Continuous models of population-level heterogeneity incorporated in analyses of animal dispersal

Eliezer Gurarie, James J. Anderson, Richard W. Zabel

January 31, 2008

Abstract

Behavioral heterogeneity among individuals is a universal feature of natural populations. However, diffusion-based models of animal dispersal implicitly assume homogeneous dispersal parameters. Heterogeneity can be accounted for by dividing a population into subpopulations, but this method is limited by the increasing number of parameters needed to characterize each subpopulation. Here, we describe an alternative approach that describes heterogeneity with a continuous distribution of a given parameter within a population. The approach provides novel insights into the interaction between population-level heterogeneity and randomness in animal movement. We demonstrate the application and tractability of the approach for spatial distributions and first passage time data on resident freshwater chub *Nocomis leptocephalus*) and migrating juvenile salmonids *Oncorhynnchus spp.*) respectively.

1 Introduction

Diffusion models in ecology (Skellam, 1951; Turchin, 1998; Okubo and Levin, 2001) owe their development in large part to statistical mechanics in physics, which deals with large numbers of indistinguishable particles. Consequently, a common implicit assumption behind

diffusion models is that the individuals in a population are identical. However, animals are not particles: there exist genotypic and phenotypic differences between individuals which have an effect on dispersal rates and distributions(Fraser et al., 2001; Waples et al., 2004).

A common result of empirical studies show that dispersing organisms commonly exhibit spatial distributions with positive kurtosis (Price et al., 1994; Kot et al., 1997; Skalski and Gilliam, 2000; Coombs and Rodríguez, 2007). Several recent investigations have proposed heterogeneity as an explanation of leptokurtic distributions. Notably, in their analysis of data collected on the dispersal of bluehead chub (*Nocomis leptocephalus*), Skalski and Gilliam (2000, 2003) suggest that the observed shape of the spatial dispersal can be explained by a superposition of two or more Gaussian dispersals corresponding to faster and more slowly dispersing fish. Their model readily identifies the relative proportions and parameter values for several sub-groups within a population. However, the cost in statistical power of subdividing a population can be high and limits the ability of this technique to fully characterize the variability within a population.

A solution to this problem is to represent the heterogeneity in movement within the population as a continuous distribution with relatively few parameters. The application of this technique to the velocity and diffusion parameters yields several tractable analytical results. An estimation of the complete set of parameters predicted allows for a relatively straightforward quantification of the role that population-level heterogeneity can play in characterizing the eventual observed distributions of dispersing organisms.

It is generally a difficult and resource intensive endeavor to obtain a series of snapshots of spatial dispersal. Measuring fluxes of dispersing organisms through a boundary, on the other hand, is often much simpler. The heterogeneity framework developed in this report can be implemented in an exactly analogous manner to interpreting first passage time distributions.

To illustrate these methods and their application, we analyzed two data-sets that consider two fundamentally different kinds of movement processes: dispersal of a resident species and directed migrations. In the first, we revisit the bluehead chub data, in which a large number

of individuals are released at a single location and gradually disperse both up and downstream over a period of month in the search for new habitat. The range of the dispersal is governed by population-level variation in the diffusion rate parameter in a readily estimable way. In the second application, we consider the seaward migration of juvenile salmonids (*Onchorhynchus spp.*), which is largely governed by variation in the travel velocities. An explicit accounting for heterogeneity in the travel velocities allow us to identify species-specific modes of migration as well as distinct responses to environmental cofactors; distinctions which are obscured using more standard diffusion models.

2 General heterogeneity framework

We confine the discussion to one-dimensional movement for the sake of simplicity and because this conforms roughly with the stream-bound movement discussed in the applications. The modelling framework is, however, adaptable to any number of dimensions.

We consider an organism as having spatial displacement in time $X(t)$ expressed as a temporally evolving probability distribution function $f(x|t, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents the vector of movement parameters. A homogenous population of n identical organisms has population distribution function of the population $P(x) = n \times f(x|t, \boldsymbol{\theta})$. This product is very commonly used as the transition between description or derivations based on an individual's movement and the distribution of an ensemble of individuals and contains within it the assumption of homogeneous behaviors.

If, however, the ith individual is characterized by it's own parameter set θ_i, the total population distribution is given as the sum of all the individual distributions

$$P(x|t, \theta) = \sum_{i=1}^{n} f(x|t, \theta_i) \tag{1}$$

For large n, equation (1) is approximated as

$$P(x|t) = n \int_{\boldsymbol{D}} f(x|t,\theta)\, g(\theta) d\theta \tag{2}$$

where $g(\theta)$ is the distribution of θ and $\boldsymbol{D}$ is its domain. Dividing both sides by n yields the pdf $h(x|t)$ for the location of dispersed individuals

$$h(x|t) = \int_{\boldsymbol{D}} f(x|t,\theta)\, g(\theta) d\theta. \tag{3}$$

For two independently distributed parameters of movement, the expression becomes:

$$h(x|t) = \int_{\boldsymbol{D}_2} \int_{\boldsymbol{D}_1} f(x|t,\theta_1,\theta_2)\, g_1(\theta_1)\, g_2(\theta_2) d\theta_1 d\theta_2 \tag{4}$$

The principle can be extended for any number of parameters.

This method applies equally well to boundary-flux or first-passage time problems, where the distance x is known and the arrival time t is the random variable. For this class of problems, equation (3) is expressed as

$$h_T(t|x) = \int_{\boldsymbol{D}} f_T(t|\boldsymbol{\theta}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{5}$$

where $h_T(t|x)$ is the flux of organisms arriving over time at some fixed distance x and $f_T(t|x)$ is the arrival time distribution of a single organism.

As an example, the simplest model of individual movement is deterministic advection where $X(t) = v\, t$ and $X(0) = 0$. In terms of the formulation above, the spatial distribution of a population with heterogeneous velocities is expressed as a Dirac delta function, such that

$$h(x|t) = \int_{-\infty}^{-\infty} \delta_0(x - vt)\, g(v)\, dv \tag{6}$$

and the first passage time distribution of a population at a fixed distance x is

$$h(t|x) = \int_{-\infty}^{-\infty} \delta(t - x/v)\, g(v)\, dv \tag{7}$$

where $g(v)$ is the density of the velocities within a population.

