2nd **REFRESHER COURSE OF THE INTERNATIONAL COURSE ON PLANT PROTECTION**

Development of IPM

Dr. S.M. Raterink
Dr.Ir. R.A. Daamen
Ir. P.D. de Jong
Dr.Ir. J.A.B.M. Theunissen

ZAMORANO, HONDURAS

November 28 - December 10, 1994

ZAMORANO ZAMORANO ESCUELA AGRICOLA PANAMERICANA
DPV DEPARTEMENTO DE PROTECCION VEGETAL

IAC INTERNATIONAL AGRICULTURAL CENTRE

DEVELOPMENT OF IPM 1995

Content: Page

PREFACE

The lecture notes on "Development of IPM" are intended to provide the basic scientific background information for the lectures as well as the practicals in a summarized form.

The underlying new philosophy of IPM is that we need to spend far more effort defining the problem in the field, and only then search for appropriate solutions. Integrated Pest Management has, therefore, also been denounced as "Intelligent Pest Management", meaning that we need a lot of knowledge for a good functioning pest management.
Of course, an IPM system can be structurized in different ways. We concentrate here on the analysis of the decision making level: What kind of information is required (and how to obtain it) in order to be able to make a correct decision on the farmer's side about a yes/no control.

In the toolbox scheme (at the very end of these lecture notes) you'll find the various steps and tools necessary for reaching the decision making level: The point of departure for all IPM activities is the correct identification of the pest or constraint, the knowledge of its biology and the knowledge of possible interactions with the host plant (already treated elsewhere in this course). Using this knowledge as a tool, we then can execute crop loss assessments and forecasting and we can study the behaviour and development of the pest (epidemiology and population dynamics) in order to find out the best moment of control and to predict possible losses. Since we cannot study or analyze the entire pest and plant population, we need statistically sound procedures to ascertain that our information gathered from the field is correct (sampling, statistical data analysis).

The individual chapters of these lecture notes elaborate on these tools and steps in more detail, following a logical sequence within the fact-finding process of IPM, while the principles of developing an IPM system are treated in a more general way. The objective is to discuss and train the employment of basic methodologies and tools which enable the students to work out and follow up an IPM programme of their own.

Since the lecture notes are more of a generalized and scientific nature, examples from the practice will be presented in a separate reader as well as in additional field excursions.

1. INTRODUCTION TO IPM

1.1 Goals of IPM

The aims of Integrated Pest Management are twofold:

1. to reduce the risk of disease and pests outbreaks by **preventive** measures. The time-span of preventive measures is usually long, it comprises the use of sound crop rotations in large areas, and e.g. the choice of resistant cultivars. To emphasize this long time span, actions taken in this context are called **strategical**. A "low risk of disease and pest outbreaks" is synonymous with "a high resistance of the agroecosystem".
2. to **control** outbreaks of pests and diseases preferably by mechanical and biological methods. If these methods are not available, selective pesticides are used only then when their use is economically justified. Measures taken affect one field in one cropping period. To emphasize this short time-span and small scale, actions taken in this context are called **tactical** or **operational**.

IPM is part of Integrated Farming Systems, in which the goal of agriculture is shifted from a **single goal**, maximal production, to a **multiple goal**, cost reduction and improvement of quality of both products and production ways, through substituting expensive and potentially noxious inputs, such as fertilizers and pesticides, by both agricultural and ecological knowledge, labour and non-chemical husbandry techniques. In regions were nature is rare, agriculture should also encourage and conserve the flora and fauna in and around fields, to stabilise the agro-ecosystem as a major preventive measure against outbreaks of pests, weeds and pathogens.

Thus the word **integrated** refers to the fact that agricultural activities are directed to more goals than maximal production alone. Now what is **management**? Consider the crop and disease/pest system X, T and Y, which characterize the complete production system. X --> T --> Y are causal relations, thus if X changes then T changes and Y changes consequently. X are the variables which represent the environment (e.g. weather, fertilizers, pesticides, labour; the input variables), T are the variables which determine the state of the black box (soil, cultivar, growth stage, diseases/pests, antagonists/predators/parasites), and Y are the variables which determine the output (yield, waste products, income):

```
              ----------
             I          I
X --------> I     T    I  ---------> Y
             I          I
              ----------
```

in which X are the input or driving variables
T are the variables which determine the state of the black box
Y are the output variables.

The system can be analyzed in three different ways:

Problem	Given	Find
Synthesis, identification	X, Y	T
Forecasting (prediction)	X, T	Y
Management (control)	T, Y	X

Synthesis. Research activities may concentrate on identification and synthesis. Thus, given the phenomenon X --> Y, what is then the structure of the blackbox. This is often done by the so called effect-response (X-Y) research. A model will be the end result of research and the model replaces the blackbox. The state of the model can then be characterised by the actual state of the variables T.

Forecasting. Given the model, the output of the system (Y) can be predicted or forecasted if we have information on the state of the model T and the environment X.

Management. In management we think the other way around. Given the desired output (Y) of the system and given the state of the model (T), what should then be the input (X) to reach the goals we are striving at. In integrated pest management the goal is to optimize net yield (Y_1) while minimising the waste of agrochemicals (Y_2), to save biological resources. Net yield is gross financial yield minus costs of inputs (X). The waste of agrochemicals can be minimized by not using them at all or by a very selective use.

In simple systems, e.g., some X variates, some T variates and one Y variate, the model is rather simple, and forecasting and management are simple applications of results of effect-response research. However, in more complex situations, e.g. many X variates, many T variates and some Y variates, usually no complete model is made during research. Instead, many models, each describing a part of the total model, are to be used. Then the task of synthesis (the compilation of all the models into a complete model) shifts from the level of identification to forecasting and management. Thus different simple models are evaluated and combined during forecasting and management. For example the effect of fertilizers on yield becomes increasingly complex if the effect of fertilizers on diseases and pests and their effects on yield has to be taken into account. In many situations no formal models are yet available to handle complex problems. Nevertheless, the farmer does.

In research, identification and synthesis are the main goals. In IPM, forecasting and management are the goals. During forecasting we predict the output Y. **Comparison of this predicted Y with the desired Y, reveals whether measures should be taken.**

1.2 Strategical measures

Consider the generalised life-cycle of a pest. Usually it contains periods (or life stages) during which:
1. the pest disperses;
2. the insect pests establishes or the pathogen infects the crop;

3. it grows, usually vegetative or asexual;
4. it multiplies, usually generative or sexual;
5. it is in its resting stage, during which the pest survives adverse (weather) conditions and the host-free period.

Each of these periods have a duration and a success (survival or multiplication).

The longer the duration (multiplication excepted) and the lower the success of each life stage, the lower the rate of increase and thus the risk of a pest outbreak. Moreover, at low rates of increase we can see a gradual build-up of the pest which allow enough time to take action.

The shorter the durations (multiplication excepted) and the higher the successes of each life stage, the higher the rate of increase of the pest and thus the higher the risk of outbreaks. At high rates of increase (instable host/pest systems) we often miss the build-up phase of the epidemic and consequently the crop may be taken by surprise. From management point of view this is unacceptable. However sometimes (not often!) catastrophes are needed to point to the importance of support in developing IPM.

Most strategies of pest prevention are directed to desynchronize the life-cycles of the host versus those of the pests, e.g. by choosing sound crop rotations and planting dates. In addition, measures are encouraged which retard the duration of life stages of the pest and decreases their successes. This will be detailed in chapter 2 (Epidemiology).

These strategical measures all require a vast amount of agricultural and ecological knowledge. As each agro-ecosystem is different in different regions (coastal, valleys, plains, altitudes etc.), the knowledge available is often too general or too anecdotal to design biological sound agroecosystems. To reach the multiple goal of agriculture, the need to increase and acquire this knowledge by researchers, advisors and farmers is therefore stressed. In this course principles are treated, which can be used to design IPM schemes tailored to specific cropping systems and regions.

1.3 Tactical measures

After the pest appeared it will develop in the field. Elements of a tactical scheme to forecast the epidemic and to decide whether actions should be taken are:

Elements of tactical advisory systems

---

1. actualise the state of the field by pest monitoring
2. determine the forecasting period
3. forecast epidemic
4. forecast damage based on the expected epidemic
5. determine the product price to forecast loss
6. determine the efficacy of the control method
7. calculate the costs of applying the control method

8. evaluate the expected costs and benefits of the control action
9. determine when a new observation (monitoring) is desired

These variables determine the economic injury level (EIL), which is used to trigger the control of a pest. It should be clear that **no exact EIL exists for a certain host-parasite relation**. First, it depends on the crop developmental stage. (Question, which of the elements listed is/are affected most by the crop developmental stage?). Second, many of the elements listed are not exactly known, they are uncertain to some extend. E.g. weather affects the cropping period and thus the forecasting period.In addition, weather affects the course of the epidemic and the efficacy of the control method. So the exact EIL for a specific field is unknown, as we do not know the future weather. Nevertheless we can estimate the expected EIL for a field, because we know field specific characteristics (e.g. soil type, cultivar, fertilization etc.) and by assuming an averaged cropping period, an averaged course of the epidemic and an averaged efficacy of the control method.

An example of a tactical advisory system is Blitecast, which uses weather data to forecast potato late blight during 7-14 days. These forecasts are used to recommend the necessity of a treatment and the spraying schedule. Another example is EPIPRE, which uses disease and aphid incidences of winter wheat in relation to cultivar and crop growth stage, to forecast epidemics and to evaluate cost/benefit of control. In entomology and nematology many tactical advisory systems are operational by use of economic injury levels, if the pest intensity is above that level, a control is recommended.

Pesticides

In addition to IPM strategies, Pesticide Management is used, to determine the most selective pesticides and to alternate them. Sometimes this is abusively named Integrated Pesticide Management in literature. Pesticide choice is then based on selectivity in terms of natural enemies, toxicity, persistence. Pesticides with different modes of action are alternated to retard the build-up of pesticide resistance. In addition, the amount of pesticides used per area (hectare) can be reduced by technical means. For example by use of low volume spray equipment and row treatments instead of full field treatments, the amount of active ingredients per hectare may be reduced considerably. These aspects will not be treated in this IPM Toolbox.

1.4 Decision tools

Decision tools are used to identify, structurize and analyze a problem. By that the effort used to solve a problem may be reduced. For example if people work on a pest management problem, they may start roughly from two different viewpoints and will reach two different solutions.

Scientific viewpoint. From a scientific point of view a pest management problem is used to improve our knowledge. Underlying biological and ecological processes are identified and specific methods are developed. However, the intrinsic goal of science is to obtain (general) new knowledge and not to solve the problem.

From viewpoint of **research and development** the goal is to implement an acceptable control method, whether it is a new or old method. A new method may even have a disadvantage compared to an old and approved method as its effectivity is often not tested in many situations. The intrinsic goal to solve a problem may therefore stagnate development of new control options.

Usually the problem is just there, but, it is necessary to identify the problem and the control options, both from scientific as from development point of view. An entry for identification is to describe and analyze the history of cropping and crop protection practices. It may reveal that the problem is man made e.g. a pest may become important after an unselective pesticide was used to control a primary pest. The option is then not to use a second pesticide for this secondary pest but to search for a selective pesticide to control the primary pest. In many situations cropping practices change gradually (e.g. crop rotations practised, fertilization, soil preparations, cultivars, seed rates, sowing and planting dates), and, gradually the cropping system may become locked in a system which uses more and more pesticides. Another entry to identify the problem and solutions is to assess the relationship the pest problem has with other factors, e.g. regional and cropping characteristics, predators/parasites etc. Based on these relations an interaction matrix can be made which depicts all important relations. These are qualitative assessments which help to focus research and development to solve the problem. At the same time information gaps are detected.

The information gap is the difference between the information that farmers have and that which they should have to make good decisions. The information gap is usually broad, covering fields of fundamental and field research, training and extension. It consist of:
1. **Research gap**. Not all information is available, appropriate research is necessary to close the gap.
2. **Synthesis/interpretation gap**. Research results are available, but they are fragmentary and not analyzed and interpreted in the context of the farmers problem. A compilation and analysis/interpretation of the information in the context of the problem should be undertaken.
3. **Dissemination gap**. Though the desired information is available in some places, it is not available to the farmers or extension officers. By appropriate extension programs and services this gap may be closed.
4. **Reception gap**. Although the information necessary to make good decisions is available to the farmers, they cannot use it because they don't place it in the context of the problem or they need background information. This implies that special trainings are necessary to improve the perception of the farmers.

Suppose that after an intensive research, development and extension program a new control method (e.g. new cultivars, other crop rotations or cropping practices, a new selective pesticide) is available to the farmers. Whether this new method will be adopted does not only depend on its biological performance. Usually the new has to replace (compete with) an old accepted control method. That method will be adopted which is:

---

1. practically feasible
2. technically possible
3. environmentally sound
4. economically desirable
5. politically advantageous

To improve research and development efforts in terms of their impact, these constraints have to be identified for the existing control methods and the control methods in development. By that, research and development can be focused. Also the constraints may be reduced by e.g. special training programmes, improved information dissemination, demonstration projects and governmental regulations.

1.5 Questions

1. Draw a relation diagram of fertilizers (X), crop (T) and yield (Y). Add + in these relations to show that these relations are positive (more fertilizers -> more crop biomass -> more yield). Add an insect pets (T2) in the figure. Add (-) in the diagram to show that more insects cause less crop biomass. Assume that fertilizers (X) stimulate (+) the insect pest. Adapt the figure. Add the relations of a pesticide (X2) and its waste (Y2), indicate + and - in the figure for positive and negative relations. Which elements are to be minimized and which to be maximised in IPM? How would you manage this (according to the figure) and which are the conflicting options?

2. Will crop rotation in general be an effective strategy in managing soil-borne diseases or in managing seed-borne diseases? Will crop rotation in general be an effective strategy in managing weeds or in managing insects? Is crop rotation a strategical or a tactical method in IPM?

2. EPIDEMIOLOGY, POPULATION DYNAMICS AND FORECASTING IN IPM

2.1 Introduction

The basic biological processes are growth and multiplication. **Growth** is the increase in biomas, **multiplication** is the increase in individuals or biological entities. Usually a biological entity cannot grow unlimited. After reaching a certain size (age, ripening) the accumulated biomass is not invested anymore in growth but in multiplication (spores, eggs). By that growth and multiplication are to some extend related. Multiplication and **reproduction** are here used as synonyms. Biological organisms perform growth and reproduction during their **life cycle**. During their life cycle they pass different developmental stages: **development**. During certain developmental stages growth dominates, during others reproduction dominates. Moreover, many developmental stages exist during which the organism disperses or rests to survive adverse environmental conditions. The time needed to pass the whole life cycle is the **generation time**, e.g. the time from egg to egg or from spore to spore. This time is often named the latency period in fungal diseases.
The life stages often have different functions, durations and names.

