Guillermo R. Chantre
José L. González-Andújar *Editors*

Decision Support Systems for Weed Management

MOREMEDIA

Springer

Decision Support Systems for Weed Management

Guillermo R. Chantre • José L. González-Andújar
Editors

Decision Support Systems for Weed Management

Springer

Editors
Guillermo R. Chantre
Departamento de Agronomía
Universidad Nacional del Sur - CERZOS/CONICET
Bahía Blanca, Buenos Aires, Argentina

José L. González-Andújar
Department of Crop Protection
Institute for Sustainable Agriculture
Córdoba, Córdoba, Spain

ISBN 978-3-030-44401-3 ISBN 978-3-030-44402-0 (eBook)
https://doi.org/10.1007/978-3-030-44402-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This book is dedicated to León

Foreword

It is about time! Or, perhaps more accurately, it is about "timing." Safe and successful short- and long-term weed management is highly dependent upon when weed seeds lose or gain dormancy, when they germinate, when seedlings emerge, how fast plants grow, when flowers and seeds form, differential sensitivities to disruption of growth and development during all phases of plant life cycles, and the fickle nature of herbicide fate. No farmer can understand all of these dependencies for even a single weed species. Nor, for that matter, can any individual weed scientist. Failure to comprehend and predict these dependencies helps explain why weeds remain common and usually unwanted residents of agricultural fields even after decades of intense efforts at controlling them. Indeed, by the year 2020, many species of weeds have evolved resistance to various forms of weed control, and they now are not just common, but rampantly abundant in some fields. The sheer volume of literature in Weed Science published during the past two decades pertaining to resistance underscores the fact that this problem is increasing, not diminishing.

Even though no individual person understands all of the variables that affect any weed, groups of weed scientists can come close to doing so. These groups of scientists can collaborate, conceptualize, experimentally test, and develop models that attempt to mimic weed behavior and control. Although models of weed growth and management were initiated many years ago, only some scientific groups continued pursuing this line of research to the present. Many other groups, however, curtailed modeling activities with the advent of genetically modified herbicide tolerant crops. Creation of herbicide tolerant crops represented truly remarkable scientific achievements, and these achievements revolutionized weed management beginning in the mid-1990s in countries that allowed GM crops to be grown. Unfortunately for farmers and weed scientists in those same countries, evolution also is quite remarkable. Selection for weed resistance to herbicides used in GMO-based cropping systems occurred faster and was more widespread than anyone had anticipated. This was, indeed, a sobering development for Weed Science.

Weed resistance to herbicides is not confined to GMO-based cropping systems. Weeds evolve resistance to herbicides whenever and wherever overreliance on herbicides occurs, even in countries that banned GM crops. Consequently, the need for understanding weed biology and management is worldwide in scope, and it is never-ending, as weeds will continue evolving as new cropping systems and weed control techniques are developed and implemented.

Fortunately, small pockets of weed scientists scattered across the globe recognized the continued need for weed models even during the GMO revolution. The continued efforts, intellect, and dedication of those groups are reflected in this book, Decision Support Systems for Weed Management. The book is divided into four parts, each with multiple chapters: (1) Modelling: A Brief Introduction to Decision Support Systems, (2) Bio-Ecological and Site-Specific based models, (3) Environmental Risk Modelling, and (4) Weed Management Decision Support Systems: Study Cases. These parts explain to readers the general and technical aspects of modeling and its utility in Weed Science; historical and recent advances in the modeling of weed behavior and dynamics, crop–weed interactions, and site-specific phenomena; assessments of unintended consequences of weed management, especially herbicide fate and effects; the utility of several highly functional DSS models developed in Australia, Europe, and Latin America. These are truly exciting developments.

In my view, the individual chapters, its sections, and the book as a whole represent the twenty-first-century basis for integrated weed management. In other words, adoption of the concepts, if not the specific models, described in this book will help lead to the sustainable cropping systems that agriculture must have in the future. It is about time!

University of Minnesota
St Paul, MN, USA

Frank Forcella,

Preface

Weed management decision support systems (DSS) are increasingly important computer-based tools for modern agriculture. Nowadays, extensive agriculture has become highly dependent on external inputs, and both economic costs and the negative environmental impact of agricultural activities demand knowledge-based technology for the optimization and protection of nonrenewable resources. In this context, weed management strategies should aim to maximize economic profit by preserving and enhancing agricultural systems resources. Although previous contributions focusing on weed biology and weed management provide valuable insight on many aspects of weed species ecology and practical guides for weed control, no attempts have been made to highlight the forthcoming importance of DSS in weed management. This book is a first attempt to integrate “concepts and practice” providing a novel guide to the state of the art of DSS and the future prospects, which hopefully would be of interest to higher-level students, academics, and professionals in related areas.

Buenos Aires, Argentina — Guillermo R. Chantre
Córdoba, Spain — José L. González-Andújar

Contents

Part I
Modelling: A Brief Introduction to Decision Support Systems

Chapter 1
Mathematical Models

Niels Holst

Abstract Decision support systems (DSSs) rely on computational machinery in which mathematical models often constitute an important part. In this chapter, it is discussed which kinds of models are best suited for different kinds of DSSs. The practical steps involved in model construction are outlined, keeping in mind that model construction is a process that must be integrated into the larger software development project launched to construct the whole DSS. You are invited into the modeller's workshop, as you follow the considerations involved in formulating a simple model of weed emergence. Two case studies close the chapter, demonstrating models of the population dynamics of annual weeds in a crop rotation and of an invasive weed. R scripts for all models can be found in the book's online appendix. It is concluded that weed modellers must be prepared to work in multidisciplinary teams and that they should be better at considering the needs of the DSS users. For purposes of quality control, the mathematical models should be published open-source, while the DSS itself might be proprietary.

Keywords Decision support systems · Model construction · Software development · Weed population dynamics · Invasive weed · Weed modeller · R scripts

---

Electronic supplementary material: The online version of this chapter (https://doi.org/10.1007/978-3-030-44402-0_1) contains supplementary material, which is available to authorized users.