For many biological populations, distributions of a movement parameter within a population can be hypothesized or experimentally measured. Mathematically, these distributions are analogous to prior distributions used in Bayesian inference and there is benefit in modeling the parameter distribution $g(\theta)$ with complementary functions. For example, normally distributed velocities and gamma distributed Wiener variances yield analytical solutions (table 1 and figure 1). In the following section, we present two applications of the heterogeneity framework that provide novel insights into the processes that control animal disperal and migration.

3 Model applications

3.1 Spatial distribution of chub in a stream

3.1.1 Data

Skalski and Gilliam (2000) performed a mark-recapture experiment on several freshwater fish species in a creek in Tennessee. They released 190 marked bluehead chub (Cyprinidae: *Nocomis leptocephalus*) at a single location and recaptured them over 50 detection sites at monthly intervals, reporting the moments of the subsequently observed spatial distribution (table 2). Chub dispersal displayed a slight mean shift and a variance, both of which increased linearly in time consistent with Gaussian models. The distributions also displayed a constant, positive kurtosis. The authors propose that the kurtosis was the result of fish displaying two or more modes of movement: a "fast" diffusion and a "slow" diffusion. Their proposed model is essentially a mixture of two Gaussians with different means and variances (Skalski

and Gilliam, 2000, 2003). While the model is generalizable to any number (n) of movement modes, it is limited by the number of parameters that need to be estimated: a velocity, a diffusion parameter and a proportion for each mode of movement makes $3n - 1$ estimates leading to what the authors refer to as the "spectre of parameter explosion".

3.1.2 Heterogeneous variances model

Within the framework defined in equation 3, we can define the movement of an individual fish as

$$X(t) \sim f(x|t, v, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(x - vt)^2}{2\sigma^2 t}\right) \tag{8}$$

where v is the advective velocity and σ is the Wiener variance. This is the standard travelling, widening Gaussian distribution that arises from the unconstrained solution of the diffusion equation (Okubo and Levin, 2001). This solution also arises from any movement process X in which $\Delta X = X(t+\tau) - X(t)$ are iid random variables for all t with mean $v\tau$ and variance $\sigma^2\tau$ where the time-scale τ characterizes the length of the auto-correlation of the movement. The sum of such processes will approximate normality according to the central limit theorem, regardless of the nature of the distribution of ΔX. From a biological point of view, the latter derivation is more satisfactory as the parameters can be related to measurements of individual movements.

We can account for heterogeneity in the Wiener process by hypothesizing a gamma distribution for σ^2

$$\sigma^2 \sim g(\sigma^2|\alpha, \beta) = \sigma^{2(\alpha-1)} \frac{e^{-\sigma^2/\beta}}{\beta^\alpha \, \Gamma(\alpha)} \qquad (\sigma^2 > 0) \tag{9}$$

where α and β are the shape and scale parameters. The gamma distribution is unimodal and positively skewed, taking shapes ranging from extremely rapid decay and long tails ($\alpha \leq 1$) to an exactly exponential shape ($\alpha = 1$) to approximate normality ($\alpha \gg 1$). Applying

formulation 3 to equations 8 and 9,

$$\begin{aligned} h(x|t) &= \int_0^\infty \frac{1}{\sqrt{2\pi t}\beta^\alpha \Gamma(\alpha)} \sigma_w^{2\alpha-3} \exp\left(-\frac{(x-vt)^2}{2\sigma^2 t} - \sigma^2/\beta\right) d\sigma^2 \\ &= \frac{1}{\Gamma(\alpha)}\sqrt{\frac{2}{\pi t \beta}}\left(\frac{|x-vt|}{\sqrt{2tb}}\right)^{\alpha-\frac{1}{2}} K_{\frac{1}{2}-\alpha}\left(\frac{|x-vt|}{\sqrt{(\beta t)/2}}\right) \end{aligned} \tag{10}$$

where $K_n(x)$ is the modified Bessel function of the second kind. The modified Bessel functions exist in the positive domain and decreases monotonically; the absolute value in the argument leads to a peak at $x = vt$ with a symmetric decrease on both sides. Thus, (10) is a unimodal, symmetric pdf on x that displays advection at rate v and a characteristic widening typical of diffusion processes. We refer to equation (10) as the gamma variance process (GVP).

The statistical properties of this distribution were first described by Teichroew (1957). Yamamura (2002, 2004) independently derived a virtually identical distribution in an ecological context with fundamentally different assumptions. In these studies, the author considered the travel time of dispersing pollen to be gamma distributed. Tufto et al. (2005) applied a similar approach to describe two-dimensional dispersal of passerine birds.

3.1.3 Estimating parameters

The gamma variance process has surprisingly simple expressions for centralized higher moments:

$$\begin{aligned} &\text{Mean} : \mu = v\,t \\ &\text{Variance} : \sigma^2 = \alpha\,\beta\,t \\ &\text{Kurtosis} : \kappa = \frac{3}{\alpha} \end{aligned}$$

These moments can be compared directly to the results reported by Skalski and Gilliam (table 2) and used to obtain method of moments type estimators for the three GVP parameters.

Specifically, $\hat{v}$ is obtained by regressing the reported means $\bar{X}$ against time, $\hat{\alpha} \cdot \hat{\beta}$ are obtained by regressing the reported sample variance S^2 against time, and the measured kurtosis $\bar{\kappa}$ is used to separate $\hat{\beta}$ and $\hat{\alpha}$. Simulated 95% confidence intervals for all estimates were obtained by performing this estimation procedure 10,000 times over measurement values drawn from Skalski and Gilliam's reported standard errors.

Other techniques such as maximum likelihood can also be used to obtain parameter estimates for this distribution [see Yamamura (2002, 2004); Tufto et al. (2005)]. However the simplicity of the MME method make it the most attractive for direct application to Garrick and Skalski's published results.

