstens soll die theoretische Grundlage der Erfassung von Vorkommen und Verbreitung, die Binomialverteilung, am Beispiel von Vögeln beschrieben werden. Die statistische Wissenschaft erlaubt oft die Formalisierung von Dingen, die man «im Bauch» einigermaßen zu verstehen glaubt. Durch diese Formalisierung erlangt man ein sehr viel besseres und klareres Verständnis des Beobachtungsprozesses, der derart wichtig ist bei der Interpretation unserer Daten sowohl als Amateur- als auch als professionelle Ornithologen. Insbesondere werde ich den Zusammenhang zwischen wirklicher und beobachteter Verbreitung darstellen, mit der Antreffwahrscheinlichkeit als ihrem wichtigen Bindeglied. Zweitens werde ich sogenannte Site-Occupancy-Modelle (MacKenzie et al. 2002, 2003, 2006) vorstellen. Diese in den letzten Jahren entwickelte statistische Methode erlaubt, Verbreitungsstudien nach unvollständiger Beobachtbarkeit zu korrigieren und somit die effektive Grösse eines Verbreitungsgebietes zu schätzen sowie die Verbreitung in Abhängigkeit vom Habitat zu modellieren. Diese wichtige Modellklasse ist jedoch noch längstens nicht so bekannt, wie sie es verdient. Danach folgen zuerst anhand von simulierten Daten eine Veranschaulichung des Beobachtungsprozesses von Verbreitungen sowie ein Vergleich der herkömmlichen, direkten Interpretation von Vorkommensbeobachtungen (d.h. ohne Korrektur nach Antreffwahrscheinlichkeit) und der modernen Interpretation von Vorkommensbeobachtungen mittels Site-Occupancy-Modell. Anschliessend gebe ich vier praktische Beispiele, welche die Theorie der Bestandserfassung sowie die Anwendung von Site-Occupancy-Modellen illustrieren.

Mein Artikel ist nicht mit Literaturhinweisen überladen und bewusst essayhaft geschrieben. Damit und durch die eher *légère* Sprache hoffe ich, das unvermeidlich etwas quantitative, aber wichtige Thema der Binomialverteilung als Grundlage der Vogelbestandserfassung sowie der Antreffwahrscheinlichkeit im Kontext von Vorkommen und Verbreitung möglichst vielen Ornithologen näher zu bringen. Obwohl ich einige Formeln nicht vermeiden konnte oder wollte, arbeite ich viel mit Zahlenbeispielen

und auch mit simulierten Datensätzen. Die meisten Ornithologen (wie ich selber auch) sind nicht formal statistisch ausgebildet. Für sie stellt die experimentelle Betrachtung (oder, anders gesagt, das Ausprobieren) von Dingen wie dem Beobachtungsprozess von Vögeln anhand von Simulationen die beste und eigentlich oft einzige Möglichkeit dar, ein System zu begreifen, dessen mathematische Beschreibung ihnen nicht genügend zugänglich ist.

Dieser Essay ist die erste deutsche Einführung in Site-Occupancy-Modelle im Kontext von Vogelerfassungen (s. auch Schmidt 2008, im Kontext von Kriechtieren; für eine leicht leserliche französische Einführung ins Konzept der Antreffwahrscheinlichkeit s. Kéry 2007). Ich bringe vor allem Beispiele aus der Schweiz, oft auch aus dem mir besonders vertrauten Monitoring Häufige Brutvögel (MHB). Ich hoffe, dass möglichst viele Leser meine Beispiele in eine für sie interessante Situation zu übersetzen vermögen und glaube, dass sie anschliessend besser in der Lage sind, Vogelerfassungen zu planen, durchzuführen, zu analysieren und zu interpretieren.

## 1. Was ist Vorkommen und Verbreitung?

Vorkommen (an einem oder wenigen Orten) oder Verbreitung (Vorkommen an vielen Orten) bezeichnet die Präsenz einer Art an einem oder vielen «Orten», bezieht sich also immer auf eine bestimmte räumliche Einheit. Die räumliche Einheit kann politisch definiert sein (z.B. Länder, Kantone, Gemeinden), geometrisch (z.B. Quadrate mit 50, 10, 2, 1 oder 0,25 km Seitenlänge) oder biologisch (z.B. mehr oder weniger einheitliche Habitate wie Feuchtgebiete, Waldparzellen, Felsobjekte, Hecken, Felder, Brachen). Man beachte, dass diese räumliche Einheit (z.B. ihre Grösse) immer durch eine biologische Fragestellung bzw. einen Biologen und nicht «durch die Statistik» bzw. den Statistiker definiert wird. In meinem Essay bezeichne ich eine räumliche Einheit entsprechend jener im MHB oft einfach als «Quadrat». Zumindest implizit hat die Definition von Vorkommen und Verbreitung auch eine zeitliche Dimension, die meist durch das Datum und die

Dauer der Feldarbeit gegeben ist. So dokumentiert beispielsweise der zweite Schweizer Brutvogelatlas (Schmid et al. 1998) das Vorkommen von Brutvögeln in der Schweiz auf einer 10- und 1-km-Skala zu Beginn der 90er-Jahre des letzten Jahrhunderts.

Verbreitungsstudien verlangen auch eine Definition des Zustands, den wir unter Präsenz verstehen. In vielen Untersuchungen setzen wir Vorkommen mehr oder weniger gleich mit permanentem Aufenthalt einer Art an einem Ort zur Brutzeit zwecks Fortpflanzung und verlangen z.B. die Präsenz in geeignetem Habitat. Eine solche Definition muss in der Praxis konkretisiert werden, um zu entscheiden, welche Evidenz ausreicht für den Schluss auf ein so definiertes Vorkommen (z.B. Beobachtung eines einzelnen Vogels, eines Paares oder unselbstständiger Jungvögel oder eben das Kriterium der Beobachtung in einem potenziellen Habitat). In Brutvogelatlasen werden Beobachtungen oft mit dem Atlascode klassifiziert, der unterschiedlich starke Evidenz für Brutvorkommen codiert. Einer gegebenen Feststellung (z.B. Beobachtung eines Paares) wird für unterschiedliche Arten unterschiedliches Gewicht als Evidenz zugemessen. So würde sie für ein Auerhuhn *Tetrao urogallus* als sehr starker Hinweis auf ein Brutvorkommen gelten, nicht aber bei vielen Entenarten. Es gibt natürlich auch andere Arten von Vorkommen. So ist es absolut erlaubt, Vorkommen als «gelegentlicher Aufenthalt im Quadrat» zu definieren. Die Definition des Zustands des Vorkommens sollte der untersuchten Art (z.B. in Bezug auf ihre Reviergrösse), der Fragestellung sowie den logistischen Möglichkeiten einer Studie angepasst sein.

In diesem Essay geht es um Vorkommen und Verbreitung, doch Verbreitung und Bestand sind in gewisser Weise äquivalent, denn erstens entspricht Vorkommen an einem Ort ganz einfach dem Fall, dass der Bestand dort grösser als Null ist. Ein zweiter Zusammenhang besteht via die Grösse der räumlichen Einheit, auf die Vorkommen bezogen wird: Machen wir ein Quadrat so klein, dass höchstens ein Vogel oder ein Paar darin vorkommen kann, so entspricht die Anzahl der besetzten Einheiten genau dem Bestand.

Zusammenfassend verlangt also die Beschreibung von «Vorkommen» eine räumliche und zeitliche Einheit, eine Definition des Zustands des «Vorkommens» sowie der Art von Beobachtung, die als Evidenz für ein solcherart definiertes Vorkommen gelten darf. Vorkommen/Verbreitung kann dann entweder in Form einer Zählung (z.B. «die Art wurde in 45 von 468 Quadraten gefunden») oder durch ein Verhältnis ausgedrückt werden («wurde auf 9,6 % der Gesamtfläche gefunden»). Schliesslich besteht eine gewisse Äquivalenz zwischen Bestand und Verbreitung.

## 2. Vögel zählen ist fast wie Münzen werfen

In der Regel können nicht alle anwesenden Vögel beobachtet werden, d.h. ihre Antreffwahrscheinlichkeit  $p$  (auch Entdeckung- oder Beobachtungswahrscheinlichkeit genannt) ist kleiner als 1 (d.h.  $p < 1$ ). In diesem Sinne sind unsere Beobachtungen eigentlich nie perfekt. Sobald  $p < 1$ , ist eine Vorkommensbeobachtung («beobachtet» oder «nicht beobachtet») keine fixe Grösse mehr, sondern eine sogenannte Zufallsvariable, die am besten mit der Analogie eines Münzwurfs veranschaulicht werden kann. Wenn ich in mein MHB-Quadrat gehe und aus bisheriger Erfahrung weiss, dass ich in einem Quadrat mit Vorkommen des Schwarzspechts *Dryocopus martius* die Art mit einer Wahrscheinlichkeit von  $p = 0,41$  festzustellen vermag, so ist das genau so, wie wenn ich eine gezinkte Münze werfe, die im Durchschnitt in 41 % der Würfe mit dem Kopf nach oben landet; wir können schreiben  $P(\text{Kopf}) = 0,41$ . Wenn ich Kopf werfe, sehe ich einen Schwarzspecht in meinem Quadrat, und wenn ich Zahl werfe, sehe ich keinen. (Genau genommen sind Beobachtungen auch Zufallsvariablen, wenn  $p = 1$ , aber dieser Grenzfall ist weniger interessant, weil er im Feld wohl nur sehr selten vorkommt.)

Diese Münzwurfanalogie ist zentral für mein Essay, aber auf den ersten Blick mag sie hinken, denn erstens gibt es ja auch den Fall  $p = 0,5$ , und da kann man genau genommen nicht von einer gezinkten Münze sprechen. Und zweitens ist eine Münze, die immer mit Kopf



(entsprechend  $p = 1$ ) oder immer mit Zahl (entsprechend  $p = 0$ ) nach oben landet, schwer vorstellbar, und dennoch sind in der Natur zumindest Fälle mit  $p = 1$  denkbar. Und drittens kann kritisiert werden, dass Vögel und ihre Beobachter ja lernfähige Lebewesen und keine Maschinen (oder Münzen) seien und dass daher diese Analogie grundsätzlich unzulässig sei.

Zur technischen Seite der Analogie ist zu sagen, dass eine «faire» Münze bei der Hälfte der Würfe mit Kopf oben und bei der anderen Hälfte mit Zahl oben landet, d.h.  $p = 0,5$ . Wenn eine faire Münze nun durch Druck in ihr Zentrum zumindest ein wenig in die Form eines Suppentellers gebracht wird, so landet sie häufiger auf der einen als auf der anderen Seite, d.h. sie ist gezinkt. Auf diese Weise können wir uns vorstellen, dass wir zumindest im Prinzip eine physikalische Analogie für jeden beliebigen binären Zufallsprozess mit allen möglichen Wahrscheinlichkeiten  $p$  zwischen 0 und 1 besitzen. Wir betrachten dann eine faire Münze einfach als Spezialfall des allgemeinen Falles einer gezinkten Münze und stellen uns vor, wir könnten eine Münze dermassen in der Mitte eindellen, dass wir dadurch selbst extreme Fälle simulieren könnten, z.B. Zufallsprozesse mit  $p = 0,03$  oder  $p = 0,99$  (s. Kap. 9, Beispiel 4).

Zur philosophischen Seite der Münzwurfanalogie bzw. der Kritik, dass Lebewesen keine Maschinen oder Münzen seien, ist zu sagen, dass das Verständnis eines komplizierten Systems einzig und allein über dessen Abstraktion, d.h. Vereinfachung, möglich ist. Die Münzwurfanalogie ist ein simples Modell für eine komplizierte Wirklichkeit. Ein taugliches Modell behält die für eine bestimmte Fragestellung wesentlichen Grundzüge des zu erklärenden Systems bei, lässt aber alle für eine bestimmte Fragestellung irrelevanten Details aus und verschafft uns genau dadurch ein tieferes Verständnis in das System. Die wesentliche Frage ist also nicht, ob wir Lebewesen, Maschinen oder Münzen seien, sondern ob die wesentlichen Eigenschaften des Beobachtungsprozesses von Lebewesen (z.B. Vögel) durch andere Lebewesen (z.B. Ornithologinnen und Ornithologen) im Wurf einer oder mehrerer gezinkter Münzen beibehalten sind oder nicht. Und diese Frage würde ich emphatisch bejahen.

Mit der wichtigen Münzwurfanalogie lassen sich nämlich verschiedene und oftmals verkannte Eigenschaften von Beobachtungen wunderbar illustrieren. Wenn  $p = 0,41$ , finde ich den Schwarzspecht auch in Quadraten, in denen er vorkommt, häufiger nicht, als dass ich ihn feststelle. Falls z.B. 100 MHB-Quadrate mit Schwarzspechtvorkommen je einmal bearbeitet würden (d.h. jeder Beobachter einmal die Münze wirft), so würde die Art bloss in etwa 41 Quadraten auch gefunden. Und je häufiger ich innerhalb eines Quadrats die Münze werfe, desto wahrscheinlicher wird es, dass ich mindestens einmal Kopf werfe, also den Schwarzspecht in einem besetzten Quadrat auch feststelle. Auch lässt sich vorgängig über eine einzelne Begehung eigentlich nichts Sicheres aussagen, d.h. der Schwarzspecht kann gefunden werden oder auch nicht, es ist im Fall mit  $P(\text{Kopf}) = 0,41$  bloss so, dass er etwas häufiger übersehen als gefunden wird. Man kann bloss etwas über viele wiederholte Begehungen oder über viele unterschiedliche Ornithologen aussagen, nämlich, dass der Schwarzspecht in besetzten Quadraten in etwa 41 % der Fälle nachgewiesen wird.

In Realität wirft man natürlich nicht immer dieselbe Münze. Für eine Begehung in einem Quadrat mit einem besetzten Höhlenbaum würde man vielleicht eine Münze mit  $P(\text{Kopf}) = 0,63$ , in einem Quadrat, welches bloss zum Aufenthaltsgebiet des Schwarzspechts gehört, eine mit  $P(\text{Kopf}) = 0,28$  und in einem Quadrat, in dem gar zwei besetzte Höhlenbäume stehen, vielleicht gar eine mit  $P(\text{Kopf}) = 0,72$ . Weiter werfen erfahrene Beobachter eine andere Münze als Anfänger, oder eine 6-stündige Begehung entspräche dem Wurf einer anderen Münze als eine 2-stündige Begehung. Ein Lerneffekt wird simuliert durch wiederholte Würfe von Münzen mit immer höherem  $p$ , und das Gegenteil, wenn eine Art störungsverursacht immer heimlicher wird, durch solche mit immer kleinerem  $p$ .

Für das Verständnis der theoretischen Grundlage der Bestandserfassung ist es aber vorteilhaft, sich zunächst vorzustellen, alle würfen immer dieselbe Münze. In den statistischen Methoden, mit denen man solche Daten analysieren kann, ist es möglich, auch Würfe mit verschiedenen Münzen und viele weitere Kom-

plexitäten zuzulassen. Und das Wichtigste ist einfach die Erkenntnis, dass bei Bestandserfassungen fast immer alle irgendeine Münze werfen, nur in einem einzigen, aussergewöhnlichen Fall nicht; nämlich dann, wenn von einer Art in einem Quadrat immer alle vorhandenen Individuen sicher festgestellt werden, also wenn  $P(\text{Kopf}) = 1,00$ .

### 3. Formale Beschreibung der theoretischen Grundlage der Bestandserfassung: die Binomialverteilung

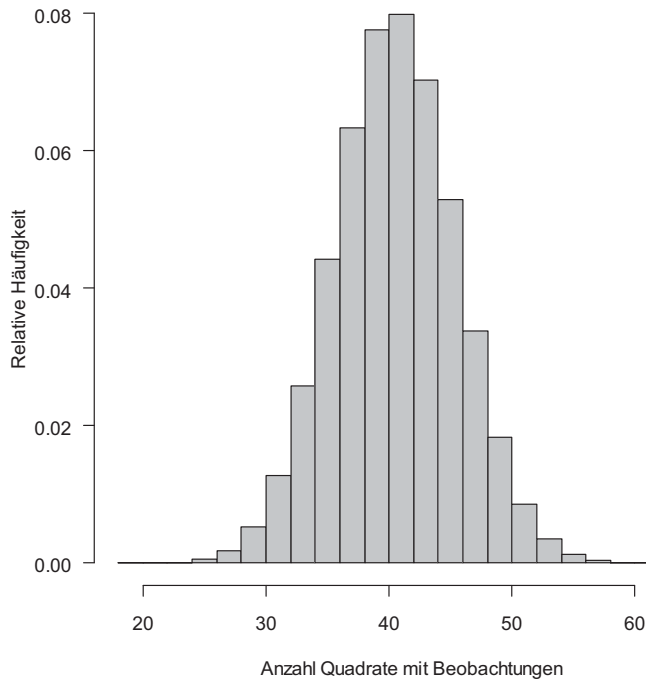
In erster Näherung entsprechen Bestands- und Vorkommensfeststellungen (bei gegebenem Bestand und Vorkommen) ebenso wie Münzwürfe formal einer binomial verteilten Zufallsvariablen. Das Modell der Binomialverteilung in Worten und auf unser Schwarzspechtbeispiel angewandt lautet etwa so: «Ist in N (z.B. 100) MHB-Quadraten mit Vorkommen des Schwarzspechts in jedem die Chance, einen Schwarzspecht bei einer einzelnen Begehung festzustellen, unabhängig und identisch gleich p (z.B. 0,41 = 41 %, s. Kéry & Schmidt im Druck), so ist die Anzahl der Quadrate, in denen der Schwarzspecht dann auch tatsächlich festgestellt wird, binomial verteilt». Man schreibt auch,  $n \sim \text{Binomial}(N, p)$ : Eine Zählung n von besetzten Quadraten ist binomial verteilt mit Index oder Stichprobengrösse N und Antreffwahrscheinlichkeit p.

Hier ist ein wichtiger Einschub: Genau genommen stammen alle Bestands- und Vorkommensbeobachtungen aus zwei hierarchischen (geschachtelten) Zufallsprozessen (Royle & Dorazio 2006, 2008). Der erste ist der biologische Prozess, der Individuen in Raum und Zeit verteilt und dadurch ein bestimmtes, wahres Muster von Bestand und Vorkommen generiert. Dieser Prozess ist nicht genau vorhersagbar und daher bereits ein Zufallsprozess. Sein Resultat (man spricht auch von der Realisation eines Zufallsprozesses), der wahre Bestand oder die effektive Verbreitung ist das, woran wir normalerweise Interesse haben. Doch die Realisation dieses ersten Zufallsprozesses lässt sich fast nie exakt beobachten; unsere Beobachtungen entstehen, indem die Realisation aus

dem biologischen Prozess durch einen zweiten Zufallsprozess hindurch müssen, den Beobachtungsprozess. Meine Beschreibung der Theorie der Bestandserfassung in diesem Essay ist in erster Linie eine Beschreibung des Beobachtungsprozesses und geht fast immer von einer gegebenen Realisation des biologischen Prozesses aus, d.h. man geht von einem bestimmten Bestand oder einer bestimmten Anzahl besetzter Quadrate aus und untersucht die möglichen Beobachtungen. Eine Ausnahme sind die Site-Occupancy-Modelle (s. unten), die sowohl den biologischen Prozess (die effektive Verbreitung) als auch den Beobachtungsprozess (die Beobachtungen) mittels Binomialverteilungen beschreiben.