N. Holst (✉)
Department of Agroecology, Aarhus University, Slagelse, Denmark
e-mail: niels.holst@agro.au.dk

G. R. Chantre, J. L. González-Andújar (eds.), *Decision Support Systems for Weed Management*, https://doi.org/10.1007/978-3-030-44402-0_1

1.1 Introduction

We build models to grasp the world and to manage our lives and surroundings. Whether in science or in everyday life, we express ourselves, we rationalise and we communicate by concepts that reflect our perspective on reality. We all have models of the world in our minds, whether we are humans—or bats (Nagel 1974). When we express models in the language of mathematics, we take our more or less fluffy concepts and dip them in the acid of mathematics. Whatever is left stands clearly written in equations. Then truly, *what can be said at all can be said clearly* (Wittgenstein 1922).

The disciplines of mathematical modelling and software engineering are essential to any decision support system (DSS). Mathematical models, which are constructed from mathematics and algorithms, constitute the wisdom of the DSS, while the DSS user interface makes that wisdom accessible in a language and operational mode that is convenient to the user. New DSSs are created in research and development environments by teams comprising experts on the problem domain (e.g. weed control), together with modellers and software engineers.

In a professional setting, the whole software development process is played out according to a well-defined software development protocol, such as agile development (Martin 2006). Ideally, in the early design phase of a DSS, the users and their needs are defined. Once the user problem domain has been delineated, the next step is to identify the modelling approach that will enable the development of models, which can provide information helpful to the user.

An unfortunate but common déroute in DSS development is to let the whole construction process take place in a closed forum of researchers and modellers, who believe that there is a real need for the DSS that they have in mind. When the finished DSS ultimately attracts little interest, they will blame the end users (e.g. the farmers for being too lazy to count weed seedlings and enter those numbers into the DSS). I wish to reiterate what has been said many times yet seems a surprisingly difficult advice to follow: *Before developing a model, make clear what its purpose is*. Add to that: *Before making a DSS, make certain there is an actual need for the guidance it will offer, and that end-users will pay the price in time and money needed to use the DSS*. Private companies would call it a business plan.

To make certain that an initial brainstorm will reveal the full range of possible DSS designs, the matrix in Table 1.1 can be used as a guide. The matrix is defined

Table 1.1 A design matrix for decision support systems (DSSs) with different scopes

	Query	Q&A	Scenarios
	What's the status?	What should I do?	What if?
Tactics	2.1	2.2	2.3
Strategy	3.1	3.2	3.3
Policy	4.1	4.2	4.3

Numbers refer to the subsections which explore these nine types of DSS further
Question and answer (Q&A)

by Conway's (1984) typology which classifies decisions at either tactical, strategic or policy level vs. the kind of the support needed, whether it's a status query, a question and answer session or an exploration of *what-if* scenarios. In the following, I will discuss which modelling approaches are most appropriate for the nine classes of DSS resulting from the combination of the two typologies in Table 1.1. Two case studies and a few recommendations conclude the chapter.

1.2 Models to Support Tactics

Tactical decisions are often the easiest to support with a DSS and also the easiest for which to confirm that a model provides accurate advice. Tactical decisions define a short time frame and a narrow spatial scale (e.g., weed management decisions within a given season and for a specific field). Long-term and larger-scale consequences of one's actions are deliberately ignored. The typical decision maker is a farmer or technical advisor.

1.2.1 Tactical Queries

Basic queries concern weed status: Which species have emerged? At which densities? In which fields? Where in the fields? Previously, these questions were difficult to address except by personal observation, but with the advent of artificial vision and multispectral imaging, weed maps can now be drawn with increasing precision from videos captured by Global Positioning System (GPS)-enabled field equipment or from more or less autonomous rolling or flying drones. The development of mathematical models to extract patterns, such as weed species distributions, from digital images is a ripe research field driven by demands outside agronomy (e.g. military intelligence). This means that a DSS should be designed with future changes in mind; it should be easy to plug in new methods for pattern recognition as they become available.

1.2.2 Tactical Q&A

When the current state of weed pressure has been assessed, whether through high-tech monitoring, visual scouting or personal experience from earlier growing seasons, the question is what to do about it? Thus, a farmer may ask whether weed control is necessary, and, if so, which herbicide/s and dosage/s will provide efficient control or minimise ecotoxicological side effects?

The model to answer such questions would be based on a database of herbicide efficacy for different weed species, maybe even parameters for dose-response relations and corrections for weed growth stage and crop. The simultaneous optimisation on several criteria might be addressed best by optimising each separately, and then let it over to the farmer to take the final decision, weighing the options.

The models defining the optimisation problem might be based on simple regression models that describe dose-response-price-environment relations. However, with all the possible combinations of weed species, crops, herbicides and non-chemical treatment options, these models quickly turn very data hungry. Sensible ways of cutting down on this combinatorial explosion should be addressed early in the process of model development.

In precision agriculture (PA), questions must be addressed at a fine spatial resolution within each field. This will make the optimisation problem more difficult, maybe difficult even to define. Numerical optimisation in itself is a classical discipline within mathematics, physics and computer science. Please, see Chap. 3 of this section for a detailed description on numerical optimisation.

1.2.3 Tactical Scenarios

We most often think of scenarios as something distant and far reaching, but even within the scope of a single field in a single season, different scenarios can be envisaged at the time of weed control. Thus, a farmer may ask, which among the available control options will give the highest yield, by grain or by net income? If we get a dry spring and I do not control the weeds, what will the yield loss be? How much should the price of grain change to make one control tactic economically better than another?

With scenarios, DSS models become more demanding. Maybe the total range of possible outcomes cannot be described by regression models alone. More complex simulation models might become necessary. This will incur additional costs in terms of model development and assessment of model reliability. A DSS in scenarios mode easily gets more speculative, and the user interface more difficult to design to strike the right level of detail and functionality.