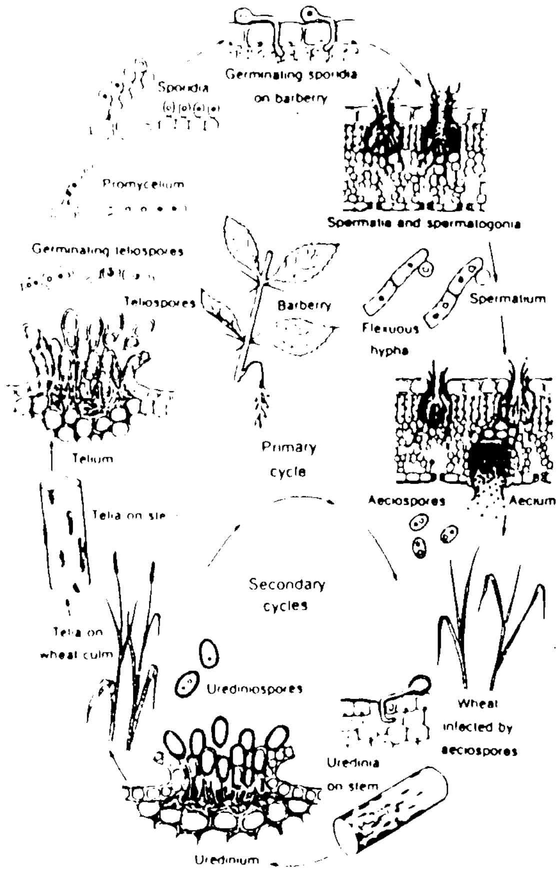
Life cycle of an insect:
1. egg
2. larval (nymphal) stages
3. pupa
4. adult

Life cycle of a fungus:
1. spore
2. mycelium
3. resting body
4. fruiting body (many different types)

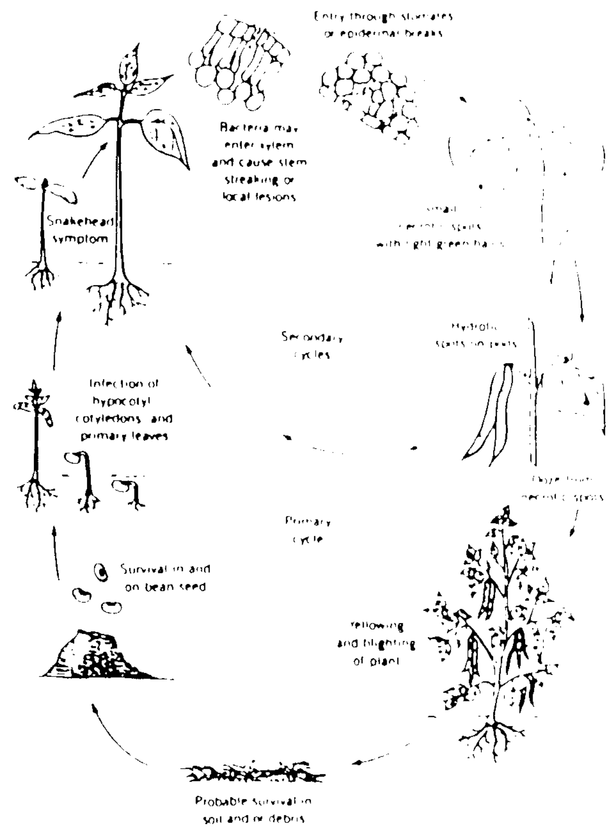
Life cycle of a nematode:
1. egg
2. juvenile stages
3. adult

Examples of life cycles are given in Fig. 8.1a-d. In fact each organism may have different life stages. Some people like to give them all different and difficult names. This hampers the understanding of the similarities. Important is that the life cycle of the organism is known and that the stages and functions are exactly described.
Mono-cycle (univoltine) pests have only one generation per year, e.g. blister beetles, loose smut of barley and wheat, many weeds.
Several-cycle (bi or trivoltine) pests have few generations per year, e.g. codling moth, many stem-borer species.
Poly-cycle (plurivoltine) pests have many generations per year, e.g. aphids, mites, sorghum midge, fruit flies, mealy bugs, powdery mildew, rusts, bacteria, virus.
Pests with several-year cycles e.g. cockchafer, sugarcane white grub, striped click beetle.

Growth, multiplication and development depend on the environment of the organism.

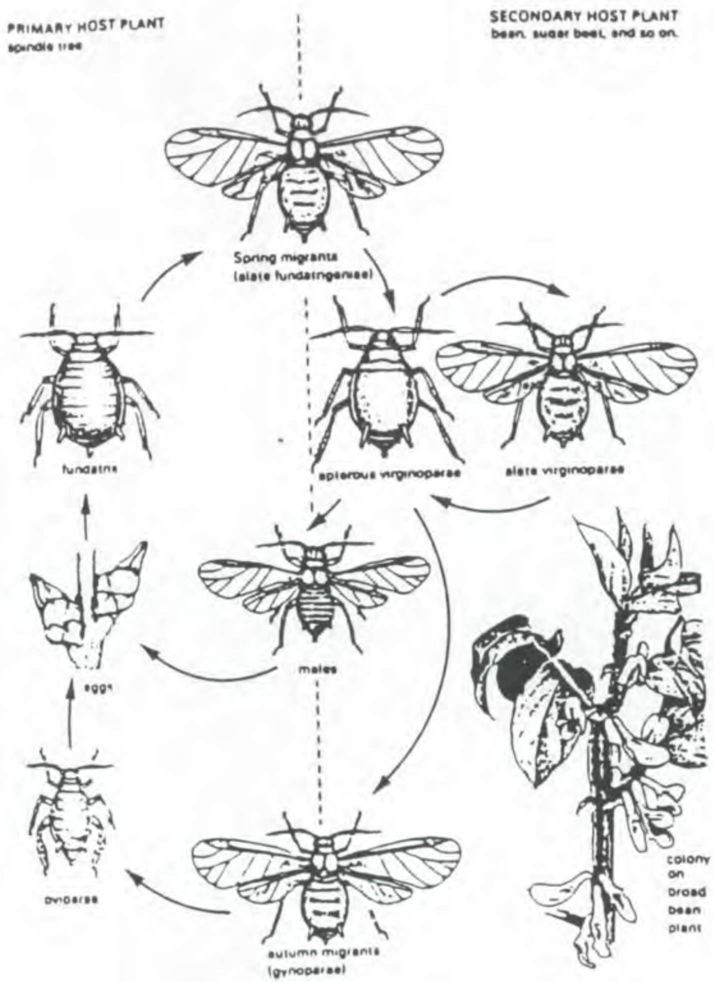


a: life cycle of stem rust of cereals.

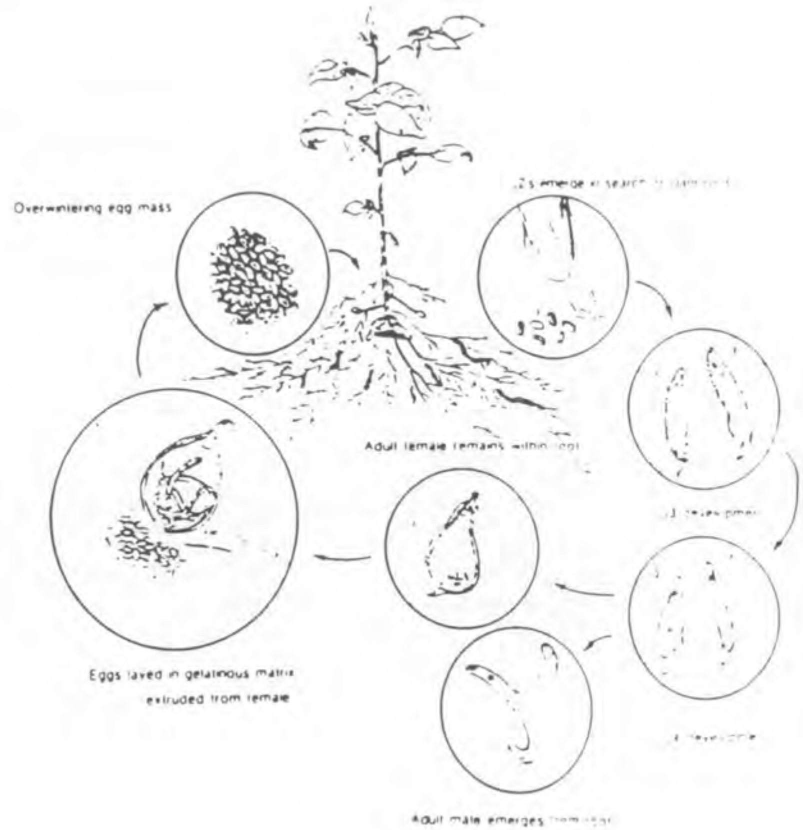


b: Life cycle of Pseudomonas phaseoloca, the cause of halo blight of bean.

Fig. 2.1 a + b: Examples of life cycles of pathogenic fungi



c: Life cycle of the black bean aphid, Aphis fabae



d: Life cycle of Meloidogyne spp.

Fig. 2.1 c + d: Examples of life cycles of an insect pest and a nematode

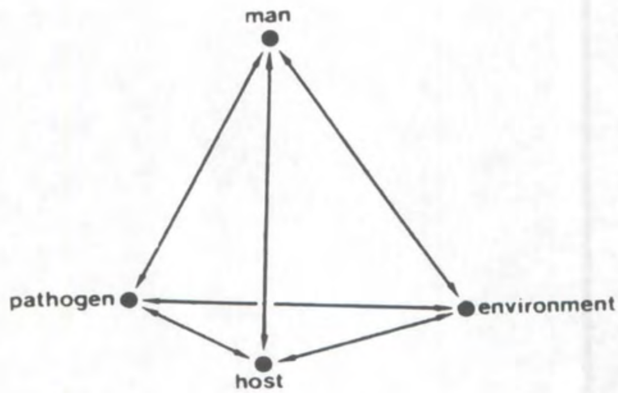


Fig. 2.2: The pest tetrahedron. The base symbolizes the interaction of host, pathogen, and environment. On each of these, man has various effects that are important to the development and control of epidemics.

Usually growth and multiplication are enhanced by favourable environments, while development may be enhanced by unfavourable environments.
E.g. a favourable environment for plants is well fertilised and watered. Plants then become big with many leaves (growth dominates) and stay longer vegetative as flowering and seed ripening is delayed (delayed development). If water stress or nitrogen stress occur, plants become smaller with fewer leaves (growth is less) and flowering and seed ripening is enhanced (enhanced development). Also in insects and fungi a favourable environment enhances the vegetative phases of the life cycle, while harsh conditions stimulate development and the formation of resting stages (pupa, asci, sclerotia).

Important elements of the environment of a pest are the host (crop) the **environment** (e.g. weather) and **man** (Fig. 2.2). The pest tetrahedron illustrates this. **Epidemiology or population dynamics** is the study of the interrelations which exist between the pest population and its environment. The environment (host and man included) affect the pest population and the pest population affects the environment. Epidemiology in the context of IPM, concentrates on the question how to manipulate the environment, to manage pests to acceptable levels. Therefore in epidemiological research conceptual and mathematical models are constructed. As described in the introduction (chapter 1) we may consider strategical and tactical measures to manage populations:

2.2 Epidemiology and strategical measures

Remember:

Strategical measures have a long time-span and are directed towards pest prevention. Tactical measures have a short time-span and are directed towards pest control.

Elementary strategies of pest prevention are:
1. to desynchronize life cycles of the host versus the disease/pest;
2. to increase the resistance of the agroecosystem. This is done by retarding and by decreasing the success of one or more of the life stages of the pest, preferably the most critical stage.

Desynchronization of the cycli in time and place may be obtained by:

---

1. sound crop rotation schemes;
2. growth of different cultivars of the host which differ in their origin of resistance. Different cultivars may be sown at the same time on different fields or cultivars may alternate over time;
3. clean field period (host-free or crop-free period) during a certain time over a vast area, during which plant rests (e.g. stubbles) and volunteer plants are eliminated;
4. synchronisation of the life cycles of the host (e.g., same optimal planting dates) over a vast area;
5. seed or plant sanitation and certification;
6. others: ?

In fact the strategy of desynchronization resembles that of playing hide and seek with the host and the pests. If many pest entities fail to find or to land on a new host, many crops may escape from an epidemic. By desynchronizing the life-cycles, the pest has to pass more **bottle-necks**. It cannot grow continuously but it has to disperse to find a new host or it has to rest during a certain time before a new host is available on which it may grow.

Desynchronization of the life-cycles of the host and the pest is especially effective for crops which have short growing periods (annuals, e.g., maize, rice, potatoes) and less for crops with long growing periods (perennials, e.g. fruit trees). Then cultural practices need to be directed to block or retard the development of pests, and to spread risk.

The **resistance of the agro-ecosystem** is increased by:

---

1. growth of different crops (mixed cropping) and (resistant) cultivars;
2. enhancement of antagonists, predators, parasitoids and parasites;
3. not too big scale fields to enable a rapid immigration of beneficials from refugia (other crops, nature) after a disturbance;
4. balanced fertilization (N,P,K, trace elements);
5. proper timing of cultural methods (e.g., pruning, weeding, irrigation, sanitation).
6. others?

2.3 Epidemiology and tactical measures

Tactical measures have a short time-span and are directed towards pest control. Here a short time-span is defined as one cropping period (annuals) or one growth season (perennials). Control measures may be biological (parasites, sterile males, virus), mechanical (ploughing, harrowing) and chemical (pesticides). In general a tactical measure is taken if economical damage is expected in a certain field. Economic damage will occur if there is an outbreak of the pest/disease. A forecasting system gives a prediction of the development of a disease/pest (Box 2.1, elements 1-3). Advisory systems also perform an economical evaluation of forecasted damage and of the possible control actions. The elements listed are those within the blackbox which determine the economic injury level (EIL). See also chapt. 1.

Box 2.1 Elements of tactical advisory systems

****

1. monitor the initial size of the population;
2. determine the forecasting period;
3. forecast epidemic;
4. forecast damage;
5. estimate loss depending on product price;
6. estimate the efficacy of the control method;
7. estimate the costs of the control method;
8. evaluate whether a control action is economically justified;
9. determine when a new observation (monitoring) is desired.

****

2.4 Forecasting in IPM

Forecasting is the application of knowledge to predict the development or outbreak of an epidemic. In warning systems the occurrence or development of diseases/pests is forecasted. If such forecasts are followed by forecasts of consequent damage and if an evaluation of a control method is made we usually speak of an advisory system (see chapter 1, elements of forecasting/advisory systems).

Recall the three views on a problem:

```
                    ----------
                   I          I
X -------->        I    T     I  --------->  Y
                   I          I
                    ----------
```

in which X are the input or driving variables
T are the variables which determine the state of the blackbox
Y are the output variables.

During forecasting we predict Y, given the situation of the inputs X and state of the system T. Thus warning systems give a prediction of population development or epidemics. In management, the forecast (Y-forecasted) is used to determine what we should do (change X) in order to reach a desired output (Y-desired). Thus advisory systems give an evaluation of possible control options to avoid undesired forecasted outputs e.g. an undesired epidemic. Forecasts are based on the knowledge we have on the actual state of the model of the blackbox T (e.g. cultivar susceptibility, presence of parasites) and/or on the delay of the black box (e.g. temperature sum, days after the rains started, fertilizer or pesticide application).

Types of forecasts. Forecasts are predictions over a certain time-span in future. This forecasting period can be any period. Usually three types are named: **long range forecasts** (forecasts over seasons and cropping periods), **medium range forecasts** (forecasts during one season or cropping period) and **short range forecasts** (forecasts during a life cycle or life stage of the pest).
Not all forecasts are real predictions of the future. E.g. in the degree-days model below, temperature is measured and if accumulated temperature reaches a certain value, a certain pest is expected (at that right moment). Though there is no fixed terminology, forecasts are usually more used for predictions in future, while predictions is a more general term for an expectation (now or in future).
We may forecast when a pest will appear, this is called a **phenological forecasts**. By this only the moment of appearance is forecasted, not the intensity of the pest. If the intensity of the pest is forecasted this is called **population forecast**.

What can be forecasted in crop protection?

1. Forecasts which are already feasible:
   - whether or not a certain pest will exceed the damage threshold in the forthcoming season;
   - first date at which the control threshold will be reached, e.g. "negative prognosis";
   - course of population development for various applications (e.g. for the determination of control thresholds as well as for risk periods);
   - necessity of control measures, timing and/or dose of next treatment, or kind of treatment. If the number of sprays cannot be reduced, any further forecasts should make them more effective. this requires a more flexible spraying;
   - crop losses.

