Chapter 2

Optimal movement between patches under incomplete information about the spatial distribution of food items

Raymond H.G. Klaassen, Bart A. Nolet, Jan. A van Gils & Silke Bauer
Theoretical Population Biology, *in press*

*Feeding in a one-dimensional lineage. Oscar Langevoord, Bart Nolet and Thijs de Boer (from left to right). In this example each patch (soup cup) contains food.*
Abstract
If the food distribution contains spatial pattern, the food density in a particular patch provides a forager with information about nearby patches. Foragers might use this information to exploit patchily distributed resources profitably. We model the decision on how far to move to the next patch in linear environments with different spatial patterns in the food distribution (clumped, random, and regular) for foragers that differ in their degree of information. An ignorant forager is uninformed and therefore always moves to the nearest patch (be it empty or filled). In contrast, a prescient forager is fully informed and only exploits filled patches, skipping all empty patches. A Bayesian assessor has prior knowledge about the content of patches (i.e. it knows the characteristics of the spatial pattern) and may skip neighbouring patches accordingly by moving to the patch where the highest gain rate is expected. In most clumped and regular distributions there is a benefit of assessment, i.e. Bayesian assessor achieve substantially higher long-term gain rates than ignorant foragers. However, this is not the case in distributions with less strong spatial pattern, despite the fact that there is a large potential benefit from a sophisticated movement rule (i.e. a large penalty of ignorance). Bayesian assessor do also not achieve substantially higher gain rates in environments that are relatively rich or poor in food. These results underline that an incompletely informed forager that is sensitive to spatial pattern should not always respond to existing pattern. Furthermore, we show that an assessing forager can enhance its long-term gain rate in highly clumped and some specific near-regular food distributions, by sampling the environment in slightly larger spatial units.
Introduction
Food items tend to be heterogeneously distributed throughout the environment (Sparrow, 1999; Wiens, 2000). This patchiness in resource abundance allows a forager to concentrate foraging effort in areas which are, on average, richer in resources (Pyke, 1984; Valone and Brown, 1989). Optimal foraging theory investigates the problem of how to exploit patchily distributed food in order to maximize long-term energy gain rate (MacArthur and Pianka, 1966; Schoener, 1971; Stephens and Krebs, 1986). This is not a trivial problem to solve, especially since a forager is usually incompletely informed about the content of patches prior to exploitation (Clark and Mangel, 1984; Valone and Brown, 1989; Valone, 1991). Animals can optimize resource use by modulating patch use and/or inter-patch movement (Walsh, 1996). In this study we are interested in the optimal solution to move between patches while maximizing long-term gain rate, for various amounts of information about the spatial distribution of food.

When a forager is omniscient about the content of all patches (full or empty), optimal movement becomes a “travelling salesman problem”: what is the shortest path that leads along full patches only (Anderson, 1983)? A simple near-optimal solution of this problem is to always move to the nearest full patch (Gross et al., 1995). When a forager is totally uninformed about the content of patches, the optimal movement policy is to minimize movement costs by always travelling to the nearest unexploited patch. In many natural situations it is not likely that a forager is omniscient because food items are hidden or cryptic (Oaten, 1977). However, it is also unlikely that a forager is totally uninformed about surrounding patches because normally food items are not randomly distributed, but a spatial pattern exists (Li and Reynolds, 1995; Wiens, 2000). In a spatially patterned environment, information about the content of a given patch provides information about the content of nearby patches (Mangel and Adler, 1994; Fortin, 2003). Incompletely informed foragers that know the spatial characteristics of the food distribution (e.g. the degree of aggregation of prey), but not the exact position of full and empty patches before exploitation, should be sensitive to spatial pattern. The optimal movement pattern for such incompletely informed forager is more complex.

Models on animal movement pattern show that a forager should achieve a higher long-term gain rate if movement is related to the forager’s recent foraging experience (Knoppien and Reddingius, 1985; Benhamou, 1992). In these models foragers move in an intensive searching mode (also referred to as an area-restricted search) or in an extensive searching mode. Intensive search is triggered by an encounter with a food item, whether extensive search is triggered by a period without any encounter with food. In spatially one-dimensional models (Knoppien and Reddingius, 1985; Krakauer and Rodríguez-Gironés, 1995), intensive search is represented by a slow, and extensive search by a fast movement, assuming a trade-off between movement velocity and the probability to detect a food item (Gendron and Staddon, 1983). In two-dimensional models also the sinuosity of a path (i.e.
the degree of turning) is different for intensive and extensive search (high and low, respectively) (Benhamou, 1992; Fortin, 2002; Hill et al, 2003; Nolet and Mooij, 2002). The fundamental difference between these models is that in one-dimensional models foragers encounter high and low food densities in proportion to their availability, whereas in two-dimensional models strategically moving foragers encounter high densities more often, and low densities less often, than the proportions at which these densities are available (Fortin, 2002).

