

# Bounding seed loss from isolated habitat patches

Benjamin Hafner<sup>1\*</sup> and Katherine Meyer<sup>1\*</sup>

<sup>1</sup>Math and Statistics Department, Carleton College,  
One North College Street, Northfield, 55057, MN, USA.

\*Corresponding author(s). E-mail(s): [benrhafner@gmail.com](mailto:benrhafner@gmail.com);  
[kjmeyer@carleton.edu](mailto:kjmeyer@carleton.edu);

## Abstract

Dispersal of propagules (seeds, spores) from a geographically isolated population into an uninhabitable matrix can threaten population persistence if it prevents new growth from keeping pace with mortality. Quantifying propagule loss can thus inform restoration and conservation of vulnerable populations in fragmented landscapes. To model propagule loss in detail, one can integrate dispersal kernels (probabilistic descriptions of dispersal) and plant densities. However, one might lack the detailed spatial information and computational tools needed for such integral models. Here we derive two simple upper bounds on the probability of propagule loss—one assuming rotational symmetry of dispersal and the other not—that require only habitat area, habitat perimeter, and the mean dispersal distance of a propagule. We compare the bounds to simulations of integral models for the population of *Asclepias syriaca* (common milkweed) at McKnight Prairie—a 13.7 hectare reserve surrounded by agricultural fields in Goodhue County, Minnesota—and identify conditions under which the bounds closely estimate propagule loss.

**Keywords:** model, propagules, seeds, dispersal kernel, habitat fragmentation, spillover

**Acknowledgments.** We are grateful for feedback on a draft of this manuscript from the Shaw Lab and associates at UMN, including Allison Shaw, Dennis Kim, Martha Torstenson, Naven Narayanan, Jessica Valiarovski, and Matt Michalska-Smith. Carleton College supported KM’s work on the project during a sabbatical.

# 1 Introduction

This paper concerns the probability that propagules (seeds, spores) from a population in an isolated habitat patch disperse outside of that patch. In the right conditions, dispersal may help individuals reach environments with lower conspecific competition, fewer natural enemies, and new opportunities for establishment (Beckman & Sullivan 2023, Levin et al. 2003). Dispersal can also support metapopulation persistence despite local extinctions by allowing a species to (re)colonize patches and bet-hedge against fluctuating habitat quality (Levin et al. 2003, Hanski 2001). But for a small site surrounded by a large matrix of unviable habitat, dispersal out of the patch threatens local population persistence when emigration outweighs the arrival of propagules from external sources (Hanski 2001). Cheptou and colleagues (2008) documented a striking urban example of this dispersal risk in the weed *Crepis sancta*, whose dispersing seeds departed from tree plantings an estimated 55% of the time, resulting in selection for a nondispersing seed variant.

The proportion of propagules that leave viable habitat emerged as a key parameter for predicting population crashes in the colonization-mortality model of Cooney et al. (2022)—a reformulation of Tilman’s 1994 spatially implicit extinction debt model that includes spatial dispersal processes. Quantifying propagule movement from inside to outside of habitat patches could thus help managers to identify either isolated remnant populations that are at risk of decline or sites that are suitable to support restored populations.

Propagule loss from a habitat patch can be modeled in detail by integrating functions that represent the density of propagule sources and the probabilistic patterns of propagule dispersal (the dispersal kernel of the propagule source) (Cooney et al. 2022, Klausmeier 1998). In particular, the density of propagules arriving at a location  $(x_0, y_0)$  in  $\mathbb{R}^2$  is given by (Nathan et al. 2012)

$$\text{propagule arrival density at } (x_0, y_0) = \iint_{\text{habitat}} \left( \begin{array}{c} \text{propagule} \\ \text{source density} \\ \text{at } (x, y) \end{array} \right) \left( \begin{array}{c} \text{dispersal kernel} \\ \text{evaluated at} \\ (x_0 - x, y_0 - y) \end{array} \right) dA \quad (1)$$

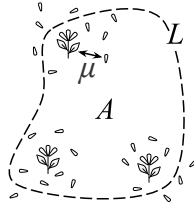
and total propagule loss can be obtained by integrating propagule arrival density over the complement of the habitat patch:

$$\text{propagule loss} = \iint_{\substack{\text{not} \\ \text{habitat}}} \left( \begin{array}{c} \text{propagule} \\ \text{arrival density} \end{array} \right) dA. \quad (2)$$

This approach incorporates information about the spatial distribution of both propagule sources and propagule dispersal. In many studies, knowledge of these attributes may be incomplete.

Here we derive an estimate of propagule loss that could be especially useful when one has only rough information on the location of propagule sources and the patterns of propagule dispersal. For brevity, we refer to propagule sources as plants and propagules as seeds. We make the simplifying assumptions that

- (i) plants are uniformly distributed within a viable habitat and



**Fig. 1** Schematic of quantities used to bound seed loss, with the habitat boundary depicted as a dashed line

- (ii) the dispersal kernel associated with a plant is independent from its location in the habitat.

In this context, we can bound the probability of seed loss  $p$  from a population in an isolated habitat patch by

$$p \leq \frac{\mu L}{2A} \quad (3)$$

where  $A$  is the habitat area,  $L$  is the perimeter, and  $\mu$  is the mean seed dispersal distance (illustrated in Figure 1). If in addition we assume that dispersal patterns are identical in each direction—that is, rotationally symmetric—then we can decrease this upper bound to

$$p \leq \frac{\mu L}{\pi A}. \quad (4)$$

Given that patch area  $A$  appears in the denominator of these expressions, while patch perimeter  $L$  appears in the numerator, we expect the seed loss bound to decrease as we scale up a habitat region while maintaining its shape (since  $L$  grows linearly while  $A$  grows quadratically). Loosely speaking, the ratio  $A/L$  can be interpreted as the habitat’s characteristic length scale, which must be kept larger than  $\mu$  in order to limit seed loss.

In addition, we observe that these bounds are tight, in two senses. In a practical sense, when applied to an ecologically relevant plant population, the bounds are close to the true value of  $p$ , especially in the rotationally symmetric case when the habitat is large relative to the mean dispersal distance (see Section 5). And from a mathematical perspective, they are tight in the sense that the constant factors  $1/2$  and  $1/\pi$  cannot be improved (see Appendix A).

Though simple to state, inequalities (3) and (4) require some creative vector calculus to prove. We begin by establishing a probabilistic modeling framework to precisely define  $p$  in Section 2. Then, in Section 3, we translate the spatial geometry of seed movement from dispersal kernels to vector fields and use the divergence theorem to reframe seed loss as a flux of seeds across the habitat boundary. In Section 4 we bound seed flux across the habitat boundary. We put all the pieces together in Theorem 7, which formalizes the claims (3) and (4).

Once the proof of Theorem 7 is complete, we provide an example of how it could be applied in a conservation context in Section 5. Specifically, we compute seed loss bounds for a population of *Asclepias syriaca* (common milkweed) located at McKnight Prairie, an isolated reserve surrounded by agriculture, and show it closely matches an integral model of seed loss. In Section 6 we discuss methods for estimating mean dispersal distance as well as future theoretical directions.

## 2 Modeling framework

Throughout this text, given sets  $U$  and  $V \subset \mathbb{R}^2$  we use  $\partial U$  to denote the boundary of  $U$ ,  $U^c$  for set complement, and  $U \setminus V$  for the set difference  $U \cap V^c$ .

We represent viable habitat for a plant population of interest as an open subset  $H \subset \mathbb{R}^2$  with finite area  $A$  and perimeter  $L$ . We treat plant positions, seed displacements, and seed landing locations as continuous random vectors with values in  $\mathbb{R}^2$ .