Die Binomialverteilung ist in gewisser Weise das Naturgesetz für den Beobachtungsprozess und somit für die Vogelbestandserfassung. Wenn nicht fälschlicherweise eine andere Art als Schwarzspecht bestimmt wird, wenn der «Bestand» an besetzten Quadraten N und die Antreffwahrscheinlichkeit p konstant und die Beobachtungen voneinander unabhängig sind, so wird man von N besetzten Quadraten immer n zählen, und die Zählung n sowie ihre Variation zwischen wiederholten Versuchen wird bestimmt durch Bestand N und Antreffwahrscheinlichkeit p. Über eine einzelne Zählung kann man vorgängig nichts Genaues aussagen, denn sie kann – auch unter identischen Bedingungen – theoretisch in den gesamten Bereich zwischen 0 und 100 Quadraten mit Schwarzspechtbeobachtungen fallen. Gewisse Werte für die Zählung sind jedoch wahrscheinlicher als andere. Die statistische Verteilung beschreibt nun genau, welche Werte (Anzahl Quadrate mit Schwarzspechtbeobachtungen) wie wahrscheinlich auftreten.

Man beachte, dass ich in diesem Essay unter «Bestand» N entweder die Anzahl Individuen oder Paare oder aber die Anzahl besetzter Quadrate oder anderer räumlicher Einheiten verstehe. Welche Art von «Bestand» gemeint ist, sollte aus dem Kontext heraus klar werden. Obwohl die beiden Arten von «Bestand» biologisch völlig andere Dinge bedeuten, sind sie konzeptionell fast identisch. Insbesondere gelten die Gesetzmässigkeiten des Zählens für beide analog.



**Abb. 1.** Die Binomialverteilung als theoretische Grundlage der Vogelbestandserfassung am Beispiel des Schwarzspechts *Dryocopus martius* (s. Text). Je eine Begehung in 100 besetzten Quadraten mit einer Antreffwahrscheinlichkeit von 41 % wurde 1 Mio. mal simuliert. Das Histogramm gibt die relative Häufigkeit jeder Beobachtung («Schwarzspecht in X Quadraten entdeckt») an, d.h., in wie vielen Fällen (auf der y-Achse) der Schwarzspecht in wie vielen Quadraten (auf der x-Achse) festgestellt wurde. – *Illustration of the binomial distribution as the theoretical basis of bird surveys in the Black Woodpecker *Dryocopus martius* example. In this simulation, a single survey was conducted in each of 100 occupied quadrats, each having a probability of detection of 41 %. This was repeated 1 M times. The histogram shows the relative frequency with which each count (i.e., number of observed, occupied quadrats) was obtained.*

In Abb. 1 habe ich dieses Szenario mit 100 besetzten Quadraten, je einer Begehung und  $p = 0,41$  mit Hilfe der Binomialverteilung 1 Mio. mal im Programm R (R Development Core Team 2007) simuliert. Man erkennt, dass auch unter identischen Bedingungen die Zählungen völlig unterschiedlich ausfallen können. Stellt man sich vor, dass das Szenario in Abb. 1 zustande kam, indem man 1 Mio. exakt gleich gute Beobachter je einmal in 100 besetzte Quadrate schickte, so ist es möglich, dass die ersten zehn Beobachter mit folgenden Zählungen besetzter Schwarzspechtquadrate zurückkämen: 42, 46, 47, 40, 38, 44, 53, 32, 36, 34. Es ist schon beachtlich: Obwohl ich diese Zahlen so generiert habe, dass alle Beobachter genau gleich gut sind, gab es trotzdem schon unter den ersten 10 Beobachtern Unterschiede von bis zu 19 besetzten Quadraten! Aus einer einzelnen Beobachtung wie «Beobachter A hat Art B in X Quadraten entdeckt» kann man also nichts über die Qualität von Beobachter A schließen.

Ebenso wenig kann man bei der Qualitätsüberprüfung in einem Inventar einen zweiten

Beobachter in jedes Quadrat schicken und dann erwarten, dass beide Beobachter dieselben Zählungen erzielen werden. Wie häufig dieselben zwei Zahlen resultieren, hängt in erster Linie von der Anzahl  $N$  besetzter Quadrate und der Antreffwahrscheinlichkeit  $p$  ab. Immer wenn  $p < 1$  wird man auch mit identischen Beobachtern viel häufiger unterschiedliche als zwei identische Zählungen erhalten. Das ist in Abb. 2 gezeigt, wo ich angenommen habe, dass zwei Beobachter unabhängig voneinander 100 Quadrate bearbeiten, von denen eine gewisse Anzahl  $N$  vom Schwarzspecht besetzt sei. Beide haben genau die gleiche Wahrscheinlichkeit  $p$ , ein besetztes Quadrat ausfindig zu machen, d.h. einen Schwarzspecht zu sehen. Dieses Szenario habe ich 100 000-mal simuliert. In nur rund 16 % der Fälle kamen diese exakt gleich kompetenten Beobachter mit einem identischen Zählresultat (z.B. «Schwarzspecht in 20 Quadraten gefunden») heim.

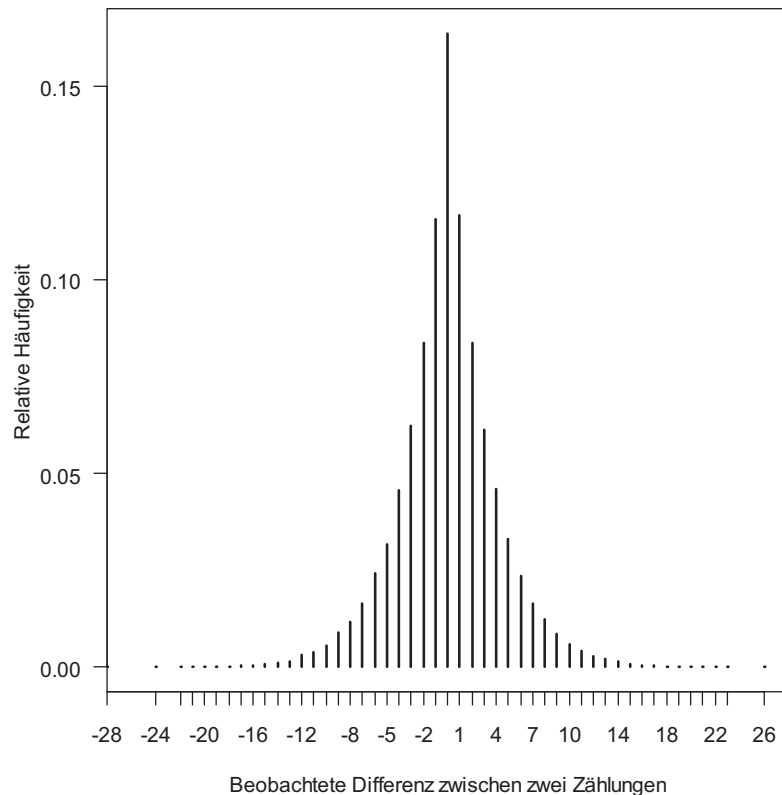
Es ist zwar schon so, dass Zählungen zusätzlich noch unterschiedlicher werden, wenn die beiden Beobachter unterschiedlich kompetent sind, doch man muss klar erkennen, dass man



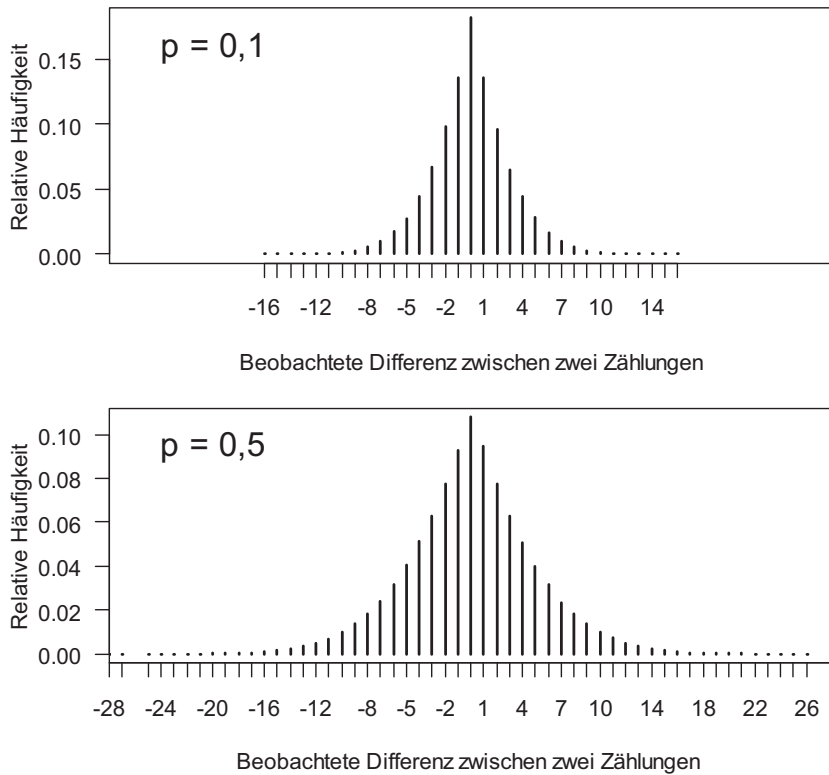
auch bei exakt gleich guten Beobachtern aus reinem (binomialem) Zufall in den meisten Fällen unterschiedliche Zählungen erhalten wird. Es macht also wenig Sinn, bei Qualitätsüberprüfungen von Monitoringprogrammen darauf zu schauen, ob man zweimal dieselbe Zählung erhält. Viel besser wäre es,  $p$  zu messen: Ein hohes  $p$  zeigt bei gleichen Kosten ein besseres, weil sensitiveres, Monitoringprogramm an als eines mit niedrigerem  $p$  (zur Messung von  $p$  s. Kap. 6f.).

Es ist sogar so, dass eine Fokussierung auf die Variation von wiederholten Zählungen oder

auf die Variation von Differenzen zwischen wiederholten Zählungen als Qualitätsmerkmal von Monitoringprogrammen zu irreführenden Ergebnissen führen kann. Das ist in Abb. 3 gezeigt, wo ich je 100 000 zweifach wiederholte Zählungen besetzter Quadrate aus zwei Monitoringprogrammen simuliert habe. Programm A hat eine «Empfindlichkeit» (im Sinne eines Messinstruments) von  $p = 0,1$ , d.h. es werden durchschnittlich 10 % aller Vorkommen gefunden. Programm B hat eine 5-mal höhere Empfindlichkeit, und es werden bei  $p = 0,5$  im Durchschnitt die Hälfte aller Vorkommen ge-



**Abb. 2.** Entstehung von rein zufälligen Differenzen zwischen zwei wiederholten Zählungen besetzter Quadrate unter völlig identischen Bedingungen, d.h. bei identischem  $N$  (= wahre Anzahl besetzter Quadrate) und Antreffwahrscheinlichkeit  $p$  in beiden binomialen Zählungen. Angegeben ist die relative Häufigkeit der beobachteten Differenzen unter 100 000 zufälligen Wiederholungen, wobei  $N$  und  $p$  für jedes Paar zufällig aus den Bereichen 0–100 bzw. 0–1 gewählt wurden. Nur in etwa 16 % erzielen die identischen Beobachter dasselbe Zählresultat. – *Owing to chance, two repeated counts of occupied quadrats will normally differ even when conducted under fully identical conditions, i.e., with identical  $N$  (= true number of occupied quadrats) and identical detection probability  $p$ , in both binomial counts. This is shown in this simulation that shows the distribution of 100 000 differences between two repeated counts that were obtained by randomly sampling  $N$  and  $p$  from 0–100 and 0–1, respectively. Only in c. 16 % of cases were repeated counts identical.*



**Abb. 3.** Die Betrachtung der Streuung von Differenzen zwischen wiederholten binomialen Zählungen als Qualitätsmerkmal von Monitoringprogrammen kann irreführen: Diese Differenzen sind bei Programm A mit  $p = 0,1$  (oben) automatisch viel kleiner als bei Programm B mit  $p = 0,5$  (unten), obwohl Programm A eine 5-mal geringere Empfindlichkeit (im Sinne einer Messmaschine) hat als Programm B. Diese Simulation beruht auf 100000 zweifach wiederholten Zählungen besetzter Quadrate. – *In the quality control of monitoring programs, a focus on the variation of differences between repeated binomial counts can be misleading. The simulation shows that, as a simple consequence of the «inbuilt variation» in binomial counts, these differences are much smaller in program A (above,  $p = 0.1$ ) than in program B (below,  $p = 0.5$ ), although program A has a sensitivity (= detection probability) 5 times less than program B. The simulation was conducted with 100000 doubly repeated counts of occupied quadrats.*

funden. Da die Streuung von Zählungen aufgrund der Binomialverteilung aber von  $p$  (und  $N$ ) bestimmt ist (s. Formel für die binomiale Standardabweichung unten), haben die Zählungen in Programm A (Abb. 3 oben) eine viel geringere Streuung als jene in Programm B (Abb. 3 unten). Doch viele Leute würden wohl so empfinden, dass Programm B besser sei, obwohl dort bloss 10,7 % aller wiederholten Zählungen identisch sind (gegenüber 18,2 % in Programm A) und die mittlere absolute Differenz zwischen wiederholten Zählungen 3,8 beträgt (gegenüber 2,3 in Programm B).

Hier ist eine weitere wichtige Folgerung für die Praxis aus der Beobachtung, wonach Zählungen selbst unter identischen Bedingungen in der Regel unterschiedlich ausfallen werden. Regelmässig werden in Inventaren oder Monitoringprogrammen pro Quadrat und Jahr zwei oder mehr wiederholte Beobachtungen gemacht, danach wird aber bloss die grösste Zählung behalten und die kleineren werden sozusagen weggeschmissen. Sofern die Bedingungen zum Zeitpunkt der beiden Zählungen etwa vergleichbar waren (wenn also nicht beispielsweise während einer Zählung dichter

Nebel lag, es stark windete oder regnete) und somit davon ausgegangen werden kann, dass sich  $p$  nicht zu stark unterschied, so sollten beide Zählungen als gleichermassen gültige Beobachtungen aus demselben binomialen Zufallsprozess betrachtet werden. Beide Zählungen, sowohl die grössere als auch die kleinere, sind gleichermassen gültige Indices für  $N$  und enthalten dieselbe Information über Muster in Vorkommen oder Häufigkeit. Alle Beobachtungen ausser die grösste wegzuerwerfen stellt eine Ressourcenverschwendung dar und verkennt die fundamentale Eigenschaft von Zählungen als Realisationen einer Zufallsvariable (vgl. auch Lindley 2006: 153).

Die Variation zwischen wiederholten Zählungen ist zwar im Einzelfall nicht vorhersagbar, folgt aber durchaus gewissen Gesetzmässigkeiten, d.h. sie kann im Durchschnitt vorhergesagt werden. Beispielsweise reichten im in Abb. 1 simulierten Fall die Zählungen unter identischen Bedingungen zwar von 18 bis 64, doch die Abbildung zeigt klar, dass gewisse Zählungen viel wahrscheinlicher sind und in der Simulation viel häufiger vorkamen als andere. Und die Fälle jenseits von 18 und 64 sind dermassen unwahrscheinlich, dass sie bei einem «kleinen» Stichprobenumfang von bloss 1 Mio. Wiederholungen nie beobachtet wurden, obwohl sie theoretisch durchaus möglich sind. Insgesamt sind die Fälle in der Gegend von etwa 40 Quadraten mit Schwarzspechtbeobachtung geklumpt und werden nach oben und unten immer weniger häufig. In der Tat ist der Erwartungswert einer Binomialverteilung, also der Durchschnitt, den man erhielte, wenn man den identischen Zählprozess sehr viele Male wiederholen würde, gleich  $N * p$  und liegt in unserem Fall bei  $100 * 0,41 = 41$  Quadraten. Auch die Standardabweichung, ein Mass für die Streuung über viele Wiederholungen, ist im Schwarzspechtbeispiel aufgrund der Theorie der Binomialverteilung bekannt; sie beträgt  $\sqrt{N * p * (1 - p)}$  oder  $\sqrt{100 * 0,41 * (1 - 0,41)}$ , also 4,9 besetzte Quadrate. Aufgrund der Formel kann man auch sehen, dass es nur unter drei Bedingungen überhaupt möglich ist, dass wiederholte Zählungen immer identisch sind. Das ist dann der Fall, wenn die Standardabweichung gleich Null wird: (1)  $N = 0$ , also gibt es

nichts zu zählen, (2)  $p = 0$ , wir möchten also unsichtbare Vögel zählen, und (3)  $p = 1$ , wenn wir jeden Vogel immer entdecken. Alle drei sind entweder uninteressante oder unwahrscheinliche Fälle. Daher muss man in der Natur selbst unter identischen Bedingungen immer mit unterschiedlichen Zählungen rechnen. Und man kann auch sehen, dass die binomiale Streuung von Zählungen bei  $p = 0,5$  am grössten ist und bei  $p = 0,2$  gleich gross wie bei  $p = 0,8$  sein wird.

In der Realität werden Zählungen ( $n$ ) selten genau durch eine simple Binomialverteilung beschrieben werden, denn  $p$  wird kaum je für alle  $N$  Quadrate und allfällige wiederholte Zählungen identisch sein. Vielmehr wird  $p$  von Quadrateigenschaften wie nicht zuletzt dem Bestand in jedem Quadrat abhängen (s. Formel 2 unten, und auch Royle & Nichols 2003), vom Habitat im Quadrat sowie von vielen weiteren Dingen, die sich zwischen Quadraten und wiederholten Zählungen unterscheiden mögen. Das bedeutet, dass  $n$  sogenannten bedingt binomial verteilt sein wird, d.h. dass gewissermassen in jedem Fall eine andere Münze geworfen wird. Ferner kann es sein, dass Beobachtungen nicht unabhängig voneinander sind, und das wird sogenannte Überdispersion in die Zählungen bringen, d.h. zusätzliche Variation, über das Mass hinaus, das durch die Binomialverteilung via  $\sqrt{N * p * (1 - p)}$  bereits vorgegeben ist.

Dennoch ist es für das konzeptionelle Verständnis der Bestandserfassung unumgänglich, sich auch in diesen Fällen Zählungen als direkte Realisationen oder als Funktionen (z.B. Summen) von binomial- oder bedingt binomialverteilten Zufallsgrössen vorzustellen. Für die Analyse von solchen Daten können viele zusätzliche Komplexitäten (Variation in  $p$ , Überdispersion) berücksichtigt werden. So nimmt beispielsweise das Modell für die Wanderfalken- und Kolkrabenanalyse (s. Kap. 9, Beispiel 1) nicht an, dass die Antreffwahrscheinlichkeit immer identisch sei, sondern lässt zu, dass sie sich als Funktion von Datum und Dauer der Beobachtung unterscheidet. Je nach Datenmenge, Art des Beobachtungsprotokolls oder biologischer Fragestellung könnte man eine Vielzahl weiterer Komplexitäten in die Analyse einbauen.