3.1.4 Results

The MME estimates for the parameters of the GVD process suggest that the Wiener variance of the population can be modelled with a gamma distribution with shape parameter $\hat{\alpha} = 0.47$ [95% CI: (0.43, 0.51)] and scale parameter $\hat{\beta} = 1.15$ (0.50, 1.81). This distribution has a median value around 0.26 (0.11, 0.41), with a faster drop and a longer, fatter tail than the exponential distribution. This result is consistent with the original authors' separation of the population into two roughly equal groups of "slow" fish, with a diffusion coefficient of 0.008, and "fast" fish, with a diffusion coefficient of 0.41. Comparisons of our three parameter model indicate qualitatively good agreement with Skalski and Gilliams five parameter model (figure 2). The form of the distribution of the variances in the population suggest that a few extremely fast fish can have a great effect on eventual dispersion rates, a conclusion corroborated by an extensive literature related to the impact of long-distance dispersal (LDD) on invasion rates (Clark et al., 2003).

3.2 Migration times of outmigrating salmonids

3.2.1 Data

Many of the rivers used as migratory corridors for Pacific salmon (*Oncorhyncchus* spp.) have been heavily impounded, leading to many populations being listed as threatened or endangered under the Endangered Species Act (USNMFS, 1998). Survival of outmigrating juveniles has been shown to be related to migration timing and speeds: Slowed rates of migration increase predation risk, higher temperatures lead to greater bioenergetic stresses, arrival time in the estuary have significant impacts on ocean phase survival (Walters et al., 1978; Zabel and Williams, 2002; Anderson et al., 2005). Consequently, there has been much interest in studying migation timing and dynamics. Since the 1990's, hundreds of thousands of juvenile salmonids in the Columbia River Basin have been implanted with individually identifiable PIT (passive integrated transponder) tags and detected at hydroelectric projects (Prentice et al., 1990). Release and detection times along with physical covariates are available on a large, public database (Pacific States Marine Fisheries Commission, 1996).

We analyzed data from spring-run Chinook salmon (*O. tschawytcha*) and steelhead trout (*O. mykiss*) released in groups throughout their migratory season from 1996 through 2005. Details of the capture and tagging methods for each of the four groups are documented elsewhere (Buettner and Brimmer, 1998). We focus on these two species because they are of similar size (100 to 230 mm) and display similar peaks of migration timing. We further focus our analysis on travel times between Lower Granite and Little Goose dams, a distance 59.9 km, and on fish traveling between April 10 and May 20, thereby capturing the largest number of both species in all years.

3.2.2 Heterogeneous velocities model

The standard approach to analyzing travel times is to assume a Gaussian diffusion with advective velocity v and diffusion rate σ_w^2 (Steel et al., 2001; Zabel, 2002), for which the first

passage time is given by the inverse Gaussian (IG) distribution

$$f_T(t|x,v,\sigma_w) = \frac{x}{\sqrt{2\pi t\sigma_w^2 t^3}} \exp\left(-\frac{(x-vt)^2}{2\sigma_w^2 t}\right) \tag{11}$$

While this model captures the basic features of a travel time distribution (positive, unimodal, right-skewed), it has a tendency to miss peaks and fat tails when fit to travel-time data (Zabel and Anderson, 1997). We account for heterogeneity by assuming that individual fish velocities are distributed normally within the migrating population:

$$g(v|\mu_v,\sigma_v) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left(\frac{(v-\mu_v)^2}{2\sigma_v^2}\right). \tag{12}$$

where μ_v and σ_v^2 are the mean and variance of the velocities in a heterogeneous population. Applying equations (11) and (12) into formulation (3) yields

$$\begin{aligned} h(t|x,\mu_v,\sigma_v,\sigma_w) &= \int_{-\infty}^{\infty} f_T(t|x,v,\sigma_w)\, g(v|\mu_v,\sigma_v)dv \\ &= \frac{x}{\sqrt{2\pi(\sigma_v^2 t+\sigma_w^2)t^{3/2}}} \exp\left(-\frac{(x-\mu_v t)^2}{2(\sigma_v^2 t+\sigma_w^2)t}\right) \end{aligned} \tag{13}$$

Two limiting cases are worth considering here. If the population is homogenous ($\sigma_v = 0$) (13) reduces to the IG model (11). If each individual moves determinisitcally with it's own fixed velocity ($\sigma_w = 0$), (13) reduces to the reciprocal normal (RN) distribution

$$h_T(t|x) = \frac{x}{\sqrt{2\pi}\sigma_v t^2}\exp\left(\frac{(x-\mu_v t)^2}{2\sigma_v^2 t^2}\right) \tag{14}$$

The RN distribution has a significantly sharper peak and fatter tail than the inverse Gaussian process (figure 3). Because equation (13) mixes features of both the IG and RN distributions, we refer to it as the IGRN distribution.

The variable velocity process corresponds to a widening, travelling Gaussian with a spatial variance $\sigma_x^2 = \sigma_v^2 t^2 + \sigma_w^2 t$. Over long times, the heterogeneity on the velocities, which scales

linearly with time, contributes far more to the total spatial dispersal of a population than the variation due to random movements, which scales with the square root of time. An important result of this analysis is that for advective processes, population-level heterogeneity will swamp the effect of diffusion in the long run.