1.3 Models to Support Strategy

While tactical decisions are taken in the season as a reaction to imminent weed problems, strategic decisions can be made off-season usually as a simulation exercise. The scale of strategic decisions extends into weed management over several years in the same field and across all the fields belonging to a farm or a landscape. Invasive weeds and the management of weeds in natural habitats are problems that necessitate strategic (and policy) level decisions. When we are developing a DSS for strategic planning, we should be careful to recognise that weed management forms only a small part of farm management and the whole-farm organisation. We should always think carefully about the interface between the DSS and other farm management software to achieve a smooth integration and convenience of use. The typical user is an agricultural consultant.

1.3.1 Strategic Queries

The necessity of a strategy, rather than just simple tactics, for weed management becomes obvious when weed problems escalate above the norm. Common causes are a reduced diversity in crop rotation (in the extreme case, monoculture), an over-reliance on a small subset of herbicides with similar modes of action and, ultimately, the advent of herbicide resistance. For example, a DSS could help by identifying and predicting imminent weed outbreaks. If monitoring data on weed occurrence were logged, together with a log of field activities, then an ideal DSS could issue early warnings which could then inspire changes in weed management strategies. Models for such a DSS would incorporate weed population dynamics analysed either statistically or numerically through simulation. However, it is doubtful whether farmers/advisers really need an early warning system for weeds. Field infestations are obvious to the naked eye, and weed problems will usually announce themselves in a few hot spots before large areas suffer from the infestation. At landscape level, a DSS taking input from remote sensing could point out patches of invasive weeds.

1.3.2 Strategic Q&A

An aspect of weed status that is important for strategic planning yet remains difficult to ascertain is an answer to the query: What is the current prevalence of herbicide resistance? It still seems far into the future that a DSS, fed with drone-collected biomolecular characteristics of weeds, could provide this information. The models underpinning such a DSS would be in the reign of bioinformatics.

A more approachable strategic question might be: will this crop rotation control this weed? Or, if I choose this crop rotation, which weed species would be prevented and which would be promoted? Or, with this rotation of herbicides, will I prevent herbicide resistance building up? A model to answer these questions could be a rather simple simulation model working in time steps of cropping seasons. The model would consist of difference equations describing the mechanisms at a rather coarse level. Even so, it might prove difficult to find solid empirical data to estimate all model parameters. The best course is then to include parameter uncertainty in the model (e.g. by supplying min-max values for all parameters) and use proper methods to derive the resulting uncertainty in model outputs (Saltelli et al. 2008).

1.3.3 Strategic Scenarios

A DSS could provide tools to design a complete weed management strategy, including crop rotations and herbicides, or the full complement of methods used in organic farming. Outputs could include economic performance, yields, weed densities, herbicide-resistance prevalence and environmental side effects. Such a DSS might acquire the flavour of a computer game, in which the user tries to win by fulfilling as many goals as possible, accepting trade-offs according to personal preferences. The model underlying this DSS will be more complex than the previous. A simulation model is clearly called for, and even more detail is needed, reflecting the detail of the scenarios and the outputs.

1.4 Models to Support Policy

Policy models are for decision makers at the highest organisational level. They might be decision makers at international, national or regional levels, or decision makers working for the interest of non-governmental organisations (NGOs), such as farmer organisations or nature conservation societies. For a modeller, it can be a frightening experience to develop models that will feed into decision processes affecting society at large, even though economist modellers seem less challenged by this prospect. Policy models play such a powerful role in modern society that they have been put in their own category dubbed *post-normal models* (Funtowicz and Ravetz 1993). For policy models, it is of particular importance to include uncertainty in model inputs (and consequently, in model outputs) to prevent abuse of the models by overzealous policymakers. In a democratic society, the models should be open-source since they are used to formulate arguments in the public debate.

1.4.1 Policy Queries

At policy level, queries are made to identify problems and motivate the formulation of policies. Thus, one may ask, what is the current distribution of an invasive weed? What is the current use of herbicides per year, and in which crops are they applied? Queries such as these can often be translated into queries into databases. Thus, the underlying model is within the range of software engineering, maybe overlaid with descriptive statistics.

1.4.2 Policy Q&A

Weed management policies are formulated with an eye to political goals, preferentially those that can be formulated in terms of performance indicators: production quantity and quality, farmer economy, environmental side effects, etc. Political instruments are foremost economical (taxes, subsidies) but also include indirect measures such as education and research. This means that even the simplest question (e.g. if herbicide taxes were increased in proportion to their ecotoxicity, what are the consequences on farmer economy and the environment?) will involve several fields of knowledge (agronomy, ecology, economics, sociology). The corresponding models will tend to be rich in assumptions and parameters estimated by expert opinion. Model outputs will be equally rich and challenging to condense into information useful to the decision maker.

It is very difficult to construct a policy model with a clear rationale for which components and mechanisms should be included and at which level of detail. Model uncertainties must be included in the DSS outputs, but there is a high risk of uncertainty being caused by structural faults (i.e. the exclusion or misrepresentation of key elements and key processes), which cannot be diagnosed by formal methods but only by scientific argument. Structural faults might lead to biased outputs, as will be pointed out soon enough by political combatants. The modeller must be prepared to defend in public the scientific base of a policy model.

1.4.3 Policy Scenarios

When even the simplest policy question leads to models of high complexity, the modelling of policy scenarios will lead to even higher complexity. We quickly reach the limit of what can be modelled with some confidence—and within a weed research budget. The conscientious modeller confronted with a demand for a model of such immense complexity should consider to decline the order.

A problem domain bordering that of weed management is pesticide legislation and regulation. Legislators and land-use administrators are in need of information on the fate of herbicides in the environment (e.g. persistence, leakage to ground and surface water) and on the magnitude of their unwanted side effects (ecotoxicological and human toxicological). A DSS to support these policymakers would incorporate models of the physicochemical pathways of herbicides in air, soil and biota and the derived effects on exposed populations (by necessity including only a few key species from selected taxa). The information provided by such a DSS could be used to formulate laws and regulations on herbicide use, including the possible banning of a specific herbicide. Due to the vast economical interest in herbicides, represented by farmers and pesticide companies, and the skepticism of NGOs representing a variety of interests, the modeller should be prepared that the DSS will be playing part in a complex political theatre.

