Appendix C. Choosing a model for the parameter *essential uptake*, $U_e(s)$

C.1 Summary of the analysis

This appendix explores the sigmoidal functions that can be used to fit data on nitrogen used by plants at different times. Used N was chosen to estimate essential uptake $U_e(s)$. The analysis is presented as follows.

1. Data description.
2. Preliminary selection. Five types of functions were considered: Gompertz, logistic, Richards, Morgan-Mercer-Flodin (MMF), and Weibull. Qualitative evaluation of the fit showed that the Weibull and MMF were not appropriate models. Several variations of each of the remaining models were fit to the pooled plant as an initial evaluation of the models. The *F*-statistic, R^2, *t*-statistic for the parameters, and residuals plots were examined. This led to the selection of the logistic and Gompertz functions. The parameter values estimated from the pooled data were used as initial values when the models were fit to individual plants.
3. Model evaluation and selection for individual plants. Several versions of logistic and Gompertz functions were considered. For each plant, the parameters were estimated using least squares (Newton-Gauss also known as Marquardt method). To estimate the quality of the selected models the following were calculated for each individual plant and for all models: (a) curvature measures (intrinsic and parameter-effects); (b) Box's percentage bias in LS estimates; (c) parameter correlation matrices. The statistical analysis and biological considerations led to the conclusion that the model used in section 3.3.2 was appropriate.

C.2 Data description

To estimate the essential uptake $U_e(s)$, the amount of N used by plants between harvests (see Section 3.3.2) was used. The harvest was collected at 7 time points (see

Section 3.3.1). The data contained values of used N for 20 plants, at seven time points. Data for one of the plants was not complete, and therefore only 19 plants were used in the analysis. The difficulty with the data was the small number of points for each plant, a common difficulty with biological data.

C.3 Preliminary selection

Plant uptake as a function of plant development stage has a sigmoidal shape (Grime *et al.* 1988). Several functions are commonly used to model sigmoidal growth (Ratkowsky 1983): logistic, Gompertz, Richards, Morgan-Mercer-Flodin (MMF), and Weibull (Table C.1). In this work, several forms of each model were considered. For logistic, Gompertz, and Richards four forms were considered: two with three parameters (models numbered 1 and 3), and two with four parameters (models numbered 2 and 4). Models numbered 1 and 3 were re-parameterized versions of each other, as were models numbered 2 and 4. Similarly for MMF models 1 and 2.

All of the models were fit to the data. Since the data for each individual plant were sparse, the data for all plants were pooled (so that for each time point there were 19 observations for uptake). The pooled data were used to evaluate the qualitative adequacy of the models. For two models, MMF and Weibull, the fit was not satisfactory, as the models did not approach maximum at the last observation point (Figure 1). For the remaining models, logistic, Gompertz, and Richards, the statistics characterizing the fit are summarized in Table C.2. All models had very similar R^2. The value of y_o was estimated to be zero for the Richards models 2 and 4. Therefore, their fit did not differ from Richards 1 and 3.

Table C.1 Common models for sigmoidal growth (Ratkowsky 1983). Several forms of each model were considered.

Gompertz	Gompertz 1	$f(s)=ae^{-e^{s_0-bs}}$
	Gompertz 2	$f(s)=y_0+ae^{-e^{s_0-bs}}$
	Gompertz 3	$f(s)=ae^{-e^{-\frac{s-s_0}{b}}}$
	Gompertz 4	$f(s)=y_0+ae^{-e^{-\frac{s-s_0}{b}}}$
Logistic	Logistic 1	$f(s)=\frac{a}{1+e^{s_0-bs}}$
	Logistic 2	$f(s)=y_0+\frac{a}{1+e^{s_0-bs}}$
	Logistic 3	$f(s)=\frac{a}{1+e^{-\frac{s-s_0}{b}}}$
	Logistic 4	$f(s)=y_0+\frac{a}{1+e^{-\frac{s-s_0}{b}}}$
Richards	Richards 1	$f(s)=\frac{a}{\left(1+e^{s_0-bs}\right)^{1/c}}$
	Richards 2	$f(s)=y_0+\frac{a}{\left(1+e^{s_0-bs}\right)^{1/c}}$
	Richards 3	$f(s)=\frac{a}{\left(1+e^{-\frac{s-s_0}{b}}\right)^{c}}$
	Richards 4	$f(s)=y_0+\frac{a}{\left(1+e^{-\frac{s-s_0}{b}}\right)^{c}}$
Morgan-Mercer-Flodin (MMF)	MMF 1	$f(s)=\frac{dc+as^b}{c+s^b}$
	MMF 2	$f(s)=y_0+\frac{as^b}{c^b+s^b}$
Weibull	Weibull 1	$f(s)=a\left(1-e^{-cs^d}\right)$