2. Amongst forecasts which are desirable though until now unexplored are:
   - dynamics of pathogen races as essential information for strategies based on sowing resistant varieties;
   - potential danger from pests not yet occurring in an area (important for quarantine, and in areas with vulnerable crops);
   - behaviour of newly introduced pests in changing agro-ecosystems;
   - the economic return of crop protection or non-protection;
   - needed supply of material for crop protection, including resistant varieties, for gene management, etc.

As a matter of fact, application of forecasts in crop protection seemed to be on a decline in many developed countries during the past years. Certainly, a growing disinterest in forecasting was until recently to be attributed to a favourable cost/benefit ratio of pesticides, which allow rather freely their use for "insurance" or just for convenience.

On the other hand, research on forecasts fell short of expectation. The number of forecasts which eventually proved feasible for implementation is fairly small. Some of them, though operational, have been overruled by the above mentioned development in crop protection or have not been adopted for technical or other reasons.

This limited success in research and implementation of forecasts in crop protection is partly due to too much reliance on climatic factors only, and deficient application of modern epidemiology and population dynamics. The most important deficiency in the development of forecasts is their inadequate testing for validly and general application. However, prospects for more efficient and practical forecasting methods have brightened in past years because of advanced concepts originating from research in epidemiology and population dynamics) and technology (computers for research and implementation, appliances for measurements, remote sensing, etc.). It is also increasingly realized that forecasting is indispensable for integrated pest control and other more rational, i.e. more efficient, economic and less hazardous crop protection.

On what criteria can forecasts be based in crop protection?

All forecasts developed must be adequate, that means valid and sufficiently accurate, reliable, and as simple as possible.

Basically, forecasting methods can be based on criteria from:
- the host plant biology, like growth stages or age (based on knowledge about correlations with pest outbreak and development);
- the biology of insect pest or pathogens, and states of disease or pest development;
- climatic and other environmental factors;
- or a combination of two or three of the above criteria.

(see also chapters 5 - 7).

Ideal forecasting methods are the ones to be operated on farm or village level without the necessity of expensive equipment or skilled professional staff. On the other hand, there are tendencies for centralized warning systems with a central operational unit, e.g. the computer with a model that predicts with data provided by individual subscribers, observers or scouts over telephones or terminals.

<u>Usefulness of forecasts</u>

Forecasts are useful when:

---

1. The pest causes considerable damage in a certain area;
2. The development and intensity of the pest varies considerably; e.g. in one year it is important in another year not;
3. The pest can be controlled by a control method;
4. We have enough knowledge on the dynamics of the pest to make reliable forecasts;
5. The forecasting period is long enough to allow a control action.

By this forecasts are mainly used in tactical decision making. Strategical control actions to prevent pests (see introduction, e.g. crop rotations, resistant cultivars) have usually a time-span which is too long.

It is obvious that forecasts are irrelevant when a pest never exceeds the economic damage threshold, or no means of control are available, or if a cheap routine measure takes care of the problem. But if the economic threshold is exceeded (see also chapt. 8), and control is possible and required, the feasibility of forecasting still depends on:

1. Adequate simplicity of the method to implement it in a given field.
2. Satisfactory accuracy:: A 80% level of correct forecasts should be aimed at, to satisfy risk aversion of farmers.
3. The pest should not be controlled already by measures directed against another pest.
4. The notice of a forecast must not be too short, in order to take the reaction time of farmers or advisers, farm routine, delays in the communication network, and the action of the control measure into account.

<u>No monitoring and forecasting of critical periods</u>

In some situations monitoring is impossible. This because the pest is invisible or because the damage threshold is too low. It is then often assumed that:
1. the pest is present (from a certain crop developmental stage onwards).
2. the pest always causes economical damage if the pest multiplies.

This approach is often used in fungal diseases which cause high economical losses and as consequence monitoring is not possible. Moreover, pesticides used to control these disease are rather ineffective, e.g. only inhibit germination/infection during a small time-span. Examples are late blight of potatoes and apple scab. The strategy practised is to inhibit multiplication. Weather data and other factors (e.g. crop susceptibility) are evaluated to assess whether the environment during a certain period has been favourable for multiplication or not. For late blight and scab (and for many other diseases) a certain duration of leaf wetness combined with certain periods of high humidity are critical (sporulation and infection). If favourable, a control action is advised. If the weather was unfavourable no control action is advised; this is usually called a **negative forecast**. Here like in the phenological degree-day model, no real future forecast is made.

From viewpoint of IPM, it should be clear that this approach to manage pathosystems is not desired. Effort should be directed to increase the resistance of the pathosystem, e.g. by breeding for resistance and sound crop rotation schemes.

The entomological EIL

The concept of the economical threshold level was developed in entomology. It is the cornerstone of IPM. The basic assumptions are that:

---

1. The pest has one generation per growing season;
2. Each individual insect causes a certain amount of damage during its life-cycle;
3. It is possible to monitor a life stage of the insect, which damage it already caused is negligible;
4. Pesticides are highly effective.

The strategy followed is that the amount of insects can be monitored and thus the expected damage can be forecasted. As pesticides are effective, nearly all insects can be killed and no economical damage occurs. To determine the EIL, the relation between damage and the amount of insects is empirically determined. E.g. in field experiments different population intensities of insects are made by use of pesticides and field releases of insects, and subsequent yields are determined. Another approach is to estimate the damage caused by one insect, from the leaf surface it consumes or the amount of flowers or pods it damages. From such damage relations the EIL can be derived.
This concept was very successful and it is now broadened to apply it to other pests and diseases, which usually do not meet the basic assumptions. E.g. if the insect has two generations per year we have to forecast the second generation in addition of the amount insects monitored (the first generation).

<u>Uncertainty in forecasts</u>

Every forecast will be wrong. Due to errors and unknown factors we cannot forecast accurately. However, without a forecast we are also making wrong decisions and the rationale is to improve decisions by using forecasts. In case with sampling (chapter 4) we obtain an estimate of the density of the population and we may obtain also information about its error (= uncertainty). The same approach can be applied with forecasting. We may forecast the course of the epidemic and its associated error. Unfortunately, analysis of errors in forecasts are seldom performed. Some aspects of errors in forecasts are therefore discussed shortly.

Suppose a forecast is based on a linear model. Y is forecasted based on information of a, b and x:

$$Y = a + b.x + e$$

in which e denotes the residual error. Suppose that the model was based on research, during which the linear relation was established and estimates of a and b and their associated errors and e were obtained.
If we use this model for a new forecast, the error of the new predicted Y, is usually computed as:

$$Var(Y) = Var(a) + (x-\bar{x})^2 . Var(b) + Var(e),$$

in which $\bar{x}$ is the average x which was used in the research to estimate a, b and e.
Often it is not possible or easy to obtain one equation to compute the error of prediction. In research a simulation approach is used to assess the error of prediction. To illustrate this, suppose that a disease/pest grows exponentially:

$$Y_t = Y_0 . e^{rt}$$

and by monitoring Y_0 was 1.5. Due to sampling error there is a probability of 10% that Y_0 is 1 or lower. An analysis of historical epidemics revealed that the disease/pest has a relative growth rate of 0.15 for this cultivar, but there is a probability of 10% that the growth rate is 0.10 or lower. The expected number of Y_t after 10 days would be 6.7. If we combine the low values of Y_0 and r and substitute them in the equation then Y_t is 2.7. The probability Yt being 2.7 or lower is 1%, as values for both estimates were at 10% (thus both combined gives 0.1 x 0.1 = 0.01).
In case of p=10% that Y_0 is 2 or higher and r is 0.20 or higher, there is 1% probability that Y_t is 14.8 or higher. Thus, average Y_t is expected to be 6.7 and 2% bounds (both tails have p=0.01) could be 2.7 and 14.8. Note that the average Y_t is more close to the lower bound than to the upper bound. This is caused by the non linear (exponential) model. But you can give it a meaning in the context of crop protection. As diseases/pests usually grow exponential, the risk of an outbreak is high when a combination of factors is favourable. This also points to the main strategy of IPM, to take many cultural practices which are unfavourable for the disease/pest.

3. USE OF MATHEMATICAL MODELS IN IPM: (EPIDEMIOLOGY, POPULATION DYNAMICS AND FORECASTING)

3.1 Temporal models

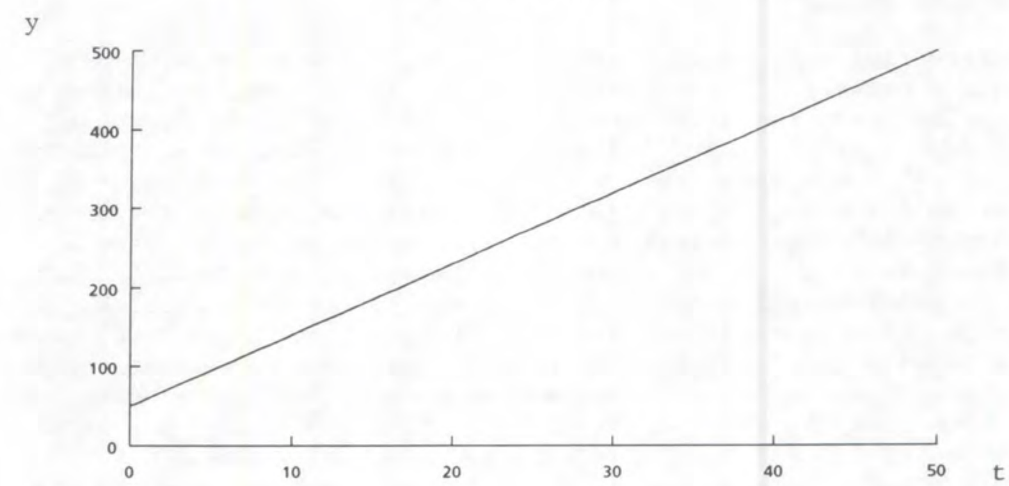
When considering individual plants, animals, bacteria etc., a distinction can be made between qualitative and quantitative characteristics. Quantitative characteristics are those that we can count or measure: e.g. biomass or weight of an individual, length of an individual. Qualitative characteristics are those that we can not express in numbers, e.g. the life stage of an individual. Both quantitative and qualitative characteristics will change with time. **Growth** is the increase in biomass, whereas **development** is used for the change in life stage: e.g. from juvenile to adult. On the population level, all qualitative characteristics such as life stage can be considered in a quantitative way, by just counting all the individuals in a certain life stage. Thus, on the population level, every characteristic can be expressed in numbers and these numbers will change with time. In IPM, we are mostly interested in changes at the population level, as one individual will seldom cause damage.
Sometimes, clear patterns can be distinguished in these changes and then it is very tempting to use mathematical functions to describe the change of a characteristic with time. Also, biologists often like to simplify the complex reality by making all kind of assumptions on the underlying processes of the change in a population characteristic. This simplification is necessary, to make it plausible, that a particular mathematical function represent the changes in some population characteristic. The actual mathematical function, used as this simplified representation of reality, is called a **mathematical model**. A mathematical model makes it possible to **forecast** characteristics of the population. It is then necessary to assume, that the population will behave according to the model.
An example of a simple mathematical model, that can be used to describe the change of the number of individuals in a population with time, is the **linear growth model**. With the linear growth model, the number of individuals increase per unit of time with a constant number:

$$dY/dt = c$$

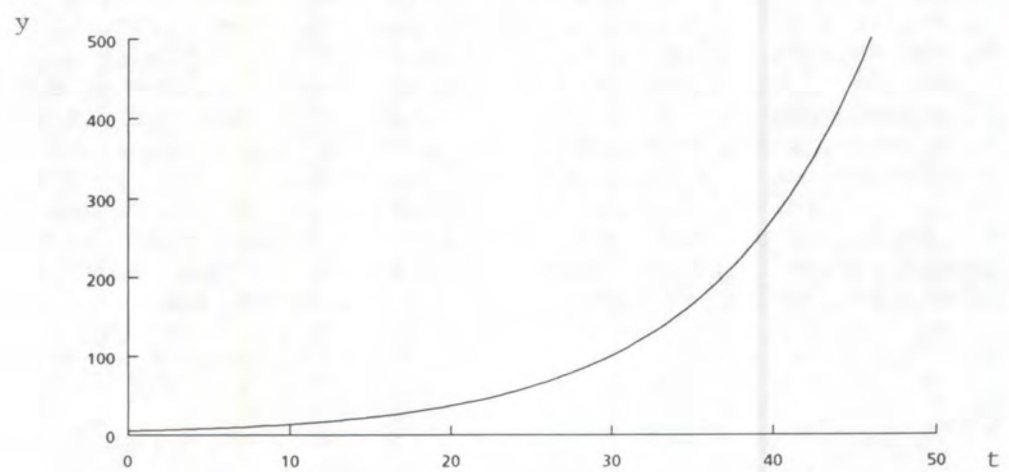
$$Y_t = Y_0 + c.t$$

Y: number of individuals [N]
t: time [days]
c: linear growth rate [$N.day^{-1}$]
The dimensions of the variables are given between [] and it is a good habit to check the consistency of equations by checking the dimensions.
The linear growth model may be used just as a description of observed numbers of individuals, but its use can also follow from some biological reasoning. For instance, the following set of assumptions may justify the use of a linear growth model:
1. we start with a large batch of eggs, there are no individuals in their reproductive stage
2. every unit of time, a constant number of eggs is hatching.

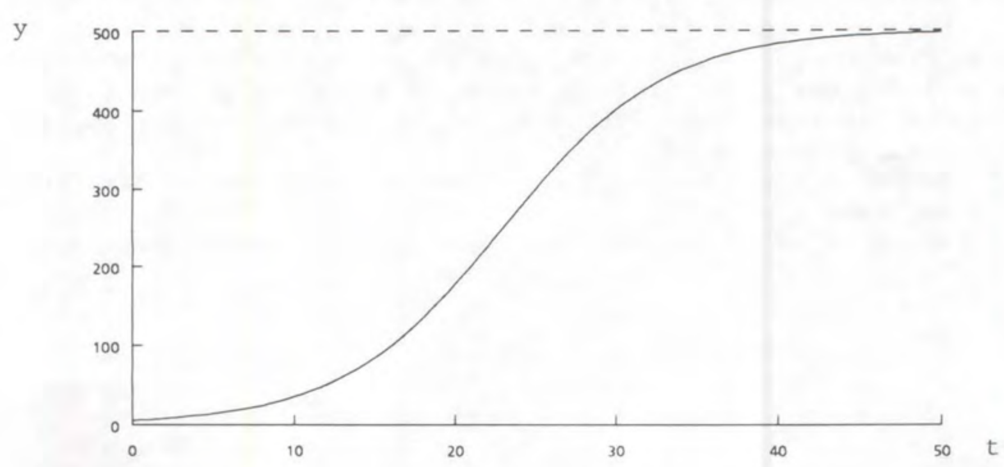
linear growth



exponential growth



logistic growth



Exercise 1

A biologist finds, that the number of some species of a social insect (e.g. bees) is increasing linearly, as the queen is producing a constant number of eggs per day. At the same time, these insects are eaten by a particular species of bird. Every bird eats a constant number of insects per day. The biologist presents the following temporal model for the number of insects (all life stages):

$$Y_t = Y_0 + c.t - p.B$$

Y: number of insects [N_{insect}]
c: linear growth rate [$N_{insect}.day^{-1}$]
t: time [days]
p: constant predation rate [$N_{insect}.N_{bird}^{-1}.day^{-1}$]
B: number of birds [N_{bird}]
<u>Task</u>: Check the dimensions.

