Chapter 3. Individual-based model of diverse populations

3.1 Introduction

The results of the mean-field model showed that the maintenance of diversity in a community depended on particular relationships between the basic physiological traits of individuals in the community. Moreover, the nature of the competition between individuals strongly affected the dynamics of the system. If relationships between physiological traits and interactions between individuals are crucial for community diversity, a more realistic description of individuals and their interactions should provide a more accurate representation of the mechanisms leading to diversity. Individual-based models provide a convenient framework for the description of individuals in terms of their physiological traits. In this work, an individual-based simulation model is used to complement the results of the mean-field model and to explore the effects of local interactions on diversity. The structure of this model was developed elsewhere and is fully described in Bown (2000) and Bown *et al.* (*in preparation*).

Defining individuals in terms of quantifiable physiological traits allows experimental data to be used to parameterise the model. This establishes an important link between the model and the modelled biological community, as has been noted in the introduction. The model can then provide a more realistic representation of the studied community and offer quantitative results. This chapter describes a parameterisation of the model, using experimental data to represent communities of the grassland species *Rumex acetosa* (Bown *et al. in preparation*). The data were obtained from a study that examined individual variation in species-rich grassland. The species *R. acetosa* was chosen for parameterisation because it is common throughout the study sites – *R. acetosa* is a perennial that coexists with many other species in grazed pastures – and had been subject to detailed physiological analysis (Bausenwein *et al.* 2001).

3.2 Model description

'Biologists must consider the mathematicians view of logic, low dimensionality, and simplicity; mathematicians must recognize the biologist's tendency to believe that everything is important.'

Gross *et al.* (1992), p. 513.

The model was formulated to represent an isolated patch of plants (*i.e.* without input of seed from outside) on a small scale (1-10 m^2). Space in the model is represented explicitly (*i.e.* plants have locations that define their position with respect to other plants). An explicit representation of space was used to explore the effects of local interactions on the dynamics. The resource was also explicitly defined. Plant interactions occurred through competition for resource and space. A description of plant interactions mediated by resource is easier to parameterise than direct interactions between plants. Measuring the effect of one plant on another, or one species on another, is quite difficult; while measuring plant resource uptake and request (*i.e.* interactions with resource) is feasible and more accurate. The version of the model described here assumes that plants self-replicate (*i.e.* produce offspring identical to themselves), to ignore the effects of genetic coupling. The aim was to explore specifically the physiological, *i.e.* non-genetic, mechanisms generating diversity.

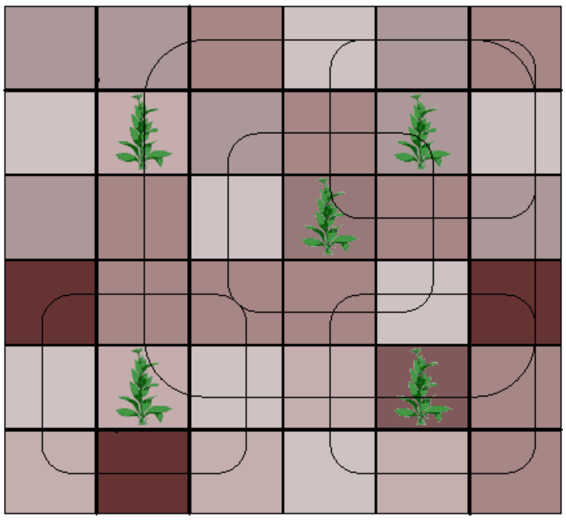

Figure 3.1 A representation of a part of the lattice simulated in the individual-based model. The squares of different shades of brown represent lattice cells with different level of substrate. The rounded squares around the plants cover the lattice sites from which the plants take up resource. Where the rounded squares overlap, the plants compete for resource.

The model simulates plant communities on a 2-D spatial lattice, with at most one plant possible at each lattice cell (Figure 3.1). Space and time in the model are discrete. It is assumed that plant growth is limited by a single resource which is distributed over the lattice. The amount of resource is defined individually for each lattice cell. Competition occurs for resource and for space, as described below. Plants are defined by traits describing essential physiological functions: resource uptake, development and reproduction. Plants develop by progressing along development stages. Those parameters that change with plant development are described as functions of the development stage. The 12 plant traits are listed, along with their numerical values, in Table 3.1 and described below.

Table 3.1 Parameter values and distributions. For parameters which are represented by distributions, sign '±' separates the means and the standard deviations of the distributions (see text for derivation of the values).

Parameter	**Values**
Essential uptake, $U_e(s)$	A sigmoidal curve $y = y_0 + \frac{\alpha}{1+e^{-(s-s_0)/\beta}}$ where s is the development stage. y_0 = 0.24 ± 0.310, β = 8.12 ± 2.65, α = 10.52, s_0 = 29.09.
Requested/essential uptake ratio, r_u	1.1 ± 0.58
Spatial distribution of uptake, D_u	see Table 3.3
Resource storage partition trait, P_s	0.8
General storage release proportion, r_g	0.28 ± 0.143
Surplus storage release proportion, r_s	0.40 ± 0.153
Time dependent reproduction relation, R_t	45 ± 8.6 time steps
Storage-fecundity relation, R_f	S_r / R_{min} where S_r is the plant's storage available for reproduction, and $R_{min}=U_e(1)$;
Seed dispersal pattern, D_p	randomly dispersed in an area 5 lattice cells away from parent plant;
Survival threshold, V_t	0.1*$U_e(s)$ where s is plant's current development stage;
Survival assessment period, V_p	5 time steps
Probability of plant death, P_d	0.001

