

THE USE OF BANDING RECOVERY DATA TO ESTIMATE DISPERSAL RATES AND GENE FLOW IN AVIAN SPECIES: CASE STUDIES IN THE RED-WINGED BLACKBIRD AND COMMON GRACKLE¹

WILLIAM S. MOORE²

Department of Biological Sciences, Wayne State University, Detroit, MI 48202

RICHARD A. DOLBEER

*Denver Wildlife Research Center, United States Department of Agriculture,
6100 Columbus Avenue, Sandusky, OH 44870*

Abstract. Dispersal resulting in gene flow strongly affects the evolution of genetic structure in populations. This report describes statistical estimators of dispersal parameters based on USFWS banding recovery records. Finite-area studies of avian species yield estimates of root-mean-square (RMS) dispersal along a transect of about 1 km per generation. In contrast, estimates of RMS dispersal for the Red-winged Blackbird (*Agelaius phoeniceus*) and Common Grackle (*Quiscalus quiscula*), based on USFWS banding recovery records, are 94.6 and 111.4 km per generation, respectively. Distributions for both species are extremely leptokurtic, and confidence intervals based on jackknife statistics are large because the estimators are sensitive to outlying values. Dispersal rates can also be estimated from gene frequency data. Although all three kinds of data are not available for any one avian species, genetic-based estimates for several species are consistent with our estimates for Red-winged Blackbirds and Common Grackles in inferring that gene flow is generally high in North American birds—probably closer to 100 km than 1 km per generation. High gene flow also implies that where geographic variation is observed, such as plumage patterns across hybrid zones, selection plays a role in maintaining the pattern of geographic variation.

Key words: *Avian dispersal; gene flow; population genetics; hybrid zone; cline; Red-winged Blackbird; Common Grackle.*

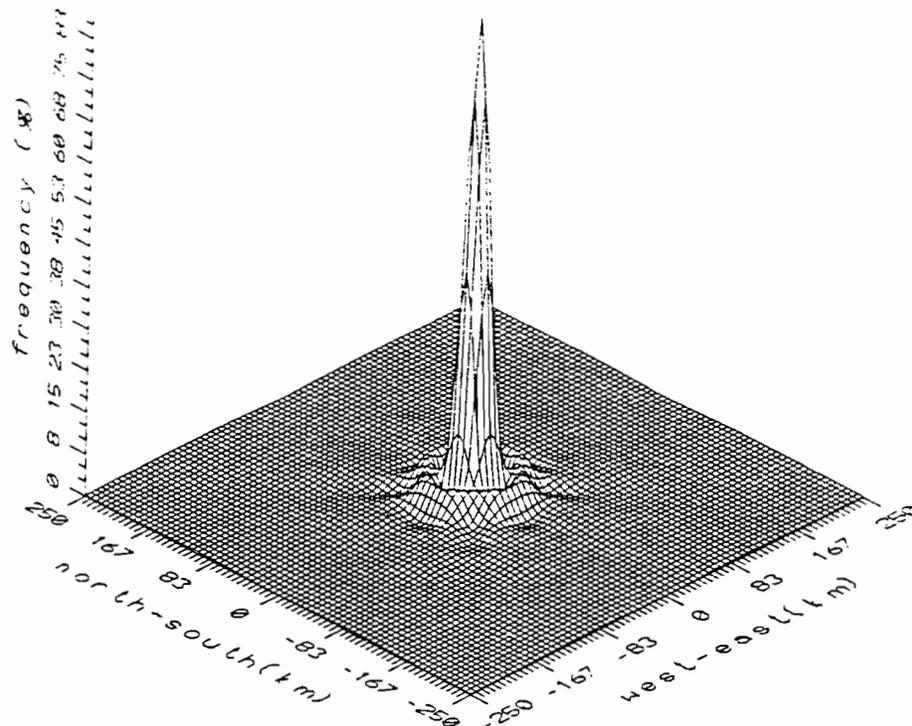
INTRODUCTION

Gene flow is an important process that governs geographic variation in natural populations and, hence, the evolution of diversity (for reviews see Nagylaki 1975, Endler 1977, Wright 1978, Slatkin 1985b). Numerous population genetics models illustrate how geographical patterns of genetic divergence evolve as functions of local selection pressures, population size, and gene flow between local and regional populations. Small local populations, or demes, will rapidly diverge by genetic drift if gene flow between them is restricted. At the other extreme, large populations interconnected by substantial numbers of dispersers will not diverge in the absence of strong local selection that offsets the homogenizing effect of gene flow.

Although determining the roles of selection and drift in the evolution of genetic structure in natural populations depends upon having at least a coarse estimate of dispersal, few dispersal data are presently available. Dispersal (or gene flow) can be estimated in two rather different ways. The first is to mark individuals and measure the distances from where they were born to where they breed. The second is to infer the level of gene flow by using mathematical models that interrelate gene flow and various measures of genetic structure. Each method has advantages and shortcomings. In this paper we concentrate on methods for measuring dispersal based on marked individuals. Among the disadvantages of this general approach are that the studies are long-term, laborious, logistically complex, and the data are usually disappointingly meager. In animals, individuals must be marked and followed throughout their lives from their birthplaces to subsequent breeding grounds. Mortality is usually high and thousands of individuals must be marked to expect an acceptable number of

¹ Received 9 September 1988. Final acceptance 3 January 1989.

² Present address: Division of Biotic Systems and Resources, National Science Foundation, 1800 G Street, N.W., Washington, DC 20550.



Distribution of dispersal distances

FIGURE 1. An idealized bivariate probability density function for dispersal distances in birds. This particular figure was generated from the data for adult male Common Grackles (Table 3) dispersing from 0–260 km. The apparent warping in the surface results from grouping data into classes, sampling error, interpolation of the surface between input data curves, and cubic-spline smoothing.

recoveries. More problematic is the fact that the search for marked individuals is usually limited to the small area of the species range where the animals were marked and released; this can substantially bias the estimate on the low side because individuals that disperse beyond the study boundary are not included in the estimate and long-distance dispersers make a disproportionately large contribution to gene flow (May et al. 1975, Moore and Buchanan 1985). Because of these impracticalities, dispersal studies are unlikely for most kinds of animals.

An exceptional group of animal species for which dispersal data are potentially available is the breeding birds of North America and Europe because they have been subject to extensive, governmental, banding programs for decades. Our objective is to describe methods for obtaining dispersal data from the United States Fish and Wildlife Service (USFWS) banding recovery data base and to describe statistical estimators of the dispersal parameters that appear in population genetics models such that these models can be applied to the study of the evolution of geographic variation in avian populations. We have chosen as examples the Red-winged Blackbird (*Age-*

lais phoeniceus) and Common Grackle (*Quiscalus quiscula*), but the methods should apply to other noncolonial species for which there are sufficient recovery records. The statistical methods are applicable to any species for which comparable data are available. Analyses of some of these data have been reported previously (Dolbeer 1978, 1982), but here we report on larger sample sizes and analysis that provides estimates of dispersal parameters as they appear in the population genetics literature.