The random vector  $\mathbf{X}_p$  models the position of plants in the population. Recall that  $\mathbf{X}_p$  relates to its probability density function  $f_p$  as follows: given a set  $U \subset \mathbb{R}^2$ ,

$$P(\mathbf{X}_p \in U) = \iint_{\mathbf{x} \in U} f_p(\mathbf{x}) dA. \quad (5)$$

Here  $dA$  refers to an infinitesimal area of integration (e.g.  $dx_1 dx_2$ ) and does not imply a relationship to total habitat area  $A$ . We assume plant positions are uniformly distributed over  $H$ , so

$$f_p(\mathbf{x}) = \begin{cases} 1/A & \text{if } \mathbf{x} \in H \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

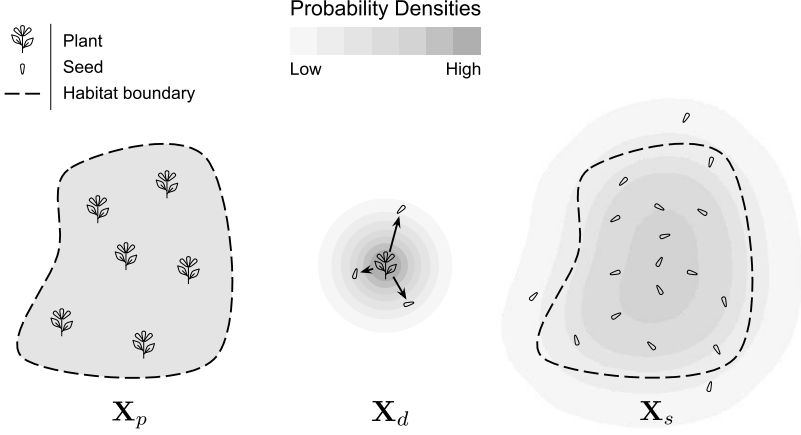
The seeds released by a plant are displaced by a random dispersal vector  $\mathbf{X}_d$  from the source plant before they land. The probability density function  $f_d$  associated with  $\mathbf{X}_d$  is also known as a dispersal kernel (Nathan et al. 2012). Importantly, we assume  $\mathbf{X}_p$  and  $\mathbf{X}_d$  are independent: the position of a source plant within the habitat does not influence the dispersal pattern of its seeds (see Section 6 for discussion of this assumption). The mean dispersal distance of seeds is given by

$$\mu = E[|\mathbf{X}_d|] = \iint_{\mathbf{x} \in \mathbb{R}^2} |\mathbf{x}| f_d(\mathbf{x}) dA. \quad (7)$$

We assume that  $f_d$  decays rapidly enough as  $|\mathbf{x}| \rightarrow \infty$  to make  $\mu$  well-defined. Commonly used functional forms of  $f_d$  include the inverse Gaussian for wind-dispersed species and the exponential power function (Bullock et al. 2017, Nathan et al. 2012).

We define the random vector  $\mathbf{X}_s$  by

$$\mathbf{X}_s = \mathbf{X}_p + \mathbf{X}_d \quad (8)$$



**Fig. 2** Visualization of the probability densities associated with random variables for plant locations  $\mathbf{X}_p$ , seed dispersal  $\mathbf{X}_d$ , and seed landing locations  $\mathbf{X}_s$

to model the landing position of a seed that originated from the plant population in  $H$ . We assume that  $H^c$  is not viable habitat for the population of interest, so seeds that originate in  $H$  but land in  $H^c$  are lost. Our main quantity of interest, the probability  $p$  of seed loss from the habitat, is given by

$$p = P(\mathbf{X}_s \in H^c) = \iint_{\mathbf{x} \in H^c} f_s(\mathbf{x}) dA \quad (9)$$

where  $f_s$  is the probability density of  $\mathbf{X}_s$ .

As a first step towards bounding seed loss probability  $p$ , we unpack  $p$ 's defining integral from (9) in terms of habitat area  $A$  and dispersal kernel  $f_d$ .

**Proposition 1** The seed loss probability  $p$  defined in equation (9) is equivalent to

$$p = \frac{1}{A} \iint_{\mathbf{x}_p \in H} \iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) dA dA. \quad (10)$$

*Proof* Because  $\mathbf{X}_p$  and  $\mathbf{X}_d$  are independent, the density of their sum  $\mathbf{X}_s$  is the convolution of their densities:

$$f_s(\mathbf{x}) = \iint_{\mathbf{x}_p \in \mathbb{R}^2} f_p(\mathbf{x}_p) f_d(\mathbf{x} - \mathbf{x}_p) dA. \quad (11)$$

Replacing  $f_p$  in (11) with its definition from (6) yields

$$f_s(\mathbf{x}) = \frac{1}{A} \iint_{\mathbf{x}_p \in H} f_d(\mathbf{x} - \mathbf{x}_p) dA. \quad (12)$$

By substituting  $f_s$  from (12) into the definition (9) of  $p$  and reversing the order of integration, one obtains the expression (10).  $\square$

The representation of  $p$  in Proposition 1 can be interpreted as seed loss probability for an individual plant at  $\mathbf{x}_p$  (the inner integral), averaged uniformly over the habitat  $H$  (the outer integral). This form is useful for calculating  $p$  numerically from a known habitat  $H$  and dispersal kernel  $f_d$ , as done by Cooney et al. (2022).

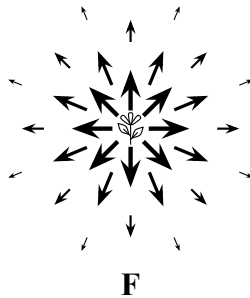
### 3 Seed loss as the flux of a vector field across habitat boundary

To bound  $p$  in terms of habitat perimeter  $L$ , we transform the inner integral over  $H^c$  in equation (10) into a line integral around the habitat boundary  $\partial H$ . Specifically, we express seed loss as a flux of seeds crossing the habitat boundary. To do so, we must describe seed flow in the language of vector fields. Subsection 3.1 constructs a vector field describing dispersal from a single plant and subsection 3.2 integrates over all plant locations to obtain a vector field representing total seed dispersal.

#### 3.1 Vector field of a single plant

The dispersal of seeds from a single plant is typically represented as a probability density function  $f_d$ , but the same information can also be represented as a vector field, which we call  $\mathbf{F}$ . The vector field  $\mathbf{F}$  describes the motion of seeds as they disperse—like a flowing fluid, whereas  $f_d$  describes their final resting positions.

The mathematical link between  $\mathbf{F}$  and  $f_d$  is the divergence operator. Intuitively, a positive divergence of  $\mathbf{F}$  indicates a source (seeds being released) and a negative divergence indicates a sink (seeds landing). Our definition of  $\mathbf{F}$  below yields a positive divergence singularity at the origin, representing the plant as a point source of seeds, and yields  $\nabla \cdot \mathbf{F} = -f_d$  everywhere else, encoding seed landing positions. For more physical interpretation and motivation of  $\mathbf{F}$ 's definition, see Appendix B.



**Fig. 3** The vector field  $\mathbf{F}$  represents the spatial flow of seeds away from a source plant

**Definition 1** Given a dispersal kernel  $f_d$ , the corresponding vector field  $\mathbf{F}$  is

$$\mathbf{F}(\mathbf{x}) = \frac{\mathbf{x}}{r_x^2} \int_{r_x}^{\infty} f_d(r, \theta_x) r dr. \quad (13)$$

**Lemma 2** For all non-zero  $\mathbf{x} \in \mathbb{R}^2$ ,

$$\nabla \cdot \mathbf{F}(\mathbf{x}) = -f_d(\mathbf{x}). \quad (14)$$

*Proof* We work in polar coordinates with unit vectors  $\hat{\mathbf{r}} = \mathbf{x}/|\mathbf{x}|$  and  $\hat{\boldsymbol{\theta}}$  the rotation of  $\hat{\mathbf{r}}$  by  $\pi/2$ . In these coordinates,  $\mathbf{F}(\mathbf{x}) = |\mathbf{F}(\mathbf{x})|\hat{\mathbf{r}} + 0\hat{\boldsymbol{\theta}}$  and by the polar divergence formula

$$\nabla \cdot \mathbf{F} = \frac{1}{r_x} \frac{\partial}{\partial r_x} [r_x |\mathbf{F}|] + \frac{1}{r_x} \frac{\partial}{\partial \theta} [0] \quad (15)$$

$$= \frac{1}{r_x} \frac{\partial}{\partial r_x} \left[ \int_{r_x}^{\infty} f_d(r, \theta_x) r dr \right] \quad (16)$$

$$= \frac{1}{r_x} (-f_d(r_x, \theta_x) r_x) \quad (17)$$

$$= -f_d(r_x, \theta_x). \quad (18)$$

□

Using Definition 1 and Lemma 2, one can convert from a dispersal kernel  $f_d$  to a vector field  $\mathbf{F}$  and back again. The two objects carry the same information about seed dispersal, simply in different forms.

The remainder of Section 3.1 concerns the flux of  $\mathbf{F}$  across the habitat boundary, which represents a single plant's probability of seed loss.