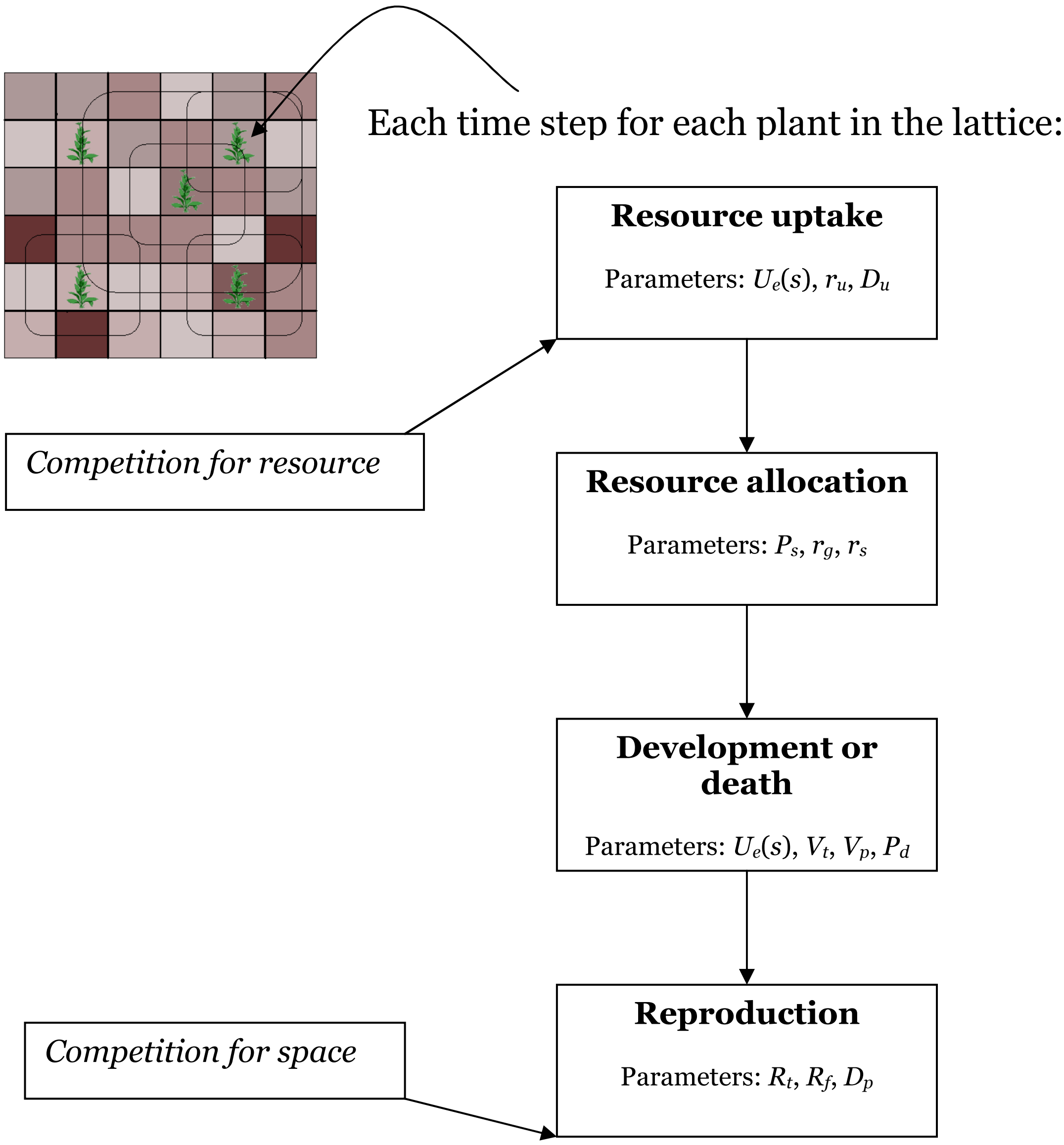

Figure 3.2 A diagram of processes that occur in the model for each plant at each time cycle.

A single time step of the model consists of plant resource uptake, resource storage, development and reproduction (a diagram of the processes that occur for each plant at each time cycle is shown in Figure 3.2). A plant acquires resource from its location and from the cells in its neighbourhood (Figure 3.3). The lattice neighbourhood represents the spatial spread of roots and leaves that the plant uses to acquire resource. The area and the distribution of uptake within this neighbourhood is described by the parameter D_u. Competition for resource occurs when more than one plant demands resource from the same lattice cell, and the demand exceeds the available resource. In this case, the resource available at the lattice cell is distributed among plants in proportion to their requests. More precisely, suppose n plants make demands on resource at a lattice cell, and each plant i makes a demand D_i. Suppose that T is the total amount of resource at the lattice cell. If $\sum_{j=1}^{n} D_j > T$, then each plant i is assigned $T\left(D_i \big/ \sum_{j=1}^{n} Dj \right)$. This represents the competition that occurs between plants' roots for example. For example, if one plant has few roots in the area, and another plant has most of the roots in that area, the second plant will be able to extract more resource from that site.