THE MODEL

The model considers the probability that a resident bird at geographical locale (x, y) , in breeding season t will be a resident at locale $(x, y)_{t-1}$ in breeding season $t + 1$ (x and y could be longitude and latitude). The probability density function would be bivariate, considering the east-west and north-south axes, and might be bell-shaped and symmetrical as idealized in Figure 1. It is unlikely that the distributions of actual bird dispersal distances and directions from specific locales are symmetrical. However, the USFWS banding recovery data would not be useful for estimating locale-specific distributions, and

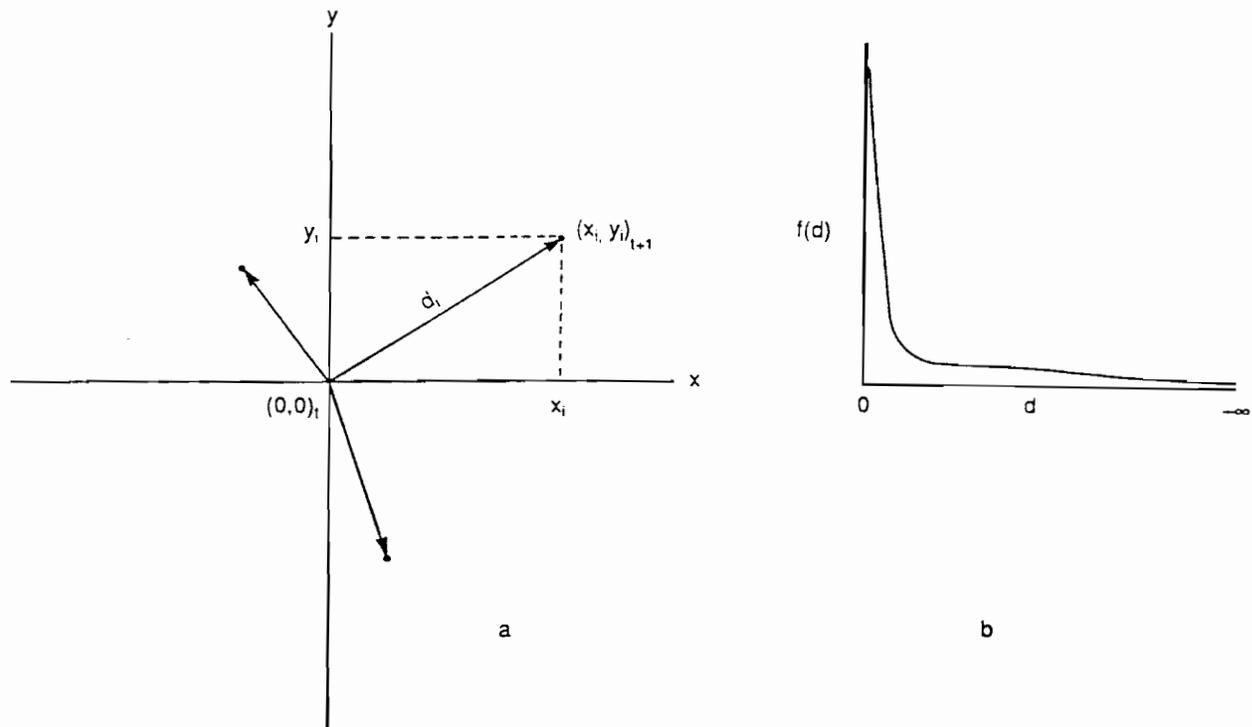


FIGURE 2. A model for dispersal distances: (a) Dispersal distances over a two-dimensional geographical area; d_i is the random variable representing a dispersal distance, x_i and y_i are the corresponding x- and y-axis components of the d_i vector. (b) The frequency distribution of dispersal distances, d_i , collected on a single axis as in the case of the USFWS banding recovery data.

at this juncture even crude approximations of dispersal distributions would substantially advance our understanding of the evolutionary forces responsible for geographic variation in birds. Therefore, we assume radial symmetry and analyze the dispersal data relative to a standardized projections $[x]$ of dispersal distances $[d]$ unless noted otherwise.)

Cases can be selected from the USFWS banding recovery data where a bird was banded (either as a hatchling or an adult) in one breeding season and recovered in a subsequent breeding season. The distance (d) between the banding and recovery locales is the value of a random variable that can be used to estimate the dispersal distribution for the species (see Fig. 2a). Banding locales vary throughout North America and few birds emanate from the same locale: thus, the data are most useful when pooled and conceptualized as half a dispersal distribution collected on a single axis as illustrated in Figure 2b.

Population genetics models of geographic variation are usually simplified by considering only a single dimension (e.g., a transect through a hybrid zone, Slatkin 1973, May et al. 1975, Barton 1979). The distribution that is needed, then, is the projection of d onto a single axis, say x , as illustrated in Figure 2a. Moreover, important parameters of the single-axis distribution are easier

to derive than those of the two-dimensional distribution d , because the expected value, $E[x]$, of the single-axis distribution is zero whereas $E[d] \neq 0$. (Throughout the remainder of this paper symbols and formulations are for the single-axis projections $[x]$ of dispersal distances $[d]$ unless noted otherwise.)

These population genetics models are usually based on diffusion equations, and the dispersal parameter is root-mean-square (RMS) dispersal, $\sigma_x = \sqrt{E[(x - \mu_x)^2]}$. For the single-axis component, RMS is actually the standard deviation in position along the x-axis for birds originating at the origin because $\mu_x = 0$. This is not the case for the distribution of d_i .

Dispersal distributions are usually leptokurtic (Bateman 1950, Levin and Kerster 1974, see also below): therefore, estimates of kurtosis are of interest. Kurtosis is defined as $k = \mu_4/(\sigma^2)^2$, where μ_4 is the fourth moment about the mean of the distribution and σ^2 is the variance. Single axis parameters and their estimators are summarized in Table 1 along with the formulas for the mean and mean-square dispersal distance in two dimensions. Key derivations are provided in the Appendix.

Yet another consideration is that most population genetics models of geographic variation

TABLE 1. Summary of formulas for dispersal-distance parameters. E is the expected value operator, d is the random variable, the dispersal distances of birds in the x-y (latitude-longitude) plane, and f(d) is the fraction of individuals dispersing distance d. The symbols subscripted with x represent parameters and statistics derived from the projection of d onto a single axis.