#### 4. Formale Beschreibung des Zusammenhangs zwischen Zählung und Bestand

Im Schwarzspechtbeispiel, das einem Münzwurf mit  $P(\text{Kopf}) = 0,41$  entspricht, haben wir gesehen, dass eine Zählung besetzter Quadrate nicht das Gleiche ist wie der wahre Bestand bzw. die wahre Anzahl besetzter Quadrate. So ist es bei der Antreffwahrscheinlichkeit von  $p = 0,41$  fast unmöglich, den Schwarzspecht in allen 100 besetzten Quadraten mit je einer einzigen Begehung nachzuweisen. Aufgrund der theoretischen Grundlage aller Vogelzählungen, der Binomialverteilung, kann man aber den Zusammenhang zwischen Zählung und Bestand im Durchschnitt beschreiben. Im Durchschnitt wird eine Zählung der besetzten Quadrate ( $n$ ) gleich dem Erwartungswert  $E(n)$  der Zufallsvariable (der Zählung) und somit gleich dem Produkt von Bestand ( $N$ , 100 besetzte Quadrate) und Antreffwahrscheinlichkeit ( $p$ , also entsprechend  $P(\text{Kopf}) = 0,41$ ) sein und 41 Quadrate betragen.

$$E(n) = N * p \quad \text{Formel 1}$$

Diese Formel ist so fundamental für Vogelbestandserfassungen, dass ich ihr einen Namen gegeben habe: Sie kann ein wenig augenzwinkernd als die «wichtigste Formel für die Feldornithologen» bezeichnet werden, weil sie den Zusammenhang zwischen einer Zählung und der Realität via den Erwartungswert einer binomialen Zählung  $n$  als Produkt aus  $N$  und  $p$  beschreibt.

Eine zweite wichtige Formel, die sich ebenfalls direkt aus der Binomialverteilung als Modell für Vogelzählungen herleiten lässt, ist diese:

$$P_{\text{gesamt}} = 1 - (1 - p)^W \quad \text{Formel 2}$$

Diese «zweitwichtigste Formel für die Feldornithologen» beschreibt die gesamte Antreffwahrscheinlichkeit  $P_{\text{gesamt}}$  als Funktion der Antreffwahrscheinlichkeit für einen einzelnen «Fall» ( $p$ ) und der Anzahl «Wiederholungen» ( $W$ ). Diese Formel hat mehrere mögliche Anwendungen und Interpretationen.

Einerseits kann sie dazu dienen, den Zusam-

menhang auszudrücken zwischen der Antreffwahrscheinlichkeit eines einzelnen Individuums einer Art ( $p$ ) und der Wahrscheinlichkeit  $P_{\text{gesamt}}$  die Art in einem Quadrat bei einem Besuch festzustellen, wenn dort  $W$  Individuen leben. Die Wahrscheinlichkeit, die Art im Quadrat bei einem Besuch festzustellen, ist gleich der Wahrscheinlichkeit, mindestens eines der  $W$  Individuen zu entdecken. Das ist gleich der Gegenwahrscheinlichkeit (1 minus der Wahrscheinlichkeit) des Falles, gar keines der  $W$  Individuen zu entdecken. Ein Zahlenbeispiel verdeutlicht das: Nehmen wir an, die Antreffwahrscheinlichkeit ( $p$ ) für jedes der  $W = 4$  Individuen des Mittelspechts *Dendrocopos medius* in einem Quadrat betrüge unabhängig voneinander identisch  $p = 0,185$ . Die Wahrscheinlichkeit für einen einzelnen Mittelspecht, bei einer Begehung nicht entdeckt zu werden, beträgt somit  $q = 1 - p = 0,815$ . Die Mittelspecht-Antreffwahrscheinlichkeit bei einem Besuch eines besetzten Quadrats ist dann gleich der Gegenwahrscheinlichkeit des Ereignisses, dass keines der  $W = 4$  Individuen entdeckt wird. Die Wahrscheinlichkeit, dass keines entdeckt wird, beträgt bei unabhängiger Entdeckung der Individuen  $q^W = q^4 = (1 - p)^4$ , also gleich  $0,815^4 = 0,44$ . Die Wahrscheinlichkeit, mindestens einen Mittelspecht zu entdecken, und somit die quadratbezogene Antreffwahrscheinlichkeit der Art, ist dann gleich  $1 - 0,44 = 0,56$ . (Diese Zahlen sind übrigens gerade so gewählt, dass sie die mittlere quadratbezogene Antreffwahrscheinlichkeit des Mittelspechts im MHB 2001–2004 ergeben; s. Kéry & Schmidt im Druck.) Wenn hingegen bloss  $W = 3$  Mittelspechte im Quadrat leben, so beträgt die mittlere quadratbezogene Antreffwahrscheinlichkeit  $1 - (1 - 0,185)^3$ , also 46 %, und bei  $W = 2$  beträgt sie 34 %. Die Antreffwahrscheinlichkeit bezogen auf die Besetzung eines Quadrats hängt somit immer auch vom Bestand in diesem Quadrat ab.

Eine zweite wichtige Anwendung der «zweitwichtigsten Formel für die Feldornithologen» besteht darin, die gesamte Antreffwahrscheinlichkeit für ein besetztes Schweizer Mittelspechtquadrat  $P_{\text{gesamt}}$  bei  $W$  wiederholten Begehungen mit konstanter Antreffwahrscheinlichkeit  $p$  auszudrücken (man beachte, hier be-

zeichnet  $W$  nicht die Anzahl Individuen). Bei  $W = 2$  Begehungen in einem durchschnittlichen Quadrat ergibt sich beispielsweise  $P_{\text{gesamt}} = 1 - 0,44^2 = 0,81$ . Bei zwei Begehungen eines besetzten Quadrats erwartet man, dass im Durchschnitt in 81 % der Fälle bei mindestens einer der beiden Begehungen auch Mittelspechte festgestellt werden. Bei  $W = 3$  Begehungen mit konstantem  $p$  ergibt sich  $P_{\text{gesamt}} = 1 - (1 - 0,56)^3$  oder  $P_{\text{gesamt}} = 1 - (1 - 0,44)^3 = 0,91$ . Im MHB werden normalerweise drei Begehungen pro Brutsaison durchgeführt. Dann erwartet man also, dass der Mittelspecht in 91 % der Fälle in einem besetzten Quadrat mindestens einmal auch festgestellt wird.

Schliesslich ist eine Kenntnis der «typischen Antreffwahrscheinlichkeit» einer Art in einem gewissen Gebiet bezogen auf ein bestimmtes Feldprotokoll nützlich für die Untersuchungsplanung. Zusammen mit der zweitwichtigsten Formel erlaubt sie nämlich die Bestimmung des minimal nötigen Stichprobenumfangs ( $W_{\text{min}}$ ) für den «Negativnachweis» einer Art (Kéry 2002). Betrachten wir ein Zahlenbeispiel: Wir möchten wissen, wie oft man ein MHB-Quadrat besuchen muss, um «sicher» zu sein, einen allfällig vorkommenden Mittelspecht nicht zu übersehen. Wir wissen, dass die durchschnittliche Antreffwahrscheinlichkeit in einem besetzten Quadrat im MHB bei etwa 56 % liegt. Wir haben im vorhergehenden Abschnitt gesehen, dass die totale Antreffwahrscheinlichkeit  $P_{\text{gesamt}}$  mit zunehmender Anzahl Begehungen  $W$  zunimmt. Wenn man nun wie oft üblich eine Wahrscheinlichkeit von 95 % mit «Sicherheit» gleichsetzt, so kann man folgende Überlegung machen: Wir möchten  $W$  in der zweitwichtigsten Formel so wählen, dass  $P_{\text{gesamt}}$  mindestens 95 % beträgt. Die folgenden Umformungen führen uns zum gewünschten Rezept, das uns erlaubt, für unsere Definition von «Sicherheit» (z.B. 95 %, d.h. 0,95) und eine bekannte Antreffwahrscheinlichkeit  $p$  die notwendige Anzahl Begehungen zu bestimmen, um eine Art «sicher» nachzuweisen.

Wir beginnen mit  $P_{\text{gesamt}} = 1 - (1 - p)^W$ , addieren  $(1 - p)^W$  auf beiden Seiten und erhalten  $P_{\text{gesamt}} + (1 - p)^W = 1$ . Dann subtrahieren wir  $P_{\text{gesamt}}$  auf beiden Seiten und erhalten  $(1 - p)^W = 1 - P_{\text{gesamt}}$ . Logarithmieren ergibt

$\log((1 - p)^W) = \log(1 - P_{\text{gesamt}})$ , und das ist das gleiche wie  $W * \log(1 - p) = \log(1 - P_{\text{gesamt}})$ . Die minimale Anzahl Wiederholungen errechnet sich also nach Division beider Seiten mit  $\log(1 - p)$  so:

$$W_{\text{min}} = \frac{\log(1 - P_{\text{gesamt}})}{\log(1 - p)} \quad \text{Formel 3}$$

$\log$  bezeichnet die logarithmische Transformation, egal ob zur Basis 2,  $e$  oder 10. Wenn man sich ein  $P_{\text{gesamt}}$  von mindestens 95 % wünscht, so ergibt das im Zähler  $\log(0,05)$ , und das Resultat für den Mittelspecht im MHB lautet dann:

$$W_{\text{min}} = \frac{\log(1 - 0,95)}{\log(1 - 0,56)} = 3,65$$

Das runden wir auf und sagen, dass ein Mittelspechtvorkommen in einem durchschnittlichen MHB-Quadrat nach spätestens vier Besuchen vom Typ einer MHB-Begehung gefunden wird, wenn die Antreffwahrscheinlichkeit konstant gleich 0,56 ist.

Wenn  $p$  zwischen den Begehungen variiert, so wird die Verwendung einer konstanten Antreffwahrscheinlichkeit in Formel 3  $P_{\text{gesamt}}$  über- und  $W_{\text{min}}$  unterschätzen (Kéry 2002). Tendenziell ergibt die Formel 3 also eine untere Grenze für den nötigen Stichprobenumfang für den Negativnachweis.

## 5. Haben solche Zahlenspielerien Praxisrelevanz ?

Hier wird sich vielleicht mancher Leser denken: «Ist ja interessant, doch bringt diese Zahlengymnastik einem Feldornithologen wirklich etwas?». Ich glaube schon. Ich bin überzeugt, dass es für jeden zählenden Ornithologen bedeutsam ist, ob er die binomiale Theorie der Bestandserfassung in ihren Grundzügen verstanden hat. Ich würde sogar behaupten, dass man erst dann in der Lage sein wird, Felduntersuchungen adäquat und effizient zu planen, durchzuführen und die Resultate von Bestandserfassungen kompetent zu interpretieren, wenn man sich ein gewisses Verständnis für die ge-



schilderten Sachverhalte erarbeitet hat. Und natürlich führt nicht jeder Feldornithologe seine eigenen Untersuchungen durch, doch Zählungen irgendeiner Art machen die meisten von uns, und wir interpretieren sie dann auch.

Die grosse Bedeutung des bisher Dargelegten liegt also darin, dass Zählungen oder Vorkommensbeobachtungen nicht direkt mit wirklichen Beständen oder wirklichem Vorkommen gleichgesetzt werden können, denn Beobachtungen sind in Bezug auf wirkliche Bestände immer mit der Antreffwahrscheinlichkeit kontaminiert; das zeigt uns Formel 1. Es gibt nur drei Ausnahmefälle, wann eine Zählung  $n$  generell, d.h. wiederholt und nicht bloss in einem Einzelfall, gleich dem Bestand  $N$  ist: (1) wenn der Bestand  $N$  gleich Null ist, (2) wenn die Antreffwahrscheinlichkeit  $p$  gleich Null ist oder (3) wenn  $p$  gleich 1 ist. Das sind exakt die gleichen drei Bedingungen, unter denen auch die Variation einer binomialen Zufallsgrösse Null wird. Die ersten beiden Fälle sind ziemlich uninteressant, und der dritte ist unwahrscheinlich für mobile Tiere mit relativ grossem Aufenthaltsgebiet, denn selbst Pflanzenvorkommen werden auch in spezialisierten Inventaren regelmässig übersehen (Kéry et al. 2006). Wenn man also eine Zählung als wirklichen Bestand interpretiert, so macht man faktisch die sehr restriktive Annahme  $p = 1$ . Ich meine, diese Annahme sollte getestet werden, d.h. es sollten Zahlen produziert werden, die belegen, dass  $p$  wirklich sehr nahe bei 1 ist. Im Minimum sollte diese Annahme explizit deklariert werden.

In Realität werden die wirklichen Bestände selbst in gut geplanten und durchgeführten Bestandserhebungen leicht um einen Faktor 2 oder mehr unterschätzt, wie das Beispiel des Eichelhäfers *Garrulus glandarius* im MHB zeigt (Royle et al. 2007). Das Problem besteht einerseits darin, dass die 1 km<sup>2</sup> messenden Quadrate bei den einzelnen Rundgängen nicht vollständig kontrolliert werden können, teilweise weil einzelne Bereiche gar nicht zugänglich sind; die Transekttrouten decken also nicht die gesamte Quadratfläche ab. Andererseits werden einzelne Reviere natürlich auch entlang der Transekte übersehen, also im durch die Transekttrouten eigentlich abgedeckten Bereich. Da somit nicht alle Reviere gefunden werden,

erlauben die Ergebnisse keine effektive Dichteschätzung. Aus diesem Grund wird im zweiten Schweizer Brutvogelatlas (Schmid et al. 1998) von «Atlasdichten» gesprochen, also von der Anzahl gefundener Reviere pro Kilometerquadrat, einem eigentlich nur schwer fassbaren Mass für den realen Bestand in einem Quadrat. Und im Falle von Vorkommen und Verbreitung kann im MHB die Verbreitung einzelner Arten durchaus um den Faktor 2 unterschätzt werden (Kéry & Schmidt im Druck; s. auch unten).

Sobald  $p < 1$ , so ist ein weiterer wichtiger Punkt, dass es keine richtigen oder falschen Zählungen gibt. Wenn Arbeitsgruppe A den Schwarzspecht in Gebiet A in 41 von 100 Quadraten entdeckt und eine andere in einem gleich grossen Gebiet B in 56 von 100, so sind beide Zählungen gleichermaßen «richtig». Hingegen ist die Interpretation, wonach der Schwarzspecht in Gebiet B 1,37-mal (56/41) weiter verbreitet ist als in Gebiet A, falsch, denn der Unterschied von 15 (d.h. 56–41) Quadraten kann durch eine real unterschiedliche Verbreitung in den Gebieten, eine unterschiedliche mittlere Antreffwahrscheinlichkeit oder aber durch «binomialen Zufall» zustande gekommen sein.

Nun hört man oft den Einwand, dass systematische Fehler in einer Schätzung (also eine Zählung bei  $p < 1$ ) gar nicht so bedeutungsvoll seien, sofern immer nur gleich falsch geschätzt werde («Immer gleich falsch ist auch wieder richtig»). Es stimmt, dass das Interesse bei Bestandsaufnahmen sehr oft weniger bei den absoluten als vielmehr bei relativen Zahlen liegt. Man fragt also nicht «wie viele sind es genau?», sondern einfach «sind es mehr oder weniger als im Vorjahr oder im anderen Gebiet?». Man möchte beispielsweise wissen, ob die Anzahl der besetzten Quadrate einer bedrohten Art abnimmt, um gegebenenfalls früh genug Schutzmassnahmen einleiten zu können. Man kann richtigerweise argumentieren, dass es in solchen Fällen egal sei, ob man die Art in allen Quadraten festgestellt hat, in denen sie vorkommt, oder sie bloss im Durchschnitt immer in etwa 80 % der besetzten Quadrate findet. In beiden Fällen wird man einen Arealschwund trotzdem erkennen, auch wenn die Zählung mit  $p < 1$  mit einem grösseren «Rauschen» daher-

kommt, d.h. mit einer grösseren Schätzunsicherheit behaftet ist. Ein anderes Beispiel wäre, dass man die wichtigsten Habitattypen für eine seltene Art schützen möchte und die Anzahl besetzter Quadrate pro Habitattyp als Mass für die Habitateignung interpretiert. Auch hier stellt sich natürlich die Frage, ob denn die ganzen Komplikationen mit  $p$  überhaupt irgendeine Relevanz haben? Man möchte ja bloss wissen, in welchem Habitat die Vogelart öfter vorkommt, und der absolute Wert des Vorkommens ist einem egal.

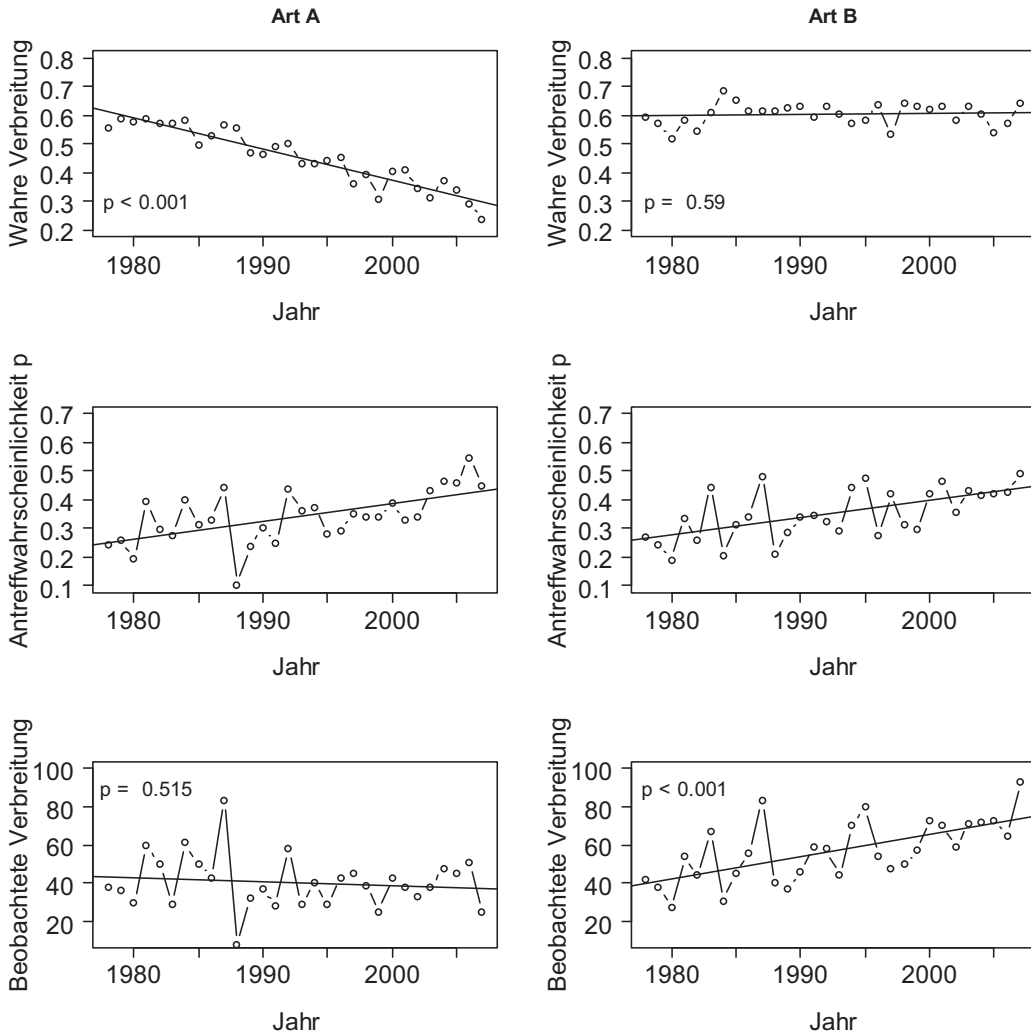
Der springende Punkt ist jedoch, dass man normalerweise keine Ahnung hat, ob man immer gleich falsch misst oder nicht. Wenn man einmal 20 % falsch misst und ein andermal 50 % daneben liegt, so hat man selbst dann ein Problem, wenn man nur an relativen Zahlen interessiert ist. Und so hat die «wichtigste Formel des Feldornithologen» eben auch bei relativen Zahlen eine Bedeutung. Hier ist ein Zahlenbeispiel: Angenommen, wir wollen die Vorkommenshäufigkeit des Waldlaubsängers *Phylloscopus sibilatrix* im Buchen- mit jener im Eichenwald vergleichen und bestimmen in je 50 Quadraten mit 250 m Seitenlänge, ob er vorkommt oder nicht. Im Eichenwald finden wir 22 besetzte Quadrate und im Buchenwald deren 15. Damit wären wir wohl versucht zu sagen, der Eichenwald sei fast 50 % beliebter ( $22 / 15 = 1,47$ ) als der Buchenwald.