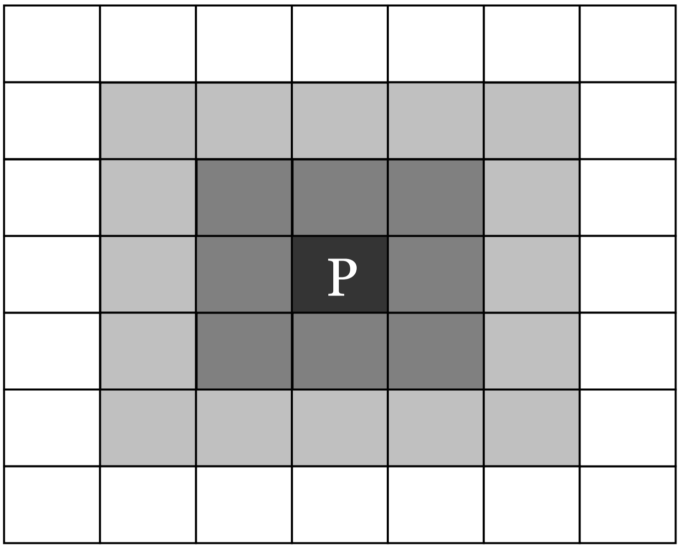

Figure 3.3 An example of a plant's capture area. The squares represent lattice cells. The cell labelled 'P' is the location of the plant. The cells shaded grey represent the area from which the plant acquires resource. The white is the area from which the plant does not request resource. The intensity of grey corresponds to the proportion of the total demand that the plant requests from the environment. For example, at the cell shaded the darkest grey (plant's location), the plant may request 0.5 of its total demand; from the lighter grey area around it 0.3; and from the lightest grey 0.2.

The plant's resource demand is described by two parameters, the essential uptake $U_e(s)$ (a function of the development stage s), and the requested/essential uptake ratio r_u. The essential uptake $U_e(s)$ describes the resource amount necessary for the optimal growth of the plant. A plant places a demand for an amount greater than $U_e(s)$ to increase its chances in the case of competition. Specifically, the amount that a plant demands is $r_u U_e(s)$.

The resource that plants obtained from the environment is allocated to development (*i.e.* growth of structure) and storage. Plants store resource in roots and leaves which can be translocated for use in development and reproduction at a later time. To represent this in the model, the acquired resource R is used by each plant for: a) maintenance of structure, b) development and c) reproduction. If a plant acquires more resource than is necessary for its optimal growth, the excess is allotted to the surplus storage. The proportion of R that is allocated for structure maintenance is described by the parameter P_s, the structural storage proportion. The resource, which is not used for structure, is stored in general storage. The resource in general storage can be used for reproduction and development. Any resource that is obtained in excess of the essential uptake $U_e(s)$ is allocated to the surplus storage.

After resource assimilation, the development stage of each plant is updated. A plant may either develop or die. A plant develops by progressing to the next development stage. A plant is transferred from a development stage s to a development (s+1), if it has acquired an amount of resource $\sum_{i=1}^{s+1} U_e(i)$. A plant can die due to lack of resource or due to some other factor (*e.g.* disease). A plant dies due to lack of resource if it has not acquired enough resource over a specified period of time. The period of time is defined in time steps by the parameter V_p, and the minimum amount of resource required for survival is described by the parameter V_t. Death occurs if a plant's uptake over the last V_p time steps was less than $V_t U_e(s)$ where s is plant's current development stage. Death due to external causes, other than lack of resource, is incorporated by introducing a probability P_d that a plant dies. When a plant dies, the storage it carried is added to the resource level of the lattice cells from which it acquired resource.

Plants in the model reproduce with a frequency defined by the parameter R_t – the number of time steps between reproduction events. Plants produce offspring using storage resource allocated for reproduction. The resource for reproduction is drawn from the general and the surplus storage. The amount of storage used for reproduction from each storage is limited by the general storage release rate r_g, and the surplus storage release rate r_s. That is, the proportion r_g of the resource in the general storage plus the proportion r_s of the resource in the surplus store can be used for reproduction at a single reproduction event. The number of offspring produced is described by the storage-fecundity relation R_f. The parent plant's parameter values are inherited by its offspring. The offspring are distributed on the lattice according to the dispersal pattern of the plant, described by the parameter D_p. To represent an isolated patch, the boundaries of the lattice are absorbing, *i.e.* if an offspring lands outside of the lattice, it effectively dies. Each offspring carries to its site enough resource to progress to the first development stage. This represents the resource storage contained within plant seeds. If the lattice cell where the seed lands is occupied, the offspring dies, and the resource it carried is added to the substrate level of the lattice cell where it landed. After reproduction, plants often lose structure, biomass, and resource store, and therefore their requirement for uptake falls. To represent this, the development stage of the parent plant is re-evaluated and reduced if the plant does not have the stored resource required by its current development stage.

This model is an effective framework for studying the effects of individual traits on community dynamics, since the simulated communities can be understood in terms of individual physiological parameters. The individual physiological traits can be parameterised using experimental data. The model was designed with a framework flexible enough to allow many plant strategies. The need for defining the competitive effect of plants on each other (which is difficult) is avoided by introducing explicit competition for resource and space. Environmental heterogeneity can be easily introduced, since the resource level is defined separately for each cell. Community diversity is also convenient to simulate, since each plant has its own set of parameters associated with it.

3.3 Model parameterisation

> 'The intrinsic rate of increase of crop plants has been known since Biblical times to be between 30-fold and 100-fold per generation (St. Matthew 13).'
>
> Crawley (1990), p. 127.