1.5 Model Development

For models that consist of a few regression equations or other statistical measures, the modelling procedure falls inside common research practice. The only challenge will be to communicate with software developers on how to embed the statistics in a DSS. For models that are simply queries into a database, the software engineer is in command and will need the weed modeller only as a consultant to assist in the proper interpretation of the data.

The really demanding models are simulation models. Since they will need to be embedded in dedicated DSS software, their implementation will become an integral part of a commercial-scale software development project. The best software design will ensure a loose coupling (Seemann 2012) between the DSS user interface and the simulation model, both kept in separate modules. This will allow independent development of the DSS and the model. Furthermore, it will allow the DSS code to be proprietary (i.e. owned by a company or institute) and the model code to be open-source and thereby open for scientific publication and public scrutiny.

Model development goes through a series of steps, generally acknowledged in the modelling community and outlined in the following: formulation, parameter estimation, verification, testing, validation, uncertainty analysis and sensitivity analysis.

1.5.1 Formulation

Simulation models are formulated in the language of mathematics and logic. They should be based on the theoretical concepts of the topic and should re-use earlier models or sub-models when possible. If the model contains many interacting components, consider software engineering methods to manage the complexity (reviewed by Holst and Belete 2015).

An important part of model formulation is parameterisation. This term is most often used in the wrong sense to mean ‘parameter estimation’ (see next subsection). What it means properly is to ‘formulate in terms of parameters’. For example, you may need a hump-shaped curve to represent a process such as seedling emergence rate through time (Fig. 1.1 top). You can formulate that as a parabolic curve using the standard parameterisation

$$y = ax^2 + bx + c \tag{1.1}$$

This parameterisation, however, has the problem that none of the parameters represent a biological feature of seedling emergence. A better parameterisation describes the curve by its start (x_{begin}) and end (x_{end}) on the x-axis and by its maximum on the y-axis (y_{max}). Equivalent to Eq. (1.1), we get

$$y = 4y_{max} \frac{(x - x_{begin})(x - x_{end})}{(x_{begin} - x_{end})(x_{end} - x_{begin})} \tag{1.2}$$

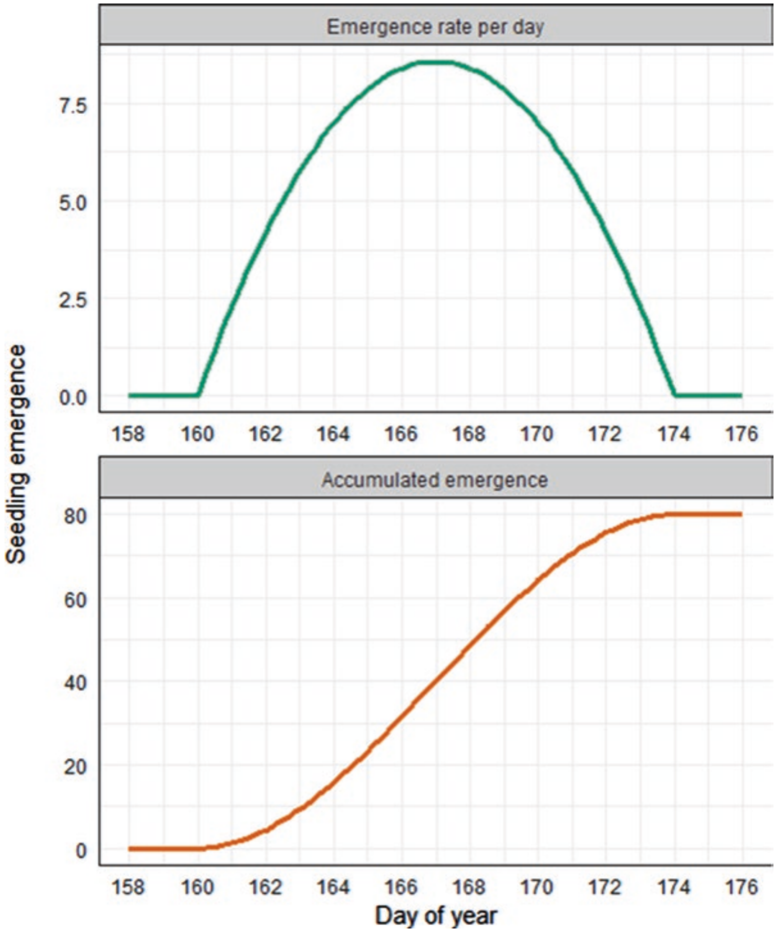

Fig. 1.1 A hump-shaped curve to describe seedling emergence rate (Eq. 1.3, top) and accumulated seedling emergence (Eq. 1.4, bottom) with $x_{begin} = 160$, $x_{end} = 174$ and $y_{total} = 80$. Implemented in the dsswm-1-1.R script

Yet, as you begin to use this equation, you may realise that y_{max} is not a convenient parameter. The area under the curve, expressing total emergence (y_{total}), would be a much better parameter. Hence, you proceed to integrate Eq. (1.2) and replace y_{max} with y_{total} and finally get

$$y = 6y_{total} \frac{(x - x_{begin})(x - x_{end})}{(x_{begin} - x_{end})(x_{end} - x_{begin})^2} \tag{1.3}$$

The benefits of this parameterisation are plenty. The user of the model (foremost yourself) can now estimate, communicate and change the parameters with a clear rationale. Moreover, in a sensitivity analysis, the uncertainty induced by parameters

x_{begin}, x_{end} and y_{total} will have a direct biological interpretation. Compare that to model uncertainty caused by a, b and c (Eq. 1.1) which would be difficult to interpret.

Note that both Eqs. (1.1) and (1.3) have three parameters. Thus, they have the exact same level of complexity (in fact, they are equivalent). You might want to add an additional parameter to Eq. (1.3) to obtain a skewed emergence curve, but with every parameter you add to a model, you incur an increasing debt of parameter estimation. If the curve is used to describe the course of seedling emergence in the field, there will be so many mechanisms not accounted for (weather and soil being the most important ones), that further detail is not merited. The detail of model formulation should match the detail in the information available about the real system. Modelling of the more intricate details of seed bank dynamics is dealt with in Chap. 4.