**Lemma 3** Let  $D$  be a disk shaped habitat with a single plant at its center, the origin. Then the flux of  $\mathbf{F}$  across the habitat boundary  $\partial D$  is equal to the probability that the plant's seeds are lost (land outside  $D$ ). That is,

$$\oint_{\mathbf{x} \in \partial D} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n} ds = \iint_{\mathbf{x} \in D^c} f_d(\mathbf{x}) dA \quad (19)$$

where  $\mathbf{n}$  is the outward pointing unit vector normal to the boundary.

*Proof* Suppose the disk has radius  $R$ . Then applying Definition 1 and simplifying yields

$$\oint_{\mathbf{x} \in \partial D} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n} ds = \oint_{\mathbf{x} \in \partial D} \frac{1}{R} \int_R^{\infty} f_d(r, \theta_x) r dr ds. \quad (20)$$

Replacing  $ds$  with  $R d\theta_x$  and parameterizing the loop integral by  $\theta_x$  gives

$$\oint_{\mathbf{x} \in \partial D} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n} ds = \int_0^{2\pi} \frac{1}{R} \int_R^\infty f_d(r, \theta_x) r dr R d\theta_x \quad (21)$$

$$= \int_0^{2\pi} \int_R^\infty f_d(r, \theta_x) r dr d\theta_x \quad (22)$$

$$= \iint_{\mathbf{x} \in D^c} f_d(\mathbf{x}) dA. \quad (23)$$

□

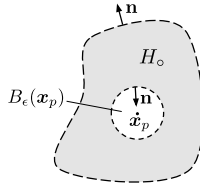
In Proposition 4, we use the divergence theorem to generalize Lemma 3 from a disk  $D$  to a generic habitat shape  $H$  and allow the plant to reside at any point  $\mathbf{x}_p$  in  $H$ , not just the origin.

**Proposition 4** For all  $\mathbf{x}_p \in H$ ,

$$\iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) dA = \oint_{\mathbf{x} \in \partial H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} ds. \quad (24)$$

*Proof* We begin with the right hand side and derive the left. To apply the divergence theorem, care is required around the singularity of  $\mathbf{F}(\mathbf{x} - \mathbf{x}_p)$  when  $\mathbf{x} = \mathbf{x}_p$ . Fix  $\mathbf{x}_p \in H$  and choose  $\epsilon > 0$  such that the ball  $B_\epsilon(\mathbf{x}_p)$  of radius  $\epsilon$  around  $\mathbf{x}_p$  lies entirely within  $H$ . Let  $H_\circ = H \setminus B_\epsilon(\mathbf{x}_p)$ . Note that the boundary of  $H_\circ$  consists of the boundary of  $H$  and the boundary of  $B_\epsilon(\mathbf{x}_p)$ . Let  $\mathbf{n}$  represent the unit normal to each boundary, with orientations shown in Figure 4. We have that

$$\oint_{\mathbf{x} \in \partial H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} ds = \underbrace{\oint_{\mathbf{x} \in \partial H_\circ} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} ds}_I - \underbrace{\oint_{\substack{\mathbf{x} \in \\ \partial B_\epsilon(\mathbf{x}_p)}} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} ds}_{II}. \quad (25)$$



**Fig. 4** Orientation of normal vectors used in the proof of Proposition 4

Because  $H_o$  excludes the singularity at  $\mathbf{x}_p$ , we may apply the divergence theorem to integral I of (25). This gives

$$\begin{aligned} \oint_{\mathbf{x} \in \partial H_o} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, ds &= \iint_{\mathbf{x} \in H_o} \nabla \cdot \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \, dA \\ &= - \iint_{\mathbf{x} \in H_o} f_d(\mathbf{x} - \mathbf{x}_p) \, dA, \end{aligned} \quad (26)$$

where the second equality follows from Lemma 2. Integral II is handled by Lemma 3, since  $B_\epsilon(\mathbf{x}_p)$  is a disk:

$$\oint_{\mathbf{x} \in \partial B_\epsilon(\mathbf{x}_p)} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, ds = - \iint_{\mathbf{x} \in B_\epsilon(\mathbf{x}_p)^c} f_d(\mathbf{x} - \mathbf{x}_p) \, dA. \quad (27)$$

Note the translation from  $\mathbf{x}$  in Lemma 3 to  $\mathbf{x} - \mathbf{x}_p$  in equation (27), along with the re-orientation of the unit normal from outward to inward, resulting in negation.

Substituting (26) and (27) for integrals I and II in (25) gives

$$\oint_{\mathbf{x} \in \partial H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, ds = - \iint_{\mathbf{x} \in H_o} f_d(\mathbf{x} - \mathbf{x}_p) \, dA + \iint_{\mathbf{x} \in B_\epsilon(\mathbf{x}_p)^c} f_d(\mathbf{x} - \mathbf{x}_p) \, dA. \quad (28)$$

Because  $H_o \subset B_\epsilon(\mathbf{x}_p)^c$ , the difference of integrals in (28) is the integral over their set difference  $B_\epsilon(\mathbf{x}_p)^c \setminus H_o$ , which is  $H^c$ . This yields our desired result

$$\oint_{\mathbf{x} \in \partial H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, ds = \iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) \, dA.$$

□

### 3.2 Integrating seed flux from all plant locations

Next we develop a second vector field  $\mathbf{G}$  that aggregates seed movement from all plants in the habitat. Using Proposition 4, we can rewrite the inner integral in the expression (10) for seed loss probability  $p$  in terms of seed flux from a single source at  $\mathbf{x}_p$  across the habitat boundary. This gives

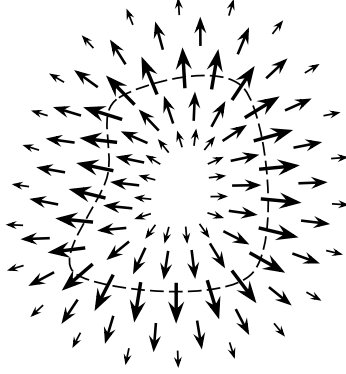
$$p = \frac{1}{A} \iint_{\mathbf{x}_p \in H} \oint_{\mathbf{x} \in \partial H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, ds \, dA. \quad (29)$$

Reversing the order of integration and moving  $\mathbf{n}$  out of the inner integral yields

$$p = \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \iint_{\mathbf{x}_p \in H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \, dA \cdot \mathbf{n} \, ds. \quad (30)$$

The inner integral in (30) now represents the effective flow of seeds contributed from all plants in the habitat at a point  $\mathbf{x}$  on the habitat boundary. We call this the total dispersal field and denote it

$$\mathbf{G}(\mathbf{x}) \equiv \iint_{\mathbf{x}_p \in H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \, dA, \quad (31)$$



**Fig. 5** The total dispersal field  $\mathbf{G}$  models the net flow of all seeds originating from the plant population in  $H$  (eq. (31)). The outward flux of  $\mathbf{G}$  across the habitat boundary, depicted here as a dotted line, represents seed loss (eq. (32))

so that

$$p = \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \, ds. \quad (32)$$

## 4 Bounding seed loss

Based on equation (32), in order to bound  $p$  it suffices to bound the magnitude of  $\mathbf{G} \cdot \mathbf{n}$ . We derive two upper bounds on  $\mathbf{G} \cdot \mathbf{n}$ : one with no assumptions on the symmetry of the dispersal kernel  $f_d$  (Lemma 5) and another when  $f_d$  is rotationally symmetric (Lemma 6).