3.3.1 Physiological data

The model was parameterised using physiological data for the plant species *Rumex acetosa*. Data derived from 20 plants were used, with 10 plants taken from each of two sites. The first site was a lowland grassland site near Cleish in Fife, Scotland, OS map location NT082934. The other site was a hillside grassland site in Kirkton near Crianlarich in the west Perthshire, Scotland, OS map location NT360284. The collected plants were cloned for use in nutrient labelling experiments in a glasshouse. During the experiment, the plants were allowed to take up ^{15}N before winter, then grown on ^{14}N next year. The contributions of N translocated from storage and N taken up during growth of various tissues was then assessed. This information was collected at 7 different times from January to September 1997: January 29, March 20, April 17, May 6, May 26, June 16, and September 9. For each of the plant organs listed in Table 3.2, the following measurements were taken: dry weight, content of C, N, and the proportions of N translocated from storage or taken up. Full details of the sites and the experiment are given in Bausenwein *et al.* (2001).

Table 3.2 Plant parts about which information was collected. H1 – harvest 1; H2 – harvest 2, etc.

Leaves	Dead leaves Old leaves at H1 New leaves between H1 and H2 New leaves between H2 and H3 New leaves between H3 and H4 New leaves between H4 and H5 New leaves between H1 and H5 New leaves between H6 and H7 Total new leaves
Reproduction	Stem (flower stock) Flowers (and forming seeds for females)
Roots	Fine root Tap root

3.3.2 Parameterisation

Most of the parameters describing plants were estimated using the experimental data for *R. acetosa* plants described above. In a few cases where the experimental data were lacking, evidence from existing literature along with general biological considerations were used. Diversity in communities was represented by allowing variation in the parameters described by experimental data. This was accomplished by defining the parameters in terms of probability distributions of possible values. Normal distributions were defined in terms of the means and standard deviations derived from the observed distribution of parameter values for the 20 plants. The observed probability distributions were tested for normality with the Shapiro-Wilk and Anderson-Darling tests. In all but one case, the hypothesis that the distributions do not differ from normal could not be rejected at the 5% significance level. Two exceptions were the distribution of values for the time of reproduction and the y_o parameter of the essential uptake curve. Due to the lack of information to the contrary, their distributions were assumed to be normal also. Some exploration of the behaviour of the model was conducted when other distributions (uniform, lognormal) were used, but no difference in the behaviour of the model was noted. In cases when the model parameter distributions included values not physiologically possible, truncated distributions were used (as described below). The experimental data used to parameterise the model were collected for plants in close to optimal conditions. The constraints on plant growth and development in the model arose as a consequence of limitation in resource availability and competition. If parameters were not defined using experimental data, they were defined by an estimated value. The plant parameter values are listed in Table 3.1, and the estimation procedures are described below. Plants development was divided into 50 development stages. This number was chosen to represent the plant life cycle on a scale sufficiently small to capture the essential plant functions. According to the parameterisation, a development stage corresponds to 4.46 days of plant growth in perfect conditions.

The essential uptake $U_e(s)$ was approximated by the amount of N used by plants as a function of the development stage *s*. This was done since N is considered to be the main limiting resource in Scottish grasslands. $U_e(s)$ was estimated by adding the amount of N uptake to the amount of N moved from the storage (if any) at each harvest time. This defined a lower limit for $U_e(s)$, since the amount of N moved from an experimental plant's storage may have been limited by the release rate of the storage. This estimation of $U_e(s)$ was, therefore, less than or equal to the actual value. The parameters for $U_e(s)$ were obtained by fitting a sigmoidal function to the values of used N (Figure 3.4):

$$y = y_0 + \alpha/(1 + e^{-\beta(s-s_0)})$$

where *y* is the estimated used N, s is the development stage of the plant, and s_0, y_0, α and β are parameters. The choice of the function is explained in Appendix C. Parameters y_0 and β were described by distributions since they can be used to define the main properties of the curve: the height (defined by y_0) and the steepness of the rise of the curve (defined by β). The values of s_0 and α were defined by mean values. The values of s_0, y_0, α and β were scaled so that $U_e(s)$ was defined for *s* in the range 0 to 50 (the development stage range) and, for average values of the parameters, was 10 at the last development stage. The value of 10 was arbitrarily chosen as the maximum uptake value. The distribution of y_0 was truncated at 0, since the initial uptake cannot be negative.

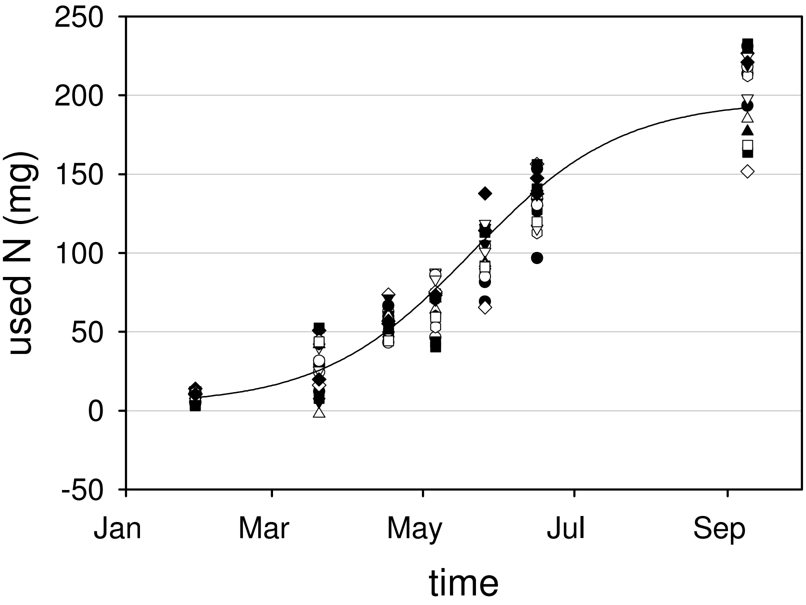

Fig. 3.4 a) Used N for *R. acetosa* plants at the seven harvests. Different symbols represent values of the used N for the experimental plants. Solid line - example of a sigmoidal curve fitted to the values of used N for one of the plants. The parameters of fitted curves, such as the curve shown, were used to define required uptake curves for plants in the model.