The increase of the number of individuals in a population often depends on the number of individuals present. The **exponential growth model** is often used to describe this situation:

$$dY/dt = r.Y$$

$$Y_t = Y_0.e^{rt}$$

Y: number of individuals [N]
t: time [days]
r: exponential growth rate or relative growth rate [day^{-1}]

Exercise 2

A population of bacteria in a culture follows the exponential growth model. The exponential growth rate is 0.2 $hour^{-1}$.
<u>Task</u>: Calculate the **doubling time**, i.e. the time necessary for the population to double its size.
(Hint: use $Y_t = Y_0.e^{0.2t}$ and the fact that during the doubling time Y_t increases from Y_0 to $2.Y_0$)

Usually, a population can not grow unlimited and the growth becomes restricted, when the population reach a certain maximal population level. This maximal population level is often called the **carrying capacity** of an habitat. The **logistic growth model** is a simple model, that combines exponential growth at a very low population level with the existence of a carrying capacity:

$$dY/dt = r.Y.(1-Y/K)$$

$$Y_t = K/(1 + b.e^{-rt}), \text{ with } b = K/Y_0 - 1$$

Y: number of individuals [N]
t: time [days]
r: logistic growth rate [day^{-1}]
K: carrying capacity [N]

Exercise 3
For populations levels much smaller than the carrying capacity, the logistic growth model is nearly identical to the exponential growth model.
Task: Prove this. (Hint: Take K>>Y, so that Y/K = 0)

With plant diseases, it is often difficult to count a number of individuals (lesions, pustules). Instead, the fraction of the leaf area covered with the disease is used as a measure for the disease intensity. A simplified version of the logistic growth function can then be applied to these data:

$dY/dt = r.Y.(1-Y)$

$Y_t = 1/(1 + b.e^{-rt})$, with $b = 1/Y_0 - 1$

Y: fraction of leaf area covered [-]
t: time [days]
r: logistic growth rate [day^{-1}]

Exercise 4
Task a: Check the dimensions.
Task b: Give the value for the carrying capacity.

3.2 Parameter estimation in temporal models

When a model is used to describe observed changes in a population, we want this model to describe these changes as accurate as possible. Or in other words, we want the differences between the model and the observations as small as possible. Therefore, we will choose the parameters of the model in such way that these differences are as small as possible. This procedure is called "fitting the model to the observations".
There are several mathematical techniques of fitting models to observations. Most of these models search for a minimal value for the **sum of squares of the differences**:

$\Sigma(Y_t - F(t,a,b))^2$

Y_t: observed population size at observation time t
F(t,a,b): the model value at time t and with values a, b etc. for the parameters

Exercise 5
Task: Explain why $\Sigma(Y_t - F(t,a,b))$ is not a useful measure of how close the model is to the data.

A much used fitting method for linear models is **linear regression** and it also uses the sum of squares of the differences. This method is available in most graphical computer packages, but is also easy to perform on a pocket calculator. The exponential growth model can be transformed into a linear model by logarithmical transformation of the observations:

$\ln Y_t = \ln Y_0 + r.t$

Y: number of individuals [N]
t: time [days]
r: exponential growth rate or relative growth rate [day^{-1}]

Linear regression can be used to estimate Y_0 and r from logarithmically transformed observations. There are also non-linear fitting techniques to fit non-linear models, e.g. the exponential growth model, to observations.

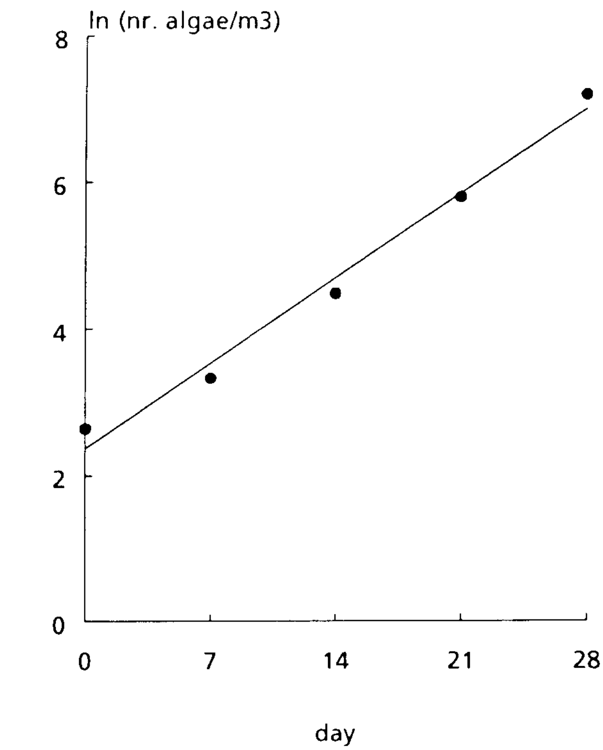
The following exercise demonstrates, that the method of fitting and parameter estimation can influence the forecasts of a model.

Exercise 6
You are manager of a large fish farm, producing for export. In one of the ponds, there are algae, that are poisonous to the fish. The fish will die, if there are more than 20,000 algae per m^3 of water. You measured the number of algae every week and found the following numbers:

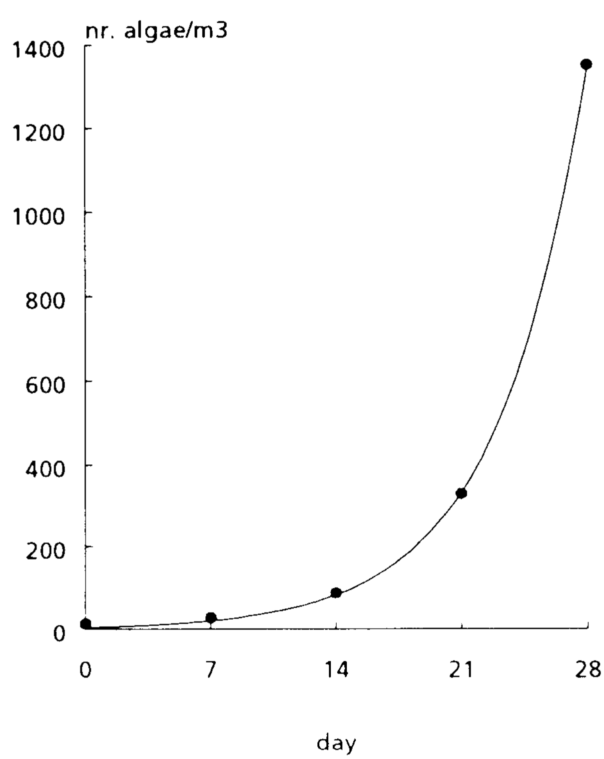
day 0	14 algae/m^3
day 7	28 algae/m^3
day 14	88 algae/m^3
day 21	330 algae/m^3
day 28	1350 algae/m^3.

Today, day 28, you have an important decision to take. There are no cold-storage facilities at the fish farm and the fish has to be deep-frozen at the factory within 10 hours after catch. On the afternoon of day 29, the van of the factory will come to pick up some other fish. The next time the van will come is 14 days later. You can harvest the fish in the pond contaminated with algae now, but you want the fish to grow for another two weeks, as bigger fish mean more kilograms and also a better price per kilogram. And the number of algae is still below the critical value of 20,000 algae/m^3. Can you wait until day 43 or do you have to harvest the fish now? You expect the number of algae to increase according to an exponential growth model, so you estimate the exponential growth rate to forecast the number of algae at day 43. You apply the logarithmical transformation and perform linear regression to the transformed data.
Task a: Transform the observed number of algae into its natural logarithm. The line fitted to the transformed data has as intercept 2.37 and as slope 0.166 (see figure).
Task b: Give the values for the parameters of the exponential growth function.
Task c: Calculate the model values for day 0, 7, 14, 21 and 28.
Task d: Calculate the predicted number of algae on day 43.
Task e: Take a decision based on the above calculations, whether to harvest the fish tomorrow morning or wait another two weeks.
That evening, you meet your neighbour at the garden gate and talk about your problem and how you solved it. Your neighbour, a mathematician, likes to fit the exponential model to the data with a non-linear technique. You give your observations and your neighbour comes back within an hour to show you a nice graph (see figure) and the estimated parameters: Y_0= 5.10 and r=0.199.

fit of linear model to transformed data



fit of exponential model to non-transformed data



Task f: Calculate the model values for day 0, 7, 14, 21 and 28 with your neighbours estimate of the parameters.
Task g: Calculate the predicted number of algae on day 43 with your neighbours estimate of the parameters.
Task h: Explain the difference between the parameters estimated with the two methods.
You now remember, that your method of measuring the number of algae has a constant absolute error. The difference between the measured number of algae and the actual number of algae in the pond is never more than 20.
Task i: Reconsider your decision.

3.3 Metapopulation models

Population models describe the dynamics of single populations, e.g. a population of rice leaffolders in one particular rice field. Metapopulation models describe the dynamics of a number of populations, that are connected by each other by migration of individuals. For instance, when we are interested in the dispersal of rice leaffolders between rice fields in a region and the fraction of fields that are infested with rice leaffolders, then we need to adopt a metapopulation approach. Metapopulation models basically describe situations in which local populations become extinct and new habitats have to be colonized by that species in order to survive. This is the case with many specialized pests and diseases of agricultural crops. When a crop is harvested, this means the extinction of the local populations of the pests and diseases in that crop.
A very simple model for the metapopulation dynamics is described below. Let us imagine a pest species, e.g. rice leaffolder, living in an environment existing of many similar habitat patches, e.g. small rice fields. The size of the local leaffolder populations occupying these rice fields is assumed to be either 0 (rice field not yet infested) or K (carrying capacity, being the outcome of population growth, predation etc.). Local dynamics are thus ignored, apart from the colonization and extinction (harvest) events. A harvested field is soon replaced by a newly planted field, thus the total number of rice fields in the region is about constant. Movements from an infested field are assumed to be equally likely to all other fields, in other words the spatial arrangement of fields is ignored or is assumed to have no consequence. The rate of colonization is assumed to be proportional to P, i.e. the fraction of infested rice fields (sources of colonists), and to 1-P, the fraction of uninfested rice fields (targets for colonization). A constant proportion of the fields is harvested per unit of time and this constant proportion is equal for infested and uninfested fields. With these assumptions, changes in P in continuous time are given by:

$$dP/dt = cP(1-P) - hP$$

P: fraction of infested fields [-]
t: time [day]
c: colonization parameter [day^{-1}]
h: instantaneous relative harvest rate [day^{-1}]

Exercise 7
Task a: Check the dimensions.
Task b: Explain the similarity with the logistic growth function.
The duration of the rice crops is 120 days. There are rice fields of all possible planting dates.

Task c: Calculate the value for h.
The value(s) for P for which $dP/dt = 0$ is(are) called the equilibrium value or steady state for P.
Task d: Explain this.
Task e: Give the equilibrium value(s) of P.
Task f: Assume that $c = 1/90$. Calculate the fraction of infected fields in the region, when we also assume a steady state for P.
In another region, the farmers practice ratooning. The leaffolders survive ratooning, but are removed with the second harvest that takes place 240 days after planting.
Task g: Calculate the fraction of infected fields in the region, when we also assume a steady state for P (and $c=1/90$).

3.4 Spatial models

The individuals of a population are not always evenly distributed over a crop. Some spots in the field can have a much larger number of pests than other spots, that can still be free of the pest. On a larger scale, some fields are infected while other fields can still be uninfected. The uninfected spots or fields will eventually become colonized by individuals coming from the highly infested spots or fields. The models, that are used to study the distribution and migration of pests and diseases are said to be **spatial**, if they incorporate the distance and location of the populations or subpopulations. The metapopulation model in the previous paragraph is used to describe the spread of some pest or disease over a region, but it is not explicitly spatial, as the location and distance of the fields are not considered.
A very simple spatial model, that describes the number of spores deposited as a function of distance of the source is:

$$S = a.e^{-b.d}$$

S: number of spores deposited per surface [$N.m^{-2}$]
a: parameter [?]
b: parameter giving the slope of the dispersal gradient [?]
d: distance from the source [m]

Exercise 8
Task a: Give the value for S, when $d = 0$.
Task b: What is the meaning of parameter a?
Task c: Give the dimensions of a and b.
The parameters of this simple model can be estimated by fitting the model to observation of spore deposition. A logarithmic transformation can be used to obtain a linear model and perform linear regression.
Task d: Perform the logarithmic transformation.

At larger distances from a source of migration, the number of migrating insects or spores is often so low, that it is impossible to measure it. Still one spore or insect can be enough to start a new population. Thus at larger distances from a source, it becomes a matter of chance, whether a field is colonized or not. Models dealing with chance are **stochastic models**, and they are discussed in the following paragraph.

3.5 Stochastic models

If we examine an individual spore of a plant disease on a leaf, it is a matter of chance whether this spore will infect the leaf. But when the leaf is covered with hundreds of spores, there is no other possibility, than that the leaf will become infected. It is good to realize, that a lot of events are probabilistic, at least from some point of view. Models, that take into account, the probabilistic nature of processes are called **stochastic models**, in contrast with **deterministic** models. Deterministic models will always generate the same outcome given a set of circumstances. Stochastic models may generate completely different results, starting with exactly the same circumstances.

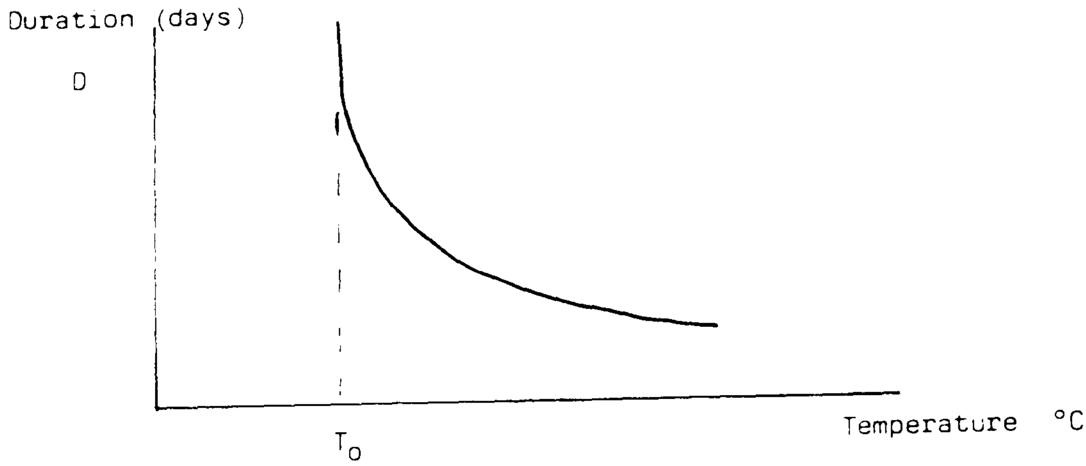
Exercise 9
Imagine a very peculiar, hypothetical insect, that we shall call the "stochastic paradox insect'. The adults of this insect mate directly after hatching from the pupa. The inseminated female keeps the sperm in a special cavity and the eggs are not yet fertilized. The female has only two eggs. If the female can find sufficient food during the first day, the eggs will both divide during the night and the female will look for more food the next day. If the female can not find sufficient food during the first day, the eggs will be fertilized with the stored sperm and they will be deposited after which the female dies. If a female, that was successful the first day, finds enough food during the second day, the four eggs will divide all four during the night, etc. If this female can not find sufficient food, the four eggs will be fertilized and deposited. In a particular habitat, the probability for a female to find sufficient food during a day is 0.5.
Let P(x) be probability of finding insufficient food during the x^{th} day after hatching, while sufficient food was found during the previous days. Let N(x) be the number of eggs laid, if the female finds insufficient food during the x^{th} day.