Parameter	Symbol	Estimator
Two-dimensional (x-y plane)		
Mean dispersal distance	$\bar{d} = E[d]$	$\hat{E}[d] = \sum_{i=1}^n d_i f(d_i)$
Mean-square dispersal distance (MS)	$E[d^2]$	$\hat{E}[d^2] = \sum_{i=1}^n d_i^2 f(d_i)$
One-dimensional (x-axis)		
Mean-square dispersal distance (MS)	σ_x^2	$\hat{\sigma}_x^2 = \hat{E}[d^2]/2$
Root-mean-square dispersal distance (RMS)		$\hat{\sigma}_x = \sqrt{\hat{\sigma}_x^2}$
Fourth moment of dispersal distance	μ_{4x}	$\hat{\mu}_{4x} = \left\{ \frac{\hat{E}[d^4]}{2} - \left(\frac{\hat{E}[d^2]}{2} \right)^2 \right\}$
Kurtosis of dispersal distance	k_x	$\hat{k}_x = \hat{\mu}_{4x} / (\hat{\sigma}_x^2)^2$

are models of species with discrete generations (e.g., annual plants) whereas most avian species have continuous, overlapping generations. The assumption of discrete generations is made to achieve mathematical tractability. Avian populations have complex demographics comprising breeding and nonbreeding classes of various ages; most importantly, the per-generation dispersal distance for an individual would be an accumulation of several dispersal events between breeding seasons.

When a bird is banded, its age class is encoded into the USFWS banding data base: U = unknown, AHY = adult, HY = immature, L = nestling and local, and SY = subadult (U.S. Fish and Wildlife Service and Canadian Wildlife Service 1976). Thus, it is possible to distinguish, to some extent, the different dispersal distributions that might obtain for first-year and adult birds. We classified all breeding-season bandings as young-of-the-year (HY and L) or adult (AHY, born in a previous breeding season). Making the simplifying assumption that a bird has two opportunities to disperse during its life, (1) from where it was born to where it first resides as a breeding season adult, and (2) from there to its "final" breeding site, it can be shown (Appendix) that

$$\sigma_G^2 = \sigma_{HY}^2 + \sigma_{AHY}^2$$

where σ_G^2 is the MS dispersal distance per generation and σ_{HY}^2 and σ_{AHY}^2 are the MS dispersals for the first-year and adult phases of the life cycle. In other words, the variances of the dispersal

distributions are additive and $\sigma_G = \sqrt{\sigma_G^2}$ (Kerster 1964; Crumpacker and Williams 1973, p. 515-516; Mallet 1985; Rockwell and Barrowclough 1987). An additional assumption in deriving this result is that the first and second dispersal distributions are stochastically independent. The two dispersal distributions that are evident from the USFWS banding recovery data reflect a minimum number of actual dispersal events, and in this respect, the sum of the two variances probably underestimates the total variance.

Similarly, the fourth moment and kurtosis (μ_{4G} , k_G) for the per-generation distribution can be derived by evaluating $E[x_G^4]$, where x_G is the dispersal distance d_G of a breeding bird, after its second dispersal event, projected onto the x axis (see Appendix). The fourth moment is:

$$\mu_{4G} = E[x_G^4] = \mu_{4HY} + \mu_{4AHY} + 6\sigma_{HY}^2\sigma_{AHY}^2.$$

The kurtosis is:

$$k_G = \frac{\mu_{4G}}{(\sigma_G^2)^2} = \frac{\mu_{4HY} + \mu_{4AHY} + 6\sigma_{HY}^2\sigma_{AHY}^2}{(\sigma_{HY}^2 + \sigma_{AHY}^2)^2},$$

where σ_{HY}^2 , σ_{AHY}^2 , μ_{4HY} , and μ_{4AHY} are the variances and means of the respective hatchling and adult dispersal vectors projected onto the x axis.

It is of interest that, for most dispersal distributions observed in nature, the kurtosis for the per-generation distribution k_G is less than the kurtosis for the component distributions (e.g., young-of-the-year and adult). The expression for k_G provides insight to the cause of this observation. For example, if we assume that the young-of-the-year and adult distributions are identical,

TABLE 2. Numerical example of jackknife statistical calculations for RMS dispersal (σ_x) for HY female Red-winged Blackbirds. (See text in Materials and Methods for explanations of symbols and formulas.)

Distance (km) x	No. observed	σ_x	I_i	I_i^2
0	5	6.52	$(-2.94) \times 5 = -14.70$	$(8.64) \times 5 = 43.22$
9.6	1	6.00	+0.70	0.49
14.3	1	5.29	+5.67	32.15
17.3	1	4.60	10.50	110.25
Sums	8		+2.17	186.11

$$I = 2.17/8 = 0.27.$$

$$\hat{\sigma}_x = 6.10; \hat{\sigma}_x \text{ corrected for bias, } \hat{\sigma}_x = 6.10 + 0.27 = 6.37.$$

$$SE = \sqrt{\frac{186.11 - (8)(0.27)^2}{(8)(7)}} = 1.82.$$

$$95\% \text{ confidence interval: } 6.37 \pm 3.57.$$

then $\mu_{4HY} = \mu_{4AHY} = \mu_4$ and $\sigma^2_{HY} = \sigma^2_{AHY} = \sigma^2$ and the kurtosis equation simplifies to $k_G = 0.5k_{HY(\text{or } AHY)} + 3/2$. Thus, $k_G < k_{HY(\text{or } AHY)}$ provided $k_{HY(\text{or } AHY)} > 3$. That is, if the distributions representing the hatchling and adult dispersal events are more leptokurtic than normal distributions, then the resultant per-generation distribution will be less leptokurtic than the component distributions. The kurtoses of the component distributions are much greater than 3 for most published dispersal distributions, including Red-winged Blackbirds and Common Grackles (see Results). This might explain the observation, noted by Bateman (1950), in dispersion studies of *Drosophila pseudoobscura* (Dobzhansky and Wright 1947), that the distributions of mutant flies released at a single point became flatter as days elapsed.

MATERIALS AND METHODS

USFWS BANDING RECOVERY DATA

Band recovery records for Red-winged Blackbirds and Common Grackles from 1924 through

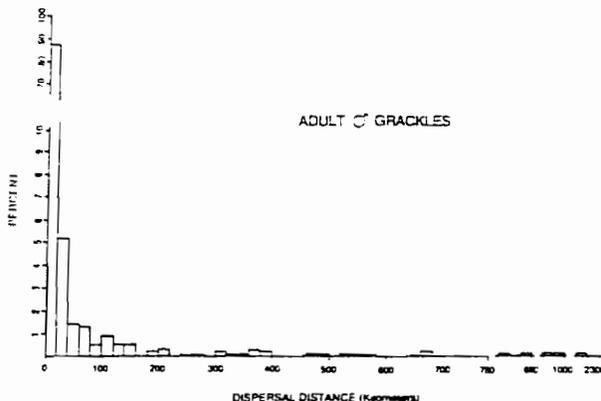


FIGURE 3. The frequency distribution for adult male Common Grackles based on USFWS banding recovery records.