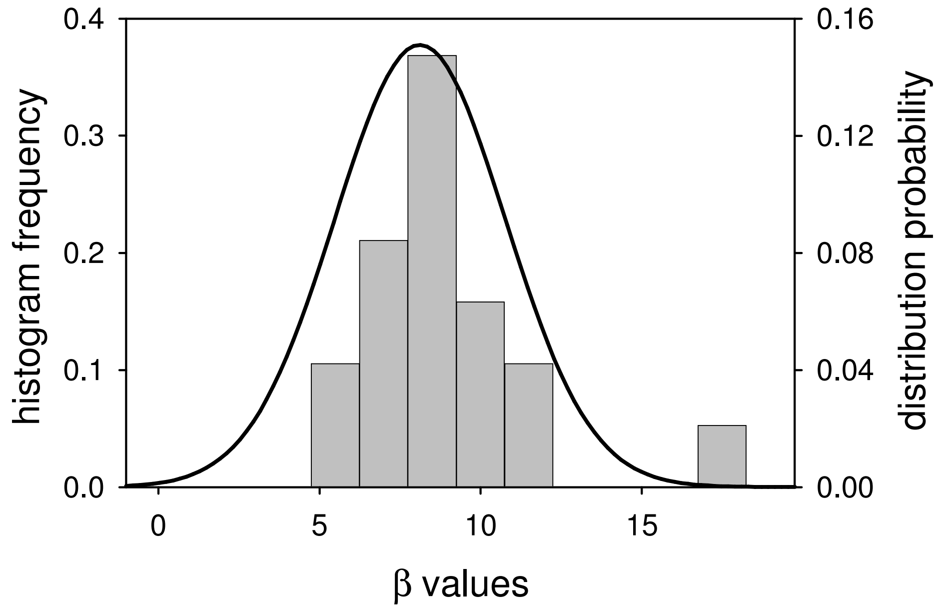

Fig. 3.4 b) The distribution of values of the parameter β, and the fitted normal distribution. The frequency of the observed values is plotted along the y-axis on the left. The distribution probability of the normalised normal distribution is plotted along the y-axis on the right.

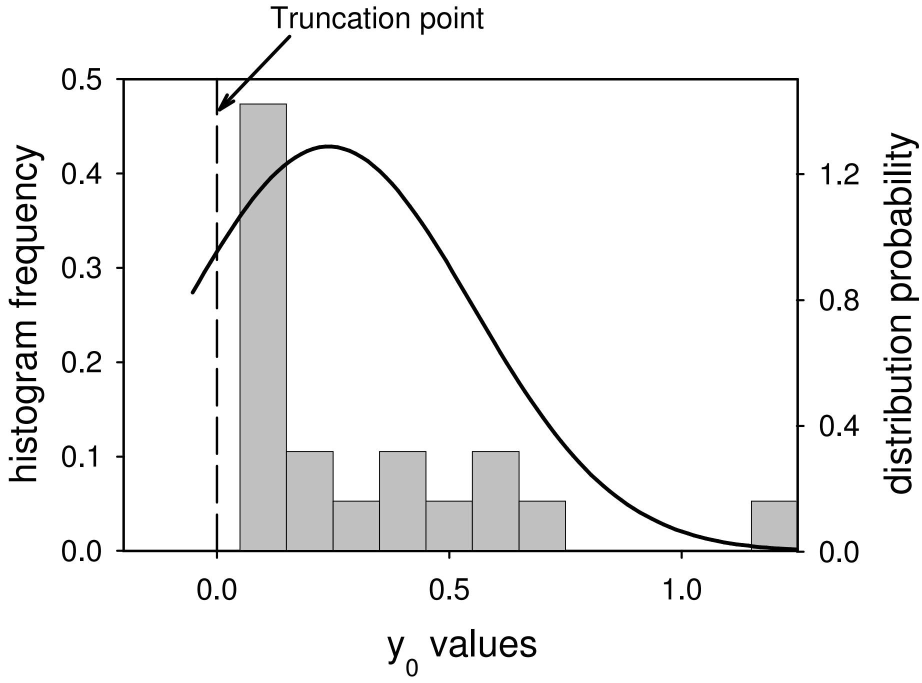

Fig. 3.4 c) The distribution of values of the parameter y_o, and the fitted normal distribution. Values of y_o below the truncation point are rounded up to 0. The truncation is at 0 since y_o describes the height of the plant uptake curve, and it cannot be less than 0. The frequency of the observed values is plotted along the y-axis on the left. The distribution probability of the normalised normal distribution is plotted along the y-axis on the right.

Figure 3.4 Experimental data and parameter distributions for plant uptake.