<u>Task a</u>: Complete the following table:

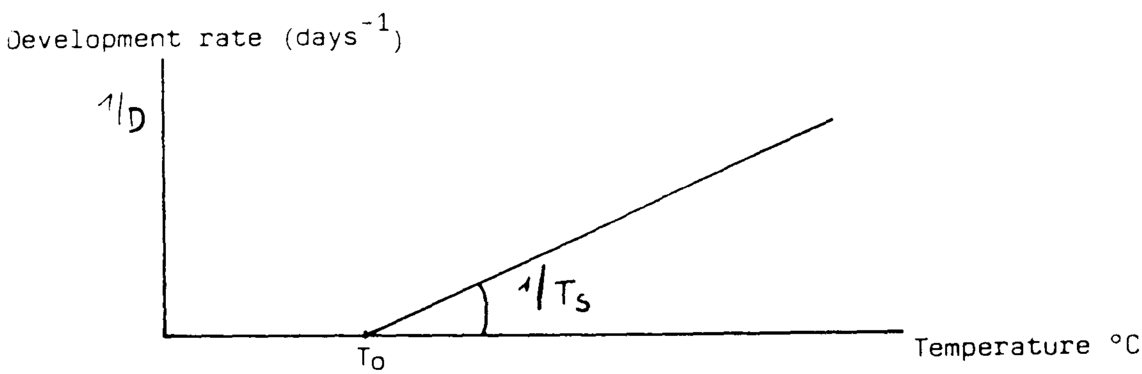
x	P(x)	N(x)
1	0.5	2
2		
3		
4		
5		
6		

The expected number of eggs deposited per female is indicated with the symbol E and can be calculated as follows:
E = P(0)*N(0) + P(1)*N(1) + P(2)*N(2) + P(3)*N(3) +...+ P(m)*N(m)
<u>Task b</u>: Explain this calculation.
An inseminated female, that finds sufficient food every day, can live for 20 days. If she found enough food during the 20^{th} day, the eggs will be divided, fertilized and deposited during that night.
<u>Task c</u>: Calculate E.
<u>Task d</u>: Perform a Monte Carlo simulation of the number of eggs deposited for one female. Use a coin, or the random number function on a calculator to simulate whether the female finds enough food during the first day, etc. The simulation results of all participants will now be used to calculate the simulated average number of eggs per female.

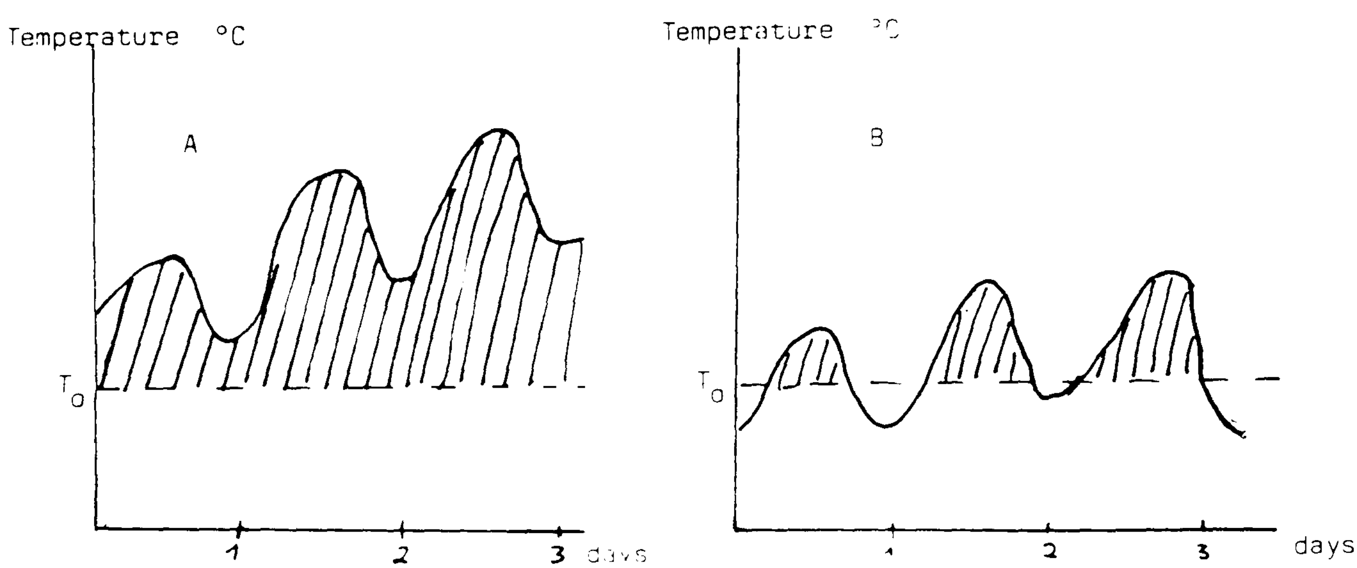
Duration of development in relation to temperature



Development rate



Shaded area is the temperature-sum (degree-days) for these three days



Task e: Explain the difference between the results of task c and task d.

Exercise 10
A population is growing linearly under favourable weather conditions. If the weather conditions are favourable during a day, the population increases with 10 individuals during that day. If the weather conditions are unfavourable during a day, the population does not increase during that day. The probability, that the weather conditions are favourable during a day is 0.5.
Task a: Calculate the expected number of days with favourable weather during 200 days.
Task b: Give the expected growth of the population during 200 days.
Task c: Perform Monte Carlo simulation for the growth during 10 days. Use a coin or the random function on your calculator to simulate the chance of favourable weather.
Task d: Consider the use of Monte Carlo simulation for the growth during one day, 10 days and 200 days.
Task e: The same exercise, but with the probability of favourable weather being 0.25.

3.6 Phenological models

After the end of the cold season or the dry season, many pests and weeds need to pass a development before they appear. The duration of this development is usually strongly dependent on temperature, because insects and plants do not regulate their own temperature. In experiments with constant temperature regimes, the following relation between the duration of the development and temperature is often found:

$D = T_s / (T - T_0)$ and $T > T_0$

D: duration of the development [day]
T: constant temperature during experiment [°C]
T_0: threshold temperature below which the organism does not develop [°C]
T_s: temperature sum [days*°C]

Why T_s is called the temperature sum can be seen by rearranging the above equation:

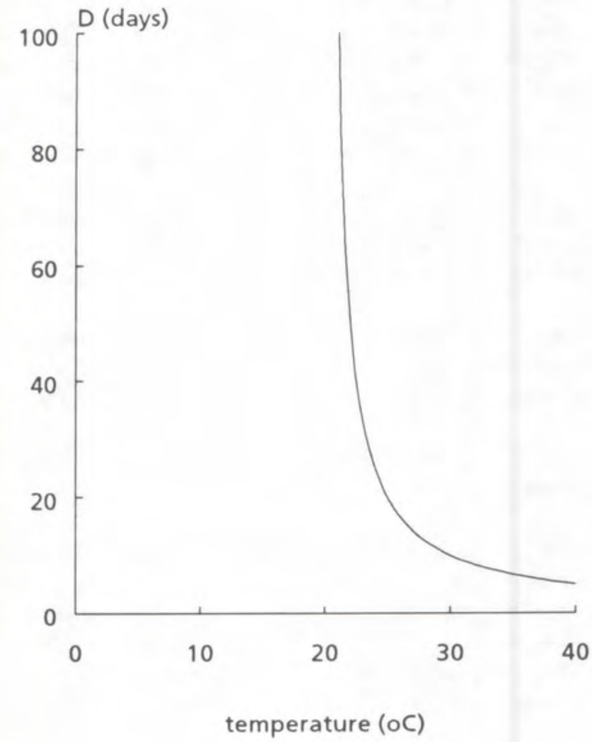
$$T_s = D*(T-T_0) = (T-T_0) + (T-T_0) + (T-T_0) + (T-T_0) + \ldots + (T-T_0)$$

To estimate T_s and T_0 from the experimental data, it is convenient to rearrange this equation again:

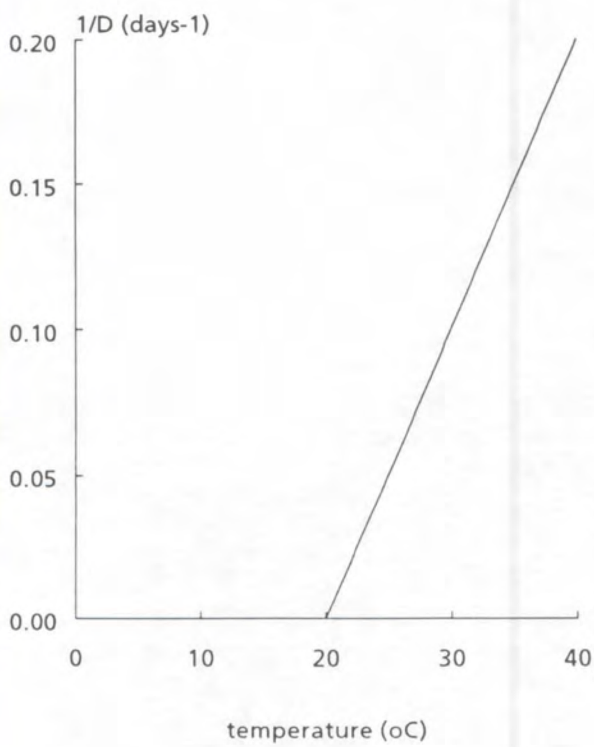
$$1/D = -T_0/T_s + T/T_s$$

By plotting 1/D against T, T_s and T_0 can be estimated from the observations by linear regression. The inverse of the development duration D, 1/D is called the development rate.

degree-day model



degree-day model



Exercise 11

Task: Check the dimensions of the above equation.

To apply the degree-day model in practice, we must assume, that the pest react exactly the same to variable temperatures as it did in the experiments with constant temperatures. Moreover, it has to be assumed, that the insects react instantaneously on change in temperature. If the temperature fluctuates strongly during a day, the average daily temperature will not be a good indication of the conditions for development of the insect. Then average hourly temperatures can be used together with a degree-hour model.

Exercise 12

By plotting 1/D (as found in an experiment with constant temperatures) against T, an intercept of -0.2 day^{-1} and a slope of 0.01 $day^{-1}.°C^{-1}$ were found.

Task a: Give Ts and T0.

After start of the development (e.g. a heavy shower) the following temperatures are observed in the field:

day	°C
1	22
2	29
3	35
4	34
5	33
6	32
7	34
8	35
9	31
10	33

Task b: Calculate the day of appearance according to the degree-day model.

****

BOX 4.1

Types of Surveys

There may be special surveys such as a countrywide check for ascertaining virus vector population at a critical time of crop development and the subsequent degree of infestation, or a (single) assessment of yield loss due to a certain pest which, when completed once, would not involve gathering any additional information.

Examples:

- *Survey on barley diseases conducted in an area of almost 2 mill. ha in England in 1967 (James, 1969),*
- *Survey on barley yellow rust conducted in the Netherlands in 1962 (Zadoks, 1966).*

Other important surveys are surveys on the presence of a new pest which might have invaded the country (or region) from an adjacent country (region), and the prevention of its spreading would be of higher interest.

Examples:

- *Surveys on coffee leaf rust in Brazil and other Latin American countries in the years that followed the discovery of leaf rust in Brazil in 1970 and in Nicaragua in 1976,*
- *Survey on witches broom in cacao in Central America after its discovery in Panama.*

****

4. SURVEILLANCE, SAMPLING AND MONITORING IN IPM

4.1 The importance of field observations for IPM

Pest management depends on decisions made by the farmer as well as on decisions made on regional and/or national level. The decision making for the farmer on whether or not to apply a control method is based, among other criteria, on the actual crop and pest situation in his field (fig. 4.1).

Crops grow and pests develop and progress in the course of the cropping season and may be more or less susceptible to control measurements. In order to be able to apply the right control method at the most appropriate time we have to get sufficient information on the crop and its developmental stage (which might be susceptible or resistant) towards a given pest. We also need information on the development of the pest (life cycle) and its progress in the field in time and space. Finally, we can obtain valuable information (which helps in decision making) from the environment which influences both crop and pest. Among the various environmental factors (climate,soil, crop husbandry, etc.) the microclimate is the most important factor because it has a direct effect on crop and pest development, which can be assessed in a relatively easy way.

The changes which take place within a crop/pest-system have to be observed or **monitored** regularly, at least during the most critical periods (e.g. when danger of infestation and subsequent yield loss is high), following a previously defined methodology which has been developed and tested in previous studies. Each crop/pest-system requires its own methodology; however, the basic principles, especially of data collection, may be similar in most of the cases.
Some of the principal methods used for assessing influential factors on pest development which are also of significance for decision making in IPM systems, like host plant population, pest infestation and environmental factors are described in chapts. 5-7.

In the relevant literature the terms **survey, surveillance, monitoring, sampling** and sometimes **scouting** are often found as synonyms or used with almost identical meanings. For a better understanding the following will give a short definition and description of the activities behind these terms and their use in the context of IPM.

4.2 Surveys

Surveys were probably one of the early forms of pest monitoring, reported first from Denmark in 1884. Since the beginning of this century, surveys were conducted on a more planned and organized scale.
The term **survey** nowadays is mostly used for an organized evaluation of pest infestation, either at a certain moment or over an extended period of time at district, regional or national level. The **objective** of a survey in plant protection usually is the detection of pest infestation, the mapping of pest distribution or crop loss assessment. Thus, surveys often just record the presence or absence of a pest, sometimes also pest intensities, and the damage they cause, whilst **thresholds not necessarily** have to be involved.

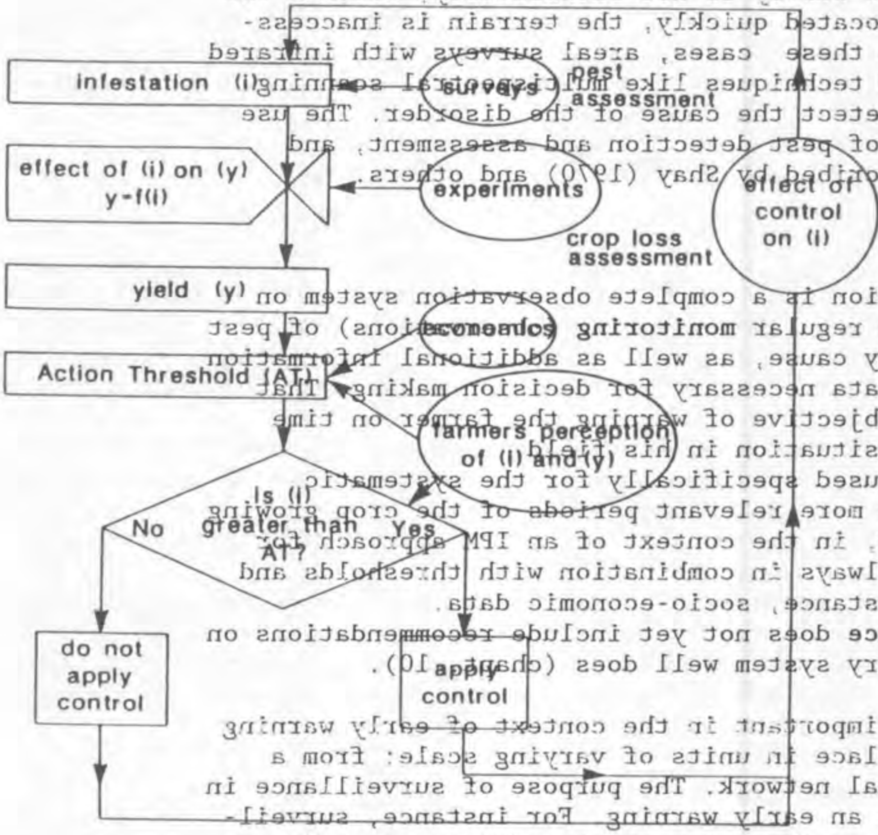


Fig. 4.1: A pest management decision system, showing the place of pest assessment and crop loss assessment

The information obtained by surveys may lead to strategical control measures, covering a longer time span, whilst **surveillance** or **monitoring** activities should lead to immediate (tactical) actions (see also chapt. 1).