To determine goodness-of-fit, the *t*-statistic of the parameters was considered for the logistic and Gompertz models with three and four parameters (models numbered 1 and 2 respectively), and for Richards 1 with three parameters (Table C.3). Since models 1 and 3 were equivalent, the quality of fit needed to be evaluated only for one of the models; similarly for models 2 and 4. Table C.3 shows that for both logistic 2 and Gompertz 2, the *t*-statistic was significant for all parameters. Note that Richards 1 was very similar in form to logistic 2 (Table C.1) with one difference: that the denominator of Richards 1 was raised to the power of $1/c$. The *t*-statistic for Richards 1 was significant for all parameters except c, for which the associated probability was 0.06. Considering that c was an extra parameter added to logistic 2, and that it was the only parameter that did not give a significant *t*-statistic, it did not add to the quality of fit. The number of observations available for each plant was small relative to the number of parameters. Therefore, unnecessary parameters were unjustifiable.

In logistic and Gompertz models with four parameters (Table C.3) the *t*-statistic for y_o was not as significant as it was for the other three parameters. However, biological considerations needed to be taken into account. The parameter y_o could be interpreted as the minimum uptake required for survival, whereas the other term of the equation could be seen to correspond to the uptake necessary for plant growth and development. The minimum uptake required for survival, y_o, was a convenient parameter that could be used to express variability between simulated plants. The other parameter that determined the intercept of the curve, a, changed not only the y-intercept but also the slope of the rise of the curve. Therefore, it did not have as clear a biological interpretation. This suggested that the models with four parameters should be used.

Residuals for all models were examined, and the associated probability that the distribution of residuals were normal was less than 0.08 for all models (according to Shapiro-Wilk and Anderson-Darling tests). This suggested that a further examination of the appropriateness of the model was required, as is described below.

Table C.2 R^2, F-statistic and associated probabilities for data fit to logistic, Gompertz, and Richards models.

	R²	**F**	**p**
Logistic 1 and 3	0.94	685.1667	<0.0001
Logistic 2 and 4	0.94	1033.365	<0.0001
Gompertz 1 and 3	0.94	667.8965	<0.0001
Gompertz 2 and 4	0.94	974.3626	<0.0001
Richards 1 and 3	0.94	689.2573	<0.0001
Richards 2 and 4	Was not considered because value y_o was 0.		

Table C.3 Parameter values and their statistical characteristics for logistic, Gompertz, and Richards models.

Models	**Parameters**	**Parameter Values**	**Std Errors**	**t**	**p**
Logistic 1	a	219.455	6.002	36.567	<0.0001
	b	0.0263	0.0016	16.275	<0.0001
	s_o	3.244	0.1487	21.8212	<0.0001
Logistic 2	a	214.442	10.453	20.514	<0.0001
	b	0.0276	0.0028	9.715	<0.0001
	s_o	3.405	0.336	10.150	<0.0001
	y_o	3.011	5.391	0.558	0.578
Gompertz 1	a	260.306	14.726	17.677	<0.0001
	b	0.0131	0.0012	10.888	<0.0001
	s_o	1.4789	0.0796	18.578	<0.0001
Gompertz 2	a	233.282	14.307	16.306	<0.0001
	b	0.0156	0.0017	9.374	<0.0001
	s_o	1.762	0.158	11.188	<0.0001
	y_o	9.685	3.996	2.424	0.0167
Richards 1	a	209.470	7.352	28.491	<0.0001
	b	0.0392	0.0144	2.718	0.0075
	c	1.934	1.0193	1.897	0.06
	s_o	5.749	2.628	2.188	0.0305