The relationship between the requested uptake U_r and $U_e(s)$ was estimated from the set of proportions that the uptake N was larger than the 'used N':

$$\frac{N_{up}(i,j)}{N_u(i,j)} \text{ if } N_{up}(i,j) > N_u(i,j) \text{ and } n \le ij .$$

where n is the number of times when $N_{up}(i,j) > N_u(i,j)$, $N_{up}(i,j)$ is the uptake of N and $N_u(i,j)$ is the 'used N' for plant i at harvest j. The values of these proportions as a function of the harvest time are shown in Figure 3.5. The dependence of the relationship between U_r and $U_e(s)$ on harvest time was unclear. Hence, U_r was assumed to be proportional to $U_e(s)$, *i.e.* $U_r = r_u U_e(s)$, where r_u was a constant and defined by the distribution with the mean and the standard deviation of the set of the above defined proportions. The distribution was truncated at 1, since the requested uptake should be more than the required uptake.

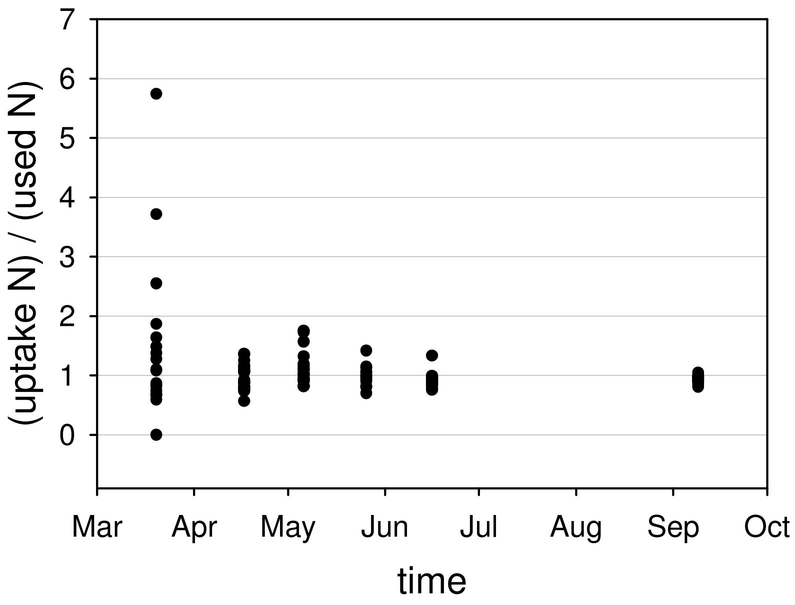

Figure 3.5 (uptake N/used N) determines whether plants absorbed more or less N than they required for growth. Dots represent values for different plants at different harvests.

The resource capture system was approximated by the root structure of *R. acetosa*. In the experimental system, competition for soil nutrients dominated, and therefore most of the competition for resource occurred in soil. Information on the actual dimensions and spatial distribution of the root structure of *R. acetosa* is lacking. Generally, however, *R. acetosa* has a cone-like root structure with a tap root more developed than the fine roots. This suggests that the resource capture area is relatively small, and its spread is generally slow and concentrated in the centre. These properties were used to define the resource capture system. It was represented in two configurations, one of which spread more slowly than the other (Table 3.3). One of the two configurations was assigned to plants with equal probability.

Table 3.3 Two types of resource capture distributions possible for *R. acetosa* in the model. B_i – band of lattice cells around location of the plant (Figure 3.3); *s* –development stage.

s	B_0	B_1
1-10	1	0
11-20	0.9	0.1
21-30	0.8	0.2
31-40	0.7	0.3
41-50	0.6	0.4

a) Type 1 resource capture distribution

s	B_0	B_1
1-10	1	0
11-20	0.8	0.2
21-30	0.6	0.4
31-40	0.4	0.6
41-50	0.2	0.8

b) Type 2 resource capture distribution

The distribution of the resource storage partition trait P_s was estimated by the proportions of the stored N that could be used between successive harvests. The distribution was defined in terms of the mean and the standard deviation of the set of proportions by which stored N in a given harvest was smaller than in the previous

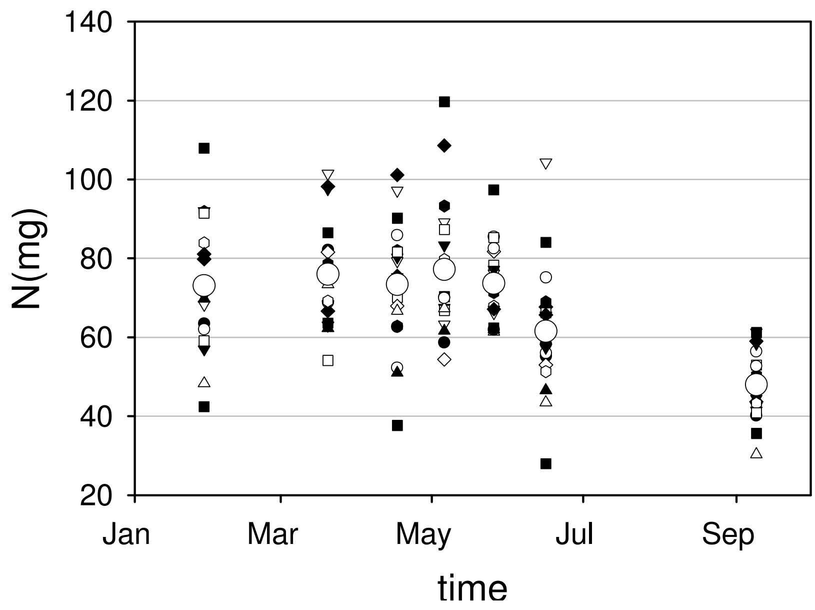

Fig. 3.6 a) Total stored N for *R. acetosa* plants at seven harvests. Different symbols represent values for different plants. Large circles are the average values for each harvest. The average proportion of stored N that remained when plants used stored N was used to approximate the resource storage partition trait P_s.