In a survey which has to cover a greater area, fields to be sampled are selected so that together they represent the whole area. The design of a survey is not easy because it requires a compromise between various requirements, such as:

- good statistical sampling, so that the sample of fields is representative for the whole area;
- transportation of personnel, since observers have to visit the sample fields;
- time, because a survey has value only when a large area is covered within a short period; and
- money, as surveys tend to be costly.

There are situations when a ground survey is not satisfactory, because the early infection foci have to be located quickly, the terrain is inaccessible or the area is too large. In these cases, areal surveys with infrared sensitive photography or advanced techniques like multispectral scanning can be helpful, but they cannot detect the cause of the disorder. The use of areal surveys for the purpose of pest detection and assessment, and yield and loss assessment are described by Shay (1970) and others.

4.3 Surveillance

Pest surveillance in crop protection is a complete observation system on pests in the field which includes regular **monitoring** (observations) of pest appearances and of the damage they cause, as well as additional information on the crop and other important data necessary for decision making. That means that **surveillance** has the objective of warning the farmer on time (=early warning) on a given pest situation in his field.
Lately, the term **surveillance** is used specifically for the systematic monitoring of pests during one or more relevant periods of the crop growing season, and, as already mentioned, in the context of an IPM approach for the control of these organisms, always in combination with thresholds and other important data, like for instance, socio-economic data.
In its strict meaning, **surveillance** does not yet include recommendations on control measures, which an advisory system well does (chapt. 10).

Spatial aspects of surveillance (important in the context of early warning systems): Surveillance can take place in units of varying scale: from a single farmer's field to a national network. The purpose of surveillance in this respect is mainly to achieve an early warning. For instance, surveillance on the farm level may focus on the hot spots to achieve an early warning of impending infestation or infection. For a regional or national network a good communication system is very important to achieve early warning through surveillance (see flow chart on surveillance systems, fig. 4.1).
The national resources and priorities of each country will determine the way and the speed in which warning services to the farmers can be developed.

Box 4.2:

Pest surveillance comprises:

1. *monitoring methods and techniques of*
   - *host plant development (chapt. 5)*
   - *pest development (chapt. 6)*
   - *environment (chapt. 7);*
2. *data bases, retrieval and analysis procedures (statistics) for long term use in decision making (chapt. 9);*
3. *warning, training and demonstration (chapt. 10);*
4. *evaluation of efficacy and demonstration of results.*

(see also IPM scheme)

Temporal aspects of surveillance (data analysis for future planning): Field data from previous years may be used to retrieve information relevant to surveillance purposes. From existing data crop loss information can be derived as well as data on the probability of occurrences of pests in the past. Such data can be used for planning, cost-benefit calculations of pest control strategies, forecasting, etc. Surveillance data thus may have a relevance beyond the present season or situation.

4.4 Sampling

Sampling is a systematic and statistically justified way of extracting information about a large population by observing and taking data of only a small part of that population. Based on this limited amount of data conclusions are drawn about the entire population. Because this is a risky business the rules of the game are rather strict.

Sampling is used to determine pest population densities, disease incidence, yield of plots etc., in general there where quantitative data are needed to get information on processes going on and the state of affairs at a given moment. The use of these data can be for research or for application in IPM during the process of decision-making. The purpose of this data collection determines the demands regarding the accuracy, simplicity, reliability etc. of the sampling results.

Sampling methods for practical application should be simple, rapid, and adequate in their reliability. For research purposes which need higher degrees of precision, more refined and elaborate methods and techniques may be required.

According to Yates (1960) a sample is a set of units or portion of an aggregate of material which has been selected in the belief that it will be representative of the whole aggregate (= population). This can either be farms, fields of parts or them, or entire host plants or a particular organ of the host plant. The smaller entities in this sequence may be used as subsamples or sampling units (which again may have "elements", e.g. individual leaves).

Sampling may be done without removal of specimens, or specimens may be removed for investigation by experts in the laboratory.

Sampling can, but does not always require the use of tools such as wire quadrate frames, catch nets, growth keys, or sophisticated appliances for remote sensing.

4.4.1 Sampling plan

After the purpose of an intended sampling operation has been formulated information on a number of aspects is necessary:

****

Box 4.3:

A good sampling plan should include information on:

1. The pest or disease: identification, sex, stages, harmful stages; the behaviour of pest or disease in relation to sampling; especially for diseases damage rating with key or illustration is very useful, within-plant distribution.
2. The crop: growth stage, sensitive stage.
3. The subject to be sampled (the pest or disease unit); stage, presence/absence or numbers, spots.
4. The size of the sampling operation: plot, field, region.
5. Sampling unit: surface units, volume, tree, branch, shoot, plant or certain number of plants in the row, leaf, tiller and its size.
6. The distribution pattern of the sampling units in the plots: even, random, clustered.
7. The start and the end of the period in which is sampled, sampling dates.
8. The frequency of sampling.
9. Sampling technique: fixed or removed unit, using hand lens, assessment keys, sweep net, traps etc., using direct or indirect counts.
10. Sampling method: random, stratified, systematic, multiple.
11. Sample size (the number of units to be sampled), to be decided upon after a plot sampling depending on requirements of accuracy.
11. Sampling procedure: what is exactly the goal of the sampling operation, who will be doing what, where and when, available transportation, gasoline money and trained manpower, design of field record forms, anticipated methods of data analysis, who will do the data analysis.

****

Information on the pest

When a new or unknown pest damages the crop and its course through time has to be followed, the first step then is to try to identify it.

a. Diagnostics

   This subject is treated elsewhere in this course.

b. If research is undertaken with a well known good identification species, all kind of biological information may be helpful to design a good sampling plan.
   - The life cycle of the organism in relation to the host may give information on the presence and number of stages (esp. the harmful stages) in the crop in a specific part of the season (host change in aphids, bugs, rusts).
   - The behaviour of pest in relation to sampling:
     Prior consideration has to be taken of the distribution and behaviour of the pest at the time of sampling.
   - The mobility of the pest:
     Sampling procedure itself, may cause highly mobile pests to scatter rapidly, e.g. white flies, midges and locusts.
     They can then only be sampled by sweeping techniques. Mobile insect pests and airborne pathogens invade fields from near-by sources, and may occur in larger number on the edges than in the interior parts of the crop. Those parts of the crop that may be endangered first, "hot spots", may therefore have to be sampled earlier. Especially in monitoring procedures these endangered zones may offer an opportunity for an early warning of a potential pest situation. But for assessment of pest densities these borders have to be avoided.

c. When it is difficult to count pest/disease individuals a damage rating may be constructed with a key or illustration. These methods are discussed in chapter 6.

Growth stages of the crop

Crop growth stages are important parameters to describe the morphological and physiological stages of the plant. They are discussed in chapter 5.

The subject to be sampled

For insect pests it is mostly simple to distinguish a specific pest unit in the field as they are mostly present as individuals. It depends on the purpose of sampling (research, advice etc.), the life cycle in relation to the damage, monitoring method, forecasting method and control method, which stage of the pest has to be sampled.
For diseases it is more complicated. It depends on the infection cycle whether it is possible to distinguish individual units. They can be found during specific stages of the disease as spores, pustules or separate spots on the leaves. But the hyphes of a fungi originating from different spores are often so mingled with each other in a host plant or in the soil that it is impossible to distinguish the individuals. In the latter case a surface unit can be taken as the disease unit. The decision which stage to be sampled depends on the same factors as in the pests.

The size of sampling plots in the field

Where large fields or orchards are to be sampled, the required accuracy can only be achieved if they are divided into several parts called "reference plots", which will be sampled separately. The size of these plots is an important consideration, which bases principally on the following two aspects:

a. The uniformity of growing conditions, with special reference to factors that may markedly affect pest populations. The more uniform the conditions and the more even the expected spatial distribution of the pest or disease, the larger the plot can be, e.g. from 5 to 25 hectares in field crops, less in orchards. However, even where plant stands appear reasonably uniform, there may be a fairly large variation, and this can be especially meaningful in crop protection work. To obviate errors due to such variability, two measures are advocated:
   - Subdivision of the reference plot into subplots (squares) which are themselves more homogenous than the reference plot as a whole.
   - Determination of the cause of variability: this may be a matter of topography (top, middle, or bottom of a slope) or of proximity to sources of infection (other crops, hedges, irrigation points or other sources).
     Once the cause responsible has been established, an attempt must be made to define the size of the area over which its influence is felt, e.g. does the influence of the hedge extend only to the rows nearest to it, or perhaps for as much as 20-30 m?

b. Various practical considerations also play a role in deciding on the size of the reference plot. The most important of these is the scale on which control operations are to be carried out: if this is a small-scale operation, e.g. plots to be treated by hand-borne equipment, smaller reference plots can be chosen than for fields to be treated by tractor-drawn equipment that may cover many rows in one passage: large reference plots will be chosen if pesticides are to be applied by aircraft, which cannot be selective in spraying some plots, but not others in the same field. Choosing large reference plots involves ignoring a certain amount of heterogeneity in the pest population.

The sampling unit

The sampling unit when using a direct absolute procedure for population density determination is a part of the habitat where the pest can be found. These may be basic units like: surface unit, volume of soil, a whole plant, plant terminals, number of plants per row, a tiller, a branch, shoot, leaf, petiole, fruit, etc. To be of value for research or in decision making these sampling units must be representative for the entire population. In other words, the statistical estimate of the sample mean should be a reliable estimate of the parametric mean. The biology of the pest gives the information to decide which units can be used best. It may be necessary to change the sampling unit throughout the season since the preference of a pest for a specific place may also change.

****

Box 4.4 *Sampling methods*

Random sampling is essential for most statistical methods and is the basis for more sophisticated sampling methods. However, pure random sampling so far has rarely been employed in crop protection work. This is not surprising. For, once adopted, the random process must be rigorously adhered to. Because of the usually not uniform distribution of disease, random sampling may even yield ill-representative samples or inferior sampling efficiency and is, therefore, rarely desirable at all levels of sampling (sample, sample units).

Stratified sampling: the lack of uniformity of disease occurrence favours stratified and systematic sampling.
Stratification divides the population (e.g. the field) into groups or strata (and substrata) of similar characters to minimize variability and increase accuracy of otherwise too heterogenic material. It thus can ensure adequate representation of existing subdivisions in the population if information on the population is available prior to the survey. One may, for example, stratify farms according to their size, soil type, crop rotation, varieties sown, and other relevant criteria. Stratification permits only estimates of sampling errors within the strata. If this is required, at least 2 units per stratum or majority of strata must be selected. In uniform strata one unit each should well represent the population. In this case special methods of estimation of sampling error have to be adopted. Sometimes stratified samples need adjustment or classification of sampling units on the basis of information obtained in a pilot sample or survey.

Provided that sampling units have been classified in the required strata, the selection of a stratified sample follows the same procedures as random sampling. If no classification was done prior to sampling but the number of units in each stratum is known, Yates (1960) suggests: select samples at random, keep a record as the selection goes on of the numbers falling on each stratum, reject any further members of a particular stratum as soon as the requisite number of that stratum has been obtained. If the numbers of units are not known, the whole population must be counted, in which case a classification as a basis for the subsequent selection of the sample may already be carried out.

A population may be stratified for two or more characteristics (multiple stratification). Here the substrata attain the rank of strata and one can proceed as described above. For instance, farms may be stratified according to size and to regions. The region-size group combinations form the substrata. For crop loss assessments, for instance, we may assume that we have enough information to do this (from censuses, surveys or other kinds, etc.). Sampling units are then selected at random until the total of every row and column of a two or more-way table for the sets of strata is at least equal to the calculated total.Based on the excess, the number of units to be rejected from the substrata groups have to be calculated, and those rejected being selected at random.

An example: determination of the yield of a crop. A large field contains two soil types: sandy and clay. To determine the average yield of the crop on this field, the field is divided into two parts: the sandy part and the clay part. These are both strata. Independent samples are taken from both strata to get better yield data. The yield of the sandy and the clay part can be compared now and differences are not lost in the total average of the entire field.

****

Sampling methods

Sampling is carried out according to a particular concept. Simple random sampling is the basic concept of probability sampling. This means that in random sampling all sampling units have the same probability of being sampled.
However, **random sampling** does not always give the most accurate results and is time consuming. Therefore, other methods of sampling have been developed which are more suitable for particular situations. An example is stratified sampling in which a field is divided in two or more strata (parts) which each differ in important respects with regards to the expected sampling results.

According to Yates (1960) the following sampling methods of how and where to take the sample in the field may be distinguished:

1. Random sampling;
2. Stratified sampling;
   a. with uniform sampling fraction
   b. multiple stratification
   c. stratification with a variable sample fraction
   d. sampling within strata with probabilities
3. Systematic sampling (from lists or other pre-emptive information);
4. Multistage sampling;
5. Multiphase sampling;
6. Sampling with probabilities proportional to size of unit;
7. Balanced sampling;
8. Line sampling;
9. Sampling on successive occasions.

It is important to eliminate any element of personal selection in locating the sampling units. This is achieved by true random sampling where each unit has the same probability of being selected. A correct random sampling takes much time to carry out. Therefore, other sampling methods are used instead to simplify matters and to make sampling more applicable for the user in the field: the extension worker and the farmer. Thus the instruction might be to select the nearest plant to the right foot after walking a specified number of paces between sampling points, and not to reject plants that are badly stunted.
In general it is better to take many small samples than to take only a few large ones to ensure a better coverage of the field to be sampled. The most commonly used sampling methods are described in Box 4.4 and Box 4.5).

Start and end of sampling

The biology of the pest in relation to injury offers the best information on the moment of start and ending of the sampling programme. There are also some important moments during the growth period of the crop (e.g. sowing date, appearance of flag leaf, start of flowering, start of heading etc.) that can be used as a biological date to start a survey.
For pests break of diapause (calculated with temperature sums), flight (observed with light or pheromone traps) may be important moments to start calculating when the injuring stages may appear in the field. Combinations of weather factors (long periods of rain plus high temperatures, or rain following a period of dryness) may be very favourable for the pest and thus trigger action. The sampling programme ends at harvesting or when the crop

****

Box 4.5 Sampling methods (continuation)

Systematic sampling is easier in practice, likely to yield more precise results, and can be less cumbersome than random or stratified sampling. Moreover, the bias of stratification is eliminated. However, problems may arise if a valid estimation of sampling error is required regularly, or if some unwanted periodicity is present. If not, systematic sampling can be highly satisfactory.
Systematic samples can be obtained from lists. In this case, of the units of a population every entry is selected, starting with a randomly selected number. This is frequently used and in line with the fact that much practical sampling in crop protection is indeed not fully random. The procedure itself approaches stratification rather than random sampling. When, however, every 20th plant in rows of potatoes is the sampling unit, and this 20th plant is chosen from random numbers between 1 and 20, then this fulfils all the conditions of a valid random sampling.
An example: determination of incidence of infestation of cabbage by caterpillars. Sample 100 plants/ha equally spaced alternatingly left and right from the spraying lanes and note the presence or absence of caterpillars.

For **multistage sampling** the population is divided into first, second, etc. stage sampling units and sampling is carried out in stages (e.g. field, tiller, flag leaf). The sampling method may be the same or different in each stage. The overall number of sampling fractions can often be kept constant by suitable choice of sampling fractions. Multistage sampling introduces more flexibility, allows for the selection of natural divisions and subdivisions as sampling units and brings about a concentration of the field work (i.e. optimization of sample size). But it is also less accurate than in one stage methods with the same number of final stage sampling units. An example: determination of mite population on leaves of water-melon. The first stage is the selection of the plant in the field, the second the vine and the third the leaf on this vine.

For **multiphase sampling** certain items of data are collected from all the units of a sample and other data only from some of these units, chosen as to constitute a subsample of the units of the original sample, etc. This incomplete use of all units make multiphase sampling different from multistage sampling. Its advantages are obvious in the case the required accuracy, or variability, differs greatly between the varieties sampled for. An example: a large sample for the determination of acreage of a crop and subsamples for the farms sampled for the yield. Very often the first phase yields the information needed to determine the size of the subsequent phase.