The curve (Fig. 1.1 top) has a superficial similarity with the normal distribution (which would also demand three parameters: mean, standard deviation and y scaling), but Eq. (1.3) has the advantage, that it has well-defined zero limits and is easily integrated if needed (as seen in the following, Eq. 1.4). In comparison, the normal distribution never reaches zero (an additional parameter would be needed), and it has no analytical integral.

1.5.2 Parameter Estimation

Often, model parameters are estimated by standard statistical procedures, such as linear or nonlinear regression. For the emergence model (Fig. 1.1 top), for example, you could regress observed cumulative emergence (Y) on the integral of Eq. (1.3) (Fig. 1.1 bottom):

$$Y = \frac{y_{\text{total}}\left(3x_{\text{end}} - x_{\text{begin}} - 2x\right)\left(x_{\text{begin}} - x\right)^2}{\left(x_{\text{end}} - x_{\text{begin}}\right)\left(x_{\text{begin}} - x_{\text{end}}\right)^2}, \tag{1.4}$$

which expands to a third-degree polynomial. The estimated polynomial coefficients can be used to calculate the three parameters: x_{begin}, x_{end} and y_{total}.

In an early stage of model development, you should consider whether the DSS ought to include uncertainty. If so, each uncertain parameter must be described by a distribution (e.g. uniform between min-max values or normal defined by mean and standard deviation). Be suspicious, in particular of parameter values that are expert opinions (guesses). Ask the expert up front for parameter ranges or distributions rather than simple point values.

In the case of parameters estimated by regression of equations such as Eq. (1.4), you cannot always use the standard error of the coefficients to generate random parameter values independently. If the standard errors of the regression parameters cannot be considered independent, you must use the regression model itself to draw random values from the predicted distribution of y given x. For the particular case of the emergence model, however, it does seems reasonable that the three parameters vary independently, though you might choose to replace x_{end} with $x_{begin} + x_{duration}$, where $x_{duration}$ designates the duration of the emergence period.

Some parameters are best estimated from the model itself, a process commonly called *calibration*. This is a somewhat dubious activity: You make the model fit your expectations, usually empirical data, by fine-tuning one or several model parameters. The more parameters you calibrate, the less confident you should feel about the general applicability and robustness of the model.

1.5.3 Verification

In model verification, you check that model behaviour makes sense. For this purpose, you define a series of parameter sets (often considered model *scenarios*) of increasing complexity and proceed, more or less formally, to check that model outputs look right. In other words, you check that model outputs could be true, that the behaviour of the model makes sense. Negative, zero or infinite weed densities are typical examples of model fragility discovered during verification. You proceed by mending the model as needed to pass verification. Do find the root cause of any problem and implement a scientifically sane solution. Do not thoughtlessly use this solution, often hiding in model code, *if (x < 0) then x = 0*, or other hacks like it.

1.5.4 Testing

Testing is an important discipline in computer science, even to the degree that the whole software development process can be centred around it (Beck 2002). Software testing is not a part of model development as such, but the testing of the DSS easily becomes intertwined with model verification. As a modeller, you should be prepared to supply the software engineers with *unit tests*: Each unit test defines the output values expected from certain input values. This makes it possible to automatise the test procedure. A less favoured method, nowadays, of software quality assurance is *debugging*, which is an unsystematic stress test traditionally carried out by the programmer.

1.5.5 Validation

Validation does not mean proof of model correctness; rather, it is the comparison of model outputs with independent field data. Thus, validation aims to convince peers that there is a robust and rather accurate match between model predictions and the real world. It is wise always to make a plan for model validation in the early phase of model development. The model design should be accommodated to make a final validation possible, often by restricting the model's scope and the modeller's ambitions.

1.5.6 Uncertainty Analysis

If some model parameters are best described not by a single estimate but by a distribution (e.g. normal), reflecting uncertainty due to statistical error or natural (irreducible) variation, then model output will be distributed as well. Users will be familiar with uncertainty from weather forecasts which may predict, for example, a 20% chance of rain tomorrow. Likewise, a weed DSS may predict, for example, that a yield loss >10% is highly improbable due to a risk level of only 1%. Models that include uncertainty make most of their assumptions transparent as they shine through in the recommendations issued by the DSS. Since DSSs are meant to be reliable tools, one should not be shy of situations which produce a very wide range of responses. Sometimes, the future may be unpredictable in essence. That information can also be useful.

1.5.7 Sensitivity Analysis

Sensitivity analysis is a step following up on uncertainty analysis. In sensitivity analysis, the uncertainty in model outputs is apportioned to the inputs thus identifying those inputs, that are most decisive for model uncertainty (Saltelli et al. 2008). Sensitivity analysis is usually an academic activity related to the scientific publication of the model, but it could potentially be useful as a DSS feature. The user could be told how much certainty would be gained in DSS outputs by giving more precise estimates of, for instance, weed density or herbicide resistance.

1.6 Case Studies in Model Development

1.6.1 A Difference Equation Model for Annual Weeds

Weed populations tend to be highly dynamic; fast establishment and rapid proliferation are part of being an *r* strategist. Once established, the soil bank of seeds or shoots becomes a constant source of potential outbreaks. The long-term management of weeds is a strategic problem raising questions such as: Which level of seedling mortality will be necessary to reduce the infestation? Or, will this change of the crop rotation help to regulate the weeds? Modellers themselves have for a long time been prolific developing models to answer such questions (Holst et al. 2007). In the following, I will go through the steps of developing a classical iterative model which moves forward in steps of 1 year:

$$S = \left(1 - \mu_{\text{soil}}\right)\left(S_{\text{prev}} - E_{\text{prev}} + P_{\text{prev}}\right) \tag{1.5}$$

The equation computes this year's seed bank (S; m^{-2}) from the previous year's seed bank (S_{prev}; m^{-2}), emergence (E_{prev}; m^{-2}) and seed production (P_{prev}; m^{-2}), while undergoing a certain basic seed bank mortality (μ_{soil}; y^{-1}).