In this study we investigate the benefit of (pre-harvest or prior) information about spatial pattern in environments with different spatial configurations of food, ranging from highly clumped to random and regular ones. For each configuration we establish the optimal movement rule (i.e. the rule that maximizes long-term gain rate), for foragers that differ in the level of information about the content of all patches (uninformed, pre-harvest information on spatial pattern, and fully informed). Subsequently, we compare the best performance of the differently informed foragers for each configuration, to evaluate in which food distributions prior knowledge about spatial pattern is beneficial to a forager, or in other words, in which food distributions there is a benefit of assessment. In order to generalize these results we consider foragers with different foraging scales.

We model the decision on how far to move between patches, in linear environments with different spatial patterns in the distribution of full and empty patches. Movement consists of steps with certain lengths, in which the length of a step depends on the forager’s recent foraging experience. The shortest possible movement is to an adjacent patch. Longer movements, to more distant patches, pass over (skip) one or several patches. Consequently, the frequency in which a forager encounters full patches can be higher than the frequency in which full patches occur throughout the environment, in contrast to other one-dimensional models (Knoppien and Reddingius, 1985; Krakauer and Rodríguez-Gironés, 1995) and in agreement with two-dimensional models (Benhamou, 1992; Fortin, 2002). Our model is not parameterized on a certain study system but represents the situation that a forager moves between distant patches without sampling the patches in between. Examples include a Grey plover Pluvialis squatarola on a mudflat hunting for invertebrate prey (Pienkowski, 1983), a bumblebee Bombus sp. nectar feeding in a field of flowers (Pyke, 1974; Heinrich, 1979) and a Tundra swan Cygnus columbianus bewickii searching for pondweed tubers in a shallow lake (Nolet and Mooij, 2002).

The model

Characteristics of the environment

We consider a one-dimensional environment consisting of cells arrayed in a linear sequence that are either full (containing 1 food item) or empty (containing 0 food items). Food items are distributed according to a first order Markov process: the probability that the adjacent cell $x+1$ contains a food item ($n_{x+1}=1$) depends on the content $n_x$ of the current cell $x$ only
inter-patch distances in a patterned world

(Krakauer and Rodríguez-Gironés, 1995). Consequently the distribution can be characterized by two transition probabilities: probability \( R(0|1) \) that the next cell is empty given that the current cell is full and probability \( R(1|0) \) that the next cell is full given that the current cell is empty:

\[
P(0 \mid 1) = P(n_{x+1} = 0 \mid n_x = 1) \quad \text{and} \quad P(1 \mid 0) = P(n_{x+1} = 1 \mid n_x = 0)
\]

(1)

Accordingly, the probabilities that the content of the current cell equals the content of the neighbouring cell are \( R(1|1) \) for full cells and \( R(0|0) \) for empty cells. Therefore:

\[
P(0 \mid 1) + P(1 \mid 1) = 1 \quad \text{and} \quad P(1 \mid 0) + P(0 \mid 0) = 1
\]

(2)

Note that \( R(0|1) \) and \( R(1|1) \) are independent of \( R(1|0) \) and \( R(0|0) \) (and vice versa). Food distributions differ in the proportion of filled \( f_1 \) and empty cells \( f_0 \):

\[
f_1 = \frac{P(1 \mid 0)}{P(1 \mid 0) + P(0 \mid 1)} \quad \text{and} \quad f_0 = 1 - f_1 = \frac{P(0 \mid 1)}{P(1 \mid 0) + P(0 \mid 1)}
\]

(3)

Furthermore, food distributions differ in their spatial pattern. A food distribution is clumped if both \( R(0|1) \) and \( R(1|0) < 0.5 \), because series of adjacent filled cells (food clumps) alternate with series of adjacent empty cells (gaps). The average size of a food clump equals \( 1/R(0|1) \) and the average size of a gap equals \( 1/R(1|0) \). Thus, small values of \( R(0|1) \) and \( R(1|0) \) result in food distributions with large food clumps and large gaps, respectively (Fig. 1a). Slightly larger values for \( R(0|1) \) and \( R(1|0) \) (but still < 0.5) corresponds to a food distribution with smaller food clumps and smaller gaps (Fig. 1b). In case both \( R(0|1) \) and \( R(1|0) > 0.5 \), filled and empty cells tend to alternate and the distribution is regular (Fig. 1c). For distributions with \( R(0|1)=R(0|0) \) the food distribution is random and contains no spatial pattern (Fig. 1d). A useful tool to describe spatial pattern is a structure function, which gives the probability that a spatial point some distance away from the current point contains resources (or not), given the resource state of the current point (Mangel and Adler, 1994). The probability that a future cell at distance \( i \) is full, given that the current cell is full, is (cf. Mangel and Adler, 1994):

\[
P(n_{x+i} = 1 \mid n_x = 1) = f_1 + (1 - f_1)\left( P(1 \mid 1) - P(1 \mid 0) \right)
\]

(4a)

The probability that a cell at distance \( i \) is empty given that the current cell is full, is:

\[
P(n_{x+i} = 0 \mid n_x = 1) = 1 - P(n_{x+i} = 1 \mid n_x = 1)
\]