One method particularly suitable for aerial surveys is sampling by taking random points on a map with probabilities proportional to their size. Here, only areas, not fields are required. This method is more accurate than others for the yield determination from known total acreage and also might be of interest in crop loss estimates. This kind of aerial survey may be combined with line sampling consisting of sets of parallel lines of strips as the sampling units.
Line sampling also may be employed for rapid and extensive surveys conducted along roads if no estimates or sampling error is needed. An example: socio-economic survey on farming data and attitudes in a region.

Sampling on successive occasions is required if variability of sampling units as well as the variability of changes in these units, etc. is to be measured. Very often changes in all units are similar in spite of the variability of the units. One must be careful here to consider this kind of successive samplings in the same plots as independent observations. Repeated samples in the same plots over time constitute a time-series which gives a trend, but the observations are not mutually independent. This has consequences for the statistical analysis.
An example is the method described for plot sampling by GOMEZ & GOMEZ (1984). If, however, information on the population means is of interest, partial replacement in the sample or subsamples will usually be preferable. One has to make sure that re-survey does not modify the units relative to the rest of the population. For sampling on successive occasions there are four alternatives: independent samples at each occasion, fixed samples, partial replacement of sampling units and subsampling.

****

has reached a stage in which it is not sensitive any more for the pest, which may occur often in the period short before harvesting.

The frequency of sampling

This depends on the rate of population growth or rate of spread of an infestation. The higher the rate the more frequent the sampling.

Certain pathogens and insect pests may initially cause severe infections on individual plants or on small loci, long before even minimal numbers of pests or symptoms can be found on neighbouring plants. Detection of such early infections is of the greatest importance and requires taking numerous samples in order to cover the field with an adequately dense grid of sampling points. On the other hand, where disease spread is rapid, as in the case of potato blight (*Phytophthora infestans*), the frequency of sampling rather than the number of samples taken must be increased.

Sampling techniques

Sampling techniques pertain to how the sampling will be carried out in terms of how the observations are taking place and what equipment might be necessary (sweep net, white board, pitfall traps etc.); this discussed in detail in chapters 5 and 6.

Sample size

Next to the choice of an appropriate sampling method a decision on the sample size is needed for a representative estimate. With the smallest possible sampling error an optimum efficiency is required. Very critical in this context are the spatial distribution of the pest and disease and the accuracy desired.
Sampling in IPM usually starts with a pilot sample, a random sample of limited sample size. The sampling results give some preliminary information on mean and variability (sample standard deviation). Based on this information and the desired requirements of a sampling result (accuracy, reliability, etc.) a sample size which satisfies these requirements can be determined. Repeated checks on the quality of the sampling results may lead to further adjustments in sample size.

Sampling procedure

The sampling plan has to contain information on the sampling procedure: what to do with the samples, where to send them, etc. A thorough briefing of the collaborators as to the scope and organization of the surveys, symptoms of pest, possibilities of misdiagnosis, handling of questionnaires, etc. is necessary, as well as instructions on the preparation of data obtained for further processing. Important are the economic and personnel constraints.

4.4.2 Sequential sampling

Strictly speaking sequential sampling is not a sampling method but a decision-making method. This method has been developed during the second world war to give rapid information on the quality of products. In monitoring pest/disease populations this "sampling" method is used for classifying populations with a variable number of samples taken. The number of samples taken depends on whether or not the results so far obtained give a definite answer to the question posed about the frequency or occurrence of an event. In practice this means that after each observation, a choice between three possible further courses of action occurs:

a) to continue sampling,
b) to discontinue sampling and take control action,
c) to discontinue sampling and refrain from control action.

The notable feature of this method is its efficiency in terms of number of observations required to reach a decision. Extremes., e.g., very high or very low infestation levels, can be established by comparatively few observations, while intermediate infestation levels will generally require more observations. This implies that the expenditure of time and effort (cost) is minimal. The sampling plan assumes that the observations behave in accordance with a well-defined distribution, e.g. Poisson or Negative Binomial distribution. Thus extensive preliminary work is necessary to establish the type of distribution.

The decision to spray or not to spray is based on the sum of observations up to and including the one actually being made. In order to help this decision, two lines are drawn (fig. 4.2), the upper indicating the cumulative number of insects or symptoms for each number of units samples, which definitely necessitates treatment, the lower indicating which cumulative number definitely does not require treatment. Actual sampling results are represented by a broken line.

The zone between these lines represents the zone of indecision: as long as observations stay within this zone, sampling has to continue until the full number of samples previously planned (using the normal calculation procedures) has been taken. These relations or decision lines have the general equation of a straight line:

$S = bn + h_0$ (upper line)
$S = bn + h_1$ (lower line)

S = cumulative number of insects or symptoms observed.
n = number of sampling units.
b = slope of the line.
h_0 and h_1 are the intercepts.

What we need further are the thresholds for the decisions:

m_0 = the upper threshold for treatment.
m_1 = the lower threshold to refrain of a treatment.

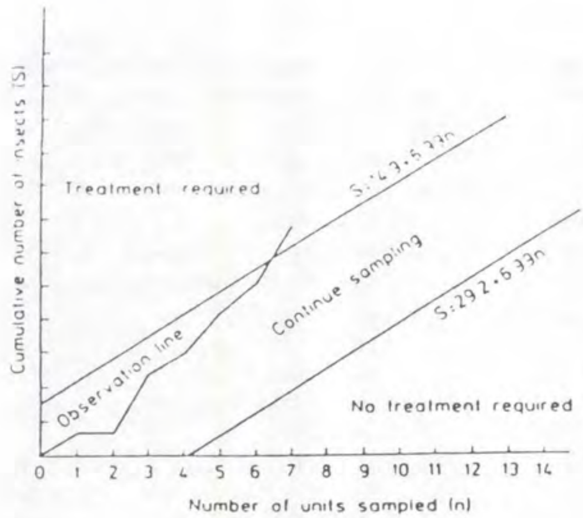


Fig. 4.2: Schematic representation of a sequential sampling plan. The sampling process proceeds step by step ("observation line") until a decision is reached.

Sequential sampling decision results for:

A - no treat situation (below threshold)

B - sample again tomorrow (near threshold)

C - treat situation (above threshold)

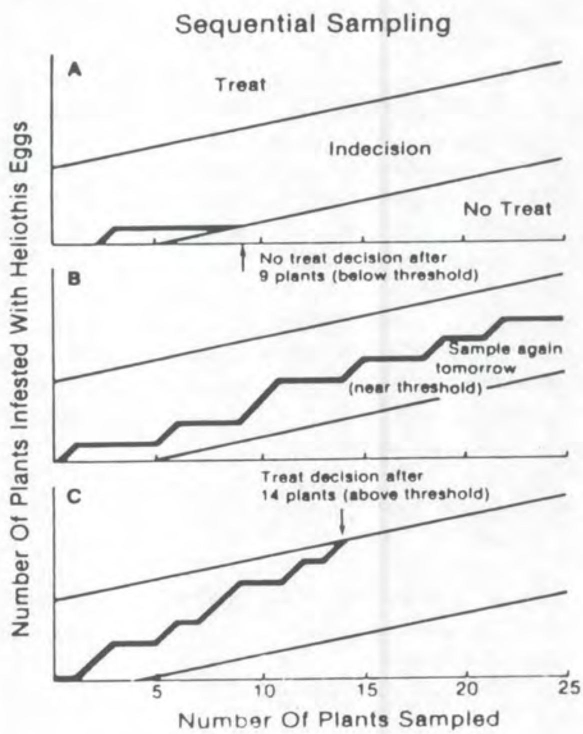


Fig. 4.3: Guidelines for Integrated Control of Cotton Pests

If the true mean of the population (m) exceeds m_0 than there exists a very small probability (α) that a wrong decision will be taken on the basis of the sample, e.i. that it will be decided not to spray when spraying is actually required (underestimation of the population). Similarly let the lower threshold m_1 be fixed so that the probability (ß) will likewise be small for an erroneous decision to be taken when m is below m_1, i.e. that it will be decided to apply a treatment which is actually unnecessary (overestimation of the population). The area between m_0 and m_1 is called the zone of indecision. Thus, for a given statistical distribution, m after definition of the values of m_0, m_1, α and ß the decision lines can be calculated.
We can change the reliability of the sampling plan by varying α and ß according to our own wishes, but remember that we pay for increasing the reliability by taking larger samples! The actual values of α and ß will be determined also based on the consequences of a wrong decision.

For more information on sequential sampling consult the paper of Onsager (1976).

Drawing up a plan for sequential sampling:

In accordance with all that has been said above, a number of steps have to be taken to draw up a sampling plan best suited for the purpose:

a) Choose an appropriate sampling method, e.g., stratified sampling, and sampling units, such as three consecutive plants, or the lower sides of leaves.
b) Define quantitatively the threshold for treatment, i.e., the mean infestation level of your reference plot that warrants treatment, and the size of the zone of indecision below this threshold where either decision (to treat or not to treat) will be acceptable.
c) Try to find, either in literature or on the basis of previous pilot sampling, the statistical distribution (e.g. Poisson or Negative Binomial) most suitable for observations on the pest or disease which is the object of your sampling.
d) Establish the size of the sample required for a fixed sampling plan, or the boundaries for decision in a sequential sampling plan, based on determined rates of error α and ß.
e) Construct the actual sampling plan, including the sampling points.

4.4.3 Statistics of sampling

A sample can be considered a statistical category. Sampling, therefore, is a mathematically based approach to obtain information from objects for specified purposes (yield, disease severity, pest population density). Its application in the field has both theoretical and practical aspects which should be considered carefully first.

Information on disease or pest situations in the field is derived from <u>data</u> which are the result of sampling. These data can be:

1. continuous (mass, weight, temperature)
   or discontinuous (number of larvae, number of diseased plants/ha)

****

Box 4.6

A good sample satisfies the following <u>conditions</u>:

1. *All elements of the sampled population should have the same chance of being sampled.*
2. *The sampling unit must be stable during the sampling procedure.*
3. *The proportion of the entire population which uses the sampling unit as a habitat must remain constant.*
4. *The sampling unit must be easily recognized.*
5. *The sampling unit must be practical in relation to available resources and desired accuracy.*
6. *Data from the sampling procedure must relate to a unit of area.*

****

2. ordered or grouped in certain categories or scales:
   a. nominal scale (classification, equivalence) (rice, corn, wheat, cassava)
   b. ordinal scale (ranking) (to size: pineapple, papaya, coconut)
   c. interval scale (size of interval) (temp. °F and °C: $F = \frac{9}{5}C + 32$)
   d. ratio scale (true zero point, independent ratio) (weight).

Parametric statistical tests for c. and d. only, non-parametric statistics for a., b. and c. (Siegel, 1956).

Important aspects of sampling:

- ***what exactly do I want to know?*** Problem definition must be critical, answer determines what to do, how to set up a trial, what information and analyses are needed.
- ***can I obtain this information without sampling?*** What is already known and can I save myself the trouble to get the information by reading the literature.
- ***use of results.*** The sampling plan is totally determined by the intended use of the results. For instance, studying the population dynamics of a certain pest may require very intensive and frequent sampling using a limited number of plants, whereas a classification of an infestation as needing treatment or not asks more extensive and perhaps less frequent sampling.
- ***resources.*** It does not make much sense to make a large sampling plan if you don't have the necessary resources at your disposal. In many cases transportation, and hence gasoline money, are vital to the ability to carry out the regular field research. The availability of trained personnel may be crucial, etc.
- ***accuracy.*** Your problem definition also dictates to a certain extent the accuracy of the data you need. It makes a difference in effort whether you need to distinguish subtle differences in pest development in two varieties or you need only to decide if a crop needs treatment or not. Remember that you pay for accuracy by having to take more samples.
- ***biology and behaviour of organism.***
  Distribution pattern:
  a. even distribution
  b. random distribution
  c. clustered distribution

Distribution: $\bar{x}$ is the **mean** of a sample and s is the **sample standard deviation**, and s^2 is the **variance**. **Standard error of the mean** is the variability among all the possible means of a data set: $s_{\bar{x}} = s/\sqrt{n}$.

Dispersion parameter for clustered distributions (neg.binomial):

$$k = \frac{\bar{x}^2}{s^2 - \bar{x}}$$

Coefficient of dispersion: $CD = \frac{s^2}{\bar{x}}$

****

Box 4.7

Sampling in practice:

* *carry out a pilot sample and calculate the mean and sample standard deviation*
* *calculate the confidence interval*
* *if necessary take more samples*
* *calculate the minimal number of samples necessary to fulfil the required degree of accuracy*
* *make a sampling plan.*

****

a. Even or regular distribution when: $\bar{x} > s^2$, CD < 1 or k < 0, according to a binomial distribution.

b. Random distribution when: $x = s^2$, CD = 1 or $k \geq 8$, according to a Poisson distribution.

c. Clustered or contagious distribution when: $x < s^2$, CD > 1 or $0 < k < 8$, according to a negative binomial distribution.

<u>Measure of variability of sampling methods</u>:

Coefficient of variation: $V = s/x$

The type of distribution of an organism in the crop is a function of its behaviour and may change at increasing age or development. It must be known to avoid gross misinterpretation of sampling results.

The spatial distribution of organisms depends on abiotic and biotic factors. Preference for a specific microclimate (temperature, humidity) and biotic factors such as reproduction, availability of food, low mobility etc. may cause clumping of organisms. The statistical frequency distribution depends on both the spatial patterns of infestation and on the size of the sampling unit. Where the pest or symptom spreads at random in the field, and its spread is independent of other pests or symptoms, the number of individuals to be found at each sampling point will follow what is termed the Poisson statistical distribution. But if a population exhibits an aggregated spatial distribution, and if samples are taken from randomly distributed points in space, the number of individuals per sample will follow one of the so called contagious (or clustered) statistical distributions. The most common one is the "negative binomial distribution". What that means for the determination of the sample size in relation to the accuracy is explained in the article "Introduction to sampling theory" (Ruesink, 1980).

Determination of the confidence interval and sample size

A set of sampling results, if sufficiently large, may give valid information on the properties or organisms sampled. When the purpose of the sampling operation is to obtain information on the pest situation in a given crop a number of factors are important (see also chapter 4.2).

If we take as example the population density of caterpillars in a crop as the object of information, we could take a number of plants at random, count the number of caterpillars on each plant, note this number per plant and calculate the mean number of caterpillars per plant (x). Our set of sampling results will show a certain variability, which is expressed by the sample standard deviation (s). This measure of variability can be calculated only when the caterpillar counts have been registered per plant.

Therefore, all sampling results must be noted per sampling unit, here the plant. The mean ($\bar{x}$) on itself does not mean so much and indicates only the magnitude of the pest population density.

It is an estimate of the true population density, symbolized by μ (= the Greek letter mu). But another set of sampling results will give another estimate because of the natural variability in the dispersion of the insects in the crop. Hence, we are not so much interested in the various means ($\bar{x}$) but only in the true population density (μ).

For practical reasons our sample size will be limited and in most cases it will be small when compared to the number of plants in the field we could not sample. Therefore, μ can be estimated only, not determined. It is estimated by combining the mean ($\bar{x}$) and the variability (s) of the sampling results to a range in which the μ can be found. This is often expressed as ($\bar{x}$) $\pm$ constant * s.

Decisions on control measures to be taken or not, often will depend on the results of our field sampling. Therefore, it is important to know what are the risks of making a wrong estimate. By taking into account the size of our sample and the probability that x will take extreme values we know how "certain" we can be that our statement about μ is correct. This "certainty" is based on the probability of extreme values of x, suggested by s, the sample standard deviation. How can we achieve this?