**Lemma 5** For all  $\mathbf{x} \in \mathbb{R}^2$  and for all unit vectors  $\mathbf{n} \in \mathbb{R}^2$ ,

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \frac{\mathbf{n}}{2} \cdot E(\mathbf{X}_d) + \frac{\mu}{2}.$$

*Proof* It follows from the definition (31) of  $\mathbf{G}$  that

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} = \iint_{\mathbf{x}_p \in H} \mathbf{F}(\mathbf{x} - \mathbf{x}_p) \cdot \mathbf{n} \, dA.$$

To simplify the following argument, we define  $\mathbf{y} = \mathbf{x} - \mathbf{x}_p$  as a shorthand for the seed dispersal vector. Additionally, let  $\hat{H} = \{\mathbf{y} \in \mathbb{R}^2 \mid \mathbf{x} - \mathbf{y} \in H\}$  so that

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} = \iint_{\mathbf{y} \in \hat{H}} \mathbf{F}(\mathbf{y}) \cdot \mathbf{n} \, dA.$$

Expanding  $\mathbf{F}$  by Definition 1, we have

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} = \iint_{\mathbf{y} \in \hat{H}} \frac{\mathbf{y} \cdot \mathbf{n}}{r_y^2} \int_{r_y}^{\infty} f_d(r, \theta_y) r dr dA \quad (33)$$

where  $\theta_y, r_y$  are the polar coordinates of  $\mathbf{y}$ . Next, we use the fact that  $\mathbf{y} \cdot \mathbf{n} \leq \frac{\mathbf{y} \cdot \mathbf{n}}{2} + \frac{|\mathbf{y} \cdot \mathbf{n}|}{2}$  to produce the inequality

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \iint_{\mathbf{y} \in \hat{H}} \left( \frac{\mathbf{y} \cdot \mathbf{n}}{2} + \frac{|\mathbf{y} \cdot \mathbf{n}|}{2} \right) \frac{1}{r_y^2} \int_{r_y}^{\infty} f_d(r, \theta_y) r dr dA. \quad (34)$$

Substituting  $\frac{\mathbf{y} \cdot \mathbf{n}}{2} + \frac{|\mathbf{y} \cdot \mathbf{n}|}{2}$  for  $\mathbf{y} \cdot \mathbf{n}$  has no effect when  $\mathbf{y} \cdot \mathbf{n} > 0$ , that is, when the seed dispersal vector points outward across the habitat boundary. But when  $\mathbf{y} \cdot \mathbf{n} < 0$ , it is replaced by zero. Thus, we ignore the contribution of any seeds flowing inward across the habitat boundary. Although at first it appears to introduce needless complication, this substitution enables the rest of the proof.

Because  $\frac{\mathbf{y} \cdot \mathbf{n}}{2} + \frac{|\mathbf{y} \cdot \mathbf{n}|}{2}$  is never negative, we are now free to expand the region of integration from  $\mathbf{y} \in \hat{H}$  to all  $\mathbf{y} \in \mathbb{R}^2$ , which we parameterize by  $\theta_y, r_y$ , with  $\theta_y$  measured relative to  $\mathbf{n}$  so that  $\mathbf{y} \cdot \mathbf{n} = r_y \cos \theta_y$ . With this, (34) becomes

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \int_0^{\infty} \left( \frac{r_y \cos \theta_y}{2} + \frac{|r_y \cos \theta_y|}{2} \right) \frac{1}{r_y^2} \int_{r_y}^{\infty} f_d(r, \theta_y) r dr r_y dr_y d\theta_y \quad (35)$$

$$= \int_0^{2\pi} \int_0^{\infty} \int_{r_y}^{\infty} \left( \frac{\cos \theta_y}{2} + \frac{|\cos \theta_y|}{2} \right) f_d(r, \theta_y) r dr dr_y d\theta_y. \quad (36)$$

We interchange the order of integration between  $dr$  and  $dr_y$  to obtain

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \int_0^{\infty} \int_0^r \left( \frac{\cos \theta_y}{2} + \frac{|\cos \theta_y|}{2} \right) f_d(r, \theta_y) r dr_y dr d\theta_y \quad (37)$$

and evaluate the innermost integral, yielding an additional factor of  $r$ :

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \int_0^{\infty} \left( \frac{r \cos \theta_y}{2} + \frac{r |\cos \theta_y|}{2} \right) f_d(r, \theta_y) r dr d\theta_y. \quad (38)$$

Let  $\hat{\mathbf{y}} = \frac{r}{r_y} \mathbf{y}$ , so that  $\hat{\mathbf{y}}$  has the same direction as  $\mathbf{y}$  but magnitude  $r$  instead of  $r_y$ . Then  $r \cos \theta_y = \hat{\mathbf{y}} \cdot \mathbf{n}$  and

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \int_0^{\infty} \left( \frac{\hat{\mathbf{y}} \cdot \mathbf{n}}{2} + \frac{|\hat{\mathbf{y}} \cdot \mathbf{n}|}{2} \right) f_d(r, \theta_y) r dr d\theta_y \quad (39)$$

$$= \iint_{\hat{\mathbf{y}} \in \mathbb{R}^2} \left( \frac{\hat{\mathbf{y}} \cdot \mathbf{n}}{2} + \frac{|\hat{\mathbf{y}} \cdot \mathbf{n}|}{2} \right) f_d(\hat{\mathbf{y}}) dA \quad (40)$$

$$\leq \frac{\mathbf{n}}{2} \cdot \iint_{\hat{\mathbf{y}} \in \mathbb{R}^2} \hat{\mathbf{y}} f_d(\hat{\mathbf{y}}) dA + \frac{1}{2} \iint_{\hat{\mathbf{y}} \in \mathbb{R}^2} |\hat{\mathbf{y}}| f_d(\hat{\mathbf{y}}) dA \quad (41)$$

$$= \frac{\mathbf{n}}{2} \cdot E(\mathbf{X}_d) + \frac{1}{2} \mu \quad (42)$$

as claimed.  $\square$

The bound in Lemma 5 remains valid when seed dispersal is asymmetric or biased in one direction, for example by prevailing winds. In contrast, Lemma 6 requires that dispersal patterns are identical in each direction, which yields a stronger bound.

**Lemma 6** Suppose  $f_d$  is rotationally symmetric; that is, there exists a function  $\hat{f}_d$  such that  $f_d(r, \theta) = \hat{f}_d(r)$  for all  $\theta$ . Then for all  $\mathbf{x} \in \mathbb{R}^2$  and for all unit vectors  $\mathbf{n} \in \mathbb{R}^2$ ,

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \frac{\mu}{\pi}.$$

*Proof* Following the proof of Lemma 5 to equation (38), we have

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \int_0^\infty \left( \frac{r \cos \theta_y}{2} + \frac{r |\cos \theta_y|}{2} \right) f_d(r, \theta_y) r dr d\theta_y.$$

But because  $f_d(r, \theta) = \hat{f}_d(r)$ , which depends only on  $r$ , this double integral can now be separated:

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \int_0^{2\pi} \left( \frac{\cos \theta_y}{2} + \frac{|\cos \theta_y|}{2} \right) d\theta_y \int_0^\infty \hat{f}_d(r) r^2 dr \quad (43)$$

$$= 2 \int_0^\infty \hat{f}_d(r) r^2 dr. \quad (44)$$

When the definition of  $\mu$  from equation (7) is re-written in polar coordinates,

$$\mu = \int_0^{2\pi} \int_0^\infty r \hat{f}_d(r) r dr d\theta = 2\pi \int_0^\infty \hat{f}_d(r) r^2 dr, \quad (45)$$

it becomes clear, upon comparison of equations (44) and (45), that

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \frac{\mu}{\pi}. \quad (46)$$

□

**Theorem 7** Suppose a plant population with mean seed dispersal distance  $\mu$  is uniformly distributed over a habitat patch  $H$  with finite area  $A$  and perimeter  $L$ . If dispersal is independent of plant position, then the probability that a given seed disperses outside the habitat is bounded by

$$p \leq \frac{\mu L}{2A}.$$

If in addition dispersal is rotationally symmetric, then the stronger bound

$$p \leq \frac{\mu L}{\pi A}$$

holds.

*Proof* From equation (32) we have

$$p = \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \mathbf{G}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) ds,$$

where  $\mathbf{n}(x)$  is the outward-oriented unit vector orthogonal to  $\partial H$  at  $\mathbf{x}$ . In the general case, Lemma 5 gives that

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \leq \frac{\mathbf{n}(\mathbf{x})}{2} \cdot E(\mathbf{X}_d) + \frac{\mu}{2},$$

and so

$$p \leq \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \left( \frac{\mathbf{n}(\mathbf{x})}{2} \cdot E(\mathbf{X}_d) + \frac{\mu}{2} \right) ds \quad (47)$$

$$= \frac{1}{A} \left( \frac{1}{2} E(\mathbf{X}_d) \cdot \oint_{\mathbf{x} \in \partial H} \mathbf{n}(\mathbf{x}) ds + \oint_{\mathbf{x} \in \partial H} \frac{\mu}{2} ds \right). \quad (48)$$

The integral of  $\mathbf{n}(\mathbf{x})$  around the closed curve  $\partial H$  vanishes, leaving

$$p \leq \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \frac{\mu}{2} ds = \frac{\mu L}{2A} \quad (49)$$

as claimed.

In the special case that dispersal is rotationally symmetric, we have from Lemma 6 that

$$\mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \leq \frac{\mu}{\pi}$$

which immediately gives

$$p \leq \frac{1}{A} \oint_{\mathbf{x} \in \partial H} \frac{\mu}{\pi} ds = \frac{\mu L}{\pi A}.$$

□

In Appendix A we show that the constant factors  $1/2$  and  $1/\pi$  appearing in Theorem 7 are the best that can be achieved under the assumptions of this paper.