(4b)

Similarly we can calculate the probabilities that a cell at distance \( i \) is empty or full, given that the current cell is empty:

\[
P(n_{x+i} = 0 \mid n_x = 0) = f_0 + (1 - f_0)\left( P(0 \mid 0) - P(0 \mid 1) \right)
\]

(5a)

and

\[
P(n_{x+i} = 1 \mid n_x = 0) = 1 - P(n_{x+i} = 0 \mid n_x = 0)
\]

(5b)

Structure functions for the distributions depicted in Fig. 1 are presented in Fig. 2a-d.
Figure 1 Examples of one-dimensional distributions, consisting of 100 cells that contain one food item (grey) or are empty (white). Cells are filled according to a first order Markov process; the content of a cell depends on the content of the (left) neighbouring cell. Food distributions differ in their spatial pattern, as transition probabilities $P(0|1)$ and $P(1|0)$ are varied. For each type of food distribution three (independent) examples are shown. (a) highly clumped ($P(0|1)=0.1$, $P(1|0)=0.05$), (b) slightly clumped ($P(0|1)=0.5$, $P(1|0)=0.25$), (c) regular ($P(0|1)=0.9$, $P(1|0)=0.45$), (d) random food distribution ($P(0|1)=P(1|0)=0.5$). In the first three distributions $\frac{1}{3}$ of cells is filled whereas the random distribution $\frac{1}{2}$ of cells is filled.

Characteristics of the forager

A forager either harvests food items in patches or moves between patches. Movement between patches is directional, which excludes re-visitation of patches. The size of a patch is equal to the foraging scale, which consists of one or multiple cells. First we consider foragers with a foraging scale of one cell, whereas in a later stage we consider foragers with larger foraging scales (multiple cells). Food items are cryptic to non-prescient foragers. We assume that a food item is always encountered when a forager is exploiting a patch and never when it skips a patch. The patch-residence time while exploiting is the same for filled and empty patches, since we assume that the time to handle a food item is negligibly small. The time $t_f$ to exploit one cell is typically longer than the time $t_m$ to travel a distance of one cell (Arditi and Dacorogna, 1988). Under this restriction, we chose arbitrary values for $t_f$ and $t_m$ (2 and 1 time unit per cell, respectively).
We evaluate the performance of three types of foragers that have different levels of information (not-, incompletely- and fully-informed) about the position of full and empty patches throughout the environment. We assume that the energy costs per time unit for foraging equal those for travelling. Consequently, ignoring energetic costs, we can evaluate different movement strategies by considering gross gain rate estimates only (Kacelnik and Houston, 1984; Stephens and Krebs, 1986).

Figure 2 Left panels (a-d) depict structure functions ($P(b|a)$: the probability that a patch contains $b$ food items, given that the current patch contains $a$ food items). Middle and right panels depict the expected gross gain (e-h) and the expected gain rate (i-l), respectively, in patches at distance $i$ from the current patch, given that the current patch is full (filled circles) or empty (open circles). Different panels refer to distributions with different spatial patterns: highly clumped (a,e,i), slightly clumped (b,f,j), regular (c,g,k) and random (d,h,l) (see Fig. 1 for parameter values). Note that series completely overlap for the random distribution.
Long-term gain rates of differently informed foragers

Ignorant forager
An ignorant forager is uninformed about the content of patches prior to exploiting them. In order to maximize gain rate an ignorant forager should minimize time costs by always moving to the adjacent patch \(x+1\), irrespective of the food density in the current patch (Fig. 3d). Ignorant foragers visit all patches (they never skip a patch) and thus encounter full and empty patches in proportion to their availability (Fig. 3h). Hence, we can calculate their long-term gain rate \(G_{\text{ignorant}}\) from the frequency in which filled patches occur in the environment and the total time it takes to exploit and to move between patches:

\[
G_{\text{ignorant}} = \frac{f_1}{f_0 \cdot (t_f + t_m) + f_1 \cdot (t_f + t_m)}
\]

(6a)

and because \(f_0 + f_1 = 1\) this can also be expressed as:

\[
G_{\text{ignorant}} = \frac{f_1}{t_f + t_m}
\]

(6b)

Prescient forager
A prescient forager instantly knows the content of a patch upon arrival in a patch (Valone and Brown, 1989). If the patch contains a food item, a prescient forager will exploit it but if the patch is empty the prescient forager moves on. Consequently, prescient foragers only exploit full patches. The length of a movement between two patches depends on the location of the nearest full patch. The long-term gain rate of a prescient forager is calculated by dividing the frequency in which full patches occur in the environment by the time it takes to exploit these full patches and to move between patches:

\[
G_{\text{prescient}} = \frac{f_1}{f_1 \cdot (t_f + t_m) + f_0 \cdot t_m}
\]

(7)

Bayesian assessing forager
The Bayesian assessing forager knows the parameters of the structure functions of the environment (for example learned during earlier visits, i.e. pre-harvest information). This knowledge is combined with information about the current patch (sample information; Valone, 1991) to derive an estimate of the quality of nearby patches (hence the term Bayesian). The assessor is thus incompletely informed about the content of (nearby) patches.