3.2.3 Estimating parameters and assessing fits

Parameters for all three models (IG, RN and IGRN) can be obtained from data using maximum likelihood estimation. For the IG model, unbiased maximum likelihood estimates for v and σ_w (Tweedie, 1957; Folks and Chhikara, 1978) are given by

$$\hat{v} = \frac{d}{\bar{t}};\ \hat{\sigma}_w^2 = \frac{d^2}{n-1}\sum_{i=1}^{n}\left(\frac{1}{t_i} - \frac{1}{\bar{t}}\right). \tag{15}$$

The RN model can be transformed into a normally distributed variable via the transformation $Y = 1/X$, such that the MLE estimates are:

$$\hat{\mu}_v = \frac{d}{\bar{t}};\ \hat{\sigma}_v^2 = \frac{d^2}{n}\sum_{i=1}^{n}\left(\frac{1}{t_i} - \frac{1}{\bar{t}}\right)^2 \tag{16}$$

For the IGRN model $\hat{\mu}_v$ can be obtained in terms of the other parameters:

$$\hat{\mu}_v = \sum_{i=1}^{n}\frac{x}{\hat{\sigma}_w^2 t_i + \hat{\sigma}_v^2} \Big/ \sum_{i=1}^{n}\frac{t_i}{\hat{\sigma}_w^2 t_i + \hat{\sigma}_v^2} \tag{17}$$

As there are no analytical expressions for the MLE's of the IGRN model variances, they are obtained by numerically maximizing the associated log-likelihood function.

We used bootstrapping to obtain confidence intervals around the IGRN estimates and report 95% empirical quantiles from the bootstrap distribution and used AIC to compare models (figure 4). Parameter estimates were regressed against mean flows using standard linear regression.

The role of heterogeneity in describing a migration process can be summarized with a

dimensionless index

$$\phi = \sqrt{\frac{\sigma_v^2 \mu_v}{\sigma_v^2 \mu_v + \sigma_w^2 d}} \tag{18}$$

This index corresponds to the amount that population-level heterogeneity in velocities contributes to the total spatial variance of the migrating population at migration distance d. In our case, we used the 59.9 km distance between dams. For a homogenous ($\sigma_v = 0$) population, $\phi = 0$. For a fully heterogenous ($\sigma_w = 0$) population of determinstic travelers, $\phi = 1$.

3.2.4 Results

According to AIC, the IGRN model fit best in all years except those (1997, 1999, 2003) where the Wiener variance $\hat{\sigma}_w = 0$ and the IGRN model collapses into the RN model (see supplementary materials). Chinook travel time distributions were often intermediate between the RN and IG models, while the steelhead are much closer to the RN distribution (figure 3). This result indicated that the IGRN is always a preferable model, but that a reciprocal normal model on velocities which is particularly simple to implement is more appropriate for steelhead than the inverse Gaussian model.

Both velocities and heterogeneity index values were consistently higher for steelhead than for Chinook (figure 4). Aggregated over all years, the model estimated velocities $\hat{\mu}_v = 21.5$ and 12.4 km day^{-1} for steelhead and Chinook respectively (Mann-Whitney $p = 0.001$), and also had higher standard error on the velocities ($\hat{\sigma}_v = 7.52$ and 3.52 km day^{-1} respectively). The values of $\hat{\phi}$ was much higher for steelhead (mean 0.665, s.e. 0.26) than for Chinook (mean 0.200, s.e. 0.16). In 1993, 1999 and 2001, $\hat{\phi} = 1$, indicating that the model could not detect any contribution of Wiener variance to the steelhead arrival times.

Simple linear regressions of these estimates against average flow between years provide a crude illustration of the relationship between these parameters and an environemtal covariate (figure 4). Chinook show no reponse to flow in either $\hat{\mu}_v$ or $\hat{\phi}$ (p-values 0.48 and 092 respectively), while steelhead parameters values varied significantly (slope $m_v = 0.0069$ km

208 day^{-1} cms^{-1} , $p \ll 0.001$) and ($m_\phi = 2.01 \times 10^{-4}$ cms^{-1}, $p = 0.037$).

209 4 Conclusions

210 Heterogeneity is widely acknowledged to be extremely important in characterizing ecolog-
211 ical processes. The main contribution of this article is to demonstrate that incorporating
212 population-level heterogeneity in models of animal dispersal does not necessarily entail an
213 increase in parameters or intractable complications in estimation. When applied to studies
214 of dispersal and migration even relatively simplistic models provide insight into the features
215 of the population.

216 In the chub example, we have shown that kurtosis is directly related to the shape of the
217 distribution of Wiener variences in the population. It has long been noted that the relatively
218 few organisms that move furthest have the greatest impact on dispersal and invasion rates
219 (Johnson and Gaines, 1990; Kot et al., 1997). The analysis allows for a quantification of the
220 role of extreme movers in a population: the low shape parameter in the gamma distribution
221 of the Wiener variance indicates that the difference between the fastest movers and the great
222 number of slower fish can be quite extreme. This kind of population-level heterogeneity can
223 be directly related to measurable phenotypic or behavioral traits. For example, (Fraser
224 et al., 2001) demonstrated that measurements of "boldness" among captive Trinidad killifish
225 *Rivulus hartii* was a useful predictor of dispersal distance when the fish are released in the
226 wild. The variability in a population is associated with genotypic variability and can provide
227 a tool for linking the evolutionary viability of a species to a genetic component.

228 We can explicitly separate the contributions of diffusion and population-level effects to the
229 total dispersal of a population by analyzing a single first-passage time distribution. The con-
230 sistency of differences between populations which are experiencing similar environments and
231 constraints provides compelling evidence that a real behavioral difference is being observed.
232 Specifically, the heterogeneity-dominated steelhead linger less in the reservoir and show little

intra-population interactions as each individual travels at a steady clip that is strongly linked to the ambient river flow. Chinook salmon, on the other hand, show more intrinsic randomness in their migration, likely milling and moving frequently, generally spending more time rearing in the riverine environment. These differences underscore a divergence in life-history strategies which different species have evolved to mitigate survival in view of environmental constraints (Waples et al., 2004).

While the models fit well in terms of parameters which appear to be biologically meaningful, a thorough interpretation of the results would take into account many sources of variability. In our formulation, heterogeneous behaviors are absorbed into the description or random movement while heterogenous populations are modeled by some simple assumptions about distributions within populations. The effects of environmental heterogeneity can be more difficult to separate. In general, a narrowing of the temporal windows of an analyzed group leads to a decrease in measured heterogeneity (ϕ) since the organisms experience a narrower range of environments (Gurarie et al. *private communication*). By comparing the rate of decrease of the heterogeneity index to the size of a temporal window it is possible to quantify more exactly the "intrinsic" heterogeneity of a population and to formulate more sophisticated models of immediate environmental response.