1985 were obtained from the USFWS, Laurel, Maryland. We selected only records of birds that were banded and recovered at least 1 year apart during the breeding season (21 April–20 July, see Dolbeer 1982). From these records we then excluded records for which the location of banding and recovery was not known to the nearest 10 minutes of latitude and longitude. (The 10-minute latitude-longitude block is the minimum geographical area reported in the USFWS banding recovery records.) In addition, we excluded recoveries made at banding stations or that were made under conditions that would make determination of date of death questionable (how obtained codes 21, 50, 51, 56, 89, 96, and 98; U.S. Fish and Wildlife Service and Canadian Wildlife Service 1976). We further eliminated records of birds whose age classification (young-of-the-year or AHY) at the time of banding was unknown and of AHY birds whose sex was unknown. Of the 12,020 Red-winged Blackbirds and 35,891 Common Grackle recovery records, 425 and 3,181, respectively, were selected as suitable for analysis.

The distance between banding and recovery sites for each record was determined by calculating the hypotenuse of a triangle formed by the latitudinal and longitudinal coordinates:

$$D^2 = [110.0(\text{BLT} - \text{RLT})]^2 + \left[111.4 \cos\left(\frac{\text{RLT} + \text{BLT}}{2}\right) \cdot (\text{BLG} - \text{RLG}) \right]^2,$$

where: D is distance (km), BLT , RLT , BLG , and RLG are the banding and recovery latitudes and longitudes, respectively, to the nearest 1/6 of a degree.

STATISTICS

The underlying distributions of dispersal distances are unknown, but they certainly are not normal, as indicated by the high values of kurtosis (Bateman 1950, Levin and Kerster 1974, also see Results below). The "jackknife" is a numerical method that provides a correction for bias and an estimate of the standard error for any estimator (T) of a population parameter (Θ), regardless of the underlying distribution. The method entails sequentially dropping one value at a time from the sample and recalculating the statistic (T_{-j}) without the value. The statistic $I_j = (n - 1)(T - T_{-j})$ is then calculated for each dropped value (n is the sample size; T is the value of the estimator calculated from all observations). The estimator corrected for bias is: $\hat{T} = T + \bar{I}$ and the standard error for T is:

$$SE = \sqrt{\frac{\sum_{j=1}^n I_j^2 - n\bar{I}^2}{n(n-1)}}$$

where \bar{I} is the average I_j . An approximate 95% confidence interval for the bias-corrected statistic is $\hat{T} \pm 1.96 SE$ (Hinkley 1983).

A numerical example of the jackknife calculations is presented in Table 2 for the RMS dispersal estimate of Red-winged Blackbird HY females. The sample size is only 8; so, this easily can be calculated by hand. Jackknife statistics are not reliable for sample sizes this small; the example in Table 2 is given only to exemplify the calculation of these statistics. Jackknife statistics were computed by digital computer for the estimates of average dispersal distance \bar{d} , RMS dispersal ($\hat{\sigma}_x$), MS dispersal ($\hat{\sigma}^2_x$), and kurtosis (\hat{k}_x) based on the formulas in Table 1.

RESULTS

The distributions of dispersal distances are tabulated in Table 3 for Red-winged Blackbirds and Common Grackles. The distance data are grouped into 20-km intervals, a distance that corresponds roughly with the height and width of a 10-minute block of latitude and longitude in the central United States.

Sex is recorded as unknown for most young-of-the-year (HY and L). For Red-winged Blackbirds, the sex is unknown for 89 of 123 young-of-the-year recoveries, 26 are recorded as male and only eight are recorded as female. These small samples for male and female seemingly do not justify separate tabulation; therefore, the data

were pooled for male, female, and unknown sex. Although the sample sizes for young-of-the-year grackles are larger, we have tabulated only the pooled data for grackles (unknown sex, male, and female). The samples are large enough to warrant calculation of dispersal estimates for these classes (see Table 4). A histogram of dispersal distances for adult male grackles is presented in Figure 3 for the purpose of quickly conveying an impression of the shape of the dispersal distribution.

The dispersal statistics are summarized in Table 4. (The statistics are calculated from the original data rather than the grouped data compiled in Table 3.) Bearing in mind that the kurtosis of a normal distribution is 3.0, the estimated kurtoses of all of the one-dimensional dispersal distributions are large, and the lower limits of the 95% confidence intervals exceed 3.0 in all cases except for young-of-the-year female grackles, which infers that the distributions are significantly leptokurtic. Biologically, this means that most individuals do not disperse at all, but a few individuals disperse great distances.

The extreme leptokurtosis of the distributions is, of course, of biological interest, but it is also of concern with regard to estimation of the dispersal parameters. The quality of the estimates cannot be inferred from normal distributions, and, although the jackknife statistics seem to be valid, the resultant confidence intervals are large because the outlying values, which characterize the leptokurtic distributions, strongly affect the standard errors (SE). The larger sample sizes often do not result in appreciably smaller confidence intervals. Considering, for example, mean dispersal distances (\bar{d}), the confidence intervals for all age-sex classes of Red-winged Blackbirds broadly overlap. Adult grackles appear to disperse less than young-of-the-year, but this inference is clouded by the fact that the two smaller samples where young-of-the-year were identified as male and female have mean dispersal distances comparable to adults. The large confidence intervals notwithstanding, it is clear that the average dispersal distances (\bar{d}) generally exceed 14 km and that the single-axis projections of the RMS-dispersal distances (σ_x) generally exceed 32 km. The single axis kurtosis (k_x) generally exceeds 23.

The estimates of RMS dispersal (σ_G) and kurtosis (k_G) for the per-generation dispersal distributions, projected onto the single axis, can be calculated from the appropriate equations given in the previous section. For Red-winged Blackbirds the variance and fourth moment for young-

TABLE 3. The distributions of dispersal distances (kilometers) for young-of-the year and adult Red-winged Blackbirds and Common Grackles based on the USFWS banding recovery records.