For every patch at distance \(i\) from the current patch, we calculate the expected number of food items (or expected gross gain \(\hat{n}\)), given the number of food items \(a\) in the current patch (\(a=0\) or 1):
\[ \hat{n}_{x+i} \mid (n_x = a) = \sum_{n_{x+i} = 0}^{1} n_{x+i} \cdot P(n_{x+i} \mid n_x = a) \quad (8) \]

where the probability that patch \( x+i \) contains a food item, given that the current patch contains \( a \) food items, can be derived from the structure functions. For clumped food distributions the expected gain decreases with distance if the current patch is full and increases with distance if the current patch is empty (Fig. 2e,f). For regular food distributions the expected gain fluctuates around the gain expected at an infinitely large distance (Fig. 2g). The expected gain is constant for all distances \( i \) for random distributions (Fig. 2h).

**Figure 3** Left panels (a-d): Optimal movement rules of assessing and ignorant foragers, represented as the optimal length of the movement to the next patch if the current patch contains 0 or 1 food items. Right panels (e-h): Relative frequency in which empty and full patches are encountered by foragers that move optimally. Different panels refer to distributions with different spatial patterns: highly clumped (a,e), slightly clumped (b,f), regular (c,g) and random (d,h) (see Fig. 1 for parameter values). Note that the movement rules for assessing and ignorant foragers are equal for the slightly clumped and the random food distribution (dots overlap), and that in these cases both the assessing and ignorant forager encounter full patches at the frequency as available.
The expected short-term gain rate $\hat{g}$ is calculated by dividing the expected gain $\hat{n}$ by the total expected time $\hat{t}_i$, which comprises the time to move to a patch at distance $i$ (equal to $i t_m$) and the time to exploit a patch ($t_f$). Total expected time equals:

$$\hat{t}_i = i \cdot t_m + t_f \quad (9)$$

The expected short-term gain rate $\hat{g}_{x+1}$ for every next patch $x+1$, given the number of food items $n_x$ in current patch $x$ equals:

$$\hat{g}_{x+1} \bigg| (n_x = a) = \frac{\hat{n}_{x+1} \bigg| (n_x = a)}{\hat{t}_i} \quad (10)$$

For a highly clumped food distribution, the highest (short-term) gain rate is expected in the adjacent patch if the current patch is full and in a more distant patch if the current patch is empty (Fig. 2i). For a slightly clumped and a random food distribution the highest gain rate is always expected in the nearest patch, irrespective of the content of the current patch (Fig. 2j,l). For a regular food distribution the maximum expected gain rate is found in a distant patch ($x+2$) if the current patch is full and in the adjacent patch ($x+1$) if the current patch is empty (Fig. 2k), which is in qualitative sense opposite to the best movement rule in a highly clumped distribution! The strategy applied by the Bayesian assessor is to opt for the maximum short-term gain rate (which in most cases maximizes long-term gain rate as we will show below). Consequently, the length of a movement depends on the spatial pattern in the food distribution and the content of the current patch (see Fig. 3a-d).

The assessor encounters full patches at a higher and empty patches at a lower frequency than the frequency in which these patches occur throughout the environment as long as movement of the assessor is different from movement of the ignorant forager (Fig. 3e-h). Those encounter frequencies are calculated by equation 3, in which the transition probabilities $R(1|0)$ and $R(0|1)$ are derived from the structure functions. Long-term gain rate $G$ for an assessor is calculated by dividing total gain by total time:

$$G = \frac{\sum_{a=0}^{1} f_a \cdot a}{\sum_{a=0}^{1} f_a \cdot \hat{t}_a} \quad (11)$$

where $f_a$ equals the frequency at which a patch with $a$ food items is encountered, and $\hat{t}_a$ equals the total time associated with a patch containing $a$ food items, which is the time to exploit the patch plus the time to move to the next patch (which is a function of the number of prey in the patch).

The optimal movement rule for the assessor can also be found graphically, analogous to the graphical solution of the Marginal Value Theorem (Charnov, 1976) (Fig. 4a). We derive the optimal travel time by plotting the expected gain $\hat{n}$ as a function of travel time $t_m$ and
plotting the (fixed) patch residence time $t_f$ at the negative part of the x-axis (Fig. 4b). The line starting at $(t_f, 0)$ and tangent to the gain curve maximizes short-term gain rate and hence reveals the optimal $t_m$. In the example of Fig. 4b this is found for a travel time of 4 time units (patch $x+4$) if the current patch is empty and for a travel time of 1 time unit (patch $x+1$) if the current patch contains a food item.