Schauen wir mal hinter die Kulissen des Prozesses, in dem die beiden Zählungen von durch den Waldlaubsänger besetzten Quadrate entstanden sein mögen: Im Eichenwald gibt es eine unbekannte Anzahl  $N$  Eichenwald besetzte Quadrate und im Buchenwald eine unbekannte Anzahl  $N_{\text{Buchenwald}}$  besetzte Quadrate. Wir sind eigentlich am Verhältnis  $N_{\text{Eichenwald}} / N_{\text{Buchenwald}}$  interessiert, aber wir können das nicht direkt beobachten. Unsere Beobachtungen hingegen entsprechen  $N_{\text{Eichenwald}} * p_{\text{Eichenwald}}$  im Eichen- und  $N_{\text{Buchenwald}} * p_{\text{Buchenwald}}$  im Buchenwald. Das heisst, wenn wir die 22 beobachteten Quadrate im Eichenwald und die 15 im Buchenwald direkt miteinander vergleichen und somit genau gleich werten, so machen wir faktisch die Annahme, dass der Waldlaubsänger in beiden Biotopen gleich gut nachweisbar sei, also beispielsweise dass  $p_{\text{Eichenwald}} = p_{\text{Bu-}}$

chenwald = 0,70. Das muss aber nicht sein, und diese wichtige Annahme sollte wenn möglich getestet werden. Es wäre ja möglich, dass wir im Eichenwald vor allem Durchzügler beobachten oder unverpaarte ♂, die sehr aktiv singen (vgl. Amrhein et al. 2007 für ein Beispiel bei der Nachtigall *Luscinia megarhynchos*). Im bevorzugten Buchenwald finden die ♂ rascher ein ♀ und beginnen sofort zu brüten. Sie singen danach viel seltener und sind deshalb weniger auffällig. Aufgrund der Beobachtungen wäre es dann möglich, dass man genau das Gegenteil von dem schliesst, was in Wirklichkeit der Fall ist.

In Abb. 4 ist ein weiteres Zahlenbeispiel dargestellt. Ich habe darin die Bestandsentwicklung von zwei imaginären Arten für die 30 Jahre zwischen 1978 und 2007 aufgetragen (Kéry & Schmidt im Druck). Die Bestandsentwicklung wurde hier ausgedrückt als Anteil der besetzten Quadrate unter 267 untersuchten Quadraten. Art A erleidet einen steilen und statistisch signifikanten Bestandsrückgang, während der Bestand von Art B konstant bleibt (Abb. 4 oben). Bei beiden Arten nimmt aber die Wahrscheinlichkeit, ein besetztes Quadrat zu entdecken, über die gleiche Periode stark zu, vielleicht als Folge einer immer grösseren Zahl von Beobachtern (Abb. 4 Mitte). Das beobachtete Resultat (Abb. 4 unten) ist eine scheinbar konstante Verbreitung für die abnehmende Art A, aber eine Zunahme bei Art B mit in Realität konstanter Verbreitung. Dieses Beispiel zeigt eindrücklich, dass auch relative Zahlen durch die Antreffwahrscheinlichkeit verfälscht werden können, wenn ähnliche Muster in der Verbreitung und in der Antreffwahrscheinlichkeit vorkommen.

Schliesslich kann das Addieren von Zählungen heikel sein. Möchte man beispielsweise im Waldlaubsängerbeispiel eine Aussage über die Verbreitung im gesamten Untersuchungsgebiet machen, so ist es problematisch, einfach die Anzahl der beobachteten besetzten Quadraten in beiden Waldtypen zusammenzuzählen. Dieses Vorgehen ist nur dann zulässig, wenn die Antreffwahrscheinlichkeit in beiden Habitaten identisch ist. Ansonsten macht man nämlich eine Rechnung vom Typ «30 Britische Pfund + 20 Türkische Lira = 50». Ja, 30 plus 20 ergibt



**Abb. 4.** Schlüsse über «relatives Vorkommen» können durch «gefährliche Muster» in der Antreffwahrscheinlichkeit  $p$  irrefleitet werden, hier am Beispiel mit zwei simulierten Datensätzen mit je 267 Quadraten dargestellt. Zeitliche Trends in  $p$  können dazu führen, dass effektiv vorhandene Ausbreitungstendenzen in den Beobachtungen maskiert werden (Art A) oder aber, dass nicht existente Ausbreitungstendenzen in den Beobachtungen irrtümlich diagnostiziert werden (Art B). – «Dangerous patterns» in detection probability  $p$  can mislead conclusions even about «relative occurrence», as shown by two simulated data sets with 267 quadrats each. Time trends in  $p$  can lead to a masking of true distributional trends (species A) or lead to the erroneous diagnosis of nonexistent distributional trends (species B).

50, aber 50 was? In diesem Beispiel ist die Unsinnigkeit einer solchen Addition von Äpfeln und Birnen offensichtlich, und dennoch machen wir sehr oft genau das gleiche bei der Untersuchung von Vorkommen und Verbreitung von Vögeln. In Bezug auf die wahren Bestände

ist die Antreffwahrscheinlichkeit gewissermaßen ein Skalierungsfaktor für die Beobachtungen, genau gleich wie es die Umrechnungskurse bei den Währungen sind! (Allerdings sind bei Vogelzählungen die Umrechnungskurse stochastisch, d.h. sie sind nur im Durchschnitt

definiert.) Site-Occupancy-Modelle stellen gewissermaßen die Analogie zu einem Wechselkursberechner dar.

Zusammenfassend muss man bezüglich der Antreffwahrscheinlichkeit  $p$  immer dann anpassen, wenn wir (1) Interesse an den effektiven Beständen oder der vollständigen Verbreitung haben oder (2), wenn es möglicherweise «gefährliche Muster» in  $p$  gibt, d.h.  $p$  mit denselben Faktoren korreliert, deren Einfluss auf das Vorkommen wir untersuchen möchten (z.B. wenn sich Bestände und Antreffwahrscheinlichkeit nach Habitat unterscheiden). Es ist hingegen nicht nötig, dass  $p$  genau konstant sei, so wie das oft behauptet wird (z.B. Mazerolle et al. 2007), damit man Zählungen legitimerweise direkt interpretieren darf. Richtig ist, dass es keine Muster in  $p$  entlang einer gewünschten Vergleichsdimension der effektiven Verbreitung geben darf. Haben wir beispielsweise Interesse an einem zeitlichen Trend im Verbreitungsgebiet der Ringdrossel *Turdus torquatus* im Jura und bearbeiten immer dieselben Quadrate, so brauchen uns allfällige Habitatunterschiede in der Antreffwahrscheinlichkeit der Ringdrossel nicht zu kümmern, wohl aber zeitliche Trends in der Antreffwahrscheinlichkeit, wie wir am vorgängigen Beispiel in Abb. 4 gesehen haben. Wenn wir umgekehrt die Habitatbevorzugung des Waldlaubsängers über mehrere Jahre untersuchen, so brauchen wir einen allfälligen zeitlichen Trend in  $p$  nicht zu fürchten, sofern dieser in allen Habitaten gleichartig verläuft.

## 6. Wie soll man vorgehen, wenn die Antreffwahrscheinlichkeit ein Problem sein könnte?

Was kann man also tun, wenn man entweder einen exakten Wert für Vorkommen/Verbreitung benötigt (z.B. für die Klassifikation einer Art nach Rote-Liste-Status), wenn man beweisen möchte, dass  $p$  nahe bei 1 oder konstant über die gewünschten Vergleichsdimensionen sei oder wenn man «gefährliche Muster» in der Antreffwahrscheinlichkeit befürchtet? Dann bleibt nichts anderes übrig, als die Antreffwahrscheinlichkeit  $p$  (und somit via Formel 1 auch den Bestand  $N$ ) formal zu schätzen. Das

kann auf verschieden komplizierte Art und Weise geschehen. Die einfachste Möglichkeit besteht darin, möglichst viele als besetzt bekannte Quadrate wiederholt zu besuchen und die Antreffwahrscheinlichkeit ganz einfach aus dem Verhältnis der Anzahl Besuche mit Feststellung der Art zur Gesamtanzahl der Besuche zu schätzen. Im Fall von verschiedenen Habitaten muss man diese Übung natürlich für alle Habitattypen durchführen und bei Befürchtung eines zeitlichen Trends in  $p$  in verschiedenen Jahren wiederholen.

Ein Zahlenbeispiel aus unserem Waldlaubsängerbeispiel: Um eine simple Korrektur der Anzahl besetzter Quadrate zu erhalten, könnten wir beispielsweise in beiden Habitattypen je 6 Quadrate, von denen wir bereits vorgängig wissen, dass sie besetzt sind, an je 5 aufeinander folgenden Tagen aufsuchen und festhalten, ob wir die Art nachweisen oder nicht. Nehmen wir an, im Eichenwald verliefen 20 dieser wiederholten Besuche erfolgreich und 10 erfolglos, und im Buchenwald seien 11 Besuche erfolgreich und 19 erfolglos. Eine einfache Schätzung der Antreffwahrscheinlichkeit für den Waldlaubsänger in beiden Habitaten ist dann  $p_{\text{Eichenwald}} = 20/30 = 0,66$  und  $p_{\text{Buchenwald}} = 11/30 = 0,37$ . Somit könnten wir die wahre Zahl besetzter Quadrate in der 50er-Stichprobe aus vorhergehendem Kapitel durch eine Umformung der «wichtigsten Formel für die Feldornithologen» wie folgt schätzen:  $N_{\text{besetzte Eichenquadrate}} = 22/0,66 = 33,3$  und  $N_{\text{besetzte Buchenquadrate}} = 15/0,37 = 40,5$ . In diesem Beispiel hat also ein «gefährliches Muster» in der Antreffwahrscheinlichkeit  $p$  zu einer Umkehr der realen Habitatpräferenz des Waldlaubsängers in unseren Beobachtungen geführt.

Auf diese Art kann man die Antreffwahrscheinlichkeit sehr einfach schätzen, und es gilt einzig folgendes zu beachten: Wenn wir anfänglich nichts über die Besetzung eines Quadrats wissen, so dürfen wir in jedem Quadrat erst jene Begehungen zur Berechnung der Antreffwahrscheinlichkeit  $p$  verwenden, die nach der ersten positiven Feststellung stattfinden (Wintle et al. 2004). Besser als diese simple Methode ist es jedoch, eine formale statistische Methode zur Schätzung der Antreffwahrscheinlichkeit  $p$  und damit zur Korrektur der

beobachteten Verbreitung für die Antreffwahrscheinlichkeit  $p$  zu verwenden (s. Kap. 7).

### 7. Formale Schätzung von Vorkommen und Verbreitung mittels Site-Occupancy-Modell

Die sogenannten Site-Occupancy-Modelle stellen eine Verallgemeinerung einer logistischen Regression dar. Sie erlauben die separate Schätzung der Vorkommenswahrscheinlichkeit  $V$ , also des Anteils der besetzten Quadrate, in der Literatur oft mit dem griechischen Buchstaben  $\psi$  (psi) bezeichnet, und der Antreffwahrscheinlichkeit  $p$  der Art in einem besetzten Quadrat, d.h. der Wahrscheinlichkeit, mindestens eines unter den Individuen unbekannter Anzahl im Quadrat zu entdecken. Dadurch korrigieren sie die beobachtete Anzahl besetzter Quadrate vollständig für  $p < 1$ . Site-Occupancy-Modelle sind eine relativ neue Familie von statistischen Modellen (MacKenzie et al. 2002) und trotz ihres enormen Potenzials noch vielerorts unbekannt. Site-Occupancy-Modelle sind die Methode der Wahl, wenn man entweder die wahre Grösse des Verbreitungsgebiets einer Art kennen, die relative Verbreitung für mögliche «gefährliche Muster» in der Antreffwahrscheinlichkeit korrigieren oder die Antreffwahrscheinlichkeit schätzen möchte, um via Formel 3 den minimalen Suchaufwand für den Negativnachweis einer Art zu bestimmen. Site-Occupancy-Modelle beruhen direkt auf dem binomialen Modell der Bestandserfassung und stellen eine Art von gekoppelter logistischer Regression dar. Die erste logistische Regression beschreibt das wahre, aber unvollständig beobachtete Muster von Präsenz/Absenz und modelliert die Vorkommenswahrscheinlichkeit  $V$ , d.h. den erwarteten Anteil der besetzten Quadrate in einer Stichprobe an Quadraten. Die zweite logistische Regression beschreibt das Muster von Beobachtung/Nichtbeobachtung in den Quadraten mit Präsenz und beschreibt die Antreffwahrscheinlichkeit  $p$ .

Hier beschreibe ich Site-Occupancy-Modelle bloss in ihren Grundzügen. Das ursprüngliche Modell und viele seiner Erweiterungen sind in einem Buch (MacKenzie et al. 2006) und in folgenden Artikeln beschrieben: Mac-

Kenzie et al. (2002, 2003, 2005a, b), Royle & Nichols (2003), Wintle et al. (2004), Dorazio et al. (2006), Dorazio (2007), Nichols et al. (2007) und Royle & Kéry (2007). Für Biologen besonders zu empfehlende Einführungen sind MacKenzie (2005a, b), Pellet & Schmidt (2005) und auf Deutsch Schmidt (2008). Wichtige Artikel über die Versuchsplanung von Studien zu Vorkommen und Verbreitung sind MacKenzie & Royle (2005) und Bailey et al. (2007). Site-Occupancy-Modelle können mit folgenden Gratis-Windows-Programmen gerechnet werden: PRESENCE 2 (<http://www.mbr-pwrc.usgs.gov/software/doc/presence/presence.html>) und MARK (White & Burnham 1999; <http://welcome.warnercnr.colostate.edu/~gwhite/mark/mark.htm>) sowie in WinBUGS (Spiegelhalter et al. 2003; <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>) oder R (R Development Core Team 2007; R- und WinBUGS-Code zum Fitten des Site-Occupancy-Modells kann bei mir verlangt werden).

Wie in jeder herkömmlichen Studie von Vorkommen und Verbreitung braucht es zunächst einmal die Definition einer räumlichen und zeitlichen Beobachtungseinheit, die z.B. aus einem Waldstück oder einem MHB-Quadrat über eine Brutsaison bestehen kann. Diese Definition muss vom Ornithologen gewählt werden; der Statistiker hat damit nichts zu tun. Sodann müssen der gewünschte Zustand der Besetzung definiert werden sowie die Art der Beobachtung, die als hinreichender Hinweis auf ein solcherart definiertes Vorkommen gewertet wird. Ein mögliches Beispiel ist «Art zur Brutzeit in einem möglichen Brutbiotop festgestellt», also Atlascode 2. Auch die Definition des Zustands der Besetzung hat nichts mit dem Site-Occupancy-Modell zu tun; der Statistiker nimmt ganz einfach die räumliche Einheit und den definierten Besetzungszustand, die der Ornithologe ihm gibt. Beispielsweise ist es für die Anwendung des Modells irrelevant, dass der Zustand der Besetzung eines 1 km<sup>2</sup> grossen Quadrats im Falle einer Art mit grossem Revier wie z.B. dem Wespenbussard *Pernis apivorus* etwas anderes bedeutet als bei einer Art mit kleinem Revier wie dem Buchfinken *Fringilla coelebs*. Die geeignete Flächengrösse und die Definition und Interpretation des Zustands



der Besetzung bleibt Sache des Biologen, und Interpretationsschwierigkeiten bei Artvergleichen können völlig unabhängig von der Korrektur für  $p < 1$  mittels Site-Occupancy-Modell auftreten. Das einzige, was ein Site-Occupancy-Modell gegenüber einer herkömmlichen Verbreitungsstudie mehr liefert, ist eben die Korrektur für  $p < 1$  sowie die Möglichkeit, die Faktoren, die das Vorkommen beeinflussen, formal zu untersuchen und zwar separat von jenen, die bloss die Antreffwahrscheinlichkeit beeinflussen.

Der Schlüssel zur Anwendung der Site-Occupancy-Modelle und somit zur Schätzung des nach der Antreffwahrscheinlichkeit korrigierten Vorkommens bilden nun innerhalb eines relativ kurzen Zeitraums wiederholte Besuche zumindest in einem Teil aller Quadrate. Erst die wiederholte Beobachtung einer «geschlossenen Population» erlaubt die Aufteilung der Beobachtungen in ihre beiden Komponenten, nämlich das wirkliche Vorkommen und die Antreffwahrscheinlichkeit.

Die Feldarbeit kann nach gewohntem Muster verlaufen, d.h. man sucht jedes Quadrat unter geeigneten Bedingungen mit einer der Zielart angepassten Methodik ab. Dabei kann es sich um eine Revierkartierung, Linientaxation, Punkt-Stopp-Zählung oder irgendeine andere geeignete Methode handeln, mit der die gewünschte Art möglichst treffsicher und effizient nachgewiesen werden kann (Bibby et al. 2000, Südbeck et al. 2005). Wie immer ist auch hier ein standardisiertes Feldprotokoll wichtig. Schliesslich sollten beobachtungsrelevante Faktoren, die nicht konstant gehalten werden können, wenigstens gemessen werden, so dass ihr Einfluss bei der Analyse rechnerisch eliminiert werden kann. Beispiele wären Datum, Beobachtungsdauer, Tageszeit, Windstärke oder Temperatur.