Distance (km)	Red-winged Blackbirds			Common Grackles		
	Young-of-the-year (%)	Adult male (%)	Adult female (%)	Young-of-the-year (%)	Adult male (%)	Adult female (%)
0-20	100 (81.3)	202 (81.5)	49 (90.7)	729 (76.3)	1,091 (87.1)	870 (89.3)
21-40	8 (1.5)	18 (7.3)	0 (0)	107 (11.2)	65 (5.2)	54 (5.5)
41-60	3 (2.4)	5 (2.0)	2 (3.7)	28 (2.9)	18 (1.4)	10 (1.0)
61-80	1 (0.8)	5 (2.0)	1 (1.9)	19 (2.0)	16 (1.3)	8 (0.8)
81-100	1 (0.8)	2 (0.8)	0 (0)	8 (0.8)	6 (0.5)	4 (0.4)
101-120	1 (0.8)	3 (1.2)	0 (0)	8 (0.8)	11 (0.9)	2 (0.2)
121-140	0 (0)	4 (1.6)	0 (0)	8 (0.8)	6 (0.5)	5 (0.5)
141-160	1 (0.8)	0 (0)	0 (0)	4 (0.4)	6 (0.5)	1 (0.1)
161-180	2 (1.6)	1 (0.4)	0 (0)	2 (0.2)	0 (0)	1 (0.1)
181-200	0 (0)	2 (0.8)	0 (0)	5 (0.5)	3 (0.2)	3 (0.3)
201-220	1 (0.8)	1 (0.4)	0 (0)	5 (0.5)	4 (0.3)	0 (0)
221-240	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	4 (0.4)
241-260	1 (0.8)	1 (0.4)	1 (1.9)	4 (0.4)	1 (0.1)	0 (0)
261-280	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
281-300	0 (0)	0 (0)	0 (0)	4 (0.4)	0 (0)	1 (0.1)
301-320	0 (0)	0 (0)	0 (0)	2 (0.2)	3 (0.2)	1 (0.1)
321-340	1 (0.8)	0 (0)	0 (0)	2 (0.2)	1 (0.1)	0 (0)
341-360	0 (0)	0 (0)	0 (0)	2 (0.2)	1 (0.1)	0 (0)
361-380	0 (0)	0 (0)	0 (0)	0 (0)	4 (0.3)	1 (0.1)
381-400	0 (0)	0 (0)	0 (0)	2 (0.2)	3 (0.2)	1 (0.1)
401-420	0 (0)	0 (0)	0 (0)	2 (0.2)	0 (0)	1 (0.1)
421-440	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (0.2)
441-460	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	0 (0)
461-480	0 (0)	0 (0)	0 (0)	2 (0.2)	1 (0.1)	0 (0)
481-500	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	1 (0.1)
501-520	0 (0)	1 (0.4)	0 (0)	3 (0.3)	0 (0)	1 (0.1)
521-540	1 (0.8)	1 (0.4)	0 (0)	0 (0)	1 (0.1)	0 (0)
541-560	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
561-580	0 (0)	0 (0)	0 (0)	3 (0.3)	1 (0.1)	1 (0.1)
581-600	0 (0)	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)
601-620	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
621-640	0 (0)	0 (0)	1 (1.9)	1 (0.1)	0 (0)	0 (0)
641-660	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
661-680	0 (0)	0 (0)	0 (0)	0 (0)	2 (0.2)	0 (0)
681-700	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
701-720	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
721-740	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)
741-760	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
761-780	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	0 (0)
821-840	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
861-880	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
921-940	0 (0)	1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)
961-980	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
981-1,000	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
1,121-1,140	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	0 (0)
1,201-1,220	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	0 (0)
1,221-1,240	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)
2,281-2,300	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)
2,621-2,640	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	0 (0)
Sample size	123	248	54	955	1,252	974

of-the-year are 4.678.7 and 1.2721×10^9 , respectively; for adults, the variance and fourth moment, averaged over male and female classes are 4,275.1 and 1.71095×10^9 , respectively. Thus, $\sigma_G = 94.6$ km ($\sigma_G = \sqrt{\sigma^2_G}$) and $k_G = 38.7$. The statistics calculated in the same way for grackles are $\sigma_G = 111.4$ and $k_G = 241.7$. As explained in the presentation of the model, the MS and RMS of the per-generation distribution are greater than those of the constituent young-of-the-year and adult distributions, but the kurtosis is less.

DISCUSSION

Data from two very different kinds of band-and-recover studies have been used to estimate dispersal in birds. The first is based on USFWS banding recovery data as described in this paper, in Dolbeer (1982) and in Moore and Buchanan (1985); the second is based on "finite-area" studies (see Barrowclough 1980 and Rockwell and Barrowclough 1987 for reviews). A finite-area study entails delimiting an area, banding birds within the area, and determining the distances between successive nesting sites over a period of years. Because of logistics, the actual area delimited is usually quite small; birds that leave the study site are not included in dispersal estimates. Barrowclough (1978) has devised a correction for this bias; nonetheless, the disparity between estimates based on the two methods is remarkably large. The per-generation, single-axis RMS dispersal estimates, determined in finite-area studies, for seven species of passerines ranged from 0.34 km to 1.68 km with an average of 1.00 km (House Wren *Troglodytes aedon*, Bewick's Wren *Thryomanes bewickii*, Song Sparrow *Melospiza melodia*, U.S.; Bananaquit *Coereba flaveola*, Grenada; Great Tit *Parus major*, England and Netherlands; Eurasian Redstart *Phoenicurus phoenicurus*, Netherlands; Common Reed-Bunting *Emberiza schoeniclus*, Finland; Barrowclough 1980). This is a substantial and significant disparity when compared to the RMS-dispersal estimates of 94.6 km and 111.4 km for Red-winged Blackbirds and Common Grackles based on USFWS banding recovery records.

What is the basis of the disparity? Unfortunately, the two kinds of studies have been done on different species, and it is possible that species-specific dispersal rates actually do differ to the extent indicated by the disparate estimates. It is likely, however, that the difference is at least in

TABLE 4. Summary of dispersal statistics for Red-winged Blackbirds and Common Grackles. Values are for kilometers, estimates are corrected for bias, and 95% confidence intervals are given in parentheses. Bias correction and confidence intervals are based on the jackknife method (see text).

Class	Sample size	Mean dispersal distances (km) (d, 95% C.I.)	One-dimensional dispersal statistics			Kurtosis (k, 95% C.I.)
			MS dispersal (σ^2)	4th moment $\mu^4/4$		
Red-winged Blackbirds						
Young-of-the year	123	32.1 (22.1-42.2)	4,678.7	1.2721×10^9	64.5 (17.9-111.0)	
Adult male	248	26.6 (19.7-33.4)	4,274.7	2.0051×10^9	131.9 (29.5-234.4)	
Adult female	54	22.1 (7.1-37.1)	4,275.5	1.4168×10^9	130.5 (85.9-175.1)	
Common Grackles						
Young-of-the-year						
Male	89	17.1 (11.9-22.3)	951.5	5.5599×10^7	77.5 (18.6-136.4)	
Female	71	29.8 (16.5-35.1)	2,392.3	1.5159×10^8	23.0 (0.3-45.6)	
All	955	33.7 (28.8-38.7)	8,441.1	2.7984×10^{10}	632.0 (57.2-1,207.0)	
Adult male	1,252	21.3 (17.8-24.9)	5,558.3	1.4709×10^{10}	749.3 (55.4-1,443.3)	
Adult female	974	14.8 (12.2-17.5)	2,385.6	3.4262×10^9	407.2 (54.4-760.0)	

part an artifact of methodology. In this context, the biases inherent in the two kinds of studies need to be examined. The most serious shortcoming of estimates based on the USFWS banding recovery data is that all that is known of a bird is that it was recovered at a specific locale during the breeding season, and one cannot be certain that the bird was breeding or likely to breed at that locale. It is possible, for example, that most dispersers are birds that failed to establish territories at their natal sites and were forced to emigrate. These "losers" may have a reduced probability of establishing territories anywhere, and their inclusion in dispersal estimates would result in overestimates of gene flow. Thus, estimates of gene flow based on the USFWS data, or on any band-and-recovery study where the breeding status of the birds is unknown, will be unbiased only if dispersers and nondispersers have the same probability of breeding.