**Larger foraging scales**

So far we have assumed a foraging scale equal to one cell. However, the foraging scale (and consequently the size of a patch) might consist of multiple cells. A forager with a larger foraging scale samples the environment in larger spatial units, and thus obtains more

---

**Figure 4** (a) Graphical solution of the problem ‘how long to remain in a patch’ (based on Charnov, 1976). The tangent from the (fixed) travelling time to the gain curve gives the optimal patch residence time. (b) Graphical solution of the problem ‘how to move between patches’. Two gain curves are shown: one for the case that the current patch is empty and one for the case that the current patch is full. The tangent from the (fixed) patch residence time to the gain curve gives the optimal travelling time. In this example the food distribution is highly clumped $R(0)=0.10$, $R(1)=0.05$. 

---

### Marginal Value Theorem

- Expected gain
- Distance $i$
- Time in patch
- Travel time
- Time in patch
- Expected gain
- Distance $i$
information about the food distribution per patch. On the other hand, might a forager with a larger foraging scale be less efficient in exploiting patchy resources, especially if patchiness occurs at a small scale (Schmidt and Brown, 1996; Klaassen et al., 2006). A forager with a larger foraging scale might thus be more or less efficient in resource use. The optimal movement rules and long-term gain rates for foragers with foraging scales consisting of two to six cells are worked out in Appendix A.

**Results**

*Comparing differently informed foragers*

Compared to ignorant foragers, prescient foragers achieve a higher long-term gain rate when food items are scarce (small clumps and large gaps). When food items are abundant (distributions with large clumps and small gaps) the advantage of the prescient over the ignorant forager is marginal, i.e. there is no penalty of being ignorant (Fig. 5a).

Compared to ignorant foragers, Bayesian assessors achieve a substantially higher long-term gain rate in clumped environments with large average gap and intermediate average food clump sizes as well as in regular distributions (Fig. 5b). Interestingly, the assessor does not perform better in near-random distributions, and in relatively rich (small gaps, large clumps) and relatively poor (large gaps, small clumps) environments. In these distributions there is no benefit of assessment! In some specific regular environments the assessor surprisingly achieves a lower long-term gain rate than the ignorant forager.

Compared to prescient foragers, Bayesian assessors perform (nearly) as well as the prescient forager in food distributions with large average clump sizes (Fig. 5c). However, in poor environments with smaller average clump sizes, the assessor achieves a much lower gain rate than the prescient forager, except if the average gap size is also small (regular distributions).

*Larger foraging scales*

The optimal movement rules for assessors with larger foraging scales are qualitatively comparable to the movement rules for assessors with a small foraging scale. For most food distributions the same step length is predicted from patches containing 1 and several food items, i.e. no distinction is made between the different number of food items in a patch, but only whether the patch contains a food item or not. The performance of an assessor in comparison to the ignorant and prescient forager is very comparable across different foraging scales, except for regular distributions. When having a large foraging scale, regular distributions are observed as homogeneous and thus Bayesian assessors cannot achieve higher gain rates than ignorant foragers.

To evaluate whether a larger foraging scale is advantageous or disadvantageous for different food distributions we compared for each food distribution the best performance of foragers with different foraging scales. For food distributions where the best movement rule
Figure 5 Comparison of the long-term gain rates achieved by foragers with three different levels of information about the food density in the next patch, for food distributions with different spatial patterns. (a) The advantage of a prescient over an ignorant forager, (b) The advantage of an assessor over an ignorant forager, (c) The advantage of a prescient over an assessor forager. Note that the z-axis is different for (b) than for (a) and (c).
of the assessor was equal to the ignorant rule (i.e. always move to adjacent patch), the gain rate is equal for assessors with different foraging scales, and no optimum scale could be determined (Fig. 6). For most other food distributions, the maximum long-term gain rate is achieved by the smallest foraging scale (1 cell) (Fig. 6). However, at clumped food distributions with large average clump and large average gap sizes and at some near-regular distributions, the highest long term gain rate is achieved by intermediate foraging scales (Fig. 6). Relatively large foraging scales (4-6 cells) never achieve the highest long term gain rate in the large range of food distributions considered.

Discussion
Adaptation of prior knowledge about the spatial distribution of food
In environments with a strong spatial pattern in the distribution of food, for example highly clumped and regular distributions, an incompletely informed forager that has prior knowledge about the spatial distribution of food items (Bayesian assessor) is able to lift its long-term gain rate by regulating the length of inter-patch movements. However, both in environments with a less strong spatial pattern and in environments that are very rich or very poor in food an assessor cannot achieve a higher gain rate than a forager that is not sensitive to spatial pattern (ignorant forager). In other words, there is not always a benefit of assessment. The presence of spatial pattern does not mean that an assessor must always respond to it.

Figure 6 The foraging scale (ranging from 1-6 cells) yielding the maximum long-term gain rate for different environments. The environments where all foraging scales yielded the same long-term gain rate are indicated by “all”.
In environments with a less strong spatial pattern, for example a slightly-clumped distribution, information about the content of the current patch still provides some information about the content of nearby patches. For the slightly-clumped distribution this means that the probability that the adjacent patch is full is slightly higher if the current patch is full and slightly lower if the current patch is empty. However, due to high unpredictability in the distribution of food items, this information is valueless since the highest short-term gain rate is always expected in the adjacent patch, irrespective of the content of the current patch. If spatial pattern is less strong, we consequently expect that Bayesian assessors act like ignorant foragers, i.e. always move to the adjacent patch. In these environments there might be a more important role for remembering rich patches, at least when the location of rich patches is constant and food is renewed at a high rate (Gill and Wolf, 1977).