At the smallest level, an individual organism presumably responds in fairly well-defined ways to its immediate environment, biophysical constraints and internal state. From its own perspective, there is presumably little that is truly "random" about its movements. When describing a system that consists of many different individuals in a complex environment moving in space and time, notions of randomness or stochasticity, heterogeneity, variance and error have a tendency to blur together. However, when precisely defined they can refer to very distinct ideas. The models presented here are tractable steps in partitioning these sources of variability.

References

Anderson, J. J., E. Gurarie, and R. W. Zabel, 2005. Mean free-path length theory of predator-prey interactions: application to juvenile salmon migration. *Ecological Modelling* **186**:196–211.

Buettner, E. W. and A. F. Brimmer, 1998. Smolt monitoring at the head of Lower Granite Reservoir and Lower Granite Dam. Annual Report 1996 to Bonneville Power Administration, Idaho Department of Fish and Game, Portland, Oregon.

Clark, J. S., M. Lewis, J. McLachlan, and J. HilleRisLambers, 2003. Estimating population spread: What can we forecast and how well? *Ecology* **84**:1979–1988.

Coombs, M. F. and M. A. Rodríguez, 2007. A field test of simple dispersal models as predictors of movement in a cohort of lake-dwelling brook charr. *Journal of Animal Ecology* **76**:4557.

Folks, J. L. and R. S. Chhikara, 1978. The inverse Gaussian distribution and its statistical application - a review. *Journal of the Royal Statistics Society. Series B (Methodological)* **40**:263–289.

Fraser, D. F., J. F. Gilliam, M. J. Daley, A. N. Le, and G. T. Skalski, 2001. Explaining leptokurtic movement distributions: intrapopulation variation in boldness and exploration. *The American Naturalist* **158**:124–135.

Johnson, M. L. and M. S. Gaines, 1990. Evolution of dispersal: Theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21**:449–480.

Kot, M., M. A. Lewis, and P. v. d. Driessche, 1997. Dispersal data and the spread of invading organisms. *Ecology* **77**:2027–2042.

Okubo, A. and S. Levin, 2001. Diffusion and Ecological Problems: Modern Perspectives. Springer Verlag, New York.

Pacific States Marine Fisheries Commission, 1996. PIT tag information system (PTAGIS). Online database (Available through Internet: www.psmfc.org/ptagis).

Prentice, E. F., T. A. Flagg, C. S. McCutcheon, and D. F. Brastow, 1990. Pit-tag monitoring systems for hydroelectric dams and fish hatcheries. In N. P. e. al., editor, *Fish-marking techniques*, pages 323–334. American Fisheries Society Symposium 7, Bethesda, Maryland.

Price, M. V., P. A. Kelly, and R. L. Goldingay, 1994. Distances moved by Stephens' kangaroo rat (*Dipodomys stephensi Merriam*) and implications for conservation. *Journal of Mammalogy* **75**:929–939.

Skalski, G. T. and J. F. Gilliam, 2000. Modeling diffusive spread in a heteregeneous population: a movement study with stream fish. *Ecology* **81**:1685–1700.

Skalski, G. T. and J. F. Gilliam, 2003. A diffusion-based theory of organism dispersal in heterogeneous populations. *The American Naturalist* **161**:441–458.

Skellam, J. G., 1951. Random dispersal in theoretical populations. *Biometrika* **38**:196–218.

Steel, A. E., P. Guttorp, J. J. Anderson, and D. C. Caccia, 2001. Modeling juvenile salmon migration using a simple Markov chain. *Journal of Agricultural, Biological and Environmental Statistics* **6**:80–88.

Teichroew, D., 1957. The mixture of normal distributions with different variances. *The Mixture of Normal Distributions with Different Variances* **28**:510–512.

Tufto, J., T.-H. Ringsby, A. A. Dhondt, F. Adriaensen, and E. Matthysen, 2005. A parametric model for estimation of dispersal patterns applied to five passerine spatially structured populations. *The American Naturalist* **165**:E13E26.

Turchin, P., 1998. Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. Sinauer, Sunderland, Mass.

Tweedie, M. C. K., 1957. Statistical properties of Inverse Gaussian distributions. *The Annals of Mathematical Statistics* **28**:362–377.

USNMFS, 1998. Endangered and threatened species: threatened status for two ESU's of steelhead in Washington, Oregon and California. Technical report, U. S. National Marine Fisheries Service.

Walters, C. J., R. Hilborn, R. M. Peterman, and M. J. Staley, 1978. Model for examining early ocean limitation of pacific salmon production. *Journal of the Fisheries Research Board of Canada* **35**:13031315.

Waples, R. S., D. J. Teel, J. M. Myers, and A. R. Marschall, 2004. Life-history divergence in chinook salmon: historic contingency and parallel evolution. *Evolution* **58**:386–403.

Yamamura, K., 2002. Dispersal distance of heterogeneous populations. *Population Ecology* **44**:93101.

Yamamura, K., 2004. Dispersal distance of corn pollen under fluctuating diffusion coefficient. *Population Ecology* **46**:87101.

Zabel, R. and J. G. Williams, 2002. Selective mortality in chinook salmon: what is the role of human disturbance? *Ecological Applications* **12**:173183.

Zabel, R. W., 2002. Using "Travel Time" Data to Characterize the Behavior of Migrating Animals. *The American Naturalist* **159**:372–387.

Zabel, R. W. and J. J. Anderson, 1997. A model of the travel time of migrating juvenile salmon, with an application to Snake River spring chinook. *North American Journal of Fisheries Management* **118**:558–560.

Table 1: Table of underlying movement processes, heteregenous parameters and resulting distributions for several simple one-dimensional cases.