A related spectre is the possibility that the long-distance dispersers were actually recovered during spring migration, en route to their breeding grounds. The potential for this bias is particularly great given that mortality rates and, presumably, band-recovery rates are high during migration. We have attempted to reduce the risk of this bias by limiting recoveries to dates well within the known breeding seasons. Moreover, if this is a significant source of bias, then a preponderance of apparent long-distance dispersers should be birds recovered south of where they were banded. To test this, we selected cases from the data base in which the bird was recovered more than 484 km (300 miles) from where it was banded. Three of seven long-distance Red-winged Blackbird dispersers were recovered south of their banding sites, three north, and one at the same latitude. For the Common Grackle the breakdown was 24, 20, and 1, respectively. These data suggest that misclassified spring migrants are not a source of bias.

The most serious potential bias in a finite-area study results from the exclusion of birds that leave the study area. Although this result is not intuitive, a few birds dispersing a long distance make an enormous contribution to RMS dispersal and, hence, potentially determine the genetic structure and evolution of the species. This is apparent from the formula for RMS dispersal (rewritten from Table 1, $\hat{\sigma}_x =$

$\sqrt{\sum_{i=1}^n d_i^2 f(d_i) / 2}$): the distance dispersed (d) is squared whereas the probability of an individual

dispersing distance d , $f(d)$, is not. Thus, one individual dispersing 100 km contributes as much to RMS-dispersal as 10,000 individuals dispersing 1 km. Another way to illustrate this is to calculate RMS dispersal excluding from the calculation all birds that disperse less than 100 km. For example, of the 1,252 male Common Grackles, 1,196 dispersed less than 100 km; excluding the contribution of these individuals to RMS dispersal yields a value of 76.1 km as opposed to 76.8 km (Table 4). It is apparent that finite-area studies that fail to detect even a few long-distance dispersers will grossly underestimate RMS dispersal.

Estimates of gene flow inferred from the genetic structure of populations could resolve the question of high vs. low dispersal rates in avian species. Slatkin (1981, 1985a) has developed estimators of gene flow based on the *conditional average frequencies* of alleles sampled from populations at different locales. Conditional-average-allele-frequency curves have been published for the Northern Flicker *Colaptes auratus* (Grudzien et al. 1987) and the Western Flycatcher *Empidonax difficilis* (Johnson and Marten 1988). Both species have concave-shaped curves characteristic of species with high dispersal rates.

Slatkin (1985a) elaborated upon his original method and developed an estimate of Nm , the actual number of dispersers between demes, based on the average frequency of "private alleles." Private alleles are alleles found in only one of the demes sampled. N is deme size and m is the dispersal rate between demes; therefore, the product Nm is an estimate of the actual number of dispersers between demes. Estimates of Nm based on protein studies have been reported for six avian species: the Northern Flicker, $Nm = 4.44$ (Grudzien et al. 1987); the California Quail *Callipepla californica*, $Nm = 9.5$, the White-crowned Sparrow *Zonotrichia leucophrys*, $Nm = 1.8$, the Fox Sparrow *Passerella iliaca*, $Nm = 4.2$ (Zink and Remsen 1986); the Yellow-rumped Warbler *Dendroica coronata*, $Nm = 9.5$ (Rockwell and Barrowclough 1987) and the Western Flycatcher, $Nm = 9.62$ (for continental populations, Johnson and Marten 1988). The statistic Nm is not directly comparable to RMS dispersal. In addition, estimates of Nm are influenced by the distances between demes; these vary among the several studies, and so the estimates of Nm are not even comparable among the studies. Moreover, the analysis assumes that populations are in "quasi-equilibrium" with regard to mutants arising and going extinct (Barton and Slat-

kin 1986). Nevertheless, all of the estimates of Nm for birds are high (see Slatkin 1985a, table 7 for comparison with other animals), and these results are consistent with the estimates of RMS dispersal derived from the USFWS banding recovery data in suggesting that gene flow is generally high in avian species populations on the North American continent.

Unfortunately, there are no published protein electrophoretic surveys of either the Red-winged Blackbird or the Common Grackle. However, Ball et al. (1988) analyzed the genetic structure of the Red-winged Blackbird species population using mitochondrial DNA (mtDNA). Although mtDNA is a rapidly evolving molecule which has revealed substantial geographic variation in deer mice *Peromyscus maniculatus*, the North American Red-winged Blackbird population had very little genetic structure, and Ball et al. (1988) thought this was a consequence of high dispersal rates. Thus, the genetic structure of the Red-winged Blackbird species population is consistent with our inference of a high dispersal rate.

The importance of accurate dispersal estimates in interpreting the evolutionary significance of geographic variation in birds can be illustrated with an example. Hybrid zones between closely related avian taxa are well-known (see Moore 1977, Rising 1983 for reviews). In fact, two subspecies of grackles, the Purple Grackle (*Q. q. quiscula*) and the Bronzed Grackle (*Q. q. versicolor*) form a narrow hybrid zone along the interface of their range boundaries in the southeastern and eastern United States (Huntington 1952, Yang and Selander 1968). In Louisiana, the more northern Bronzed Grackle inhabits pine and mixed pine-hardwood forests whereas the Purple Grackle inhabits cypress-tupelugum swamp and coastal marshes (Yang and Selander 1968). The width of the Louisiana hybrid zone varies from 24–64 km (Moore 1977). The interaction between selection and dispersal in maintaining allele frequency differences across a hybrid zone can be analyzed using cline models from population genetics (Slatkin 1973; May et al. 1975; Barton 1979, 1983; Barton and Hewitt 1985). A cline is a continuous, monotonic transition in gene frequency over a geographical range; conceptually these cline models represent one-locus-two-allele hybrid zones. The important equation that results from cline theory is: $l_c = \sigma_x / \sqrt{s}$ where l_c is the characteristic length of the cline, σ_x is RMS dispersal along a single axis and s is a measure of the selection differential between the two homozygotes across the hybrid