Das Resultat einer solchen Untersuchung ist dann eine Matrix mit sogenannten Entdeckungsgeschichten (auch Fangmatrix genannt), wobei «Fang» gleichbedeutend ist mit blosser Feststellung. Ein imaginäres Beispiel eines Mittelspechtinventars im Elsass ist in Tab. 1 dargestellt. Hier wird für jedes Quadrat und jeden Besuch festgehalten, ob der Mittelspecht gesehen wurde (1) oder nicht (0). Natürlich

**Tab. 1.** Beispiel einer Fangmatrix (auch Fanggeschichtenmatrix) aus einem imaginären Inventar am Mittelspecht *Dendrocopos medius* im Elsass. Die Matrix enthält für jeden Besuch und Ort eine 1, wenn der Mittelspecht festgestellt wurde, und eine 0, wenn er nicht festgestellt wurde. Striche bezeichnen fehlende Besuche. Auslassungszeichen deuten an, dass im Normalfall mehr Quadrate besucht würden, z.B. mindestens 20 oder 30. – *Example for a capture history matrix obtained in an imaginary inventory of the Middle Spotted Woodpecker Dendrocopos medius in the Alsace. For each visit and survey the matrix contains a one if the woodpecker was detected and a zero if it was not. Dashes indicate missing visits. Ellipses indicate that more than four sites would normally be visited, e.g., at least 20 or 30.*

Quadrat	Besuch-Nummer			
	1	2	3	4
Attenschwiller	0	1	–	0
Buschwiller	1	0	1	0
Colmar	–	0	0	1
...	...			
...	0	0	0	0
Zaessingue	0	0	0	0

kann es vorkommen, dass Werte fehlen (z.B. dass nicht alle Orte gleich oft, einzelne davon gar nur einmal besucht werden), doch solche Lücken stellen normalerweise kein Problem für die Analyse dar, sie verringern einfach den Stichprobenumfang. Man beachte, dass der Mittelspecht an (sehr) vielen Orten überhaupt nie festgestellt wurde; an solchen Orten (z.B. Zaessingue) besteht die Fanggeschichte aus lauter Nullen.

Die folgenden zwei Annahmen teilen Site-Occupancy-Modelle mit anderen Formen der Interpretation von solchen Daten wie beispielsweise normale Atlasstudien: (1) Ein Quadrat ist während der Zeitdauer der vier Besuche entweder immer oder aber nie besetzt. Das ist die Annahme der «geschlossenen Population». Es ist aber nicht fatal, wenn diese Annahme nicht ganz erfüllt ist; dann schätzt man anstatt des Besetzungszustands «permanenter Aufenthalt» einfach jenen der «gelegentlichen Nutzung». (2) Der Mittelspecht wird nicht fälschlicherweise in einem Quadrat festgestellt, wo er gar nicht vorkommt, d.h. es gibt keine Fehlbe-

stimmungen. Diese Annahme ist wichtig, denn wenn Fehlbestimmungen vorkommen, so wird das Vorkommensgebiet überschätzt. Das zeigt, dass auch bei Anwendung von modernen Analysemethoden auf die unschätzbare Arbeit von hochqualifizierten Feldbeobachtern nicht verzichtet werden kann! (Wenn hingegen ein Mittelspecht fälschlicherweise z.B. als Buntspecht bestimmt wird, so wird das durch die normale Antreffwahrscheinlichkeit «aufgefangen», d.h. es hat einfach den Effekt, dass  $p$  verringert wird.)

Mittels eines der oben genannten Computerprogramme können sodann Site-Occupancy-Modelle verwendet werden, um den wahren Anteil der Quadrate mit Mittelspechten ( $V$ ) sowie die Antreffwahrscheinlichkeit pro Besuch ( $p$ ) zu berechnen mitsamt einer Schätzung der statistischen Fehlermarge, dem Standardfehler oder einem Vertrauensintervall. Um ein Gefühl zu erhalten, woher die Information zu diesen Schätzungen kommt, soll in zwei Beispielen gezeigt werden, wie die beobachteten Daten zusammenhängen mit den Parametern für Besetzung und Antreffen, also den wahren Werten in der Grundgesamtheit aller möglichen Untersuchungsquadrate im Elsass.

Im Quadrat Buschwiller ist die Fanggeschichte 1010, d.h. der Mittelspecht wurde während der ersten und dritten Begehung festgestellt. Die Wahrscheinlichkeit der Beobachtung dieser Fanggeschichte kann man so ausdrücken:  $P(1010) = V * p * (1 - p) * p * (1 - p)$ , d.h. als Produkt der Vorkommenswahrscheinlichkeit des Mittelspechts und den Wahrscheinlichkeiten, dass ein besetztes Quadrat bei der ersten und dritten Begehung als solches erkannt wird (die zwei  $p$ ) und jenen, dass er bei den Begehungen 2 und 4 nicht festgestellt wird (die zwei  $(1 - p)$ ).

Das war einfach. Im Quadrat Zaessingue haben wir einen etwas komplizierteren Fall, denn da wurde der Mittelspecht gar nie festgestellt. Diese Beobachtung kann auf zwei Arten entstehen: Entweder kommt der Mittelspecht dort nicht vor (mit Wahrscheinlichkeit  $1 - V$ ) oder er kommt vor, wurde aber viermal übersehen (mit Wahrscheinlichkeit  $V * (1 - p) * (1 - p) * (1 - p) * (1 - p)$ ). Da sich diese beiden Ereignisse gegenseitig ausschließen, ist die Wahr-

scheinlichkeit dieser Nur-Null-Fanggeschichte gleich der Summe der Wahrscheinlichkeit beider Ereignisse, also gleich  $(1 - V) + V * (1 - p) * (1 - p) * (1 - p) * (1 - p)$ . Nun haben wir für jedes Quadrat einen solchen Ausdruck mit Parametern für Vorkommens- ( $V$ ) und Antreffwahrscheinlichkeit ( $p$ ), gewissermassen also ein Gleichungssystem mit ebenso vielen Gleichungen wie Orten und zwei Unbekannten. Ein solches Gleichungssystem lösen uns die Computerprogramme.

Im Beispiel habe ich der Einfachheit der Darstellung halber angenommen, dass die Antreffwahrscheinlichkeit  $p$  bei allen Begehungen und in allen Quadraten gleich sei. Das muss nicht so sein:  $p$  kann sich je nach Quadrat, Zeitpunkt (erste bis vierte Begehung), Zeitdauer, Wind, Datum, Beobachter und vielen weiteren Faktoren unterscheiden. Und es könnte auch sein, dass die Vorkommenswahrscheinlichkeit des Mittelspechts nicht in jedem Quadrat dieselbe ist; es wäre z.B. zu erwarten, dass der Waldanteil im Quadrat oder der Eichenanteil im Wald einen wichtigen Einfluss auf die Vorkommenswahrscheinlichkeit des Spechts hätte. Solche zusätzlichen Informationen kann man als Kovariablen in die Analyse einfügen und somit testen, ob diese Faktoren wirklich einen Einfluss auf Vorkommen oder Antreffwahrscheinlichkeit haben.

Die dargestellte Version des Modells lässt nur Beobachtungen aus einem relativ kurzen Zeitraum zu, in dem die Population als geschlossen angenommen werden darf, also typischerweise aus einer einzigen Brutsaison. Oft jedoch wiederholt man Vogelaufnahmen in zwei oder mehr Jahren. In diesem Fall benötigt man dann eine erweiterte Version des Modells, das die Veränderung im Besetzungszustand der Quadrate von einem Jahr zum nächsten als Funktion von zwei weiteren Parametern darstellt, nämlich der Aussterbens- und der Kolonisierungswahrscheinlichkeit eines Quadrats. Eine Beschreibung dieser wichtigen Erweiterung des Modells findet man u.a. in MacKenzie et al. (2003, 2006), MacKenzie (2005a) und Royle & Kéry (2007). Auch dieses Modell kann man in PRESENCE 2, MARK, WinBUGS oder R fitten und darin auch Kovariablen für alle vier Parametertypen einführen.

Das Site-Occupancy-Modell ist sehr flexibel. So können je nach Quadratdefinition völlig unterschiedliche biologische Grössen geschätzt werden. Ändert man z.B. die Zeilennamen in der Fanggeschichtenmatrix in biologisch definierte Flächen wie Wasserflächen, Waldinseln oder Meeresinseln, so erhält man im Falle des Modells für mehrere Jahre genau ein klassisches Metapopulationsmodell (s. z.B. Royle & Kéry 2007) und schätzt die Anzahl der Subpopulationen oder äquivalent den Anteil aller untersuchten Flächen, die von einer Population bewohnt sind. Und definiert man ein Quadrat als ein Gebiet, in dem maximal ein Brutpaar vorkommen kann, so ist die Schätzung der Verbreitung äquivalent mit jener des Bestands: Die geschätzte Anzahl der besetzten Quadrate entspricht dann genau der wirklichen Anzahl Brutpaare innerhalb aller untersuchten Quadrate, d.h. sie ist korrigiert für allfälliges Übersehen eines Paares. Ein Beispiel wäre die Definition eines Quadrats als ein Felsgebiet: Dann könnten Site-Occupancy-Modelle verwendet werden, um den Bestand von Steinadler *Aquila chrysaetos*, Wanderfalke *Falco peregrinus* oder Uhu *Bubo bubo* zu schätzen (s. Kap. 9). Und stehen als Zeilennamen in der Matrix Vogelnamen, so können wir damit die Artenzahl korrigiert für übersehene Arten schätzen.

## 8. Site-Occupancy-Modelle in zwei Beispielen mit simulierten Daten

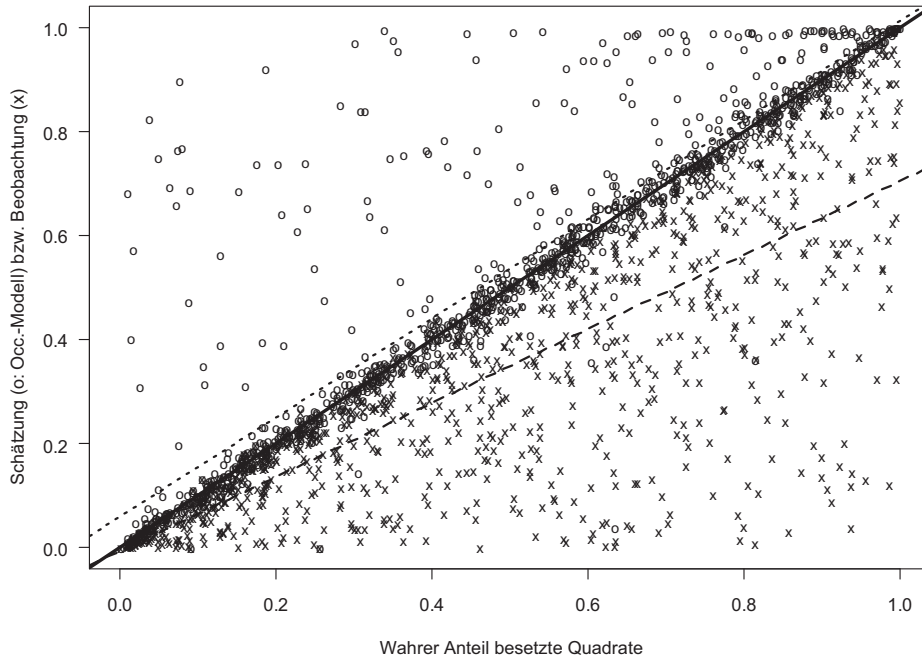
Simulierten Daten haftet auf den ersten Blick vielleicht etwas Unnatürliches an. Sie haben aber den grossen Vorteil, dass man die «Wahrheit» kennt und somit sozusagen in einem System, bei einer Tierpopulation und ihrer Erfassung, hinter die Kulissen blicken kann. Das bedeutet beispielsweise, dass man direkt nachprüfen kann, wie gut eine Methode ist. Hier zeige ich in zwei Simulationen, wie es möglich ist, dass bei einer Verbreitungsstudie die Antreffwahrscheinlichkeit die Beobachtungen so verfälschen kann, dass falsche Schlüsse gezogen werden können, und zwar sowohl für die absolute (Beispiel 1) als auch für die relative Verbreitung (Beispiel 2). Im ersten Beispiel ist das Problem, dass  $p$  kleiner als 1 ist, und im

zweiten Beispiel sind sowohl Vorkommen als auch Antreffwahrscheinlichkeit von derselben Kovariablen beeinflusst, d.h. es gibt ein «gefährliches Muster» in  $p$ .

### Beispiel 1: Schätzung der absoluten Verbreitung

Eine Vogelart besitzt in 1000 Quadraten eine gewisse Vorkommenswahrscheinlichkeit, die zufällig aus dem Bereich zwischen 0 und 1 gezogen wurde, und jedes Quadrat wurde zwischen ein- und achtmal besucht. Dieses Szenario könnte etwa den Daten entsprechen, die wir an der Schweizerischen Vogelwarte aus dem Informationsdienst (ID; Zbinden & Schmid 1995) erhalten. Die Antreffwahrscheinlichkeit der Art in besetzten Quadraten wurde als konstant über alle wiederholten Besuche und Quadrate angenommen und ebenfalls zufällig aus dem gesamten Bereich zwischen 0 und 1 gezogen. Das wurde 1000-mal wiederholt, d.h. es wurden 1000 Datensätze von diesem Typ generiert. Dann wurden die so generierten Daten auf der einen Seite mit einem Site-Occupancy-Modell analysiert und Vorkommen und Antreffwahrscheinlichkeit geschätzt. Auf der anderen Seite wurde die herkömmliche Interpretation dieser Daten durchgeführt, wo Vorkommen einfach als der Quotient ausgedrückt wird zwischen der Anzahl Quadrate, in denen die Art mindestens einmal nachgewiesen wurde, und der gesamten Anzahl Quadrate (hier 1000).

In Abb. 5 stellen die Kreise für jeden der 1000 Datensätze dar, welcher Wert des Vorkommens zufällig ausgewählt wurde und welche Schätzung des Vorkommens vom Site-Occupancy-Modell generiert wurde. Die Kreuze zeigen das Analoge für die herkömmliche Interpretation von solchen Daten (jeder einzelne der 1000 simulierten Fälle ist also sowohl durch einen Kreis als auch durch ein Kreuz repräsentiert). Die durchgezogene Linie zeigt an, wo perfekte Schätzungen zu liegen kämen, also solche, die genau dem Wert von Vorkommen  $V$  entsprechen, mit dem der betreffende Datensatz simuliert wurde. Die oberste, gepunktete Linie zeigt die lineare Regressionsgerade für alle Schätzungen des Site-Occupancy-Modells und die unterste, schraffierte Linie jene für die



**Abb. 5.** Vergleich der absoluten Verbreitungsschätzung durch ein Site-Occupancy-Modell mit der herkömmlichen Interpretation von solchen Daten anhand von 1000 simulierten Datensätzen mit je 1000 besuchten Quadraten. Die Verbreitungsschätzung mittels herkömmlicher Interpretation der Beobachtungen liegt im Durchschnitt (schraffierte Linie) viel weiter von der Realität (durchgezogene Linie) entfernt als die Schätzung des Site-Occupancy-Modells (gepunktete Linie). – Comparison of the estimates of absolute distribution based on a site-occupancy model with the conventional interpretation of distribution data in a simulation comprising 1000 data sets for 1000 surveyed quadrats each. The estimated distribution under the conventional interpretation of the data (dashed line) is much further away from truth (black line) than that under the site-occupancy model (dotted line).

herkömmliche Interpretation der Daten. Man sieht, dass die Schätzungen des Site-Occupancy-Modells (Kreise) im Durchschnitt wesentlich näher bei den wahren Werten des Vorkommens (mittlere Gerade) liegen als die beobachteten Werte (Kreuze); die oberste Gerade liegt sehr viel näher an der mittleren als die unterste. Interessanterweise liegen die Schätzungen des Site-Occupancy-Modells im Durchschnitt geringfügig zu hoch, und zwar bei kleinerem Vorkommen etwas mehr als bei grösserem. Dagegen liegen allerdings die herkömmlichen Interpretationen der Daten (Kreuze) immer zu tief. Allerdings gibt es auch beim Site-Occupancy-Modell in Einzelfällen krass zu grosse Schätzungen, wenn auch viel weniger häufig als Abweichungen nach unten bei der klassischen Interpretation der Daten. Eine genauere Unter-

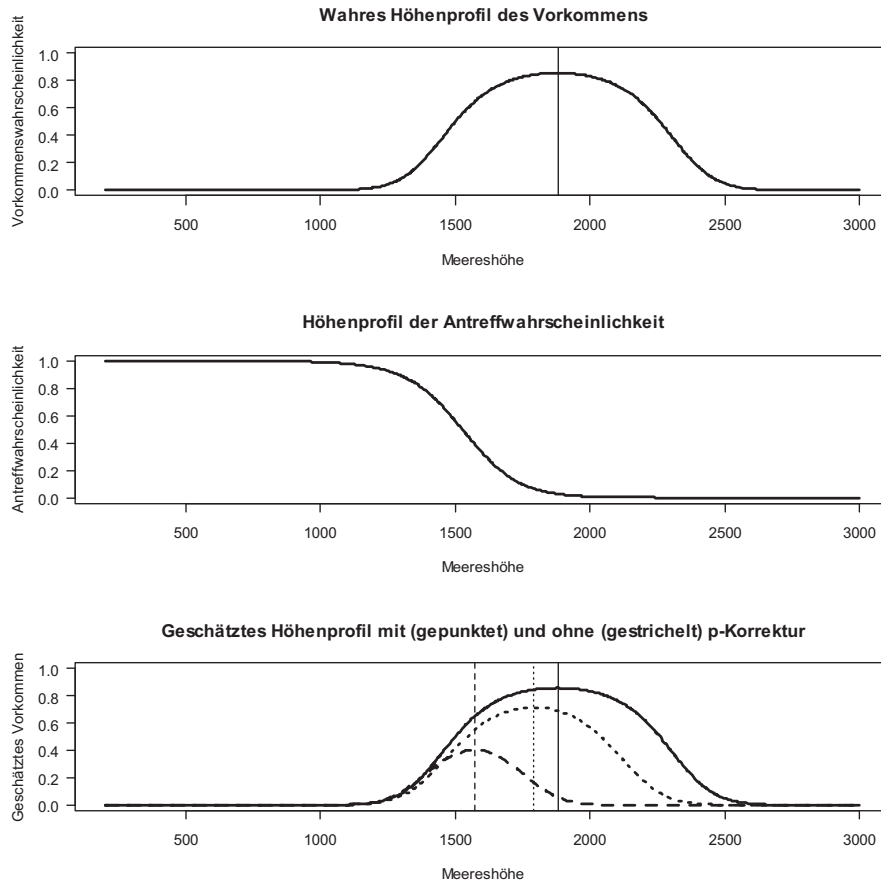
suchung zeigt, dass das jene Fälle sind, wo die Antreffwahrscheinlichkeit  $p$  weniger als etwa 0,2 betrug. Es ist bekannt, dass Site-Occupancy-Modelle bei kleinem  $p$  dazu tendieren können, das Vorkommen  $V$  stark zu überschätzen (MacKenzie et al. 2002, S. 2251). Daher sollte man auch bei Verwendung von Site-Occupancy-Modellen im Feld möglichst sauber arbeiten und alles daransetzen, die gesuchte Art so oft als möglich auch nachzuweisen, wenn sie vorkommt, also ein hohes  $p$  zu erzielen.