Underlying process	**Parameter distribution**		**Resulting Distribution**
Deterministic movement: $f(x\|t,v) = \delta(x - vt)$	homogeneous population: constant v	1a)	$h(x\|t) = \delta(x - vt)$
		1b)	$h_T(t\|x) = \delta(t - x/v)$
	heterogeneous velocity: $V \sim \mathrm{N}\left\{\mu_v, \sigma_v^2\right\}$	2a)	$h(x\|t) = \frac{1}{\sqrt{2\pi}\sigma_v t}\exp\left(-\frac{(x-\mu_v t)^2}{2(\sigma_v t)^2}\right)$
		2b)	$h_T(t\|x) = \frac{x}{\sqrt{2\pi}\sigma_v t^2}\exp\left(-\frac{(x-\mu_v t)^2}{2(\sigma_v t)^2}\right)$
Wiener process: $f(x\|t,v) = \mathrm{N}\left\{xt, \sigma_w^2 t\right\}$	homogeneous population: constant v and σ_w	3a)	$h(x\|t) = \frac{1}{\sqrt{2\pi}\sigma_w t^{1/2}}\exp\left(-\frac{(x-\mu_v t)^2}{2\sigma_w^2 t}\right)$
		3b)	$h_T(t\|x) = \frac{x}{\sqrt{2\pi}\sigma_w t^{3/2}}\exp\left(-\frac{(x-\mu_v t)^2}{2\sigma_w^2 t}\right)$
	heterogeneous velocity: $V \sim \mathrm{N}\left\{\mu_v, \sigma_v\right\}$	4a)	$h(x\|t) = \frac{1}{\sqrt{2\pi(\sigma_v^2 t+\sigma_w^2)t}}\exp\left(-\frac{(x-\mu_v t)^2}{2(\sigma_v^2 t+\sigma_w^2)t}\right)$
		4b)	$h_T(t\|x) = \frac{x}{\sqrt{2\pi(\sigma_v^2 t+\sigma_w^2)}t^{3/2}}\exp\left(-\frac{(x-\mu_v t)^2}{2(\sigma_v^2 t+\sigma_w^2)t}\right)$
	heterogeneous Wiener variance $\sigma_w^2 \sim \mathrm{gamma}\{\alpha, \beta\}$	5a)	$h(x\|t) = \frac{1}{\Gamma(\alpha)}\sqrt{\frac{2}{\pi\beta t}}\left(\frac{\|x-vt\|}{\sqrt{2\beta t}}\right)^{\alpha-\frac{1}{2}} K_{\frac{1}{2}-\alpha}\left(\frac{\|x-vt\|}{\sqrt{(\beta t)/2}}\right)$
		5b)	$h_T(t\|x) = \frac{x}{\Gamma(\alpha)}\sqrt{\frac{2}{\pi t^3 b}}\left(\frac{\|x-vt\|}{\sqrt{2\beta t}}\right)^{\alpha-\frac{1}{2}} K_{\frac{1}{2}-\alpha}\left(\frac{\|x-vt\|}{\sqrt{(\beta t)/2}}\right)$

Table 2: Mean, variance, skewness and kurtosis estimates for spatial distribution of bluehead chub (*Nocomis leptocephalus*) movements from (Skalski and Gilliam, 2000). The fish were marked and released at one site and monitored over a 4-mo period. Dispersal distance is measured in terms of number of sites.

Days	Mean	(se, n)	Variance	(se, n)	Skew	(se, n)	Kurt	(se, n)
0	0	(0,190)	0	(0,190)		na		na
30	1.13	(0.35,134)	22.29	(5.42,157)	0.44	(0.43,134)	7.34	(0.39,157)
60	1.57	(0.5,86)	27.40	(7.86,101)	0.95	(0.31,86)	6.37	(0.48,101)
90	3.19	(0.8,59)	45.64	(14.03,69)	0.87	(0.25,59)	4.58	(0.57,69)
120	2.44	(0.77,32)	66.62	(30.69,44)	1.08	(0.22,32)	7.51	(0.7,44)