## 5 Example

To illustrate a practical application of Theorem 7, we turn to McKnight Prairie, a 14 hectare (34 acre) patch of remnant prairie in southern Minnesota. The site was mostly spared from agricultural use due to its hilly terrain, and is now protected under the Minnesota Department of Natural Resources (DNR) Native Prairie Bank Program.

Since the plot is embedded in a matrix of commercial agriculture, any seeds that land outside its borders are effectively lost. Especially for a species like *Asclepias syriaca* (common milkweed), whose seeds are easily carried away on the wind, one might wonder if a large fraction of seeds are lost to the surrounding farmland, hampering the species' ability to reach available sites within the local prairie and potentially leading to population decline. In this type of scenario, Theorem 7 may provide a rough estimate and upper bounds on seed loss. We compute these bounds in Section 5.1, simulate seed loss via a full dispersal kernel model in Section 5.2, and compare the bounds to simulations in Section 5.3.

## 5.1 Upper bounds on seed loss

To compute the bounds given by Theorem 7, three inputs are required: habitat area  $A$ , habitat perimeter  $L$ , and mean dispersal distance  $\mu$ . We traced the border of McKnight Prairie from Google Maps and then used a simple Python program to estimate its area and perimeter:  $A = 137,000 \text{ m}^2$  (13.7 hectares) and  $L = 1,930 \text{ m}$ . One could also make physical measurements or use GIS software.

Mean dispersal distance is somewhat more challenging to estimate. Section 6.1 discusses several practical routes to computing  $\mu$ , including the “ballistic” model for wind-dispersed seeds,

$$\mu = \frac{v_w h}{v_T}, \quad (50)$$

which requires release height  $h$ , terminal velocity  $v_T$ , and wind speed  $v_w$  (Nathan et al. 2011). For *A. syriaca*, we take  $h = 0.866 \text{ m}$  and  $v_T = 0.219 \text{ m/s}$  from Sullivan et al. (2018). We estimate wind speed using proxy data recorded by Seeley (2021) at Cedar Creek Ecosystem Science Reserve, which is 100 km from McKnight Prairie. We estimate  $v_w = 4.23 \text{ m/s}$ , which is the maximum hourly speed over September 17–23 (a seed-release period used by Sullivan et al. (2018)) averaged over all 19 years with complete wind data. Equation (50) then yields  $\mu = 16.7 \text{ m}$ .

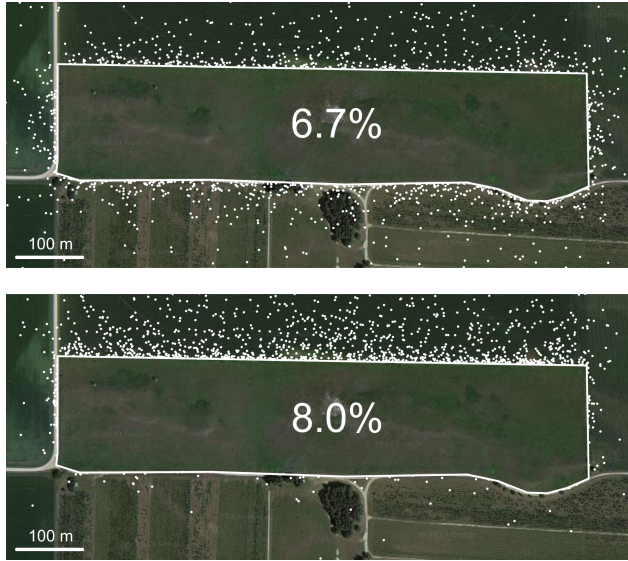
With these parameters ( $A = 137,000 \text{ m}^2$ ,  $L = 1,930 \text{ m}$ ,  $\mu = 16.7 \text{ m}$ ), Theorem 7 gives the following upper bounds on seed loss probability:

$$\begin{array}{ll} p \leq \frac{\mu L}{\pi A} = 7.5\% & p \leq \frac{\mu L}{2A} = 11.8\% \\ \text{(Symmetric)} & \text{(Asymmetric).} \end{array}$$

These relatively low bounds are a result of the habitat’s large size, on the order of 100’s of meters, which gives a characteristic length scale  $A/L$  that is much greater than  $\mu$ . Even in the asymmetric case,  $p \leq 11.8\%$ , meaning McKnight Prairie will retain almost 90% of *A. syriaca*’s seeds. While additional factors such as mortality rates within the patch may interact with seed loss to determine the ultimate outcome, the species’ high seed retention is encouraging.

## 5.2 Simulation of seed loss

In addition to using Theorem 7 to derive upper bounds on seed loss from *A. syriaca* at McKnight prairie, we computed  $p$  using a numerical simulation consistent with the probabilistic definition of  $p$  in equation 9. We implemented a simple Monte Carlo style simulation in which a plant is placed uniformly at random within the habitat and a seed disperses from there, landing either inside or outside the habitat. This process was repeated 10 million times, and the fraction of lost seeds was reported as  $p$ . We ran the simulation twice, once for symmetric dispersal and once for asymmetric dispersal. In both cases,

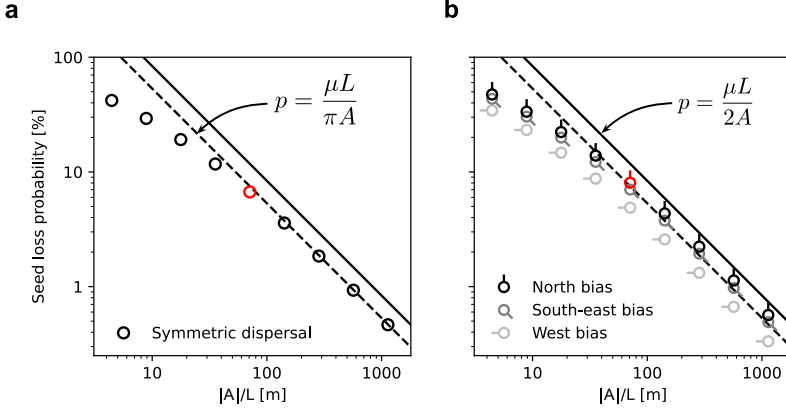


**Fig. 6** A random sample of several thousand *A. syriaca* seeds lost from McNight Prairie (outlined in white) to the surrounding agricultural fields. More seeds are lost when dispersal is biased northward (bottom panel) than when dispersal is symmetric (top panel) even though mean dispersal distance is the same in both cases. Aerial image: Google, ©2023 Airbus, CNES/Airbus, Maxar Technologies, USDA/FPAC/GEO

dispersal distance was sampled from an inverse Gaussian (Wald) distribution with mean  $\mu = 16.7$  m (as computed above) and shape parameter  $\lambda = 2.5$  m (as reported in [Sullivan et al. \(2018\)](#)). In the symmetric case, the direction of the seed dispersal vector was chosen uniformly at random, but in the asymmetric case, a northern bias was introduced. To simulate a hypothetical prevailing northern wind, the direction,  $\theta$ , measured relative to due north, was sampled according to the probability density function  $(1 + 10\theta^2)^{-1}$  (normalized so that its integral from  $-\pi$  to  $\pi$  is 1). The results of the simulations were  $p = 6.7\%$  in the symmetric case and  $p = 8.0\%$  in the asymmetric case. As expected, these values are less than the theoretical bounds of 7.5% and 11.8%.

### 5.3 Comparison of seed loss bounds to simulations

To illustrate a wider range of seed loss scenarios, we also simulated the effect of increasing or decreasing the habitat's area to perimeter ratio  $A/L$ , which is inversely proportional to the bounds on  $p$ . To change of  $A/L$ , we scaled the habitat size up or down uniformly. (Equivalently, we could have held the habitat size constant and scaled the mean dispersal distance  $\mu$  down and up.) As shown in Figure 7a, the simulated results trend closely with the theoretical bounds over a wide range of  $A/L$  values. For larger  $A/L$  values, the simulated seed loss probability for symmetric dispersal in particular shows remarkable agreement with the predicted bound. In these cases where the habitat is significantly larger than the length-scale of dispersal,  $\mu L/\pi A$  becomes not only a bound but also a close estimate for  $p$ .