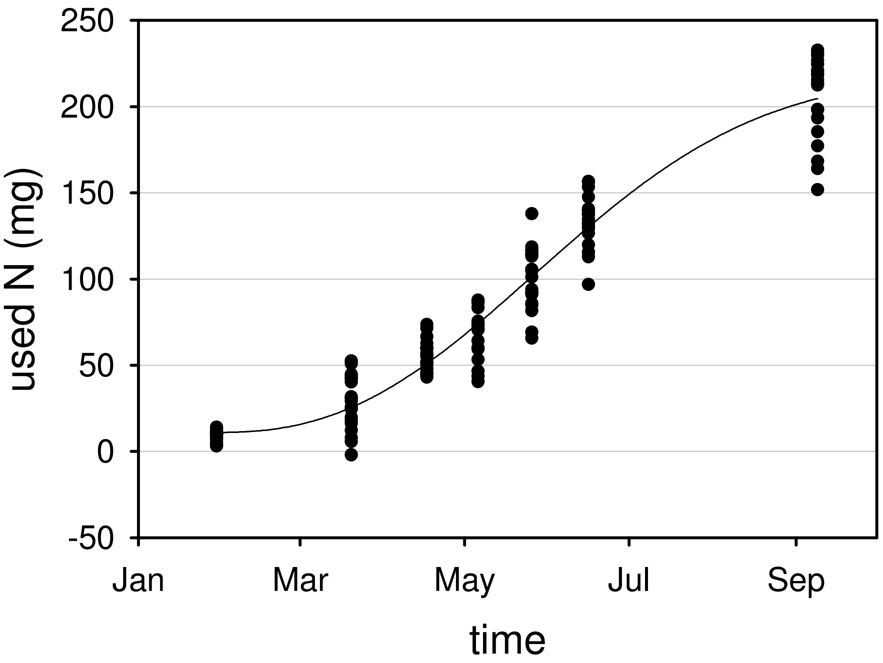

Figure C.1 Fit of Weibull model to pooled data for plants at different harvests. The model does not approach its maximum at the last harvest date.

C.4 Model evaluation and selection for individual plants

As the next step of analysis, logistic and Gompertz models were fit to data for individual plants. The fit was done using the Newton-Gauss (Marquardt) method in Mathematica. In addition, the curvature measures of nonlinearity of Bates and Watts were calculated, as well as the asymmetry measure of parameter bias (Ratkowsky 1983). When the models were fit to individual plant data, the initial parameter values were obtained from the pooled data. The F- and t-statistics and R^2 for the fits to individual plant data were similar to the fit to pooled plant data. The curvature measures of nonlinearity are presented in Table C.4. Overall logistic and Gompertz 1 had smaller values for intrinsic and parameter-effects curvatures than 2. Similarly, logistic and Gompertz 3 had smaller values than models 4. Curvatures of Gompertz models were, in general, larger than those of logistic ones.

Table C.4 shows that the intrinsic curvature for logistic and Gompertz models 1 and 3 was, in general, less than the critical value (i.e. the intrinsic curvature was not significant at the 0.05 level). Intrinsic curvature for logistic 2 and 4 in most cases exceeded the critical values. However, in most cases, the values did not exceed twice the critical value, except marked in grey in the Table where the values were very large.

This occurred in the logistic 4 model, and only when the value of y_o was estimated to be zero.

The parameter-effects curvatures were, in general, bigger than intrinsic curvatures. Moreover, they were, in most cases, larger than the critical values (i.e. the parameter-effects curvature was significant on the 0.05 level). This suggested at least one of the parameters was biased.

In order to find the parameters that caused nonlinearity in the models, Box's percentage bias was calculated for all models (Table C.5). Since the 4-parameter models were the most biologically appropriate, the biases were calculated using the 4-parameter models (numbered 2 and 4) except in those cases when the value of y_o was estimated to be zero. In these cases, the 3-parameter models were used to calculate bias. The values of parameter biases reflect the previous analysis of the parameter-effects curvatures. In particular, the bias values for Gompertz models were higher overall than those for logistic models.

In logistic and Gompertz models 1 and 2, y_o and a had high bias values, and therefore are responsible for the parameter-effects curvature of the models. The other two parameters, b and s_o, have low values of bias in these models, particularly parameter b. The biases of y_o were similar for the logistic and Gompertz models. The bias values for plants 12 and 13 were particularly high for parameter y_o in models 1 and 2, and for all the parameters in models 3 and 4. For most of the plants, the bias for a was higher in Gompertz. Comparison of models 1 and 2 versus 3 and 4 showed that bias values for b and s_o were much higher for the latter models.

Two conclusions can be drawn from the bias analysis. First, that the logistic models had lower curvature and parameter bias. Second, that parameters in models 1 and 2 are less biased than those in models 3 and 4. However, it was again necessary to consider the biological interpretation of the models. In models 3 and 4, the roles of parameters b and s_o could be clearly described. Parameter b was responsible for the steepness of the rise of the essential uptake curve. Parameter s_o was responsible for the time of the rise. In models 1 and 2, the roles of b and s_o could not be so clearly delineated. Therefore, for the purposes of using the selected model to represent the essential uptake curve, models 3 and 4 were preferable.