Also in environments that are very rich or very poor in food, a Bayesian assessor is unable to achieve a (substantially) higher long-term gain rate than an ignorant forager, despite the presence of a clear spatial pattern. In rich environments (large food clumps, very small gaps) this is caused by the fact that there is not much to gain from a sophisticated movement rule, since already the ignorant strategy achieves a gain rate very close to the maximum achievable rate (the rate achieved by the prescient forager). In other words, in rich environments there is no problem to solve, i.e. there is no penalty of ignorance (Olsson and Brown, 2006). In contrast, in poor environments (very small food clumps, large gaps) a large potential benefit from a sophisticated movement rule exists, i.e. there is a large difference in performance between the prescient forager and the Bayesian assessor. In this case the assessor fails to achieve a higher gain rate, despite spatial pattern in the distribution of food, presumably due to large unpredictability about the location of the few small food clumps. Being sensitive to spatial pattern does not seem to be an adaptive trait for a forager in environments that are very rich or very poor in food.

Assessors were assumed to always move to the patch with the highest expected gain rate. In order to investigate whether this short-term rate maximizing strategy also maximized gain rate in the long run, we ran iterative Monte-Carlo simulations. These simulations (data not shown) confirmed that foragers nearly always maximized their long-term gain rate by applying this short-term maximizing strategy. This was not the case in some specific near-regular distributions. In these distributions the ignorant forager surprisingly achieved a higher long-term gain rate than the assessor! These distributions have in common that the (maximum) expected gain rate in the subsequent patch is much lower if the current patch is full, than if the current patch is empty (see example in Table 1). The assessor visits a higher number of filled cells than the ignorant forager, but in order to do so it spends more time traveling, which more than outweighs the advantage of trying to visit full cells (Table 1). Hence, for these near-regular distributions the long-term optimal movement rule for an assessor would be the ignorant rule.
Table 1  Parameter values for an ignorant and an assessing forager in a near-regular food distribution where the ignorant forager achieves a higher long-term gain rate than the assessor. The movement rule for the ignorant forager is to always move to the adjacent patch \((x+1)\). The movement rule for the assessor is to move to the adjacent patch \((x+1)\) if the current patch is empty, and to patch \(x+2\) if the current patch is full.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ignorant forager</th>
<th>Assessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(1</td>
<td>0))</td>
<td>0.85</td>
</tr>
<tr>
<td>(P(0</td>
<td>1))</td>
<td>0.55</td>
</tr>
<tr>
<td>(\hat{n} \mid (n_x = 0))</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>(\hat{n} \mid (n_x = 1))</td>
<td>0.15</td>
<td>0.17</td>
</tr>
<tr>
<td>(f_0)</td>
<td>0.39</td>
<td>0.28</td>
</tr>
<tr>
<td>(f_1)</td>
<td>0.61</td>
<td>0.72</td>
</tr>
<tr>
<td>(\sum_{a=0}^{1} f_a \cdot a)</td>
<td>0.61</td>
<td>0.72</td>
</tr>
<tr>
<td>(\sum_{a=0}^{1} f_a \cdot \hat{f}_a)</td>
<td>3</td>
<td>3.72</td>
</tr>
<tr>
<td>(G)</td>
<td>0.20</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Applying the model to natural systems

Recently we applied the model described in this paper to predict lengths of inter-patch movements of Tundra swans feeding on Sago pondweed (\textit{Potamogeton pectinatus}) tubers (Klaassen et al., in press). Swans have to dig craters to reach tubers that are buried in the sediment, and thus their movement path consists of craters (that we considered as patches) and movements in-between craters. In accordance with model assumptions for an assessing forager, swans could only assess the quality of a (potential) patch by sampling the patch, and not during a movement (swim) between patches. Furthermore, swans had relatively straight foraging paths, which corroborates the use of a one-dimensional model.

The spatial pattern in the distribution of tuber densities was recorded just before swans arrived at the site and described by structure functions. Net rather than gross energy gain rates were calculated by incorporating extensive details on the swans’ behaviour, such as energy expenditures of different behaviours, time spent in patches as a function of initial food density, the swans’ functional response, and velocity of movement. Rate-maximizing swans were predicted to move to adjacent patches if the food density in the current patch had been high, and to a more distant patch (at 7-8 m) if the food density in the current patch had been low, which reflects the clumped fashion in which tuber densities are distributed. There was a rather good quantitative agreement between these predictions and
distances swans moved in the field, with swans moving long distances from a low density and short distances from high density patches.