**Fig. 7** Impact of area to perimeter ratio  $A/L$ , on simulated seed loss probability  $p$ . Red symbols represent the two examples shown above, while other symbols represent hypothetical scenarios where McKnight Prairie is uniformly scaled up or down by a factor of 2, 4, 8, or 16, thus adjusting  $A/L$ . As the habitat grows larger, the symmetric bound (dashed line) becomes sharp, whereas the asymmetric bound (solid line) remains an over-estimate, especially for the alternative bias directions shown in gray

We suspect that this general assertion about the sharpness of the symmetric bound could be proved, if formalized correctly—perhaps by generalizing the argument presented in Appendix A2 that  $p$  approaches  $\mu L/\pi A$  in a specific example.

In the case of asymmetric dispersal (Figure 7b), the bound  $\mu L/2A$  remains an overestimate of simulated seed loss even as  $A/L$  increases. The degree of overestimate depends on the bias direction of dispersal—for example, simulated seed loss is lower when dispersal is biased west, because fewer seeds are lost across the habitat’s long north and south edges. It is interesting to note that although  $\mu L/\pi A$  no longer serves as an upper bound in the asymmetric case, it does appear to capture the average seed loss among different dispersal directions as  $A/L$  becomes large.

## 6 Discussion

Whereas numerical simulation of  $p$  requires detailed knowledge and coding of habitat geometry and plant dispersal kernels, the bounds of Theorem 7 require only mean dispersal distance  $\mu$ , habitat perimeter  $L$ , and habitat area  $A$ . We therefore anticipate that the bound will provide an accessible upper estimate and worst-case-scenario indicator of seed loss across a variety of applications. Because mean dispersal distance is the most challenging component to estimate in practice, we discuss methods for doing so in section 6.1 before outlining directions for future research in section 6.2.

## 6.1 Estimating mean dispersal distance

Estimating mean dispersal distance is perhaps most straightforward for wind-dispersed seeds, for which mechanistic models are well-developed. In particular, the “ballistic” model

$$\mu = \frac{v_w h}{v_T} \quad (51)$$

has long been used to estimate dispersal distance  $\mu$  in terms of seed release height  $h$ , terminal velocity  $v_T$ , and wind speed  $v_w$  (for a history, see [Nathan et al. \(2011\)](#)), and it remains the mean dispersal distance in the more sophisticated probabilistic WALD model that incorporates turbulent airflow ([Katul et al. 2005](#)). Both wind speed  $v_w$  and seed release height  $h$  can be obtained from field measurements, while terminal velocity  $v_T$  can be determined in a lab. Sullivan and colleagues ([2018](#)) describe measurement methods for these quantities and report values for 50 grassland species at Cedar Creek Ecosystem Science Reserve in Isanti County, Minnesota.

In the example in Section 5, estimating wind velocity was the most difficult part of using equation (51)—particularly matching wind data to the time and location of seed release. At the time of writing, the season of seed release was 5 months in the future and the height of wind sensors at the nearest weather stations was orders of magnitude higher than the height of seed release. To overcome such obstacles, one can use proxy measurements from similar sites (as we did) or wait for the season of seed release to take on-site field measurements.

For seeds dispersed by mechanisms other than wind, existing literature offers several starting points for estimating mean dispersal distance  $\mu$ . Rough estimates can be obtained from a meta-analysis of 148 mostly empirical sources conducted by Thomson and colleagues ([2011](#)). In the course of studying relationships with plant height and seed mass, the authors condensed data on the mean dispersal distances of over 200 species into summary statistics grouped by 8 dispersal mechanisms (e.g. unassisted, water, vertebrate), presented in Appendix S3. Because the report does not provide data by species or list the primary research articles it draws upon, additional effort is required to narrow these order-of-magnitude bounds.

First, one can consult existing literature for proxy estimates. For example, Bullock and colleagues published a synthesis report in [2017](#) that lists dispersal data sources by species for 144 vascular plant species in Table S4 of the online supplement. Second, novel analyses can be performed to estimate dispersal distances. Methods for estimating dispersal across a variety of mechanisms are reviewed by [Rogers et al. \(2019\)](#). By considering a combination of dispersal estimates, one can develop a range that seems plausible for the species and site under consideration.

## 6.2 Future directions

As noted in Section 5.3, a salient area that remains to be explored is the conditions under which the quantities  $\mu L/2A$  or  $\mu L/\pi A$  serve not only as upper bounds but also as close estimates for  $p$ . If, as conjectured, the bounds are tight when dispersal is symmetric and mean dispersal distance is short relative

to habitat size, then  $\mu L/\pi A$  could be used to estimate seed loss directly in such settings.

One variable to consider while exploring tightness of bounds is the connectedness of habitat patches. Although our illustrations and example have featured a single connected habitat patch, the bounds also apply to a habitat “patch” consisting of multiple nearby components that are disconnected by roadways or other land uses. (In this case, the total perimeter and area of all components are summed to determine  $L$  and  $A$ , respectively.) Because dividing a habitat patch into two components via a narrow roadway may increase total perimeter  $L$  without substantially altering area or seed loss, the bounds of Theorem 7 may significantly overestimate seed loss for disconnected habitats. (Of course, a roadway could also alter dispersal patterns by changing animal behavior, wind, or water flow in ways that violate the assumptions of our model, as discussed in the next paragraph.)

Further theoretical efforts could work towards seed loss bounds that avoid one or several of the simplifying assumptions that underpin the dispersal model (9). First, we assume that the dispersal location random variable  $\mathbf{X}_d$  and plant location random variable  $\mathbf{X}_p$  are independent. However, dispersal patterns might vary according to plant location based on the effects of topographic slope and habitat edges on wind, water, or animal behavior. For example, in experimental patches of longleaf pine savanna embedded in a pine plantation matrix, Damschen and colleagues (2014) found significant differences in wind velocities and propagule dispersal patterns at various locations within a patch. Patch shape influenced this phenomenon, with long corridors of savanna in particular promoting uplifting and long distance dispersal. Second, we assume that plant locations  $\mathbf{X}_p$  follow a uniform distribution within a habitat, but variations in local site conditions could bias the plant distributions. In keeping with the orientation of this work towards estimates that use minimal information inputs, it would be particularly interesting to find bounds that allow for the variations described above without requiring their full characterization. Lastly, we note that the binary classification of land as habitable or not habitable may be unrealistic for some ecosystems. Indeed, dispersal out of targeted reserves can have positive spillover effects on the surrounding landscape (Brudvig et al. 2009), pointing towards a more nuanced interpretation of seed loss from a habitat patch. Combining estimates of seed loss with estimates of establishment in the surrounding matrix (e.g. Craig et al. (2011)) could yield insights into the potential benefits of dispersal out of a reserve.

## Statements and Declarations

**Competing Interests.** The authors declare no competing interests.

**Data Availability.** Our estimate of *A. syriaca*’s mean dispersal distance (see equation (50)) is based on plant height and seed terminal velocity data measured by Sullivan et al. (2018), as well as hourly wind speed data from Seeley (2021). The satellite image of McKnight Prairie used in Figure 6 is from Google Maps (Accessed 21 March 2024).

## Appendix A Sharpness of bounds

### A.1 Asymmetric dispersal

In Theorem 7 we prove that, without any assumptions on the symmetry of the dispersal kernel, the proportion  $p$  of seeds lost from a habitat patch with area  $A$  and perimeter  $L$  is bounded by

$$p \leq \frac{\mu L}{2A}, \quad (\text{A1})$$

where  $\mu$  is the mean dispersal distance of the seeds. Rearranging, this is equivalent to

$$\frac{pA}{\mu L} \leq \frac{1}{2}.$$

Let  $k = \frac{pA}{\mu L}$ , so that the bound (A1) is equivalent to  $k \leq 1/2$ . The following proposition implies that this bound is the lowest possible under the assumptions given in Section 2.

**Proposition 8** There exist habitat geometries and dispersal kernels that make  $k \equiv \frac{pA}{\mu L}$  arbitrarily close to  $1/2$ .