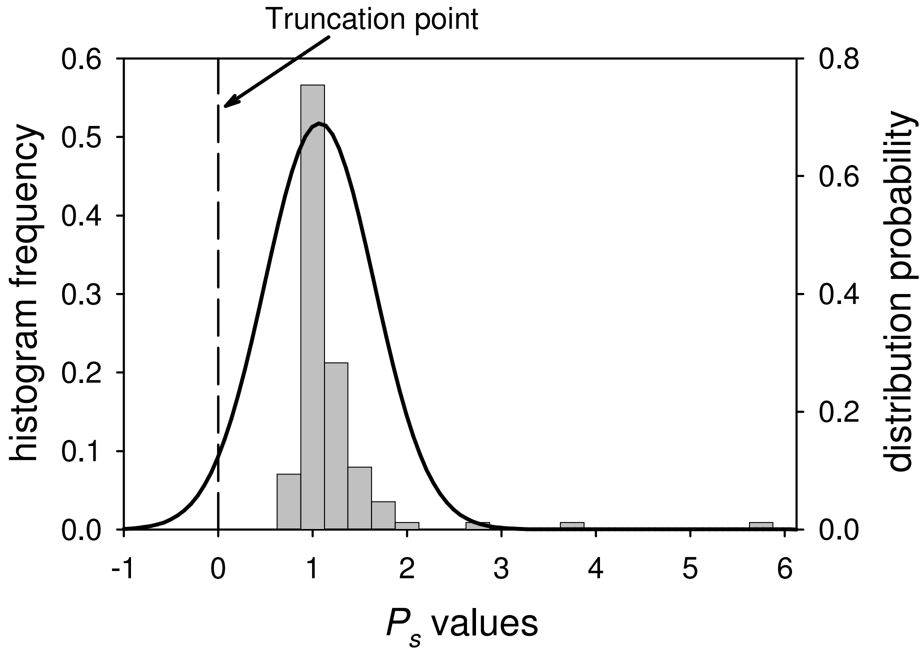

Fig. 3.6 b) The distribution of values of the parameter P_s, and the fitted normal distribution. Values of P_s below the truncation point are rounded up to 0. The truncation is at 0 since P_s is the resource storage partition trait, and it cannot be less than 0. The frequency of the observed values is plotted along the y-axis on the left. The distribution probability of the normalised normal distribution is plotted along the y-axis on the right.

Figure 3.6 Total stored N and the obtained distribution of the parameter P_s.

harvest:

$$\frac{N_s(i, j+1)}{N_s(i, j)} \text{ if } N_s(i, j) > N_s(i, j+1) \text{ and } n \le ij$$

where n is the number of times that $N_s(i, j) > N_s(i, j+1)$, N_s is the stored N that a plant i used between harvests j and $j+1$ (Figure 3.6).

The distribution of the general storage release proportion r_g was defined in terms of the mean and the standard deviation of the set of the proportions, N_l, of N lost by a plant in new leaves and fine root during reproduction. The amount of N in new leaves and fine root at different harvests is shown in Figure 3.7. For each plant, N_l was found by calculating the proportion by which the N content in leaves and fine roots at the harvest with maximum N was larger than the corresponding value at the harvest when N was minimum. In other words

$$N_l = \frac{\max\{N_{lf}(j)\} - \min\{N_{lf}(j)\}}{\max\{N_{lf}(j)\}}$$

where $N_{lf}(j)$ is the proportion of N lost from the new leaves and fine roots at harvest j, $1 \le j \le 7$. The distribution was truncated at 1 and 0, since the release proportion cannot be greater than 1 or less than 0.

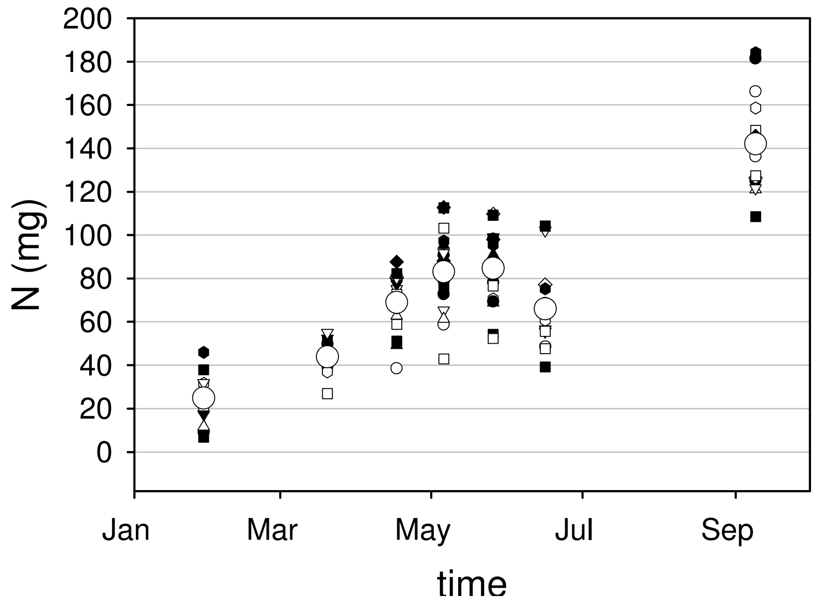

Fig. 3.7 a) N content in new leaves and fine roots in the *R. acetosa* plants. Different symbols represent values for different plants. Large circles are the average values for each harvest. The loss of N during reproduction was used to approximate the general storage release rate, r_g.