Finally, the correlation matrices of the parameters were considered (Table C.6). Presented are correlation matrices for pooled data of the logistic 4 and Gompertz 4 models. However, correlation matrices for all models and each plant were examined, and matrices presented in Table C.6 are characteristic of those not shown. For some parameters the correlation was quite high, e.g. between y_o and a, y_o and b, and a and b. However, these correlations were observed for all models, and were likely to be due to the small number of time points.

In summary, logistic model 4 appeared to be the most appropriate model from both statistical and biological points of view. This model performed statistically better than all other non-logistic models. From biological considerations, it was the best logistic model, because it had the necessary parameter structure to represent biological processes.

Table C.4 Intrinsic curvature measurements for all versions of logistic and Gompertz models. The upper and lower values show estimated maximum intrinsic (IN) and maximum parameter-effects (PE) curvatures, respectively, and CV is the corresponding confidence value. IN and PE should be less than or as close as possible to CV. The cells marked in grey are those where the parameter effects were very large.

Plant No.	**Logistic 4**				**Gompertz 4**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
CV	***0.390***	***0.331***	***0.390***	***0.331***	***0.390***	***0.331***	***0.390***	***0.331***
1	0.452 1.408	0.808 4.323	0.452 1.081	0.807 63.09	0.588 3.950	1.111 5.858	0.587 1.870	1.111 15.27
2	0.315 1.533	0.564 4.261	0.315 1.039	0.564 21.53	0.592 5.917	0.669 7.870	0.592 2.096	0.669 7.137
3	0.160 0.563	0.262 1.974	0.160 0.402	6.72 $6.5x10^{16}$	0.189 1.111	0.272 1.289	0.189 0.531	0.272 2.264
4	0.192 1.006	0.335 2.589	0.192 0.643	0.335 6.105	0.415 3.986	0.427 3.396	0.415 1.398	0.105 2.163
5	0.390 1.014	0.734 2.501	0.390 0.665	0.376 11.052	0.568 2.173	1.529 3.229	0.568 1.165	1.529 10.103
6	0.138 1.098	0.239 3.781	0.138 0.639	1160.89 $7.34x10^{18}$	0.277 5.117	0.226 7.402	0.277 1.484	0.226 5.178
7	0.186 1.499	0.343 2.835	0.186 0.861	0.342 3.583	0.321 7.794	0.394 6.460	0.321 2.161	0.0955 2.582
8	0.667 2.190	1.152 5.919	0.667 1.285	1.152 19.02	0.934 3.948	1.496 5.274	0.934 1.962	1.497 8.678
9	0.229 1.006	0.404 2.217	0.229 99.845	0.404 3.536	0.517 3.913	0.498 3.918	0.517 1.480	0.498 2.97
10	0.322 0.912	0.547 3.040	0.322 0.701	0.567 $1.9x10^{14}$	0.340 1.845	0.820 2.816	0.340 0.962	0.820 17.965
11	0.263 2.715	0.492 7.090	0.263 1.470	0.492 14.308	0.420 13.825	0.491 13.429	0.420 3.694	0.491 4.612
12	0.320 5.365	0.612 20.468	0.320 2.601	0.612 109.186	0.386 29.088	0.390 47.007	0.386 7.663	0.39 17.031
13	0.192 49.204	0.335 1.302	0.192 28.126	870.145 $6.31x10^{18}$	0.154 488.61	0.244 $5.1x10_6$	0.154 264.304	0.393 83445.2
14	0.231 2.111	0.408 7.107	0.231 1.178	0.408 70.291	0.395 8.850	0.386 11.115	0.395 2.473	0.368 5.541
15	0.310 1.197	0.475 4.712	0.310 0.860	0.147 $5.19x10^{15}$	0.386 2.602	0.524 4.045	0.386 1.180	0.524 17.307
16	0.257 2.569	0.361 21.063	0.257 1.398	35.477 $2.42x10^{17}$	0.380 7.258	0.226 17.096	0.380 2.090	0.113 54.029
17	0.457 2.011	0.898 2.511	0.457 1.418	0.898 1.859	0.915 11.684	1.292 2.072	0.915 3.747	1.292 0.955
18	0.236 1.462	0.443 2.934	0.236 0.894	0.442 3.943	0.411 6.394	0.436 6.569	0.411 1.994	0.140 3.676
19	0.275 0.928	0.468 3.072	0.275 0.677	0.47 $1.62x10^{11}$	0.337 2.101	0.688 2.735	0.337 1.013	0.688 4.191

Table C.5 Percentage bias in the least squares estimates of the parameters of logistic and Gompertz models. Lines marked in grey have very large values of the bias for at least one parameter.