#### Beispiel 2: Schätzung der relativen Verbreitung

Dieses Beispiel zeigt, dass man auch im Falle der relativen Verbreitung nicht unbedingt darauf vertrauen darf, dass «immer gleich falsch wieder richtig» sei. Es ist nämlich nicht ge-

sagt, dass man immer gleich falsch misst, d.h. dass die Antreffwahrscheinlichkeit über die gewünschte Vergleichsdimension im Durchschnitt wirklich konstant sei. Wir betrachten die Höhenverbreitung einer imaginären Art, die bevorzugt in der montanen bis alpinen Zone vorkommt (Abb. 6 oben). Unterhalb von etwa

1200 m und oberhalb von etwa 2600 m ü.M. kommt sie kaum vor. Im simulierten Fall liegt das Hauptverbreitungsgebiet (wo die Vorkommenswahrscheinlichkeit  $V$  am höchsten ist) bei 1881 m. Sodann nehme ich an, dass die Antreffwahrscheinlichkeit in tieferen Lagen etwa 1,0 beträgt, aber mit zunehmender Meereshö-



**Abb. 6.** Vergleich der relativen Verbreitungsschätzung (Höhenprofil des Vorkommens) durch ein Site-Occupancy-Modell mit der herkömmlichen Interpretation von solchen Daten anhand eines simulierten Datensatzes mit 1000 Quadraten. Die herkömmliche Interpretation (schraffierte Kurve unten) unterschätzt das Ausmass des Vorkommens stark und schätzt die mittlere Höhenverbreitung (schraffierte vertikale Linie unten) im Beispiel etwa 300 m zu tief, während das Site-Occupancy-Modell (gepunktete Kurve unten) die wirkliche Höhenverbreitung (schwarze Kurve unten) viel besser schätzt und die mittlere Höhenverbreitung in diesem Beispiel bloss um etwa 90 m (gepunktete vertikale Linie unten) zu niedrig schätzt. – *A comparison of the estimated relative distribution (altitudinal profile) under a site-occupancy model and by the conventional interpretation of such data using a simulated data set with 1000 quadrats. The conventional interpretation of the data (dashed curve in lower panel) greatly underestimates the general distribution of the organism and underestimates the mean elevation of occurrence (dashed vertical line) by about 300 m. In contrast, the site-occupancy model (dotted curve) yields a much better estimate of the altitudinal distribution. In addition, its estimates of the mean elevation of occurrence (dotted vertical line) is only about 90 m off truth (solid vertical line).*



he immer stärker abnimmt und bis ca. 2000 m praktisch auf Null abfällt (Abb. 6 Mitte).

Generiert man nun Beobachtungen aus diesem System, d.h. machen wir je 3 Begehungen in jedem von 1000 Quadraten und beobachten die Art mit der angenommenen Antreffwahrscheinlichkeit auf jeder Höhe, so erhalten wir ein geschätztes Höhenprofil ihres Vorkommens (Abb. 6 unten, schraffierte Kurve), das die reale Höhenverbreitung nur schlecht wiedergibt. Die durchgezogene Kurve in Abb. 6 (unten) zeigt die wirkliche Höhenverbreitung noch einmal. Eine herkömmliche Interpretation der Beobachtungen nimmt  $p = 1$  an. Sie unterschätzt das Ausmass der Verbreitung der simulierten Art sehr stark, d.h. die schraffierte Kurve liegt weit unterhalb der durchgezogenen. Vielleicht noch schlimmer aber wird auch die Höhenverbreitung nur sehr schlecht wiedergegeben, denn die Höhe des Hauptverbreitungsgebiets (vgl. schraffierte vertikale Linie) wird auf bloss 1572 m geschätzt und somit gegenüber der realen Situation (durchgezogene vertikale Linie) um 300 m unterschätzt.

Eine andere herkömmliche Interpretation solcher Beobachtungsdaten erlaubt zwar  $p < 1$ , macht aber die Annahme, dass sich die durchschnittliche Antreffwahrscheinlichkeit entlang dem Höhenprofil nicht verändert. Das würde einer horizontalen Linie in Abb. 6 (Mitte) entsprechen. Wer diese Annahme macht, der wäre nicht so besorgt über die absolute Unterschätzung der Verbreitung, aber er würde sagen, dass man die relative Höhenverbreitung durchaus auch dann richtig abschätzen kann, sofern man nur immer gleich falsch misst. Mein konstruiertes Beispiel zeigt, dass die relative Verbreitung in diesem Fall sehr schlecht geschätzt wird, da die Schätzung der Höhe der Hauptverbreitung um etwa 300 m daneben liegt.

Ein Site-Occupancy-Modell dagegen (vgl. gepunktete Linie in Abb. 6 unten) gibt die Höhenverbreitung viel besser wieder, sowohl in Bezug auf die absolute Verbreitung als auch auf die geschätzte Höhe der Hauptverbreitung, die im simulierten Fall bloss um 90 m daneben liegt (vgl. gepunktete vertikale Linie, Abb. 6 unten). Auch wurde die absolute Verbreitung sehr viel besser geschätzt. In der Simulation betrug die Anzahl besetzter Quadrate 245 von

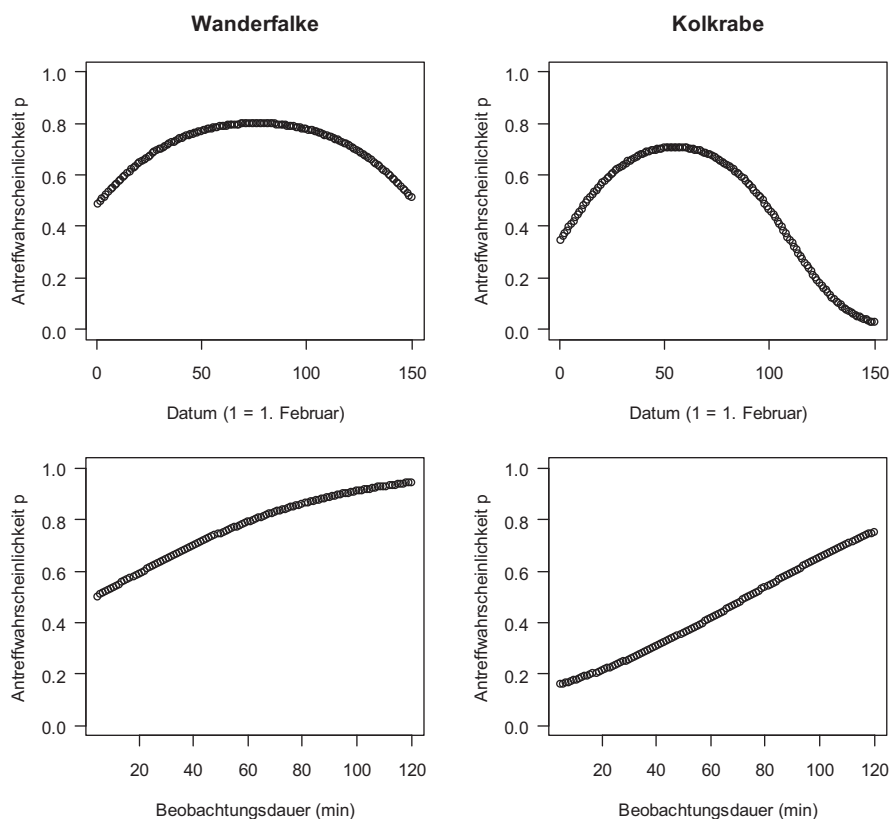
1000, aber es wurden Vögel in bloss deren 61 festgestellt, entsprechend einer durchschnittlichen gesamten Antreffwahrscheinlichkeit der Art in einem besetzten Quadrat von etwa 25 %. Das Site-Occupancy-Modell schätzte die Anzahl besetzter Quadrate viel besser auf 223, allerdings mit einem grossen Unsicherheitsbereich (95%-Vertrauensintervall), der von 78 bis 537 reichte. Mit dem Site-Occupancy-Modell kann man daher nicht nur die absolute Verbreitung einer Vogelart viel besser schätzen, sondern erreicht auch einen wesentlich besseren Schutz vor «gefährlichen Mustern» in der Antreffwahrscheinlichkeit. Man beachte, dass sich die genauen Zahlen je nach simuliertem Datensatz unterscheiden werden, dass aber die generellen Muster gleich bleiben werden.

## 9. Site-Occupancy-Modelle in vier Beispielen mit realen Daten

In diesem Abschnitt illustriere ich einige Anwendungen der Formeln 1–3 sowie von Site-Occupancy-Modellen an vier realen Beispielen. In den ersten beiden entspricht die Abundanz etwa der Besetzung, und das dritte und vierte Beispiel illustrieren klassische Anwendungen einmal für eine, einmal für viele Vogelarten.

### Beispiel 1: Felsenmonitoring im Nordjura

Seit 2005 wird nördlich einer Linie von Biaufond (Kanton Jura) nach Biel (Kanton Bern) eine möglichst grosse Anzahl von Felsobjekten auf Vorkommen von Wanderfalken, Kolkkraben *Corvus corax* und anderen Felsbrütern überprüft. Als Teil davon besuchte ich 119 Felsobjekte von Februar bis Juni 2007 zwischen 1- und 6-mal, im Mittel 1,8- und insgesamt 213-mal. Wanderfalken wurden an 66 und Kolkkraben an 49 Plätzen mindestens einmal beobachtet. Eine direkte Interpretation dieser Beobachtungen geschieht unter der Annahme, dass die Antreffwahrscheinlichkeit bei allen Beobachtungsdaten dieselbe ist. Daher benutzte ich ein Site-Occupancy-Modell, um diese Annahme bei meinen Wanderfalken- und Kolkkrabenbeobachtungen in Bezug auf die Faktoren Saison und Beobachtungsdauer zu testen.



**Abb. 7.** Abhängigkeit der Antreffwahrscheinlichkeit des Wanderfalcken *Falco peregrinus* und des Kolkraaben *Corvus corax* im Felsmonitoring Jura Nord zur Brutzeit 2007 von Saison (Tag 1 = 1. Februar) und Beobachtungsdauer (min) aufgrund einer Site-Occupancy-Analyse eigener Beobachtungen an 119 Felsobjekten. – *Relationships between the detection probability of Peregrine Falcons *Falco peregrinus* and Common Ravens *Corvus corax* in a cliff monitoring program in the Northern Jura mountains in 2007 with season (Day 1 = 1 February) and duration of observation (min) based on a site-occupancy analysis of own observations made at 119 cliff sites.*

In der Tat beeinflussten beide Faktoren die Chance, mit der Wanderfalke und Kolkraibe an besetzten Felsen festgestellt wurden (Abb. 7), und beide Arten zeigten recht ähnliche Muster. Die Antreffwahrscheinlichkeit stieg im Verlauf der Brutsaison von anfänglich etwa 50 % beim Wanderfalcken (Abb. 7, links oben) und 30 % beim Kolkraaben (Abb. 7, rechts oben) an, erreichte ihr Maximum von etwa 80 % um den 17. April beim Wanderfalcken bzw. von 70 % um den 15. März beim Kolkraaben und fiel danach bis Ende Juni auf etwa 50 % beim Wanderfalcken bzw. auf 10 % beim Kolkraaben ab. Die Beobachtungsdauer hatte bei beiden den erwarteten positiven Einfluss auf die Antreff-

wahrscheinlichkeit: Beim Wanderfalcken (Abb. 7, links unten) stieg sie von etwa 50 % auf 90 %, wenn die Beobachtungsdauer von 5 min auf 2 h erhöht wurde, während sie sich beim Kolkraaben von 20 % auf 70 % erhöhte (Abb. 7, rechts unten).

Solche Ergebnisse erlauben es, Feldbeobachtungen besser zu planen bzw. zu interpretieren. Natürlich muss das immer mit Vorsicht geschehen. So entspricht das saisonale Muster von  $p$  beim Wanderfalcken (Abb. 7, links oben) überhaupt nicht meiner subjektiven Erfahrung, wonach der Wanderfalcke zur Zeit der intensivsten Balz vor der Eiablage (um Tag 45) am leichtesten nachzuweisen sei. Zu dieser Zeit

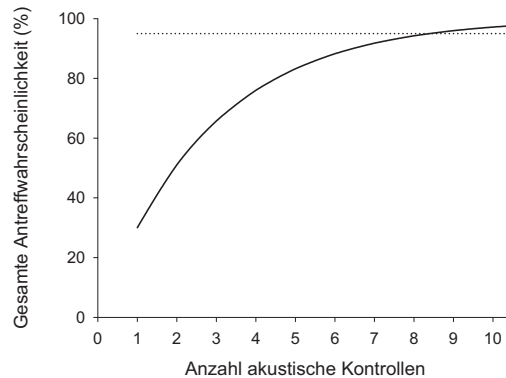
fliegen und rufen beide Partner viel und sind manchmal sehr einfach zu finden. Dagegen ist die Brutperiode eine sehr ruhige Zeit mit langen Phasen ohne Flug- oder Rufaktivität. Eine mögliche Erklärung für dieses gegenintuitive Resultat liegt darin, dass ich zur Brutzeit die brütenden Vögel auch ohne auffälliges Verhalten in vielen Felsen einfach finde, da ich die Horstnischen oft bereits kenne. In einem mir unbekanntem Gebiet oder für einen im Nordjura neuen Beobachter würde man trotzdem erwarten, dass die Antreffwahrscheinlichkeit in der ersten Hälfte März am grössten wäre.

### Beispiel 2: Uhuinventar im Oberengadin

David Jenny begann 2005 im Oberengadin (Kanton Graubünden) ein Monitoringprogramm, in dem 11 historisch bekannte Reviere (hier entsprechend einem «Quadrat») zur Balzzeit 1–5-mal aufgesucht und während je etwa 2 h auf singende Uhus kontrolliert wurden. Es wurden 8 besetzte Reviere festgestellt, 6 davon rein akustisch (Jenny 2005). Ich benutzte ein Site-Occupancy-Modell für die akustischen Feststellungen, um zu überprüfen, wie vertrauenswürdig diese Schätzung von 8 besetzten Revieren ist.

Die Antreffwahrscheinlichkeit  $p$  für ein besetztes Uhuviere wurde auf 30 % geschätzt. Die 6 Reviere, an denen Uhus tatsächlich gehört wurden, entsprechen einem beobachteten Besetzungsgrad  $V$  von 55 %. Der geschätzte Besetzungsgrad  $V$  unter Berücksichtigung der nicht perfekten Antreffwahrscheinlichkeit beim Verhören von Uhus betrug 78 %, was sich in  $11 \cdot 0,78$ , d.h. ungefähr 8,6 besetzte Reviere übersetzen lässt. Das deutet darauf hin, dass mit dem grossen Aufwand in diesem Inventar praktisch alle besetzten Reviere auch gefunden wurden, wenn auch nicht alle akustisch.

Zusammen mit dieser Schätzung der Antreffwahrscheinlichkeit liefert Formel 2 die totale Antreffwahrscheinlichkeit eines besetzten Reviers nach  $W$  Verhörbesuchen (Abb. 8).  $W_{\min}$  beträgt 8,4 Besuche, entsprechend dem Wert auf der x-Achse, wo sich die Kurve und die Gerade bei  $P_{\text{gesamt}} = 0,95$  schneiden. Um ein besetztes Revier akustisch mit 95 % Wahrscheinlichkeit nachzuweisen, müsste David



**Abb. 8.** Eine Anwendung von Formel 2 (s. Text) am Beispiel des Monitorings des Uhus *Bubo bubo* im Oberengadin 2005 (Jenny 2005) für den sogenannten Negativnachweis. Die Nachweiswahrscheinlichkeit (also die gesamte Antreffwahrscheinlichkeit  $P_{\text{gesamt}}$  über  $W$  Besuche) nimmt mit zunehmender Anzahl ( $W$ ) Verhörbesuche zu. Erst nach etwa 9 Besuchen à je 2 h kann David Jenny «sicher» sein, einen Uhu an einem besetzten Platz akustisch auch nachzuweisen. Das entspricht dem Ort, wo sich die horizontale Linie (bei 95 %) und die Kurve kreuzen. – *An application of equation 2 to the example of the Eagle owl *Bubo bubo* monitoring program in the Upper Engadine in 2005 (Jenny 2005) to estimate the minimal sampling effort for the so-called «proof of absence». The combined probability to record an Eagle owl at an occupied site (on the y axis) increases with increasing number of surveys (on the x axis). Only after c. 9 surveys lasting 2 h each can David Jenny be «certain» to acoustically ascertain an Eagle Owl at an occupied site. This represents the intersection of the horizontal line at 0.95 and the curve.*

Jenny folglich mindestens neun zweistündige Verhörbesuche (total 18 h!) an jedem Platz durchführen. Ein derart grosser Beobachtungsaufwand ist aber bei mittel- oder grossräumigen Bestandsaufnahmen kaum möglich. Um bei schwierigen Arten wie dem Uhu noch vertrauenswürdige Zahlen zur Anzahl der besetzten Reviere, oder äquivalent, dem Besetzungsgrad  $V$ , zu erhalten, sind Schätzmethoden wie das Site-Occupancy-Modell sowie zusätzliche Suchtechniken (z.B. Kot- und Rupfungsspurenuche) zu empfehlen.

### Beispiel 3: Auerhuhninventar in den Voralpen

Anlässlich einer Pilotstudie für ein grossräumiges Monitoring am Auerhuhn *Tetrao urogallus*

untersuchte Pierre Mollet (unveröff.) 2005 das Vorkommen dieser Art in einem 9 km<sup>2</sup> grossen Untersuchungsgebiet in den Voralpen anhand der Kotsuche in 36 250-m-Quadraten. Um die Antreffwahrscheinlichkeit eines besetzten (genauer, verkoteten) Quadrats abzuschätzen, suchten zwei Personen unabhängig voneinander die ganze Fläche ab. In sechs Quadraten wurde das Auerhuhn anhand von Kots Spuren von beiden Beobachtern, in je drei Quadraten nur vom ersten oder nur vom zweiten Beobachter und in 24 Quadraten von keinem der beiden festgestellt.

Eine Analyse mit dem Site-Occupancy-Modell ergab eine Antreffwahrscheinlichkeit eines Quadrats mit Auerhuhn-Kots Spuren von 67 %. Von den 36 Quadraten waren geschätzte 38 % (= 13,5) vom Auerhuhn besetzt, d.h. wiesen Kots Spuren auf, wobei die Unsicherheit relativ gross war: Das 95%-Vertrauensintervall reichte bis zu 57 % (= 20,5) besetzte Quadrate. Auerhuhnkot wurde in 12 Quadraten (33 %) festgestellt. Somit wurden die meisten besetzten Quadrate von mindestens einem Beobachter auch gefunden. Die totale Antreffwahrscheinlichkeit für beide Beobachter gemeinsam ergibt sich aus Formel 2 und betrug 89 %; dementsprechend wurden nur wenige besetzte Quadrate übersehen (1,5). Auch in diesem Fall ergab die Analyse mit Site-Occupancy-Modell, dass man bereits in die beobachteten Daten recht viel Vertrauen haben kann. Doch anstatt sich auf ungetestete Annahmen zu stützen (wie z.B. «kein besetztes Quadrat findet man sicher») erlaubte diese Analyse, diese wichtige Annahme auch zu testen.