From the seed bank, a certain proportion (ϵ; y^{-1}) will emerge as seedlings (E; m^{-2}):

$$E = \epsilon S \tag{1.6}$$

of which again a certain proportion (μ_{control}; y^{-1}) will be killed by weed control measures:

$$N = \left(1 - \mu_{\text{control}}\right)E \tag{1.7}$$

to leave some plants surviving (N; m^{-2}) to produce new seeds (P; m^{-2}):

$$P = \frac{f_\infty}{1 + \dfrac{f_\infty - f_1}{f_1 N}} \tag{1.8}$$

The two parameters describing fecundity are f_1 (seeds per plant), which is the expected number of seeds produced by one plant growing in competition with the crop only, and f_∞ (seeds per m^2), which is the maximum number of seeds produced in competition with the crop by an infinite density of weed plants.

It is not obvious from this formulation (Eqs. 1.5–1.8) that the model is composed of *difference equations*. A mathematically more concise formulation makes this clear. Equation (1.6), for example, could be written more correctly as

$$\Delta E = \epsilon S \Delta t \tag{1.9}$$

where ΔE (m^{-2}) expresses the change in E, i.e. the difference to be added to E over the time step $\Delta t = 1$ year. Note that multiplication with Δt is necessary to make the units right; Equation (1.9) corrects Eq. (1.6) also in that sense. However, here, we will maintain the slightly incorrect formulation (Eqs. 1.5–1.8) as this is commonly found in literature. The left-out multiplications with 1 will have no effect other than to annoy finicky mathematicians, who would in any case likely prefer a differential equations formulation. To continue with Eq. (1.6) as an example, this would look like

$$\frac{dE}{dt} = \epsilon S \tag{1.10}$$

If you are mathematically skilled, the option stands open for you to build a *differential equations* model, rather than a difference equations model, but for most weed modellers, this is not the case.

Equations (1.5)–(1.7) are all linear which makes them easy to comprehend. Equation (1.8) was given a nonlinear form to take into account density dependence; fecundity per plant decreases with increasing plant density until an asymptote is reached.

Always verify that the shape of your equations makes sense in the real world. For a nonlinear equation, make a plot to ascertain its shape and check its limits both graphically and algebraically. In the case of Eq. (1.8), we get meaningful boundary conditions:

$$\begin{aligned} P &\to f_\infty \quad \text{for} \quad N \to \infty \\ P &= f_1 \quad \text{for} \quad N = 1 \\ P &\to 0 \quad \text{for} \quad N \to 0 \end{aligned} \tag{1.11}$$

Not all verification turns out as successful. For instance, Zwerger and Hurle (1989) proposed an alternative to Eq. (1.8):

$$P = Nae^{-bN} \tag{1.12}$$

for which $P \to 0$ for $N \to \infty$. At high weed density, seed production P (m^{-2}) goes towards zero. A self-defeating weed!

A model consisting of Eqs. (1.5)–(1.8) is called an *iterative* model because you run a simulation by repeatedly computing Eqs. (1.5)–(1.8), thereby updating the four state variables of the model (S, E, N, P) iteratively in time steps of 1 year. It is a *stage-structured* model of *population dynamics* since the population is divided into separate life stages (S, E, N) which are simulated dynamically.

A model needs to be started from some initial state. In this case, we need initial values for S_{prev}, E_{prev} and P_{prev}. More importantly, we need to estimate the values of the model parameters. For this model, there are only few parameters: μ_{soil}, ϵ, $\mu_{control}$, f_∞ and f_1. The task of parameter estimation seems simple until you realise that some of the parameters are likely to depend on the crop. Moreover, they are all liable to differ between years and locations. To accommodate this inherent variability of the parameters, we will define their values as ranges rather than point estimates, some of them specific to the crop (Table 1.2).

The values in Table 1.2 are the expected average values, originally given without indication of their standard errors (Zwerger and Hurle 1989). However, both $\hat{f}$ and $\hat{\epsilon}$ will certainly vary markedly between fields due differences in soil and weather. To capture this uncertainty in the model, we pick values at random inside intervals defined as

Table 1.2 Weed life history parameters from Zwerger and Hurle (1989), except* from CABI (2019)

	Spring barley	Maize	Winter wheat	Any crop
Alopecurus myosuroides				
Fecundity $(\hat{f})$	*1000	*1000	*1000	
Emergence $(\hat{\epsilon})$	0.040	0.034	0.050	
Soil mortality (μ_{soil})				0.81
Avena fatua				
Fecundity $(\hat{f})$	*200	*200	*200	
Emergence $(\hat{\epsilon})$	0.240	0.240	0.230	
Soil mortality (μ_{soil})				0.87
Fallopia convolvulus				
Fecundity $(\hat{f})$	192	1855	93	
Emergence $(\hat{\epsilon})$	0.043	0.020	0.078	
Soil mortality (μ_{soil})				0.16
Galium aparine				
Fecundity $(\hat{f})$	3	100	40	
Emergence $(\hat{\epsilon})$	0.036	0.010	0.037	
Soil mortality (μ_{soil})				0.20
Lamium purpureum				
Fecundity $(\hat{f})$	32	300	280	
Emergence $(\hat{\epsilon})$	0.013	0.017	0.023	
Soil mortality (μ_{soil})				0.16
Thlaspi arvense				
Fecundity $(\hat{f})$	60	630	330	
Emergence $(\hat{\epsilon})$	0.073	0.021	0.043	
Soil mortality (μ_{soil})				0.08
Veronica persica				
Fecundity $(\hat{f})$	150	200	150	
Emergence $(\hat{\epsilon})$	0.079	0.066	0.030	
Soil mortality (μ_{soil})				0.50

$\hat{f}$: seeds per plant; $\hat{\epsilon}$: y^{-1}; μ_{soil}: y^{-1}

$$f_1 \in \left[\frac{\hat{f}}{5}; 2\hat{f}\right[$$

$$f_\infty \in \left[10\hat{f}; 100\hat{f}\right[$$

$$\epsilon \in \left[\hat{\epsilon}/5; 2\hat{\epsilon}\right[$$

Limits for random numbers are conventionally closed-open; [*a*; *b*[designates an interval including *a* and excluding *b*.