zone. If the geographical selection gradient is steep relative to RMS dispersal, then $w = 2.08l_c$, where w is the width of the transition from 20–80% frequency of the alternate alleles (May et al. 1975). This equation can then be solved for s . If, for example, RMS dispersal for the Common Grackle were of the order of 1 km per generation, then a selection coefficient of $s = 0.001$ would explain the observed width of the hybrid zone. However, if RMS dispersal is of the order of 111.4 km per generation, as inferred from the USFWS banding recovery data (Table 4), then a selection coefficient of $s = 13.11$ would be required. The latter calculation cannot be taken at face value because at least one salient assumption of the cline model was violated; specifically, s must be small. (This stems from the fact that the equation that describes the equilibrium frequency of an allele along the geographical selection gradient was derived from a diffusion model approximation to the exact equation; Slatkin 1973, May et al. 1975.) Nonetheless, the calculation does imply that the amount of selection operating across the hybrid zone is substantial if the RMS dispersal estimate of 111.4 km per generation is even approximately correct—probably of the order of $s = 0.5$ or greater. Another concern is the possibility that the model is unrealistic for highly leptokurtic dispersal distributions.

Although this model provides only a crude approximation, it serves to illustrate how the dispersal estimate influences the evolutionary interpretation of geographic variation. Resolving this contradiction in dispersal rates estimated from finite-area studies vs. USFWS banding recovery studies is imperative to an understanding of the evolution of geographic variation in North American birds; if RMS dispersal is of the order of 100 km per generation, then selection is important; if it is typically of the order of 1 km, then selection is unimportant.

ACKNOWLEDGMENTS

We gratefully thank Zhenhua Luo and John Graham for critically reading a preliminary draft of the manuscript; Professor Luo identified mathematical errors which were corrected. We thank Montgomery Slatkin and an anonymous reviewer for many helpful comments in their reviews. This work was supported in part by National Science Foundation grants BSR 83-20605 and BSR 87-05374 to William S. Moore.

LITERATURE CITED

- BALL, R. M., S. FREEMAN, F. C. JAMES, E. BIRMINGHAM, AND J. C. AVISE. 1988. Phylogenetic population structure of Red-winged Blackbirds as-

- essed by mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85:1558-1562.
- BARROWCLOUGH, G. F. 1978. Sampling bias in dispersal studies based on finite area. *Bird-Banding* 49:333-341.
- BARROWCLOUGH, G. F. 1980. Gene flow, effective population sizes, and genetic variance components in birds. *Evolution* 34:789-798.
- BARTON, N. H. 1979. The dynamics of hybrid zones. *Heredity* 43:341-359.
- BARTON, N. H. 1983. Multilocus clines. *Evolution* 37:454-471.
- BARTON, N. H., AND G. M. HEWITT. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16:113-148.
- BARTON, N. H., AND M. SLATKIN. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56:409-415.
- BATEMAN, A. J. 1950. Is gene dispersion normal? *Heredity* 4:353-363.
- CRUMPACKER, D. W., AND J. S. WILLIAMS. 1973. Density, dispersion, and population structure in *Drosophila pseudoobscura*. *Ecol. Monogr.* 43:499-538.
- DOBZHANSKY, T., AND S. WRIGHT. 1947. Genetics of natural populations, XV. Rates of diffusion of a mutant gene through a population of *Drosophila pseudoobscura*. *Genetics* 32:303-324.
- DOLBEER, R. A. 1978. Movement and migration patterns for Red-winged Blackbirds: a continental overview. *Bird-Banding* 49:17-34.
- DOLBEER, R. A. 1982. Migration patterns for age and sex classes of blackbirds and starlings. *J. Field Ornithol.* 53:28-46.
- ENDLER, J. A. 1977. Geographic variation, speciation, and clines. Princeton Univ. Press, Princeton, NJ.
- GRUDZIEN, T. A., W. S. MOORE, J. R. COOK, AND D. TAGLE. 1987. Genetic population structure of the Northern Flicker (*Colaptes auratus*) hybrid zone. *Auk* 104:654-664.
- HINKLEY, D. 1983. Jackknife methods, p. 280-287. In S. Kotz and N. L. Johnson [eds.]. *The encyclopedia of statistical sciences*. Vol. 4.
- HUNTINGTON, C. E. 1952. Hybridization in the Purple Grackle, *Quiscalus quiscula*. *Syst. Zool.* 1:149-170.
- JOHNSON, N. K., AND J. A. MARTEN. 1988. Evolutionary genetics of flycatchers. II. Differentiation in the *Empidonax difficilis* complex. *Auk* 105:177-191.
- KERSTER, H. W. 1964. Neighborhood size in the rusty lizard, *Sceloporus olivaceus*. *Evolution* 18:445-457.
- LEVIN, D. A., AND H. W. KERSTER. 1974. Gene flow in seed plants. *Evol. Biol.* 7:139-220.
- MALLET, J. 1985. Dispersal and gene flow in a butterfly with home range behaviour: *Heliconus erato* (Lepidoptera: Nymphalidae). *Oecologia* 68:210-217.
- MAY, R. M., J. A. ENDLER, AND R. E. MCMURTRIE. 1975. Gene frequency clines in the presence of selection opposed by gene flow. *Am. Nat.* 109:659-676.
- MOORE, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. *Q. Rev. Biol.* 52:263-277.
- MOORE, W. S., AND D. B. BUCHANAN. 1985. Stability of the Northern Flicker hybrid zone. *Evolution* 39:135-151.
- NAGYLAKI, T. 1975. Conditions for the existence of clines. *Genetics* 80:595-615.
- RISING, J. D. 1983. The Great Plains hybrid zones, p. 131-157. In R. F. Johnston [ed.]. *Current ornithology*. Vol. 1. Plenum Press, New York.
- ROCKWELL, R. F., AND G. F. BARROWCLOUGH. 1987. Gene flow and the genetic structure of populations, p. 223-255. In F. Cooke and P. A. Buckley [eds.], *Avian genetics*. Academic Press, New York.
- SLATKIN, M. 1973. Gene flow and selection in a cline. *Genetics* 75:733-756.
- SLATKIN, M. 1981. Estimating levels of gene flow in natural populations. *Genetics* 99:323-335.
- SLATKIN, M. 1985a. Rare alleles as indicators of gene flow. *Evolution* 39:53-65.
- SLATKIN, M. 1985b. Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16:393-430.
- U.S. FISH AND WILDLIFE SERVICE AND CANADIAN WILDLIFE SERVICE. 1976. North American bird banding manual. Washington, DC.
- WRIGHT, S. 1978. Evolution and the genetics of populations. Vol. 4. Variability within and among populations. Univ. of Chicago Press, Chicago, IL.
- YANG, S. Y., AND R. K. SELANDER. 1968. Hybridization in the Grackle *Quiscalus quiscula* in Louisiana. *Syst. Zool.* 17:107-143.
- ZINK, R. M., AND J. V. REMSEN, JR. 1986. Evolutionary processes and patterns of geographic variation in birds, p. 1-69. In R. F. Johnston [ed.]. *Current ornithology*. Vol. 4. Plenum Press, New York.