A number of other studies also provide qualitative support to our model predictions. Here we focus on two of such studies to illustrate how our model results can help to understand movement patterns observed in the field. Heinrich (1979) studied movements of nectar-feeding Bumblebees *Bombus terricola* in a field of White clover *Trifolium repens*. He observed that bumblebees successfully avoided areas with flower heads containing little nectar by making longer movements from flower heads with few or no nectar. For this situation our model would also predict longer movements from poor patches, because a positive spatial autocorrelation in the amount of nectar in flower heads was observed.

Pienkowski (1983) studied movements of Grey plovers feeding on invertebrate prey at a mudflat. Plovers try to detect prey while standing still and move variable distances between ‘waiting points’. The area scanned for prey during a wait can be considered as a foraging patch. Plovers moved a short distance from a patch if a large prey was caught and a longer distance if a small or no prey was caught. Given that these prey items are thought to be distributed in a clumped fashion, also these results are in qualitative agreement with our model predictions.

Although abovementioned examples point out that our model is potentially applicable for different biological systems, they do not allow us to put the model to the test. A rigorous test of the model would involve quantifying movements of foragers in environments with different (contrasting) spatial configurations of food densities, for example the four configurations used as examples throughout this paper. This might only be achievable in an experimental setup, which would allow for the forager to learn the spatial pattern of the food distribution. In such experiment it is important that the assessment of the content of patches comes at a (time) cost to ensure there is a penalty of visiting all patches.

*Different foraging scales*

The scale of foraging is an important feature of a forager since it determines how the environment is perceived (Schmidt and Brown, 1996; Klaassen et al., 2006). A food distribution with intermediately sized food clumps may be observed as (highly) clumped for a forager with a small foraging scale, but may be observed as slightly-clumped, or even as random or regular by a forager with a larger foraging scale. In comparison to ignorant foragers, assessors perform best in clumped distributions with large gap and intermediate food clump sizes (Fig. 5b). What is an intermediate food clump for a forager with a small foraging scale might be a small food clump for a forager with a larger foraging scale. It is consequently not surprising that in environments with large food flumps the highest gain rates are achieved by foragers with larger foraging scales (Fig. 6). Interestingly, for most other food distributions the smallest foraging scale performs best, underlining the importance of a response to small-scale heterogeneity. We conclude that the scale of
foraging is an adaptive character of a forager, except for near-random distributions. Due to the costs associated with sampling we might expect that in environments where food always occurs in large spatial units also the foragers sample the environment in large spatial units.

Models on movement

The one-dimensional movement model introduced in this paper is applicable for situations where (1) prey are cryptic, (2) the foraging path consists of patches and movements in between patches, (3) the length of inter-patch movements is variable, (4) the foraging path is relatively straight (i.e. no revisitation of patches), and (5) the environment is not sampled during movement. Although specific, this is not an uncommon situation at all for natural systems and examples include insects (Heinrich, 1979), fish (Hill et al., 2003), birds (Pienkowski, 1983; Nolet and Mooij, 2002; Klaassen et al., in press), and mammals (Fortin, 2003).

Other models on moving animals treat different situations, and consequently highlight other aspects of animal movement. Benhamou (1992) for example focussed on path sinuosity, i.e. the degree of turning between movements. In his simulations foragers did not regulate the length but the turning angle of individual movements. This two-dimensional model is applicable for Blackbirds *Turdus merula* feeding on earthworms that naturally occur in clumps (Smith, 1974). Blackbirds respond to this spatial pattern by increasing the sinuosity of their movement path after a prey capture (Smith, 1974). The length of individual movements did not change in relation to a prey capture indicating that our model is not applicable for this situation.

Some foragers neither regulate the length of movements nor the sinuosity of their path, but the velocity of their movement in response to spatial pattern. Red Knot *Calidris canutus* and Great Knot *Calidris tenuirostris* feeding in a clumped food distribution for example intensified search (lower speed; more pecks per step) after a prey capture (Van Gils, 2004). The further away the birds came from their capture point the faster they moved (fewer pecks per step). Again, since those birds did not regulate the length of their movements, our model would not be applicable here. A model that does describe this way of movement and the response to spatial pattern is provided by Knoppen and Reddingius (1985) and Krakauer and Rodríguez-Gironés (1995).

The three types of aforementioned models find their application in specific study systems. The foragers in these models respond in different ways to spatial pattern in the distribution of food, but in all cases the result of this response is that the forager biases its effort to areas that are rich in food. Benhamou (1992) only considered a very limited number of clumped distributions in which a response to pattern was favourable in all distributions. In accordance with our results, Knoppen and Reddingius (1985) and Krakauer and Rodríguez-Gironés (1995) identified distributions in which a response to pattern was not valuable. Benhamou (1992) advocated that a large range of values of movement
parameters lead to near-maximal efficiency. However, in our model and in the models of Knoppen and Reddingius (1985) and Krakauer and Rodríguez-Gironés (1995) small deviations from optimal parameter values sometimes caused a substantial decrease in foraging performance, indicating that there is not something like a robust movement rule that is near-optimal for a large range of distributions. An important advancement of our approach in comparison to the other theoretical investigations is that we defined both an upper (prescient forager) and a lower limit (ignorant forager), which allowed us to evaluate the efficiency of the performance of the Bayesian forager.