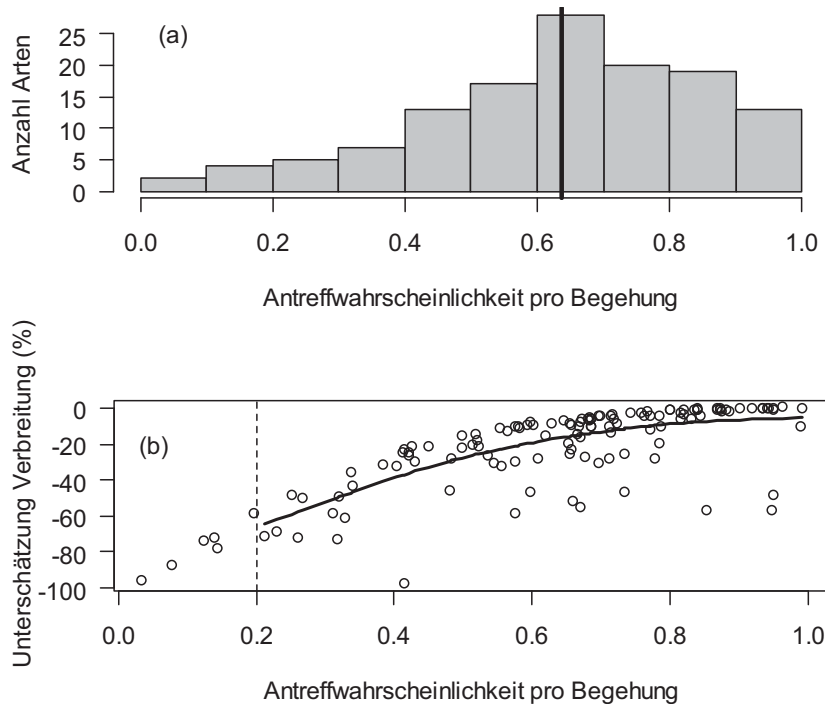
#### **Beispiel 4: Antreffwahrscheinlichkeit von 128 Arten im Monitoring Häufige Brutvögel (MHB)**

Um die im MHB erfassten Vogelarten zu vergleichen in Bezug auf die Schwierigkeit, mit der sie erfasst werden, wendete ich Site-Occupancy-Modelle auf alle 128 Arten an, die zwischen 2001 und 2004 in mindestens je einem der total 267 Kilometerquadrate gefunden wurden (Kéry & Schmidt im Druck). Ich fittete in diesem Fall die Version des Site-Occupancy-Modells für mehrere Jahre (MacKenzie et al. 2003, s. auch Royle & Kéry 2007), die für auf-

einander folgende Jahre für besetzte Quadrate Aussterbens- und für unbesetzte Quadrate Kolonisierungsereignisse zulässt. Auf diese Weise wird dafür korrigiert, dass ein Quadrat nicht unbedingt in allen vier Jahren besetzt war. Bei der Antreffwahrscheinlichkeit hingegen schätzte ich den Mittelwert über alle vier Jahre. In diese Analyse flossen nur sogenannte gültige Beobachtungen ein, also solche, die anlässlich von Begehungen nach der artspezifischen (aber national einheitlichen) Datumslimite für viele Zugvögel gemacht wurden. Auf diese Weise wurden Beobachtungen von Durchzüglern wohl weitgehend eliminiert. Ausserdem flossen viele «Nullen», die nicht «Art übersehen» sondern «Art noch nicht angekommen» bedeuten, nicht in die Analyse ein. Auf diese Weise habe ich versucht, sozusagen eine «geschlossene Population» zu schaffen (d.h. jede in einem Quadrat vorkommende Art ist während aller gültigen Begehungen präsent).

Die durchschnittliche Antreffwahrscheinlichkeit pro Begehung betrug bei 128 Arten bloss 64 % (Abb. 9a). Das bedeutet, dass im MHB-Durchschnitt bei einer einzelnen Begehung etwa ein Drittel aller im Quadrat grundsätzlich vorkommenden Arten nicht gesehen werden und dies, obwohl jede Begehung im Mittel fast 4 h dauert (Kéry & Schmidt im Druck).

Eine einheitliche Datumslimite pro Art für die ganze Schweiz und vor allem für alle Höhenlagen wird natürlich nicht überall auch eine geschlossene Population garantieren. In höheren Lagen kann es auch anlässlich von Begehungen nach einer allfälligen Datumslimite vorkommen, dass bei einigen Arten das Gros der Population bei der ersten Begehung noch nicht angekommen ist. In solchen Quadraten werden intelligente Kartierer die Rundgänge saisonal so legen, dass die Wahrscheinlichkeit einer einmaligen Feststellung einer Art bzw. eines Reviers maximal wird (gemäss Vorgaben des MHB zählt jede Feststellung innerhalb der Datumslimiten als potenzielles Revier). Beispielsweise kann in einer Probefläche im subalpinen Fichtenwald die erste Begehung schon Ende April stattfinden, wenn sie dann schon begehbar ist. Auf diese Weise können möglichst viele singende Drosseln erfasst werden, auch wenn Mönchsgrasmücken *Sylvia atrica-*



**Abb. 9.** (a) Häufigkeitsverteilung der durchschnittlichen Antreffwahrscheinlichkeit (bezogen auf eine einzelne Begehung) von 128 Schweizer Brutvogelarten im Monitoring Häufige Brutvögel (MHB) 2001–2004. Der Mittelwert von 0,64 ist durch die vertikale Linie angezeigt. (b) Grad der Unterschätzung der Verbreitung in Abhängigkeit von der Antreffwahrscheinlichkeit pro Begehung bei 128 Schweizer Brutvogelarten im Monitoring Häufige Brutvögel (MHB) 2001–2004. Für Arten mit vertrauenswürdigen Schätzungen (entsprechend etwa  $p > 0,2$ , d.h. rechts der unterbrochenen Linie) ist ein Spline-Fit des Zusammenhangs gezeigt, also eine Form von gleitendem Mittel. – (a) *Distribution of the average probability to detect a species in an occupied 1 km<sup>2</sup> quadrat (for a single survey) of 128 Swiss bird species in the Swiss BBS scheme Monitoring Häufige Brutvögel (MHB) 2001–2004. The mean value of 0.64 is indicated by the vertical line.* (b) *The degree of underestimation of the distribution in relation to the per-survey detection probability in 128 Swiss bird species in the Swiss BBS scheme Monitoring Häufige Brutvögel (MHB) 2001–2004. A spline fit of the general relationship is shown for those species (to the right of the dashed vertical line) with reliable distribution estimates (that is, with detection probability estimated at greater than c. 0.2).*

*pilla* und Zilpzalpe *Phylloscopus collybita* auf dieser Höhe möglicherweise noch nicht eingetroffen sind; diese werden in der zweiten und dritten Begehung ab Mitte Mai erfasst, wenn umgekehrt die Gesangsaktivität der Drosseln schon stark abnimmt.

In meiner Analyse habe ich sämtliche Beobachtungen nach der Datumslimite einbezogen und eine saisonal konstante Antreffwahrscheinlichkeit angenommen. Es wäre in einer raffinierteren Analyse aber durchaus auch möglich, die Resultate der Begehungen art- und quadrat- (oder höhenstufen-)spezifisch zu werten. Auf

diese Weise könnten Nullen in Quadraten, wo eine bestimmte Art noch kaum angekommen ist, eliminiert werden, wodurch die Antreffwahrscheinlichkeit und somit die korrigierte Verbreitung besser geschätzt würden. Eine alternative Lösung wäre die Einbeziehung eines höhenabhängigen saisonalen Effekts auf die Antreffwahrscheinlichkeit ins Modell (technisch: einer Höhen-Saison-Interaktion). Auf diese Weise könnte man selbst ohne Berücksichtigung von Datumslimiten voll für diese Komplexitäten korrigieren. Im hier präsentierten Modell war ich aber an artspezifischen



Durchschnittswerten interessiert und habe daher ein recht simples Modell mit konstanter Antreffwahrscheinlichkeit gewählt.

Ein zweiter wichtiger Punkt sind die beträchtlichen Artunterschiede: Am unteren Ende der Skala der Antreffwahrscheinlichkeit befanden sich Wespenbussard, Kleinspecht *Dendrocopos minor*, Habicht *Accipiter gentilis* und Sperber *A. nisus* mit 3, 8, 12 bzw. 14 % Beobachtungschancen in einem besetzten Quadrat. Demgegenüber wurden Rotkehlchen *Erithacus rubecula*, Zaunkönig *Troglodytes troglodytes*, Italiensperling *Passer hispaniolensis italiae* und Buchfink *Fringilla coelebs* in 95, 96, 99 bzw. 99 % der Quadrate mit Vorkommen bei der ersten Begehung auch gefunden. Hier sieht man, dass eine der wichtigsten Komponenten der artweisen Antreffwahrscheinlichkeit in einem Quadrat die Häufigkeit einer Art sein wird, wie dies gemäss der «zweitwichtigsten Formel der Feldornithologen» zu erwarten war.

Die herkömmliche Art der Interpretation solcher Daten beruht auf einer Kombination der Beobachtungen von allen zwei oder drei Begehungen einer Saison mit durchschnittlich fast 12 h totaler Begehungsdauer. Doch selbst mit einem solchen Suchaufwand werden nicht alle Arten gefunden (Kéry & Schmid 2006, Kéry & Schmidt im Druck), und es bleiben grosse Artunterschiede bestehen. Abb. 9b zeigt den Grad der Unterschätzung der Verbreitung in Abhängigkeit von der Antreffwahrscheinlichkeit pro Einzelbegehung für alle 128 analysierten Arten. So würde beispielsweise bei 40 % aller Arten (52 von 128) das Verbreitungsgebiet aufgrund der herkömmlichen Art der Dateninterpretation um mehr als 20 % unterschätzt. Ein gleitendes Mittel (Spline-Kurve) dieser Beziehung zeigt, wie der Grad der Unterschätzung mit der Antreffwahrscheinlichkeit zusammenhängt; diesen Zusammenhang erwarten wir natürlich bereits aufgrund des binomialen Modells der Bestandserfassung bzw. aufgrund von Formeln 1 und 2 (ausführlichere Analyse dieser Daten in Kéry & Schmidt im Druck).

Dieses Beispiel zeigt, dass die absolute Verbreitung aufgrund der reinen Beobachtungen teilweise stark unterschätzt werden kann und dass Artvergleiche riskant sind, wenn die vergleichenen Arten unterschiedlich schwierig zu

finden sind. Im Site-Occupancy-Modell hingegen wird nach diesen Unterschieden korrigiert, und die Arten werden auf diese Weise direkt vergleichbar gemacht, so wie erst die Berücksichtigung des Wechselkurses Türkliira mit britischen Pfund vergleichbar macht.

## 10. Wie interpretiert man Schätzungen von Verbreitung V und Antreffwahrscheinlichkeit p?

Typischerweise haben wir bei der Untersuchung von Vorkommen und Verbreitung irgendein grösseres Gebiet, worüber wir etwas lernen möchten. Da dieses als Grundgesamtheit oder auch statistische «Population» bezeichnete Gebiet meist zu gross ist, um vollständig untersucht zu werden, wird bloss eine Stichprobe daraus betrachtet und diese genauer untersucht. Die Gesetze der Wahrscheinlichkeitsrechnung erlauben dann, aus der Beschreibung dieser Stichprobe einen Rückschluss auf die Grundgesamtheit zu machen, unter der wichtigen Bedingung, dass die Stichprobe auf eine geeignete Weise ausgewählt wurde. Die für die Grundgesamtheit geschätzten Grössen wie Vorkommen oder Verbreitung (V) nennt man Parameter.

Hier sind zwei Beispiele für dieses typische Schema empirischer Untersuchungen: Im Felsenmonitoring Jura Nord möchte ich das Vorkommen einiger Felsenvögel im nördlichen Teil des Jurabogens untersuchen, doch wir untersuchen nicht jedes Jahr alle Felsobjekte, sondern bloss eine Auswahl (Stichprobe) davon. Im Monitoring Häufige Brutvögel (MHB) möchten wir zwar primär die Bestandstrends, sekundär aber auch die Verbreitung der Brutvögel in der ganzen Schweiz beurteilen; allerdings untersuchen wir bloss einen kleinen Teil der Schweiz, nämlich 267 Kilometerquadrate. Einer der primär interessierenden Parameter im MHB ist der Anteil der besetzten Kilometerquadrate in der ganzen Schweiz (V), doch als zweiten Parameter benötigen wir auch die Wahrscheinlichkeit, ein besetztes Kilometerquadrat als solches zu erkennen, also p. In vielen Untersuchungen existieren zumindest implizit diese zwei Ebenen: (1) die der Grundgesamtheit, an der man eigentlich Interesse hat,

und (2) die der Stichprobe, also was wir über die Grundgesamtheit als Beobachtungen in der Hand haben.

Im Idealfall ginge man in dieser Situation so vor, dass man zuerst die Grundgesamtheit definieren würde, über die eine Aussage gewünscht ist, also z.B. im MHB die ganze Schweiz. Dann würde man die möglichen Ressourcen inventarisieren und z.B. zum Schluss kommen, dass realistischerweise pro Jahr maximal etwa 300 Kilometerquadrate je 2- oder 3-mal begangen werden können. Drittens würde man diese Ressourcen auf geeignete Weise über die Grundgesamtheit verteilen. Wie man die Stichprobe auswählt, um sowohl möglichst repräsentative als auch präzise Schätzungen zu erhalten, ist sowohl eine Kunst als auch eine Wissenschaft und in der Statistik ein riesiges Gebiet, das unter die Rubrik Stichprobenerhebung (engl. «survey sampling») fällt und für alle empirischen Wissenschaften eine grosse Bedeutung hat. Nützliche Texte zu diesem wichtigen Thema sind u.a. Thompson (2002, 2004) und Williams et al. (2002). Ein Kerngebot der Stichprobenerhebung ist die Zufallsstichprobe: Nur die zufällige Auswahl der untersuchten Quadrate erlaubt es, die Grundgesamtheit überhaupt klar zu definieren, und nur die zufällige Auswahl der untersuchten Quadrate garantiert eine Stichprobe, die repräsentativ ist für diese Grundgesamtheit.

Die statistische Wissenschaft erlaubt dann, aufgrund der Beschreibung der Stichprobe (z.B. der 267 im MHB untersuchten Quadrate) eine Aussage über die Grundgesamtheit zu machen, hier also über die Verbreitung von Arten in der ganzen Schweiz, d.h. in allen etwa 41 000 solchen Quadraten, z.B. «Art X kommt auf 30 % der Landesfläche vor». Wichtig ist auch die Abschätzung der Unsicherheit in einer solchen Schätzung. Diese Schätzunsicherheit enthält mehrere Komponenten: (1) Es wird von einer Stichprobe von 267 Quadraten auf die Grundgesamtheit von 41 000 Quadraten extrapoliert, (2) das Vorkommen kann nicht perfekt beobachtet werden, d.h.  $p < 1$ , (3) die Parameter sind unbekannt und müssen geschätzt werden, (4) die Parameter können variieren, z.B. von Jahr zu Jahr, (5) verschiedene Modellvarianten ergeben unterschiedliche Schätzungen,

und wir kennen das «wahre» (oder das beste) Modell nicht. Diese und weitere Faktoren machen den statistischen Fehler aus, der mittels Standardfehler (SE) oder Vertrauensintervall quantifiziert wird. Der statistische Fehler ist ein Mass dafür, wie nahe unsere Schätzung bei dem entsprechenden Wert in der Grundgesamtheit liegen mag.

Im Site-Occupancy-Kontext unsererer MHB-Anwendung stellt die Schätzung der Vorkommenswahrscheinlichkeit  $V$  unsere beste Synthese des Anteils der im Zeitraum 2001–2004 besetzten Quadrate unter den etwa 41 000 Schweizer Kilometerquadraten dar. Wichtig ist, dass sich das Vertrauensintervall oder der Standardfehler nicht auf die 267 effektiv untersuchten Quadrate bezieht, sondern ein Mass ist für die Unsicherheit unserer Schätzung bezogen auf die ganze Schweiz. Falls man vor allem Interesse an der eigentlichen Stichprobe hat, z.B. in einer Situation wie im Engadiner Uhuinventar, so lese man zum Thema «finite-sample inference» in Royle & Kéry (2007) nach.

Die Schätzungen der Antreffwahrscheinlichkeit  $p$  der Arten gelten strikte genau für die Bestandsverhältnisse in der Schweiz 2001–2004 und sind gemittelt über die Effekte aller möglichen Einflussfaktoren im MHB. Sie beziehen sich also auf einen Pool von Beobachtern in derselben Qualität, wie wir ihn 2001–2004 hatten, auf die 2001–2004 herrschenden Wetterbedingungen, auf die Erhebungsdaten von 2001–2004, auf die typische Dauer der Begehungen von 2001–2004 usw. Die Antreffwahrscheinlichkeit einer Art wird unter anderem unter folgenden Bedingungen zunehmen oder grösser sein als hier: Wenn eine Art häufiger wird (vgl. Formel 2) oder wenn sie im Vergleich zum MHB-Durchschnitt von 2001–2004 von besseren Beobachtern, bei besseren Wetterbedingungen, zu geeigneteren Erhebungsdaten oder mittels längerer Begehungen gesucht wird. Dagegen gilt die Vorkommensschätzung  $V$  aus einem Site-Occupancy-Modell unabhängig von der Aufnahmemethode für genau jene Grundgesamtheit, aus der die besuchte Stichprobe von 267 Quadraten repräsentativ ist. Genau das ist eine der grossen Stärken der Site-Occupancy-Modelle: Indem sie für  $p$  korrigieren, stellen sie gewissermassen die grösst-

mögliche Standardisierung der Beobachtungen in Bezug auf jene Faktoren her, welche die Antrittswahrscheinlichkeit beeinflussen und die Beobachtungen verfälschen.

Leider werden die meisten realen Stichprobenerhebungen gegenüber dem Idealfall genau umgekehrt durchgeführt, d.h. irgendeine Art der Stichprobenauswahl kommt zuerst, und danach versucht man, die dazugehörige Grundgesamtheit zu identifizieren. Dieses Vorgehen ist riskant, denn es kann leicht vorkommen, dass die Grundgesamtheit, aus der die Stichprobe auf eine gar nicht formal beschreibbare Art und Weise gezogen wurde, nicht übereinstimmt mit der Grundgesamtheit, über die wir eigentlich eine Aussage zu machen wünschen. Im schlimmsten Fall existiert gar keine definierte Grundgesamtheit, wofür die vorhandene Stichprobe repräsentativ ist. In diesem Fall ist eine Stichprobe nahezu wertlos, zumindest, wenn man sie verallgemeinern möchte, was aber eigentlich immer der Fall ist.