APPENDIX

MATHEMATICAL DERIVATIONS

(refer to Appendix Figure):

(1) Mean-square (MS) dispersal along one axis, σ_x^2 , is equal to the expected value of squared dispersal distances in the x-y plane ($E[d^2]$) divided by 2; i.e., $\sigma_x^2 = E[d^2]/2$.

Derivation: Let d be the distance dispersed in a single dispersal event (e.g., one breeding season; subscripts are not required for this derivation and so d, is written, simply, as d). $d^2 = x^2 + y^2$, where d, x, and y are random variables. Taking the expected values,

$$E[d^2] = E[x^2 + y^2] = E[x^2] + E[y^2].$$

Since the mean of the distribution is (0,0),

$$E[d^2] = E[(x - 0)^2] + E[(y - 0)^2] = \sigma_x^2 + \sigma_y^2.$$

Assuming radial symmetry, $\sigma_x^2 = \sigma_y^2$; thus, $E[d^2] = 2\sigma_x^2$ (Crumpacker and Williams 1973).

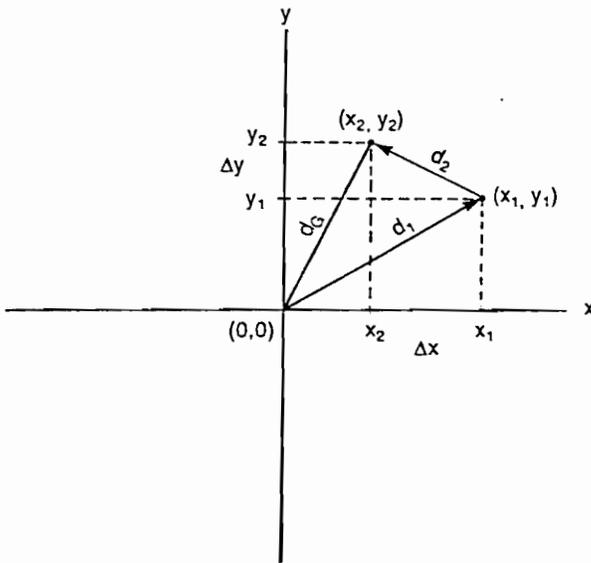
(2) The fourth moment of the single axis dispersal distribution is:

$$\mu_{4_x} = \left\{ \frac{E[d^4]}{2} - \left(\frac{E[d^2]}{2} \right)^2 \right\}.$$

Derivation:

$$\begin{aligned} d &= (x^2 + y^2)^{1/2} \\ d^2 &= [(x^2 + y^2)^{1/2}]^2 = (x^2 + y^2) \\ E[d^4] &= E[(x^2)^2 + 2x^2y^2 + (y^2)^2] \\ &= E[(x^2)^2] + 2E[x^2y^2] + E[(y^2)^2]. \end{aligned}$$

Assuming x and y are independent, i.e., the distance dispersed to the north or south (y) is independent of the distance dispersed to the east or west (x), then



$$E[d^4] = E[(x^2)^2] + 2E[x^2]E[y^2] + E[(y^2)^2]$$

Centering the distribution at the origin,

$$\begin{aligned} E[d^4] &= E[(x - 0)^4] \\ &\quad + 2E[(x - 0)^2]E[(y - 0)^2] \\ &\quad + E[(y - 0)^4] \\ &= \mu_{4_x} + 2\sigma_x^2\sigma_y^2 + \mu_{4_y} \\ &= 2\mu_{4_x} + 2(\sigma_x^2)^2 \\ &\quad \text{(assuming radial symmetry).} \end{aligned}$$

Thus,

$$\mu_{4_x} = \left\{ \frac{E[d^4]}{2} - (\sigma^2)^2 \right\} = \left\{ \frac{E[d^4]}{2} - \left(\frac{E[d^2]}{2} \right)^2 \right\}.$$

(3) Suppose that birds have two opportunities to disperse during their lives, once as hatchling-year birds (HY) and once as adults (AHY). The MS dispersal over the life span of the birds is: $\sigma_G^2 = \sigma_{HY}^2 + \sigma_{AHY}^2$; i.e., $\sigma_{x_2}^2 = \sigma_{x_1}^2 + \sigma_{\Delta x}^2$, where x_1 and Δx represent generations HY and AHY, respectively and x_2 represents the position of the bird at the end of the generation.

Derivation:

$$\begin{aligned} x_2 &= x_1 + \Delta x \\ x_2^2 &= (x_1 + \Delta x)^2 \\ E[x_2^2] &= E[(x_1 + \Delta x)^2] \\ &= E[x_1^2] + 2E[x_1\Delta x] \\ &\quad + E[(\Delta x)^2] \\ &= E[x_1^2] + E[(\Delta x)^2] \\ &\quad \text{(if } x_1 \text{ and } \Delta x \text{ are independent)} \\ &= \sigma_{x_1}^2 + \sigma_{\Delta x}^2 \\ &\quad (x_1 = x_{HY}, \Delta x = x_{AHY}). \end{aligned}$$

(4) As in (3), suppose birds have two opportunities to disperse during their lives. The single axis fourth moment of the per-generation dispersal distribution is:

$$\mu_{4_G} = \mu_{4_{x_1}} + \mu_{4_{\Delta x}} + 6\sigma_{x_1}^2\sigma_{\Delta x}^2.$$

Derivation:

$$\begin{aligned} x_2 &= x_1 + \Delta x \\ x_2^4 &= (x_1 + \Delta x)^4 \\ &= x_1^4 + 4x_1^3\Delta x \\ &\quad + 6x_1^2(\Delta x)^2 + 4x_1(\Delta x)^3 + (\Delta x)^4. \end{aligned}$$

Taking the expected values:

$$\begin{aligned} E[x_2^4] &= E[x_1^4] + 4E[x_1^3\Delta x] + 6E[x_1^2(\Delta x)^2] \\ &\quad + 4E[x_1(\Delta x)^3] + E[(\Delta x)^4] \\ &= E[x_1^4] + E[(\Delta x)^4] + 6E[x_1^2]E[(\Delta x)^2] \\ &\quad \text{(assuming } x_1 \text{ and } \Delta x \text{ are independent).} \\ \mu_{4_G} &= \mu_{4_{x_1}} + \mu_{4_{\Delta x}} + 6\sigma_{x_1}^2\sigma_{\Delta x}^2. \end{aligned}$$