Acknowledgements
We thank Rudi Drent, Marcel Klaassen, and four anonymous reviewers for valuable comments on earlier drafts of this paper and Joel S. Brown for inspiring suggestions. This is a publication of the Netherlands Institute of Ecology (NIOO-KNAW) and the Centre for Wetland Ecology.
Appendix A: Calculations on foragers with larger foraging scales

Here we formulate the optimal movement rules and long-term gain rates for different food distributions, for differently informed foragers with foraging scales consisting of two to six cells.

Foragers are not able to respond to spatial heterogeneity that exists at a smaller scale than the foraging scale (grain) (Kotliar and Wiens, 1990). We therefore assume that foragers observe the number of food items in a patch, but are not informed about the exact position of these food items within the patch. The probability that a patch at distance \( i \) (if patch is adjacent: \( i = 1 \), if one cell is located between patches: \( i = 2 \), etc) contains \( b \) food items, given that the current patch contains \( a \) food items can be calculated by combining two different probabilities:

1. the probability that the last cell of the current patch is full (or empty), given that the patch contains \( a \) food items: \( P(n_{\text{last cell}} = 1 \mid m_x = a) \)
2. the probability that a patch, at a distance \( i \) from a cell, contains \( b \) food items, given that this cell is full (or empty): \( P(m_{x+i} = b \mid n_x = 1) \)

where \( m \) is the number of food items in a patch. These probabilities can be derived from structure functions and are combined by:

\[
P(m_{x+i} = b \mid m_x = a) = P(n_{\text{last cell}} = 1 \mid m_x = a) \cdot P(m_{x+i} = b \mid n_x = 1) + P(n_{\text{last cell}} = 0 \mid m_x = a) \cdot P(m_{x+i} = b \mid n_x = 0)
\]

Subsequently we can calculate the expected gain for a patch at distance \( i \), given the number of food items \( m_i \) in the current patch:

\[
\hat{\eta}_{x+i} \mid (m_x = a) = \sum_{m_{x+i}=0}^{S} m_{x+i} \cdot P(m_{x+i} = b \mid m_x = a)
\]

in which \( S \) is the foraging scale (number of cells), which is equal to the maximum number or food item in a patch, since one cell can contain one food item at most.

The expected total time \( \hat{t}_i \) invested in moving to and exploiting a patch at distance \( i \) equals:

\[
\hat{t}_i = (S - 1 + i) \cdot t_m + S \cdot t_f
\]

The expected short-term gain rate \( \hat{g}_{x+i} \) for a patch at distance \( i \) from the current patch, given the content of the current patch, now can be calculated by dividing the expected gain by the total time (similar to equation 10).

A matrix of all transition probabilities from a certain density \( a \) in the current patch to a certain density \( b \) in the next patch is an alternative way to characterize the movement of an ignorant and an assessing forager. Transition matrices for ignorant and assessing foragers for an environment with \( P(0|1)=0.10 \) and \( P(1|0)=0.05 \) are (\( S=3 \) in this example):

<table>
<thead>
<tr>
<th>current patch</th>
<th>next patch</th>
<th>( a \backslash b )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignorant:</td>
<td></td>
<td>0.86</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.51</td>
<td>0.07</td>
<td>0.07</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.48</td>
<td>0.07</td>
<td>0.07</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.73</td>
<td></td>
</tr>
</tbody>
</table>
The expected total gain is calculated by summing the products of the frequency at which a certain food density is encountered and the number of food items. In a similar way, we calculate the expected total time by summing the products of the frequency at which a certain food density is encountered and the total time $T$ associated with that food density. The long-term gain rate now can be calculated for ignorant and assessing foragers by dividing the total gain by the total time:

$$G = \frac{\sum_{m=0}^{M} f_m \cdot m}{\sum_{m=0}^{M} f_m \cdot \hat{t}_m}$$

in which $f_m$ is the frequency in which a patch containing $m$ food items is encountered by the forager (for the ignorant forager this is equal to the frequency at which these patches are on offer), and $\hat{t}_m$ the total time related to a patch with $m$ food items.

A prescient forager knows the content of a cell prior to actual exploitation, and consequently can decide whether to exploit a patch or not. If the foraging scale consists of multiple cells a prescient forager can employ different threshold densities $c$ (number of food items per patch), below which a patch is rejected and not exploited. We calculate the long-term gain rate of a prescient forager for different threshold density $c$ (ranging from 1 to $S$) according to:

$$G_c = \frac{\sum_{m=0}^{M} f_m \cdot m}{\sum_{m=0}^{M} f_m \cdot \hat{t}_m}$$

in which $\hat{t}_m$ equals $S \cdot t_m$ if $m < c$ and $S \cdot t_m + S \cdot t_f$ if $m \geq c$. Subsequently we select $c$ that corresponds to the maximum long-term gain rate.