Diese Ausführungen sind bedeutsam für die Interpretation von Beobachtungsdaten zur Verbreitung von Vögeln und zwar unabhängig davon, ob wir die Beobachtungen herkömmlich interpretieren, d.h. ohne Korrektur für  $p$ , oder mit Korrektur für  $p$  via einem Site-Occupancy-Modell. Die Art der Stichprobenauswahl ist aber auch wichtig für die Interpretation der Parameterschätzungen für  $V$  und  $p$  in einem Site-Occupancy-Modell. Macht man beispielsweise Begehungen für ein Inventar bloss in den für eine Art besten Gebieten der Schweiz, so kann die resultierende Schätzung des Anteils besiedelter Fläche natürlich nicht auf die ganze Schweiz verallgemeinert werden, sondern gilt für eine allerdings kaum wirklich fassbare und hypothetische Grundgesamtheit der «für die Art guten Gebiete in der Schweiz». Sodann sind die resultierenden Schätzungen der Antrittswahrscheinlichkeit  $p$  natürlich ebenfalls nicht auf die durchschnittliche Schweizer Situation verallgemeinerbar; in den besten Gebieten wird die Dichte der Art höher sein, und somit ist via Formel 2 erkennbar, dass  $p$  für besetzte Quadrate dann höher sein wird, weil die mittlere Anzahl Vögel pro Quadrat ( $W$ ) in guten Gebieten eben grösser ist als im Schweizer Durchschnitt. Zusammenfassend hängt es also

stark von der Art der Stichprobenentnahme, der Durchführung der Feldaufnahmen und des angewendeten Modells zur Interpretation/Analyse der Beobachtungen ab, ob und in welcher Weise Parameterschätzungen interpretiert werden können.

## 11. Ausblick

In diesem Essay habe ich eine moderne Betrachtungsweise von Vogelbestandserfassungen am Beispiel von Vorkommen und Verbreitung dargestellt, die explizit den stochastischen, d.h. zufälligen Charakter der Beobachtung oder Nichtbeobachtung von Arten auch an Orten mit Vorkommen berücksichtigt. Wenn die Antrittswahrscheinlichkeit  $p$  einer Art nicht perfekt ist (also  $p < 1$ ), entspricht unsere Beobachtung einer binomial verteilten Zufallsvariablen und kann exakt veranschaulicht werden mit der Analogie eines simplen Münzwurfes mit einer gezinkten Münze. Nur in jenen  $n$  unter den  $N$  Vorkommen, in welchen Kopf geworfen wird, entdecken wir einen Vogel. Dieses Konzept verdeutlicht alle Kerneigenschaften von Vogelzählungen. Site-Occupancy-Modelle sind eine neue Modellklasse, die in gewisser Weise einfach eine doppelte logistische Regression darstellen und Vorkommen  $V$  korrigiert auch für allfällige Muster in der Antrittswahrscheinlichkeit  $p$  und natürlich für  $p < 1$  korrekt schätzen können.

In meiner Darstellung habe ich vieles vereinfacht, denn ich bin überzeugt davon, dass wir zuerst einmal den einfachen Fall verstehen müssen, wenn wir den komplexen Fall begreifen wollen. Unser Verständnis der komplexen Wirklichkeit bauen wir am besten schrittweise auf, indem von einem vereinfachten Fall ausgehen und ihn zunehmend komplizieren, z.B. indem wir in einer Analyse zunehmend mehr Komplexität einbauen oder restriktivere Annahmen durch weniger einschränkende ersetzen oder ganz fallen lassen. Ein wichtiger Punkt ist aber auch, dass bisherige, klassische Interpretationsmethoden von Verbreitungsdaten nicht etwa weniger, sondern im Gegenteil mehr Annahmen und Vereinfachungen machen, nur werden diese nicht explizit gemacht, son-

dern bleiben «unter dem Teppich». Solche Methoden erscheinen dadurch annahmefrei oder robust. Wenn man jedoch genau hinschaut, wird man erkennen, dass sie eigentlich genau das gleiche Set von Annahmen machen wie Site-Occupancy-Modelle plus eine weitere, nämlich entweder  $p = 1$  oder  $E(p) = \text{konstant}$ , wobei  $E$  wiederum den Erwartungswert oder Durchschnitt über interessierende Dimensionen bezeichnet.

Die explizit modellbasierte Betrachtungsweise von Beobachtungen in Bestandserfassungen ist einfach die sauberere, sicherere, generellere und ehrlichere Version der herkömmlichen Form der Interpretation von solchen Daten. Aber natürlich lösen Site-Occupancy-Modelle auch nicht alle Probleme bei der Untersuchung von Verbreitung. Speziell bei Arten, die besonders unauffällig oder selten sind (also mit kleinem  $p$  oder  $N$ ), liefern sie beispielsweise oft nur sehr unpräzise oder manchmal auch verzerrte Schätzungen (s. Kap. 8, Beispiel 1). Das ist natürlich schade, denn gerade bei solchen Arten wäre eine Korrektur der beobachteten Verbreitung nach Übersehen am interessantesten. Allerdings muss man klar sehen, dass es auch in diesem Fällen eigentlich keine Alternative zu Site-Occupancy-Modellen gibt, denn der einzige (scheinbare) Vorteil der herkömmlichen Art der Datenanalyse ist jener der Verschleierung dieser Probleme. Und die Forschung geht weiter (MacKenzie et al. 2005), so dass wir hoffen können, für immer mehr Fälle über einen methodisch rigorosen und formalen Analyseansatz zu verfügen.

Die Binomialverteilung stellt die konzeptionelle Grundlage aller Vogelbestandserfassungen dar. Ich glaube, dass ein gewisses Verständnis des nur im «Durchschnitt» vorher-sagbaren Charakters einer binomialen Zufallsvariablen mit ihrer starken Analogie des Münzwurfs sowie der Bedeutung von  $N$ ,  $p$  und der Zufallsstreuung in der Binomialverteilung geradezu eine Voraussetzung dafür ist, Studien zu Verbreitung und Vorkommen von Vögeln adäquat und effizient zu planen, durchzuführen, zu analysieren und zu interpretieren. Ich hoffe, dass mein Essay etwas zur Popularisierung des Münzwurfs als konzeptioneller Grundlage der Bestandserfassung sowie von Site-Occupancy-

Modellen als einer starken Option zur Analyse der daraus resultierenden Beobachtungen beizutragen vermag.

**Dank.** David Jenny bin ich für die Zusendung der Basisdaten seines Berichts dankbar, Pierre Mollet für die Erlaubnis zur Benutzung seiner Pilotdaten, Andy Royle und Benedikt Schmidt danke ich für Diskussionen zum Thema Bestandserfassung sowie Max Baumann, Martin Blattner, Lukas Jenni, Peter Knaus, Fränzi Korner Nievergelt, Christian Marti, Claudia Müller, David Parish, Gilberto Pasinelli, Michael Schaub, Hans Schmid und Benedikt Schmidt für wertvolle Kommentare zu diesem Essay. Fränzi Korner Nievergelt danke ich auch für die Herstellung von Abb. 6.

### Zusammenfassung

Vorkommen und Verbreitung sind für die Ornithologie zentrale Grössen. Doch jeder Ornithologe weiss, dass Vögel übersehen werden können, d.h. die Antrittswahrscheinlichkeit ( $p$ ) ist meist kleiner als 1 ( $p < 1$ ). Beobachtungen bilden wirkliche Bestände und Vorkommen daher in aller Regel nur unvollständig und möglicherweise verzerrt ab. Beobachtungen sind sogenannte Zufallsvariablen: Immer wenn  $p < 1$ , sind sie keine fixen Grössen mehr, sondern streuen selbst unter identischen Bedingungen zwangsläufig und sind nur noch im statistischen Durchschnitt vorher-sagbar. Die Binomialverteilung stellt die theoretische Grundlage für alle Bestandserfassungen dar, wenn  $N$  Einheiten (z.B. besetzte Quadrate) unabhängig voneinander alle mit derselben Chance  $p$  entdeckt werden. Die beste Analogie zum Beobachtungsprozess ist der Münzwurf: Bei Verbreitungsstudien «werfen» die Vögel in jedem der  $N$  besetzten Quadrate eine gezinkte Münze; wenn sie Kopf werfen, wird einer von ihnen gesehen und sonst nicht. Unter gezinkt verstehe ich, dass die Wahrscheinlichkeit von Kopf im Prinzip jede Zahl zwischen 0 und 1 sein kann. Am binomialen Modell von Vogelbestandserfassungen lassen sich drei wichtige Eigenschaften von Vogelbeobachtungen schön illustrieren: (1) Die Anzahl beobachteter, besetzter Quadrate ( $n$ ) wird in der Regel bei jeder Wiederholung wieder anders sein, selbst unter identischen Bedingungen, und schwankt zwischen 0 und der wahren Anzahl besetzter Quadrate ( $N$ ). (2) Über eine einzelne Zählung lässt sich also nichts Definitives aussagen, doch im Durchschnitt (d.h. über viele Wiederholungen) wird die Zählung  $N * p$  betragen und die Zufallsstreuung  $\sqrt{N * p * (1 - p)}$ . Wegen  $p \leq 1$  wird die beobachtete Verbreitung die wirkliche Verbreitung praktisch immer unterschätzen. (3) Veränderungen von Zählungen oder Unterschiede zwischen Orten oder Habitaten können durch unterschiedliches  $N$ , unterschiedliches  $p$ , beides gemeinsam oder einfach durch die binomiale Zufallsstreuung zustande kommen. Das heisst, dass z.B. Trends



in *p* reale Trends in der Verbreitung maskieren können oder bei stabiler Verbreitung fälschlicherweise Verbreitungstrends vorzutäuschen vermögen. Wer die absolute Verbreitung (*N*) benötigt oder «gefährliche Muster» in *p* (d.h. *p* und *N* korrelieren mit dem gleichen Faktor) befürchtet, muss *N* direkt schätzen, indem auch *p* geschätzt und damit für  $p < 1$  korrigiert wird. Schlüssel dazu sind kurzfristig wiederholte Beobachtungen, wobei das Muster von Beobachtung und Nichtbeobachtung einer Art die Information über *p* enthält. Eine neue Klasse von statistischen Modellen, sogenannte Site-Occupancy-Modelle (MacKenzie et al. 2002), erlaubt beim Vorliegen wiederholter Beobachtungen, die Verbreitung und ihren Zusammenhang mit externen Variablen (z.B. Habitatvariablen) für die Antreffwahrscheinlichkeit zu korrigieren. Site-Occupancy-Modelle beruhen direkt auf dem binomialen Modell der Bestandserfassung und stellen zwei gekoppelte logistische Regressionen dar: Die eine beschreibt das unvollständig beobachtete, wirkliche Muster von Präsenz/Absenz, und die zweite das Muster von Beobachtung/Nichtbeobachtung in Quadraten mit Präsenz. Kovariablen können in beide Regressionen integriert werden. In diesem Essay werden Site-Occupancy-Modelle beschrieben und es wird in zwei Simulationsstudien gezeigt, wie sie die wahre Verbreitung einer Art viel besser schätzen als die herkömmliche Interpretation von Verbreitungsdaten, und wie ihre Schätzungen von «gefährlichen Mustern» in der Antreffwahrscheinlichkeit, z.B. wenn die Meereshöhe sowohl Vorkommen als auch Antreffwahrscheinlichkeit beeinflusst, weniger stark beeinträchtigt werden als die herkömmliche Interpretation der Beobachtungen. Viele professionelle und angewandte Verbreitungsstudien an Vögeln würden von mehr konzeptioneller Strenge profitieren. Die Binomialverteilung liefert das konzeptionelle Gerüst für die Planung, Durchführung, Analyse und Interpretation von sämtlichen empirischen Studien zu Vorkommen und Verbreitung. Ich hoffe, dass ein tieferes Verständnis der in diesem Essay geschilderten Grundlagen der Bestandserfassung und die zunehmende Anwendung von Site-Occupancy-Modellen zur Verbreitungsschätzung die Qualität und Interpretierbarkeit vieler Studien verbessern wird.

## Literatur

- AMRHEIN, V., H. P. KUNC, R. SCHMIDT & M. NAGUIB (2007): Temporal patterns of territory settlement and detectability in mated and unmated Nightingales *Luscinia megarhynchos*. *Ibis* 149: 237–244.
- BAILEY, L. L., J. E. HINES, J. D. NICHOLS & D. I. MACKENZIE (2007): Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. *Ecol. Applic.* 17: 281–290.
- BIBBY, C. J., N. D. BURGESS, D. A. HILL & S. MUSTOE (2000): Bird census techniques. Academic Press, San Diego.
- DORAZIO, R. M. (2007): On the choice of statistical models for estimating occurrence and extinction from animal surveys. *Ecology* 88: 2773–2782.
- DORAZIO, R. M., J. A. ROYLE, B. SÖDERSTRÖM & A. GLIMSKÄR (2006): Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* 87: 842–854.
- JENNY, D. (2005): Uhu Bericht Oberengadin 2005. Erster Monitoringbericht über die Uhreviere im Oberengadin. Schweizerische Vogelwarte, Sempach, und Amt für Jagd und Fischerei, Chur.
- KÉRY, M. (2002): Inferring the absence of a species – A case study of snakes. *J. Wildl. Manage.* 66: 330–338. – (2007): Les deux formules les plus importantes pour le biologiste de terrain, ou petite introduction à la probabilité de détection. *Bull. Ann. Soc. Hist. Nat. Pays Montbéliard* 2007: 257–266.
- KÉRY, M. & H. SCHMID (2006): Estimating species richness: Calibrating a large avian monitoring program. *J. Appl. Ecol.* 43: 101–110.
- KÉRY, M. & B. R. SCHMIDT (im Druck): Imperfect detection and its consequences for monitoring for conservation. *Comm. Ecol.*
- KÉRY, M., J. H. SPILLMANN, C. TRUONG & R. HOLLDEREGGER (2006): How biased are estimates of extinction probability in revisitation studies? *J. Ecol.* 94: 980–986.
- KREBS, C. J. (2000): Ecology: The experimental analysis of distribution and abundance. 5<sup>th</sup> ed. Addison Wesley, Longman.
- LINDLEY, D. V. (2006): Understanding uncertainty. Wiley, Hoboken.
- MACKENZIE, D. I. (2005a): Was it there? Dealing with imperfect detection for species presence/absence data. *Aust. NZ J. Stat.* 47: 65–74. – (2005b): What are the issues with presence-absence data for wildlife managers? *J. Wildl. Manage.* 69: 849–860.
- MACKENZIE, D. I., J. D. NICHOLS, J. E. HINES, M. G. KNUTSON & A. D. FRANKLIN (2003): Estimating site occupancy, colonization and local extinction when a species is detected imperfectly. *Ecology* 84: 2200–2207.
- MACKENZIE, D. I., J. D. NICHOLS, G. B. LACHMAN, S. DROEGE, J. A. ROYLE & C. A. LANGTIMM (2002): Estimating site occupancy rates when detection probability rates are less than one. *Ecology* 83: 2248–2255.
- MACKENZIE, D. I., J. D. NICHOLS, J. A. ROYLE, K. H. POLLOCK, J. E. HINES & L. L. BAILEY (2006): Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego.
- MACKENZIE, D. I., J. D. NICHOLS, N. SUTTON, K. KAWANISHI & L. L. BAILEY (2005): Improving inferences in population studies of rare species that are detected imperfectly. *Ecology* 86: 1101–1113.
- MACKENZIE, D. I. & J. A. ROYLE (2005): Designing occupancy studies: general advice and allocating survey effort. *J. Appl. Ecol.* 42: 1105–1114.
- MAUMARY, L., L. VALLOTTON & P. KNAUS (2007): Die Vögel der Schweiz. Schweizerische Vogelwarte, Sempach, und Nos Oiseaux, Montmollin.



- MAZEROLLE, M. J., L. L. BAILEY, W. L. KENDALL, J. A. ROYLE, S. J. CONVERSE & J. D. NICHOLS (2007): Making great leaps forward: Accounting for detectability in herpetological field studies. *J. Herpetol.* 41: 672–689.
- NICHOLS, J. D., J. E. HINES, D. I. MACKENZIE, M. E. SEAMANS & R. J. GUTIERREZ (2007): Occupancy estimation and modeling with multiple states and state uncertainty. *Ecology* 88: 1395–1400.
- PELLET, J. & B. R. SCHMIDT (2005): Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence. *Biol. Conserv.* 123: 27–35.
- R Development Core Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- ROYLE, J. A. & R. M. DORAZIO (2006): Hierarchical models of animal abundance and occurrence. *J. Agric. Biol. and Environm. Stat.* 11: 249–263. – (2008) Hierarchical modeling and inference in ecology. Academic Press, Amsterdam.
- ROYLE, J. A. & M. KÉRY (2007): A Bayesian state-space formulation of dynamic occupancy models. *Ecology* 88: 1813–1823.
- ROYLE, J. A., M. KÉRY, R. GAUTIER & H. SCHMID (2007): Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecol. Monogr.* 77: 465–481.
- ROYLE, J. A. & J. D. NICHOLS (2003): Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84: 777–790.
- SCHIFFERLI, L., P. GÉROUDET & R. WINKLER (1980): Verbreitungsatlas der Brutvögel der Schweiz. Schweizerische Vogelwarte, Sempach.
- SCHMID, H., R. LUDER, B. NAEF-DAENZER, R. GRAF & N. ZBINDEN (1998): Schweizer Brutvogelatlas. Verbreitung der Brutvögel in der Schweiz und im Fürstenstein Liechtenstein 1993–1996. Schweizerische Vogelwarte, Sempach.
- SCHMID, H., N. ZBINDEN & V. KELLER (2004): Überwachung der Bestandsentwicklung häufiger Brutvögel in der Schweiz. Schweizerische Vogelwarte, Sempach.
- SCHMIDT, B. R. (2005): Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. *Aq. Conserv.: Mar. Freshw. Ecosyst.* 15: 681–692. – (2008): Neue statistische Verfahren zur Analyse von Monitoring- und Verbreitungsdaten von Amphibien und Reptilien. *Z. Feldherpetol.* 15: 1–14.
- SPIEGELHALTER, D., A. THOMAS & N. G. BEST (2003): WinBUGS User Manual, Version 1.4. MCR Biostatistics Unit, Cambridge.
- SÜDBECK, P., H. ANDERTZKE, S. FISCHER, K. GEDEON, T. SCHIKORE, K. SCHRÖDER & C. SUDFELDT (2005): Methodenstandards zur Erfassung der Brutvögel Deutschlands. Radolfzell.
- THOMPSON, S. K. (2002): Sampling. Wiley, New York.
- THOMPSON, W. L. (2004): Sampling rare and elusive species. Island Press, Washington.
- WEBER, D., U. HINTERMANN & A. ZANGGER (2004): Scale and trends in species richness: considerations for monitoring biological diversity for political purposes. *Glob. Ecol. Biogeogr.* 13: 97–104.
- WHITE, G. C. & K. P. BURNHAM (1999): Program MARK: survival estimation from populations of marked animals. *Bird Study* 46 (Suppl.): 120–139.
- WILLIAMS, B. K., J. D. NICHOLS & M. J. CONROY (2002): Analysis and management of animal populations. Academic Press, San Diego.
- WINTLE, B. A., M. A. MCCARTHY, K. M. PARRIS & M. A. BURGMAN (2004): Precision and bias of methods for estimating point survey detection probabilities. *Ecol. Applic.* 14: 703–712.
- ZBINDEN, N. & H. SCHMID (1995): Das Programm der Schweizerischen Vogelwarte zur Überwachung der Avifauna gestern und heute. *Ornithol. Beob.* 92: 39–58.

*Manuskript eingegangen 7. Mai 2008*

*Bereinigte Fassung angenommen 8. September 2008*