*Proof* Let  $a < A$  and  $b < B$  be positive real numbers. Consider a rectangular habitat

$$H = \{(x, y) \in \mathbb{R}^2 \mid 0 < x < A \text{ and } 0 < y < B\},$$

and a dispersal kernel

$$f_d(x, y) = \begin{cases} \frac{1}{ab} & 0 < x < a \text{ and } 0 < y < b \\ 0 & \text{otherwise.} \end{cases}$$

We bound the factors in  $k$  in terms of  $a$ ,  $b$ ,  $A$ , and/or  $B$ . First, we have the equality

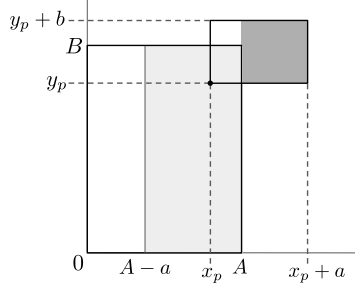
$$L = 2(A + B). \quad (\text{A2})$$

Second, from the definition (7) of mean dispersal distance we have

$$\begin{aligned} \mu &= \iint_{\mathbb{R}^2} |\mathbf{x}_d| f_d(\mathbf{x}_d) dA \\ &= \int_0^b \int_0^a \sqrt{x^2 + y^2} \frac{1}{ab} dx dy \\ &\leq \int_0^b \int_0^a (x + y) \frac{1}{ab} dx dy \\ &= \frac{a + b}{2}. \end{aligned} \quad (\text{A3})$$

From (A2) and (A3) it follows that

$$k \geq \frac{pA}{\left(\frac{a+b}{2}\right) \cdot 2(A+B)} = \frac{pA}{(a+b)(A+B)}. \quad (\text{A4})$$



**Fig. 8** Quantities used to bound the integral (A6). The light and dark grey rectangles represent the domains of integration for the outer and inner double integrals, respectively

Next we bound  $pA$ , which from Proposition 1 is equivalent to the integral

$$pA = \iint_{\mathbf{x}_p \in H} \iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) dA d\mathbf{x}_p. \quad (\text{A5})$$

By restricting the domains of integration in equation (A5), we can only decrease the integral of the non-negative quantity  $f_d$ . Let  $\mathbf{x}_p = (x_p, y_p)$ . Choosing bounds for  $\mathbf{x}$  and  $\mathbf{x}_p$  over which  $f_d(\mathbf{x} - \mathbf{x}_p) = 1/ab$  (see Figure 8), we obtain

$$pA \geq \int_0^B \int_{A-a}^A \left( \int_{y_p}^{y_p+b} \int_A^{x_p+a} \frac{1}{ab} dx dy \right) dx_p dy_p \quad (\text{A6})$$

$$= \int_0^B \int_{A-a}^A \left( \frac{x_p + a - A}{a} \right) dx_p dy_p \quad (\text{A7})$$

$$= \frac{B}{a} \int_{A-a}^A (x_p + a - A) dx_p, \quad (\text{A8})$$

which simplifies to

$$pA \geq \frac{aB}{2}. \quad (\text{A9})$$

Combining inequalities (A4) and (A9) yields

$$k \geq \frac{aB}{2(a+b)(A+B)}.$$

For habitats with  $B \gg A$  the ratio  $\frac{B}{A+B}$  can be made arbitrarily close to 1. Similarly, for dispersal kernels with  $a \gg b$ , the ratio  $\frac{a}{a+b}$  can also be made arbitrarily close to 1. It follows that  $k$  can be made arbitrarily close to  $\frac{1}{2}$ .  $\square$

Thus, no values below  $1/2$  suffice to bound  $k$  from above. It is interesting to note that the ultimate quantities compared in this argument concern the aspect ratio of the rectangular habitat  $H$  and the rectangular support for the dispersal kernel  $f_d$ . In particular, we imagine the aspect ratios becoming extreme in orthogonal directions.

## A.2 Symmetric dispersal

Following the approach in Appendix A.1, we define

$$k = \frac{pA}{\mu L} \quad (\text{A10})$$

so that the existence of habitat geometries and dispersal kernels that make  $k$  arbitrarily close to  $1/\pi$  implies that  $p$  can be arbitrarily close to its bound  $\frac{\mu L}{\pi A}$  in Theorem 7 when dispersal is rotationally symmetric.

**Proposition 9** There exist habitat geometries and rotationally symmetric dispersal kernels that make  $k \equiv \frac{pA}{\mu L}$  arbitrarily close to  $1/\pi$ .

*Proof* Define a disk habitat  $H$  of radius  $R_H$ ; that is,

$$H = \{\mathbf{x} : |\mathbf{x}| < R_H\},$$

and take the dispersal kernel  $f_d$  to be a uniform distribution over a disk support with radius  $R_d$ :

$$f_d(\mathbf{x}) = \begin{cases} \frac{1}{\pi R_d^2} & \text{if } |\mathbf{x}| < R_d \\ 0 & \text{otherwise.} \end{cases}$$

Simple geometry gives that

$$L = 2\pi R_H. \quad (\text{A11})$$

And by definition we have

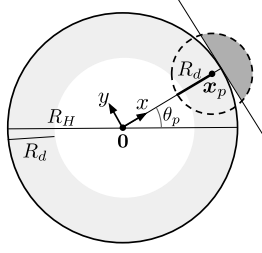
$$\begin{aligned} \mu &= \iint_{\mathbb{R}^2} |\mathbf{x}| f_d(\mathbf{x}) dA \\ &= \int_0^{2\pi} \int_0^{R_d} r \frac{1}{\pi R_d^2} r dr d\theta \\ &= \frac{2}{3} R_d. \end{aligned} \quad (\text{A12})$$

Substituting for  $pA$ ,  $L$ , and  $\mu$  in equation (A10) using Proposition 1 and equations (A11) and (A12) yields

$$k = \frac{3}{4\pi R_d R_H} \iint_{\mathbf{x}_p \in H} \iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) dA dA.$$

Because the only plant positions  $\mathbf{x}_p$  that contribute to seed loss are those at distance between  $R_H - R_d$  and  $R_H$  from the origin, we express the outer double integral in polar coordinates as

$$k = \frac{3}{4\pi R_d R_H} \int_0^{2\pi} \int_{R_H - R_d}^{R_H} \left( \iint_{\mathbf{x} \in H^c} f_d(\mathbf{x} - \mathbf{x}_p) dA \right) r_p dr_p d\theta \quad (\text{A13})$$



**Fig. 9** Quantities and coordinates used in the proof of Proposition 9

where  $\mathbf{x}_p = (r_p, \theta_p)$ . We use Cartesian coordinates  $\mathbf{x} = (x, y)$  for the inner double integral, with the  $x$ -axis aligned with the vector  $\mathbf{x}_p$  (see Figure 9). To simplify computations we restrict the domain of the inner integral to the segment of the circle of radius  $R_d$  centered on  $\mathbf{x}_p$  that is cut by the chord tangent to the habitat boundary at the  $x$ -axis. Then

$$k \geq \frac{3}{4\pi R_d R_H} \int_0^{2\pi} \int_{R_H - R_d}^{R_H} \int_{R_H}^{r_p + R_d} \int_{-\sqrt{R_d^2 - (x - r_p)^2}}^{\sqrt{R_d^2 - (x - r_p)^2}} \frac{1}{\pi R_d^2} dy dx r_p dr_p d\theta. \quad (\text{A14})$$

The  $\theta$  and  $y$  integrals can be evaluated directly, yielding

$$k \geq \frac{3}{4\pi R_d R_H} 2\pi \int_{R_H - R_d}^{R_H} \int_{R_H}^{r_p + R_d} 2\sqrt{R_d^2 - (x - r_p)^2} \frac{1}{\pi R_d^2} dx r_p dr_p \quad (\text{A15})$$

$$= \frac{3}{\pi R_d^3 R_H} \int_{R_H - R_d}^{R_H} \int_{R_H}^{r_p + R_d} \sqrt{R_d^2 - (x - r_p)^2} dx r_p dr_p. \quad (\text{A16})$$

Because  $r_p > R_H - R_d$  over the domain of integration, one can eliminate the factor of  $r_p$  from the integrand as follows:

$$k \geq \frac{3(R_H - R_d)}{\pi R_d^3 R_H} \int_{R_H - R_d}^{R_H} \int_{R_H}^{r_p + R_d} \sqrt{R_d^2 - (x - r_p)^2} dx dr_p. \quad (\text{A17})$$

We first make the substitution  $u = x - r_p$  and then reverse the order of integration:

$$k \geq \frac{3(R_H - R_d)}{\pi R_d^3 R_H} \int_{R_H - R_d}^{R_H} \int_{R_H - r_p}^{R_d} \sqrt{R_d^2 - u^2} du dr_p \quad (\text{A18})$$

$$= \frac{3(R_H - R_d)}{\pi R_d^3 R_H} \int_0^{R_d} \int_{R_H - u}^{R_H} \sqrt{R_d^2 - u^2} dr_p du \quad (\text{A19})$$

$$= \frac{3(R_H - R_d)}{\pi R_d^3 R_H} \int_0^{R_d} u \sqrt{R_d^2 - u^2} du. \quad (\text{A20})$$

Evaluating the remaining integral yields  $R_d^3/3$ , and so

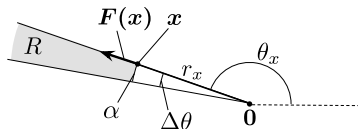
$$k \geq \frac{R_H - R_d}{\pi R_H}. \quad (\text{A21})$$

By taking  $R_H \gg R_d$ ,  $k$  can be made arbitrarily close to  $1/\pi$ , as desired.  $\square$

## Appendix B Discussion of $\mathbf{F}$

For the purpose of proving Theorem 7, the vector field  $\mathbf{F}$  was a convenient mathematical stepping stone tied to the idea of seed flow to build intuition.