For soil mortality (μ_{soil}), we will use the point estimates (Table 1.1) without any variance. We will assume that the efficacy of weed control vary quite much picking random values, $\mu_{soil} \in [0.6; 0.9[$.

Since we let four of the parameters vary randomly, our model is a *stochastic* model; it will not always give the same result. Hence, we have to run it many times to assess the uncertainty in its predictions (Fig. 1.2). During model verification, it was found that two of the weed species were dying out (ALOMY, AVEFA). Hence, the parameter values from Zwerger and Hurle (1989) were replaced with values roughly taken from CABI (2019) (Table 1.2).

The first impression of Fig. 1.2 is that the uncertainty is much larger for some species (ALOMY, AVEFA, VERPE) than for others. Note that two units on the *y*-axis correspond to variation by a factor of 100. It could be of interest to know which of the model parameters are causing this huge variation. This could be resolved by a sensitivity analysis (Saltelli et al. 2008). Some species exhibit fluctuations clearly provoked by crop rotation (FALCO, THLAR), more clearly seen for seedling than for seed bank density. This makes sense because different weed species are known to emerge either in spring or autumn sown crops, or in both.

During the 24 years covered by this simulation, most species are attaining an equilibrium density, THLAR most quickly, AVEFA most slowly. It is difficult to imagine, however, how knowledge of the equilibrium density could be interesting from a DSS perspective. We would rather like to help the farmer to achieve the situation illustrated by GALAP for which density is decreasing in this scenario; it is a weed under control. It would be wise though to consult with weed experts and discuss whether this GALAP scenario seems realistic (a belated verification of the model).

Simulation experiments with the crop rotation and a sensitivity analysis could help suggest effective control strategies for these weed species. The model could be incorporated into a DSS, allowing the farmer/advisor to address problematic weed species through strategic means, rather than the purely tactical which entails giving up on controlling the seed bank (left-hand side of Fig. 1.2) and just limiting its expression (right-hand side of Fig. 1.2), which after all is the ultimate cause of yield loss.

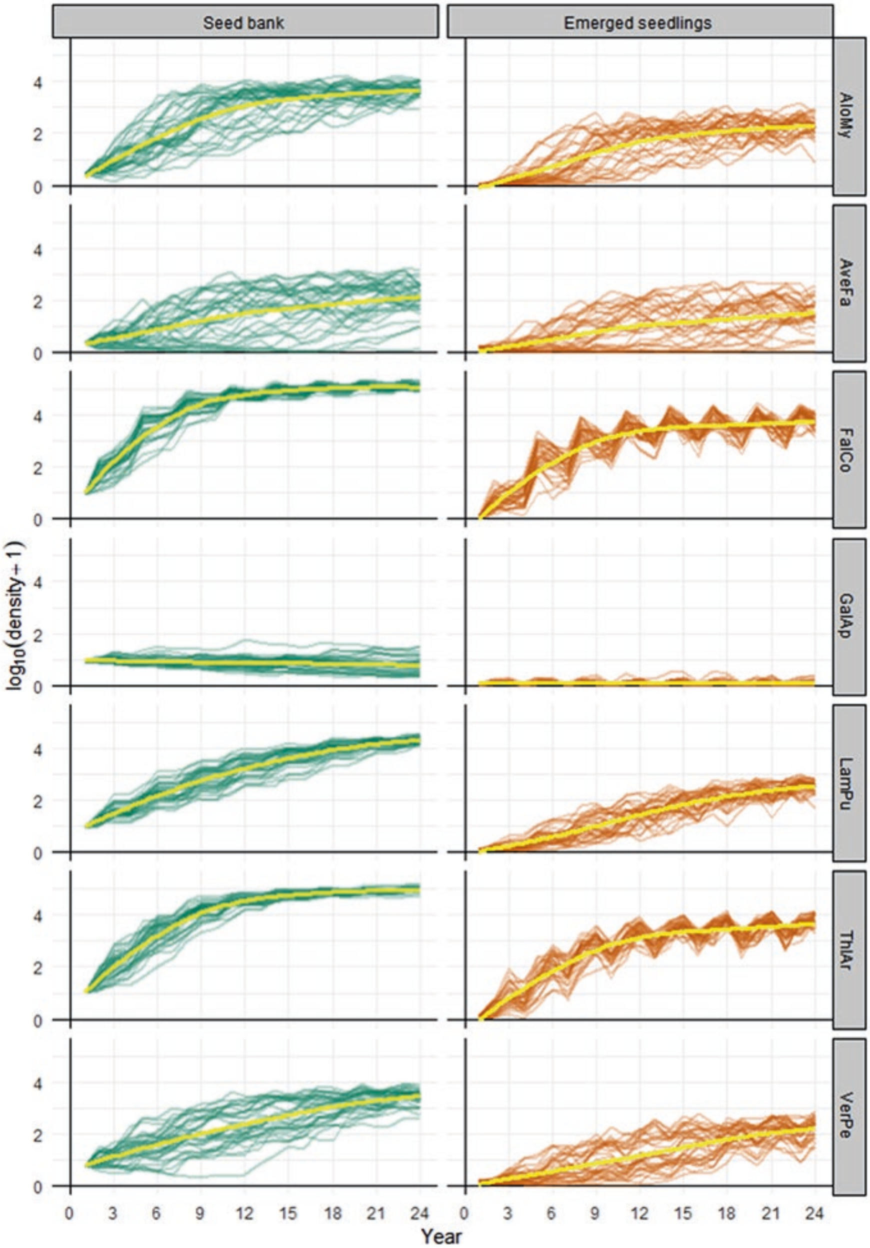

Fig. 1.2 The result of 30 simulations of the crop rotation, maize-winter wheat-spring barley. Yellow curves show smoothed averages. For full species names, see Table 1.2. All populations started with ten seeds per m^2. Model formulated in Eqs. (1.5)–(1.8) and implemented in the dsswm-1-2.R script

1.6.2 A Matrix Model for a Perennial Weed

Matrix models are a class of models which summarises the life history parameters of a population in a single matrix (Table 1.3), a so-called Leslie matrix (Leslie 1945). In the columns, you find the fate over one time step of individuals according to life stage. Column sums <1 account for mortality, and column sums >1 account for reproduction. Likewise, rows show the origin of individuals entering the different life stages. Numbers below the diagonal describe life stage progression; above the diagonal, life stage regression; and on the diagonal, life stage conservation. In this concrete matrix, the seven stages are a mixture of life stage, age and size classes.