5 Figures

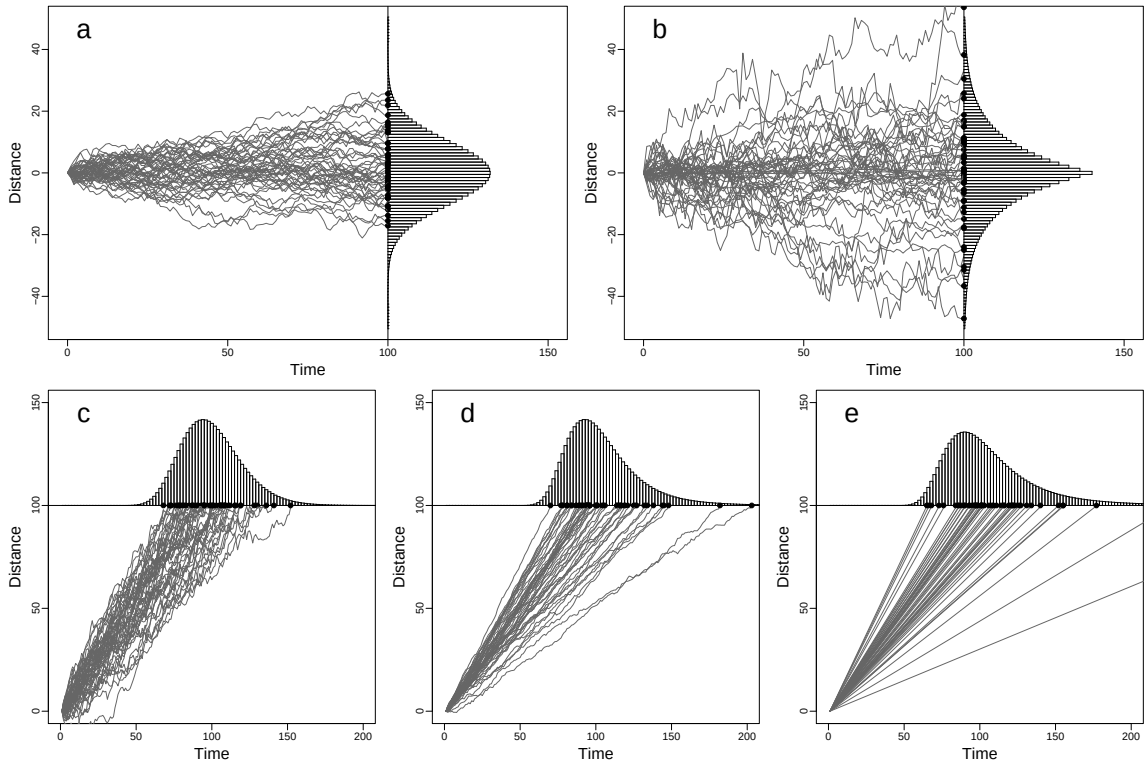

Figure 1: 50 simulated trajectories (grey lines) and theoretical distributions (bars). Spatial distribution at time 100 of (a) unbiased homogeneous random walkers ($\sigma_w = 1$) and (b) heterogeneous variance ($\sigma_w \sim \text{Gamma}\{\alpha = 1, \beta = 2\}$) walkers; arrival time distribution at fixed distance 100 for migrating walkers with mean velocity $\mu_v = 1$ of (c) homogenous random walkers ($\sigma_w = 2$, $\sigma_v = 0$); (d) a heterogeneous population of random walkers ($\mu_v = 1$, $\sigma_w = 0.5$, $\sigma_v = 0.2$) and (e) a heterogeneous population of deterministic walkers ($\sigma_v = .25$, $\sigma_w = 0$). The theoretical distributions for these cases are presented in table 1.

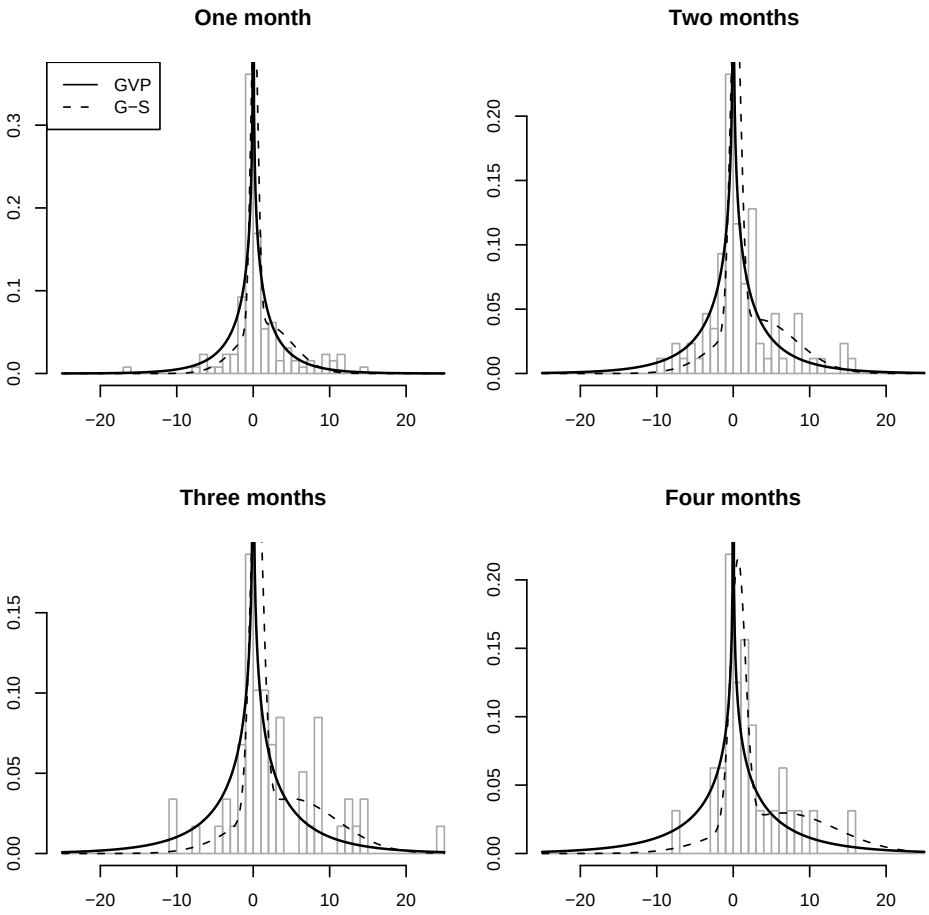

Figure 2: Comparison of the 3-parameter gamma-distributed variance model (G-V) and the five parameter Skalski-Gillam model (S-G) to histograms of bluehead chub disperal data at four months of observation.