Since the main location of the surplus storage in *R. acetosa* is in the tap root, it was assumed that the surplus release proportion r_s may be estimated from the N content in the tap root. The distribution of r_s was defined in terms of the mean and the standard deviation of the set of proportional losses of N from the tap root N_{lt}. For each plant N_{lt} was found by calculating the proportion that the N content in the tap root at the harvest with maximum N was greater than the corresponding value at the harvest

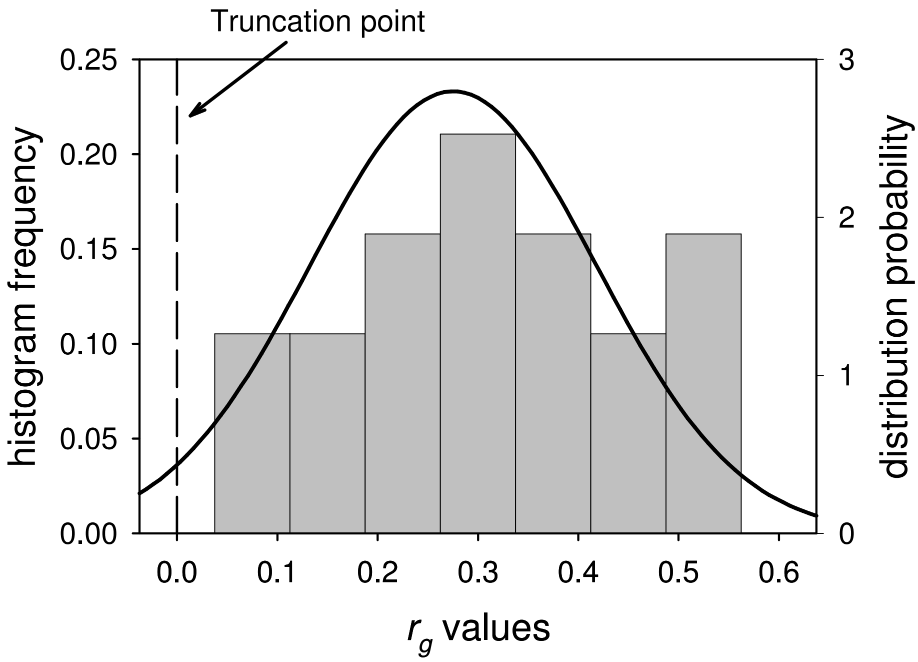

Fig. 3.7 b) The distribution of values of the parameter r_g, and the fitted normal distribution. Values of r_g below the truncation point are rounded up to 0. The truncation is at 0 since r_g is the general store release proportion, and it cannot be less than 0. The frequency of the observed values is plotted along the y-axis on the left. The distribution probability of the normalised normal distribution is plotted along the y-axis on the right.

Figure 3.7 N content in new leaves and fine roots and the obtained distribution for the parameter r_g.

when N was minimum. In other words

$$N_{lt} = \frac{\max\{N_{tap}(j)\} - \min\{N_{tap}(j)\}}{\max\{N_{tap}(j)\}}$$

where $N_{tap}(j)$ is the proportion of N lost from the tap root at harvest j, $1 \le j \le 7$. The distribution was truncated at 1 and 0, since the release proportion cannot be greater than 1 or less than 0.

The time dependent reproduction relation R_t was derived by fitting an inverted parabola to the measured dry weights of the reproductive parts:

$$y = -a(s - s_0)^2 + y_0$$

where s is time, and s_0, y_0, and a are the fitted parameters. The time when a plant started to lose weight in the reproduction parts was used to represent the time of reproduction. This time corresponded to the value of s when the parabola is at the maximum value, *i.e.* at s_0. The values of s_0, y_0, α_2 and β_2 were scaled so that *s* ranges from 0 to 50, in the same way as before. The distribution of R_t was defined in terms of a mean and a standard deviation of the s_0 values.

The dispersal of *R. acetosa* is determined by several factors, such as wind, cattle, and human activity (Grime *et al.* 1988). The dispersal of *R. acetosa* in the model was assumed to be random within a distance of 5 cells away from the reproducing plant. This corresponds to an area of 11 x 11 lattice cells centred at the plant location, or approximately 1 m^2. This area is approximately the smallest lattice size simulated (which was 10 x 10 lattice cells). An area smaller than that would contain an unrealistically small number of sites in which offspring could land given the nature of seed dispersal.

Data for the survival threshold V_t and the survival assessment period V_p were unavailable. These quantities were estimated from general biological considerations: V_t was assumed to be 1/10 of $U_e(s)$ at current development stage *s*; V_p was set to be 5 time steps (G. Squire, personal communication). The value for random death probability P_d was assumed be 0.001. P_d represents the disturbance in the system. Disturbance, in the intermediate range, was shown to promote diversity in communities (Rosenzweig 1995). Hence, the value of P_d was chosen to be in the intermediate range - higher and lower values of P_d reduced diversity in the simulated communities.

The parameter values were assumed to be independent from each other, *i.e.* when a plant was created in the model, any combination of parameter values was possible. Some of these may not be physiologically feasible (*e.g.* low uptake rate and frequent reproduction, etc.). Nevertheless, the relationships between parameters were not included in order to allow relationships between parameter values to evolve in the simulations.