Since this is a deterministic model, only one run is necessary to explore what happens after an initial introduction of ten seeds (Fig. 1.3). Notice that the model is linear which means that it gives the same result, whether we consider the simulated population dynamics pertinent to the whole population or to, say, 1 m^2.

The two fields seem clearly different (Fig. 1.3). In field L, the population is increasing, approaching exponential growth after *c.* 10 years. In field J, the population is decreasing, approaching a negative exponential decline after *c.* 5 years. In theory, these matrix models will converge towards a state in which all life stages grow (or shrink) exponentially with the same growth rate, namely the *intrinsic rate of increase* (r) known from the classical model of unlimited growth:

$$N_t = N_0 \exp(rt) \tag{1.13}$$

When r has stabilised, so has the relative proportion of the population in each stage; the *stable stage distribution* has been reached. It follows that when the stage distribution is not stable, then r is not stable either. This is obvious from the simulation (Fig. 1.3); otherwise, all the points would have fallen on a straight line. Note though that the y-axis transformation bends the exponential decrease in field J towards zero. The population density in field L initially oscillates (Fig. 1.3), but the reason behind these oscillations is different than for the oscillations in the previous model (Fig. 1.2). Here, it is due to the unstable stage distribution, and there, it was due to crop rotation.

Leslie matrix models can be analysed mathematically which was part of their original motivation. Thus, the first eigenvalue of the Leslie matrix equals exp(r), and the first eigenvector holds the stable stage distribution. For fields L and J, we get $r = 0.18\ y^{-1}$ and $r = -0.49\ y^{-1}$, respectively, which match the values arrived at by Werner and Caswell (1977).

The lines produced by these growth rates on a log scale (Eq. 1.13) are shown in Fig. 1.3, from the time about when the populations reach their stable stage distribution (again, the line is distorted on the approach towards zero for field J). The initial population size (N_0 in Eq. 1.13) was chosen to let the line pass through the average population size through years 10 to 20 for field L and through years 5–20 for field J. The stable stage distributions are computed in the dsswm-1-3.R script.

Table 1.3 Leslie matrices for *Dipsacus sylvestris* (from Werner and Caswell (1977)) estimated for two fields, L and J

	Seed0	Seed1	Seed2	RosetteS	RosetteM	RosetteL	Flowering
Field L							
Seeds0	0	0	0	0	0	0	503
Seeds1	0.43	0	0	0	0	0	0
Seeds2	0	0.97	0	0	0	0	0
RosettesS	0.01	0.021	0.005	0	0	0	0
RosettesM	0.036	0.003	0	0.19	0.253	0	0
RosettesL	0	0	0	0.07	0.105	0.15	0
Flowering	0	0	0	0	0.002	0.517	0
Field J							
Seeds0	0	0	0	0	0	0	476
Seeds1	0.423	0	0	0	0	0	0
Seeds2	0	0.987	0	0	0	0	0
RosettesS	0.024	0.009	0.006	0.007	0	0	0
RosettesM	0.044	0	0	0.05	0.158	0	0
RosettesL	0.001	0	0	0.002	0.008	0	0
Flowering	0	0	0	0	0	0.25	0

Seeds of age 0, 1 or 2 years. Rosettes of size: small, medium or large. Diagonal cells greyed for easier reading

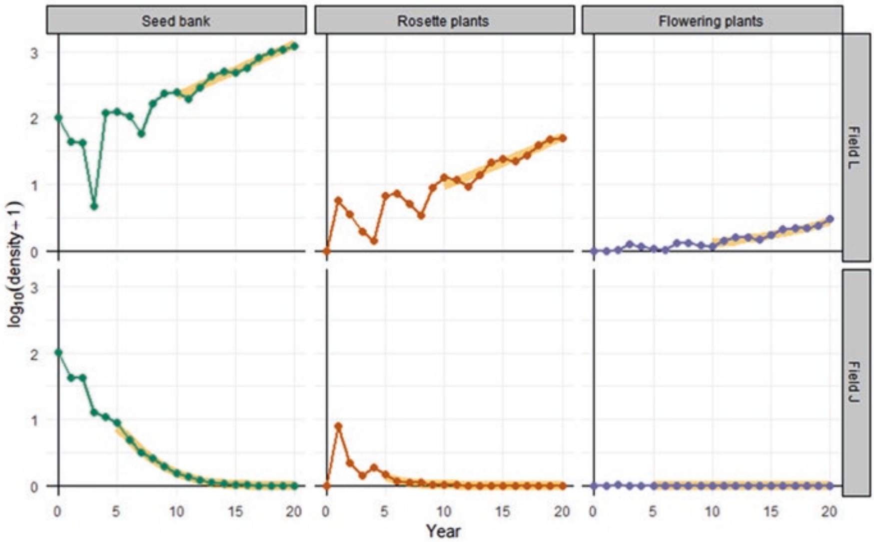

Fig. 1.3 The result of a simulation starting with ten seeds of *Dipsacus sylvestris* running for 20 years based on Leslie matrices for field L and J. Seed bank numbers are the sum of all three age classes of seeds. Rosette plants are the sum of all three size classes of plants. Orange lines show the asymptotic population growth rate. Implemented in the dsswm-1-3.R script