But a more precise physical interpretation is worth noting. Specifically, *the flux of  $\mathbf{F}$  across a curve can be interpreted as the probability that a given seed would pass through that curve if all seeds traveled in straight lines*. Here, we show that this physical description of  $\mathbf{F}$  leads logically to its mathematical definition in equation (13).



**Fig. 10** Notation used in the construction of the vector field  $\mathbf{F}(\mathbf{x})$

Consider a single plant at the origin with dispersal kernel  $f_d$ . Given a location  $\mathbf{x}$  in the landscape with polar coordinates  $(r_x, \theta_x)$ , we define an arc  $\alpha$  starting at  $\mathbf{x}$  and subtending an angle  $\Delta\theta$ . According to our interpretation above, the flux of  $\mathbf{F}$  across  $\alpha$  must equal the probability that a seed passes through  $\alpha$ . Since we are working under the hypothetical assumption that seeds travel in straight lines, the seeds that pass through  $\alpha$  are exactly those that land in the infinite polar rectangle  $R = \{(r, \theta) \mid r \geq r_x, \theta_x \leq \theta \leq \theta_x + \Delta\theta\}$  shown in Figure 10. Thus,

$$\text{flux of } \mathbf{F} \text{ across } \alpha = P(\mathbf{X}_d \in R) = \int_{\theta_x}^{\theta_x + \Delta\theta} \int_{r_x}^{\infty} f_d(r, \theta) r \, dr \, d\theta. \quad (\text{B22})$$

To be consistent with the straight-line dispersal hypothetical, we choose  $\mathbf{F}$  to point radially outward from the origin. Then since  $\mathbf{F}$  is normal to  $\alpha$ , its flux across  $\alpha$  can also be expressed as

$$\text{flux of } \mathbf{F} \text{ across } \alpha = \int_{\theta_x}^{\theta_x + \Delta\theta} |\mathbf{F}(r_x, \theta)| r_x \, d\theta, \quad (\text{B23})$$

where  $r_x \, d\theta$  serves as the differential arc length  $ds$ . After combining equations (B22) and (B23), we may shed the  $\theta$  integrals by dividing by  $\Delta\theta$  and then taking the limit as  $\Delta\theta$  approaches zero, leaving

$$|\mathbf{F}(\mathbf{x})| r_x = \int_{r_x}^{\infty} f_d(r, \theta_x) r \, dr. \quad (\text{B24})$$

Solving for  $|\mathbf{F}(\mathbf{x})|$  and then multiplying by the unit vector  $\mathbf{x}/r_x$  to recover  $\mathbf{F}$ 's direction gives our definition:

$$\mathbf{F}(\mathbf{x}) = \frac{\mathbf{x}}{r_x^2} \int_{r_x}^{\infty} f_d(r, \theta_x) r \, dr.$$

## References

- Beckman, N. G. & Sullivan, L. L. (2023), 'The causes and consequences of seed dispersal', *Annual Review of Ecology, Evolution, and Systematics* **54**, 403–427.
- Brudvig, L. A., Damschen, E. I., Tewksbury, J. J., Haddad, N. M. & Levey, D. J. (2009), 'Landscape connectivity promotes plant biodiversity spillover into non-target habitats', *Proceedings of the National Academy of Sciences* **106**(23), 9328–9332.
- Bullock, J. M., Mallada González, L., Tamme, R., Götzenberger, L., White, S. M., Pärtel, M. & Hooftman, D. A. (2017), 'A synthesis of empirical plant dispersal kernels', *Journal of Ecology* **105**(1), 6–19.
- Cheptou, P.-O., Carrue, O., Rouifed, S. & Cantarel, A. (2008), 'Rapid evolution of seed dispersal in an urban environment in the weed *crepis sancta*', *Proceedings of the National Academy of Sciences* **105**(10), 3796–3799.
- Cooney, M., Hafner, B., Johnson, S. & Lee, S. (2022), 'The effect of habitat fragmentation on plant communities in a spatially implicit grassland model'. In revision at Rose-Hulman Undergraduate Mathematics Journal.
- Craig, M. T., Orrock, J. L. & Brudvig, L. A. (2011), 'Edge-mediated patterns of seed removal in experimentally connected and fragmented landscapes', *Landscape Ecology* **26**, 1373–1381.
- Damschen, E. I., Baker, D. V., Bohrer, G., Nathan, R., Orrock, J. L., Turner, J. R., Brudvig, L. A., Haddad, N. M., Levey, D. J. & Tewksbury, J. J. (2014), 'How fragmentation and corridors affect wind dynamics and seed dispersal in open habitats', *Proceedings of the National Academy of Sciences* **111**(9), 3484–3489.
- Google Maps (Accessed 21 March 2024), 'McKnight Prairie in Goodhue County, MN [Satellite Map]'. <https://maps.app.goo.gl/TujkZDrtZUiHx32JA>.
- Hanski, I. (2001), Population dynamic consequences of dispersal in local populations and in metapopulations, in J. Clobert, E. Danchin, A. Dhondt & J. Nichols, eds, 'Dispersal', Oxford University Press, Oxford, UK, pp. 283–298.
- Katul, G., Porporato, A., Nathan, R., Siqueira, M., Soons, M., Poggi, D., Horn, H. & Levin, S. (2005), 'Mechanistic analytical models for long-distance seed dispersal by wind', *The American Naturalist* **166**(3), 368–381.
- Klausmeier, C. A. (1998), 'Extinction in multispecies and spatially explicit models of habitat destruction', *The American Naturalist* **152**(2), 303–310.
- Levin, S. A., Muller-Landau, H. C., Nathan, R. & Chave, J. (2003), 'The ecology and evolution of seed dispersal: a theoretical perspective', *Annual Review of Ecology, Evolution, and Systematics* **34**(1), 575–604.
- Nathan, R., Katul, G. G., Bohrer, G., Kuparinen, A., Soons, M. B., Thompson, S. E., Trakhtenbrot, A. & Horn, H. S. (2011), 'Mechanistic models of seed dispersal by wind', *Theoretical Ecology* **4**(2), 113–132.
- Nathan, R., Klein, E., Robledo-Arnuncio, J. J. & Revilla, E. (2012), Dispersal kernels: review, in J. Clobert, M. Baguette, T. G. Benton & J. M. Bullock, eds, 'Dispersal Ecology and Evolution', Oxford University Press Oxford, UK, pp. 187–210.

- Rogers, H. S., Beckman, N. G., Hartig, F., Johnson, J. S., Pufal, G., Shea, K., Zurell, D., Bullock, J. M., Cantrell, R. S., Loiselle, B. et al. (2019), 'The total dispersal kernel: a review and future directions', *AoB Plants* **11**(5), plz042.
- Seeley, M. (2021), 'Hourly climate data: Meterologic measurements at Cedar Creek natural history area ver 10.'. <https://doi.org/10.6073/pasta/ce85da01f57b7639a06ab0dde4447b4c>.
- Sullivan, L. L., Clark, A. T., Tilman, D. & Shaw, A. K. (2018), 'Mechanistically derived dispersal kernels explain species-level patterns of recruitment and succession', *Ecology* (11), 2415–2420.
- Thomson, F. J., Moles, A. T., Auld, T. D. & Kingsford, R. T. (2011), 'Seed dispersal distance is more strongly correlated with plant height than with seed mass', *Journal of Ecology* **99**(6), 1299–1307. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2745.2011.01867.x>.
- Tilman, D., May, R. M., Lehman, C. L. & Nowak, M. A. (1994), 'Habitat destruction and the extinction debt', *Nature* **371**(6492), 65–66.