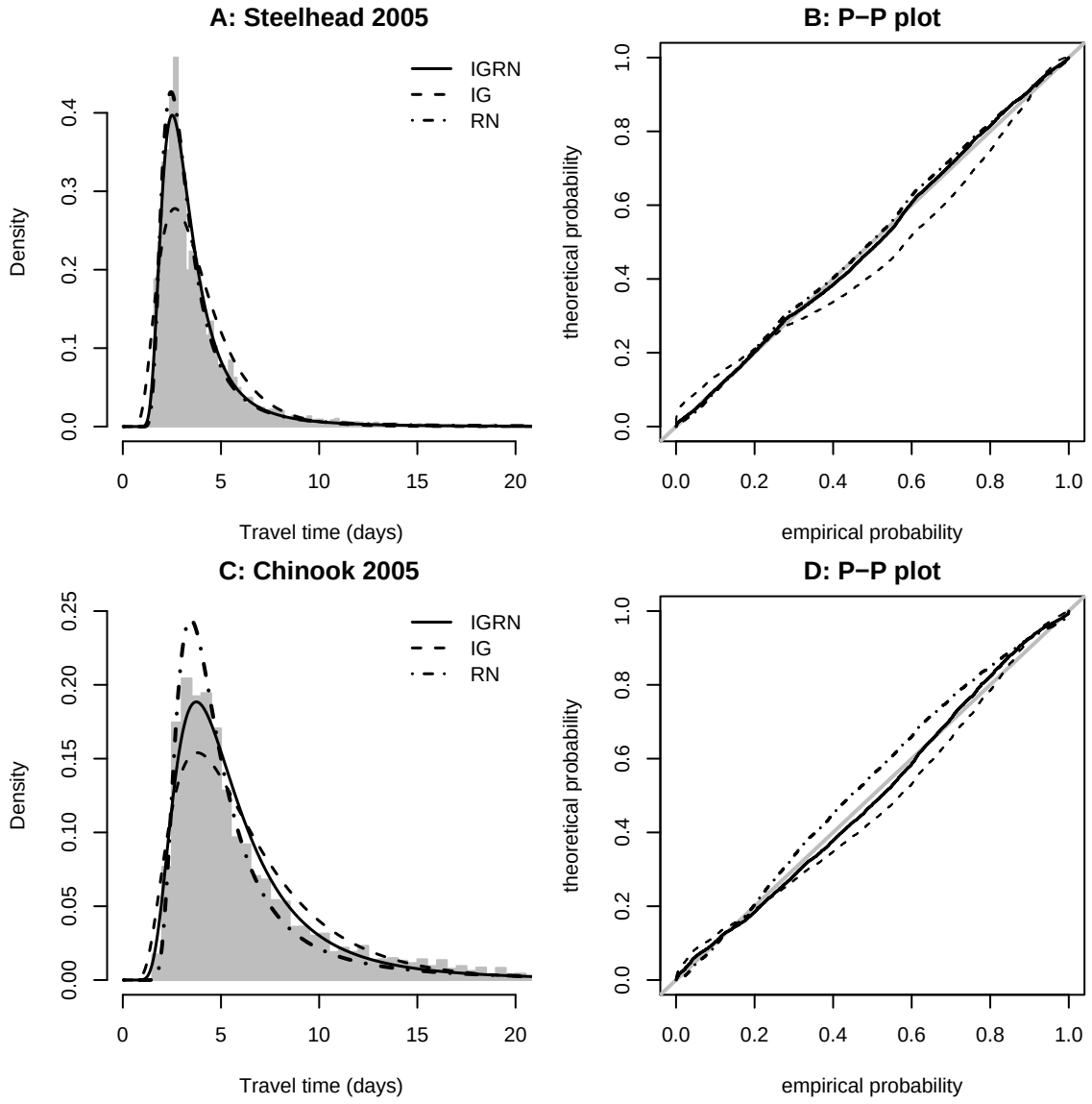

Figure 3: Examples of travel time model fits to travel times data. On all plots, the dashed line represents the homogeneous IG model (11), the half-dotted line represents the heterogeity-dominated RN model and the solid line represents the mixture IGRN distribution. The histograms represent travel times for (A) steelhead and (C) yearling chinook released at Lower Granite and detected at Little Goose dam, 59.8 km downstream, between May and June, 2005. The P-P plots (B) and (C) are a visual way to assess the fit of data to different theoretical distributions, with the 45 deg line representing a perfect fit. In all of these plots, the IGRN model is the best fit. The IG model misses the peak and overestimates the tail, while the RN model overestimates the peak for the chinook but is virtually indistinguishable from the IGRN for the steelhead.

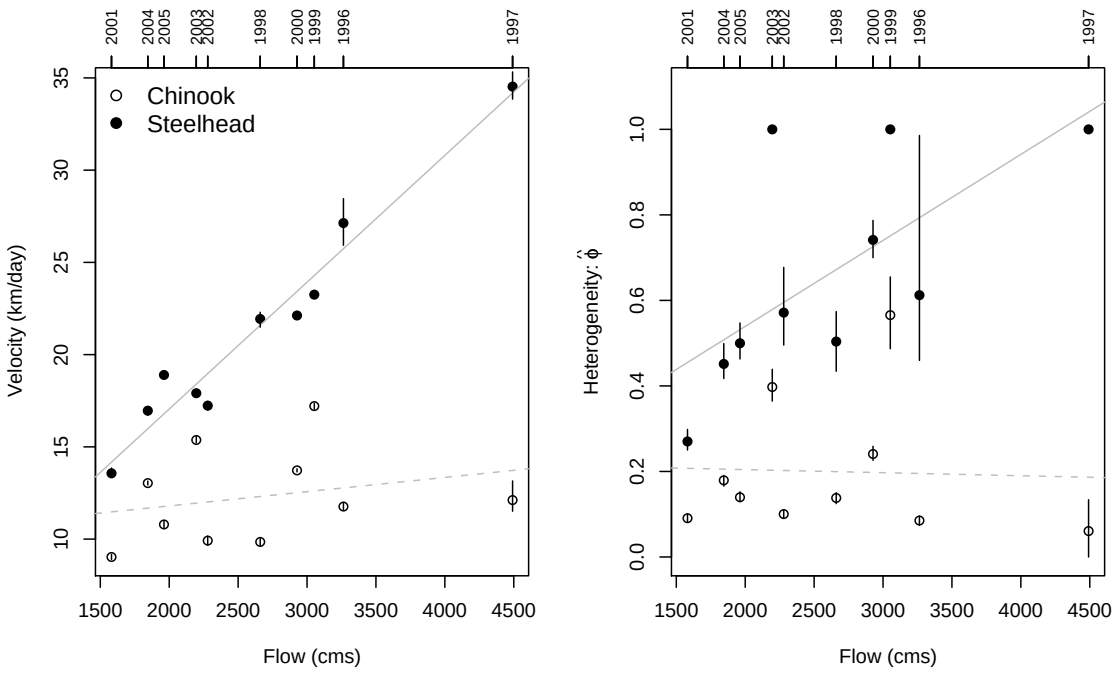

Figure 4: Plots of velocity estimate ($\hat{\mu_v}$) and (B) heterogeneity index estimates ($\hat{\phi}$) over all years against mean flow (in m^3sec^{-1}). The filled circles represent steelhead, the empty ircles represent chinook; solid and dashed lines represent linear regressions for steelhead and chinook respectively. The vertical bars are bootstrapped 95% confidence intervals.