	Logistic 1, 2				**Gompertz 1, 2**			
	y_o	a	b	s_o	y_o	a	b	s_o
1	-10.679	22.440	0.0043	0.510	-6.344	31.847	0.0031	0.357
2	-6.982	19.353	0.0015	0.211	-5.686	40.667	0.0009	0.139
3		0.769	0.0002	0.028	-0.373	1.671	0.0002	0.023
4	-3.079	6.799	0.0006	0.076	-1.914	9.461	0.0005	0.061
5	-3.798	6.049	0.0061	0.659	-3.371	10.502	0.0046	0.472
6		1.952	0.0001	0.021	-2.521	22.883	0.0000	0.021
7	-2.978	8.278	0.0005	0.076	-3.726	25.712	0.0003	0.054
8	-20.478	31.887	0.0146	1.425	-7.566	24.975	0.0105	1.004
9	-2.598	5.025	0.0012	0.133	-3.238	13.123	0.0009	0.096
10		2.082	0.0012	0.134	-1.734	7.817	0.0016	0.173
11	-8.937	34.740	0.0007	0.149	-7.029	75.483	0.0002	0.091
12	-31.507	152.888	-0.0005	0.261	-32.439	440.640	-0.0015	0.139
13		487.218	0.0001	0.629		18342.4	0.0001	0.333
14	-8.402	31.292	0.0003	0.097	-5.752	52.623	0.0001	0.053
15		3.842	0.0008	0.102	-3.500	16.450	0.0008	0.095
16		9.802	0.0004	0.080	-14.510	104.300	-0.0005	0.013
17	-3.883	7.651	0.0069	0.856	-2.599	7.134	0.0068	0.816
18	-3.379	7.659	0.0011	0.138	-5.223	25.476	0.0005	0.070
19		2.404	0.0007	0.085	-1.678	8.096	0.0011	0.124
	Logistic 3, 4				**Gompertz 3, 4**			
1	-10.678	22.437	5.665	0.674	-10.076	31.849	10.780	6.122
2	-6.981	19.350	4.611	2.516	-3.095	40.671	12.733	9.101
3		0.769	0.218	0.381	-0.311	1.671	0.568	0.290
4	-3.079	6.799	1.867	0.361	-0.497	9.462	3.528	1.865
5	-3.797	6.046	1.357	-0.628	-5.462	10.504	3.182	1.213
6		1.952	0.322	0.917	-1.167	22.883	6.817	5.516
7	-2.978	8.278	1.971	1.082	-0.735	25.706	8.132	5.584
8	-20.514	31.953	8.688	-4.824	-5.215	24.951	9.368	3.776
9	-2.598	5.025	1.467	-0.053	-0.887	13.124	5.540	2.237
10		2.082	0.667	1.016	-9.629	7.812	2.543	1.435
11	-8.938	34.742	7.329	7.373	-1.710	75.428	20.500	18.214
12	-31.514	152.932	32.290	42.345	-9.120	440.689	119.776	117.402
13		487.008	5.566	75.959		18342	218.605	588.433
14	-8.402	31.293	7.004	6.680	-1.839	52.622	15.667	12.752
15		3.843	0.919	1.762	-8.965	16.450	5.245	2.621
16		9.801	1.503	4.564	-26.391	104.290	32.380	22.226
17	-3.884	7.652	1.704	0.281	-0.366	7.136	1.814	0.504
18	-3.379	7.658	2.487	0.597	-1.230	25.470	11.636	5.756
19		2.404	0.616	1.057	-1.247	8.096	2.378	1.376

Table C.6 Parameter correlation matrices for logistic 4 and Gompertz 4 models.

Logistic 4					**Gompertz 4**			
y_o	a	b	s_o		y_o	a	b	s_o
1	-0.839	-0.799	0.172	y_o	1	-0.625	-0.563	0.145
	1	0.890	0.320	a		1	0.927	0.634
		1	0.297	b			1	0.612
			1	s_o				1