Since our sample size will be relatively small we must apply a correction factor which will also indicate the level of "certainty" (= probability) that our statement is correct. This correction factor is the statistic t, Student's t. Its value is determined by our sample size and the level of confidence we want to have that our statement is correct, for instance 95% chance. This leaves a chance of 5% that it will be incorrect. This value can be found in a table (table 1) with two entries:

1. the chance that our statement will <u>not</u> be correct (e.g. $P = 0.05$);
2. the number of degrees of freedom, abbreviated df. which is here the sample size (n) minus one: n-1.

For instance, if we want to make our statement on the 5% level (or: $P = 0.05$) and the sample size is 10 then the corresponding value for $t = 2.262$.

Instead of using $\bar{x}$ $\pm$ constant * s we are able now to introduce a sense of security in our statements by indicating the interval around x in which μ can be found under prevailing conditions of **sample size** and **confidence level**. The appropriate formula is:

$$\bar{x} \pm st/\sqrt{n}$$

where s = the sample standard deviation
t = the relevant value of the t-statistic
n = the sample size

$st/\sqrt{n}$ is often referred to as the **margin of error**. This formula gives the limits of the **confidence interval** of μ under the given conditions, according to:

$$\bar{x} - st/\sqrt{n} \leq \mu \leq \bar{x} + st/\sqrt{n}$$

If the value of $st/\sqrt{n}$ is known, for instance as a percentage of $\bar{x}$, the corresponding value for n can be calculated. When $st/\sqrt{n} = m$, then

$$n' = (st/m)^2,$$

where n' is the new sample size, adapted to our demands concerning the allowable confidence interval.

In this way the **minimal sample size** can be calculated for a given margin of error and thus a given measure of variability of the individual sampling results. Correct sample sizes can be determined based on a point estimate produced by pilot sampling.

Example: let a sample result be:

0 1 4 2 5 3 0 0 1 2 0 1 caterpillars/plant

Then: $\bar{x} = 1.58$ $s = 1.67$

$n = 12$ $df = 12 - 1 = 11$

Confidence level 5% or P = 0.05

Margin of error = $st/\sqrt{n} = (1.67 * 2.201)/\sqrt{12} = 1.06$

$\bar{x} - st/\sqrt{n} = 1.58 - 1.06 = 0.52$

$\bar{x} + st/\sqrt{n} = 1.58 + 1.06 = 2.64$

The confidence interval for μ (P = 0.05) is: $0.52 \leq \mu \leq 2.64$,
or:
"the true caterpillar population density is between 0.52 and 2.64 caterpillars/plant. The probability that this statement is correct is 95%".

N.B. The general formula to calculate the standard deviation s is:

$$s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n - 1}}$$

where: x = the values obtained from the n sampling units
n = the sample size.

The range test

Sampling results are evaluated by calculation of various parameters e.g. $\bar{x}$, s, $s_{\bar{x}}$ etc. and confidence intervals when necessary. When it is difficult to calculate s (the sample standard deviation) of a given set of sampling results another method can be used at small sample sizes (< 20). This method uses the **difference between the highest and lowest value** in the sampling results instead of $s_{\bar{x}}$ (the standard error of the mean) as a measure of sampling variability. This difference is called: the **range** and indicated by: w.

Where in the t-test the value of Student's t is: $t = \frac{(\bar{x} - \mu)}{s_{\bar{x}}}$,

in the range test the value of t is: $t_w = \frac{(\bar{x} - \mu)}{w}$.

The range test formula for the confidence interval is:

$$\bar{x} - wt_w \leq \mu \leq \bar{x} + wt_w$$

The appropriate value of t_w can be found in a table (table 2) with as entries the level of accuracy and the sample size, n.

The use of the range test is explained by some examples.

Example 1: Computation of the confidence interval:

Assume that we got the following sampling results:
1, 0, 6, 2, 0, 3, 5, 2, 3, 0 (larvae/plant)

The sample size, n = 10

The sample mean, $\bar{x}$ = 2.2 larvae/plant

The range, w = 6 - 0 = 6
For n = 10 and, for instance P = 0.05 the value of t_w in table 2 is: 0.230. To calculate the confidence interval the margin of error: wt_w can be determined now: 6 * 0.230 = 1.38. The confidence interval of our sampling results is then:

$$2.2 - 1.38 \leq \mu \leq 2.2 + 1.38 \quad \text{or} \quad 0.82 \leq \mu \leq 3.58 \ (P = 0.05)$$

Expressed in words: "the true population will be somewhere between 0.82 and 3.58 larvae/plant. The chance that this statement is correct is 95%".

Example 2: Test of significance between two sampling results:

Assume that we have the following sampling results:
plot 1: 2, 4, 3, 0, 0, 8, 0, 4, 1, 0 (aphids/leaf)
plot 2: 5, 3, 0, 0, 6, 5, 4, 7, 5, 4 (aphids/leaf).

What we want to test is whether these results differ enough to establish a significant difference. First we determine some parameters of the results:

plot 1: $\bar{x}_1$ = 2.2, w_1 = 8, n = 10

plot 2: $\bar{x}_2$ = 3.9, w_2 = 7, n = 10

The difference of the means is: $|\bar{x}_1 - \bar{x}_2| = 1.7$

The pooled range is: $w = (w_1 + w_2)/2 = 7.5$

$t_w = |\bar{x}_1 - \bar{x}_2| / w$ under H_0: "there is no difference between the populations μ_1 and μ_2 ($\mu_1 - \mu_2 = 0$)".

$$t_w = 1.7/7.5 = 0.226 \quad (t_w \text{ found})$$

For n = 10 and P = 0.05: $t_w = 0.230$ (t_w table)

→ t_w found < t_w table → there is no significant difference between both populations on the 5% level.
The confidence interval for the difference is:

$$|\bar{x}_1 - \bar{x}_2| - wt_w \leq \mu_1 - \mu_2 \leq |\bar{x}_1 - \bar{x}_2| + wt_w$$

Exercises

Exercise: Different treatments of *Plutella xylostella* larvae in heading cabbage result in the following sampling results:

plot 1: 4, 5, 10, 0, 4, 15, 0, 0, 6, 8
plot 2: 20, 16, 0, 4, 15, 0, 18, 3, 1, 0
The plant is the sampling unit, the sampling results are expressed as number of caterpillars/plant.

Questions:

1. Calculate the confidence intervals at the 5% level using the range test.

2. Test whether differences between these sampling results are significant (P = 0.05).

4.7 Monitoring

Monitoring for decision making in crop protection comprises

(a) the quantification of pests and factors which exert a certain influence on their development and
(b) the registration of qualitative and quantitative changes in time and space.

Quite often the terms **monitoring** and **sampling** are used in the literature for the very same activity. However, monitoring as it is used in crop protection has some slightly other characteristics compared to the sampling procedures in scientific research, because the objectives are different: **Sampling** is used to obtain a representative information from a population of which not every individual can be measured, whilst **monitoring** is used for decision making on whether or not to apply a control method.

Because each pest problem is different, monitoring methods for each specific insect pest, disease, weed, etc. most likely will also be different. Consequently, the methods must be tried out (usually through pilot studies) to fit the individual problem.

There are already numerous types of methods for monitoring at hand (see FAO manuals and publications). In general, all methods have as their objective to provide a number, either as an actual count or as an estimate, which reflects the density of a population or infestation. This number may pertain to any specific item, such as an insect pest, diseased or dead plant, or even to lesions on a specific leaf or plant. It is this figure which will be used in conjunction with other information to determine the conditions existing in the field and from which any recommendations for action will be made.
The accuracy of monitoring may be lower than in scientific sampling work but must fulfil the standards set by the decision rules. The risk to make a wrong decision must be low and acceptable. Taking representative samples is an essential step in monitoring crop pests.

The personnel doing the monitoring in the field has to be equipped with exhaustive information on the methodology which can be in form of an **instruction manual**, containing:

- information of the whole set-up and organisation of the surveillance system,
- instructions of how to proceed,
- sampling procedure,
- field recording sheets,
- diagnostical aids (colour photos, graphs, etc.),

****

BOX 4.8

*Monitoring methods in the field have to be cheap, non-laborious and simple to carry out. It depends on the purpose of monitoring which method will be used. Monitoring can be executed by experts (plant protection personnel, qualified extensionists) or volunteers, local promoters, local agents of the extension service or especially for the job trained and paid persons (**field scouts**) and, preferably, by the farmer himself.*

****

****

BOX 4.9

Sampling *is a mathematical approach to obtain (with a given certain degree of precision) information from the field as representative for the whole population.*

Monitoring *is a regular and organized collection of data (based on clearly defined methods) to obtain quantified information on the changes taking place in a population.*

Scouting *is sometimes used as synonym for monitoring, especially when it is done by personnel specifically trained for the job, on larger areas and executed by private counselling companies.*

****

- instructions for codification of the pests, natural enemies, growth stages, microclimatical parameters, etc.),
- descriptions of how to use the equipment for monitoring (insect traps, spore traps, meteorological instruments, etc.),
- thresholds,
- execution of field experiments,
- formate and form of transmission of the information (via computer, post-card, record sheet, etc.),
- instructions for control methods (pesticides, dosification, etc.).

Monitoring and **scouting** are the activities within a surveillance system by which data for forecasting (chapt. 2) and decision making are generated. This means that the data collected should be reliable, valid and complete. That implies in the first place strict observations of prescribed sampling techniques. Data must be collected with utmost care and the best of training possible. If one feels that adjustment to the instructions is necessary, one should report back to the authority responsible for the instructions.One should not correct them oneself and thus work on a different instruction which is not known to those accumulating and evaluating the data. The other aspect of reliability in this context is, of course, adequate knowledge of symptoms and biology of the pests involved. This implies continuous training.

Time of monitoring:

Usually, it is not required to monitor a crop the whole year long or during the entire cropping season. Pest assessment is most valuable at the time when pests have the maximum effect on yield. For a critical growth stage, the best measure of pest attack may be the first appearance of eggs or adults, the cumulative number of eggs, etc., the first small lesions. etc. The time or measure which gives the highest correlation coefficient (r) or coefficient of determination (r^2) in a yield regression will be most useful in forecasting the effect of infestation on yield. Several methods or times can be defined through a multifactorial regression (Gomez & Gomez, 1984).

5. EVALUATION OF HOST PLANT POPULATIONS

5.1 Importance and use of host information

Plants affect their own **environment** and as host plant also the environment and development of their **pests**. The genetical and phenological characteristics of the host plant influence the development of a pest (see box 5.1), so do qualitative changes within the plant (content of compounds in organs) or quantitative changes (amount of susceptible host tissue): The type of crop, the growing habit, plant height and structure, variety (susceptible, resistant, moderate resistant, etc.), the developmental stage, age of its organs, planting system and consequently crop density, foliage or amount of biomass, all these are factors or variables which influence the micro-climate within the field and the development of a given pest in time and space.

Within the concept of IPM the information gained through evaluation and monitoring of host plant populations has become an important tool for decision making, because with this information future events of pest development can be predicted (see also chapt. 9):

The stage of development reached by the crop is the criterion most convenient to link with prospective pest appearance (forecasting). In some cases the phenological stage alone can serve as such a criterion, in other cases it must be viewed in conjunction with weather and/or pest factors. **Example**: Pests may depend on flowering to enter the host plant to cause infection/infestation. This then is the risk period during or before which control must be achieved.

Quantification of the host population is important for the assessment of pest progress, especially in the case of fungal diseases: If we assess the disease at two successive times and observe equal percentage of leaf area covered by lesions, the implication is that the disease has not increased over time. If, however, the host has grown and doubled in size between the two assessments, then the amount of disease in absolute units has also doubled.

Further use of host plant criteria:

a) When assessing the injury or damage inflicted on a crop, it is important to know the normal image of the crop as reference (e.g. deformations caused by pest infestation compared to normal plant growth).

b) When assessing crop losses, critical stages for yield development are used as reference (e.g. crop loss assessment in cereals). The critical stage is the stage during which not necessarily the host is most susceptible, but during which the best correlation with yield loss can be established.

c) In the case of plant diseases, the monitoring of the crop also serves to compare epidemics on different cultivars, at different locations, or in different growing seasons.

****

BOX 5.1

Host plant influence on pest development

Examples:

- ***high crop density*** *favours humid microclimate which for its part favours infection and dissemination of fungi (e.g.* ***Septoria tritici*** *in wheat);*

- ***growth stage*** *of host plant or organs: Apple leaves that have just burst from bud are more susceptible to apple scab (Venturia inaequalis), but become more resistant when aging. The very early 'pinhead' stage of coffee berries is not susceptible to CBD (Colletotrichum coffeanun), but risk of infection increases with age of coffee berries to a maximum around the 10th week after flowering.*

- *high amounts of* ***susceptible host tissue*** *(e.g.leaves) favour and low amounts reduce or slow down pest increase: Coffee trees with dense foliage may show a high percentage of coffee leaf rust attack (Hemileia vastatrix), whilst in the following year after heavy defoliation the percentage of infection remains rather low.*

****

****

BOX 5.2

Definitions:

Growth *is the increase of size or weight of plants or plant organs over a certain period of time. Growth is a quantity, expressed as shoot length, leaf area, or dry weight, with the appropriate dimensions* $[L],[L^2]$*, or* $[M]$*. Plants grow, crops grow as the composing plants grow; plant populations grow as the composing plants grow or increase in number.*

Growth rate *is the velocity of growth, or a growth per unit of time, and the dimensions are* $[L \cdot 1/T]$*,* $[L^2 \cdot 1/T]$ *or* $[M \cdot 1/T]$*.*

Development *is the passing of the plant through successive, morphologically different phases: from seed, through vegetative and generative growth stages, to seed again. Perennial plants have similar sequences, beginning and ending with buds. The passage through these phases is environmentally affected and genetically and physiologically controlled.*
The science of the relations between climate and the periodic biologic phenomena, as observed during plant development, is called ***phenology****.*

****

5.2 Monitoring techniques of host plants

Some of the attributes concerning the host plant population are obvious and are routinely recorded. Cultivar, geographical location, soil type, time of year or planting (or year of planting with perennial crops), planting system (in rows or broadcast), time of harvest are examples of such information.

In the context of IPM research and monitoring for decision making the qualification and quantification of crop development have become very important attributes of crops which can be measured at critical times of the year (when pest infestation is most likely).

Qualitative measurement is based on the development of the crop in terms of vegetative, reproductive and ripening phase, whilst **quantitative measurement** is based on crop growth and can be expressed in plant hight, number of tillers, leaf area index, dry matter accumulated, etc. Which parameter or plant character is to be used depends on the objective of research or the monitoring instructions.
When measuring quantitative and phenological changes of the host population the time intervals between the measurements depend on the growth rate of the host and the accuracy required.

5.2.1 Measurement of crop development

Physiological susceptibility and resistance of the plant or plant parts may change in the course of its development. Therefore, it is important to be able to differentiate between the distinct developmental or growth stages of a crop.

As an indication for crop age one can use the number of days after sowing or after emergence, the latter in most of the cases makes more sense. It is a parameter relatively easy to obtain but not always provides comparable values on the actual **physiological state** of the crop as plant development may vary with climate, variety, soil, fertilization, etc.

More and exact information is obtained when using the so-called **growth stage keys** (fig 5.1) which are defined according to the physiological and phenological development of the plant. In the case of a perennial crop one can differentiate according to physiological phases like flowering or fruiting (Fig. 5.2).
For the description of the individual growth stages both numerical codes and sketches are used (see also practical instructions no. 4.2).

When phenological stages have been classified, development can be assessed quantitatively by determining the proportion of plants or plant parts that have attained a certain stage. Growth stage scales normally only indicate the physiological/reproductive status of the host, but in some cases (e.g. soybean and peanut scales) may indicate how much a host has grown.

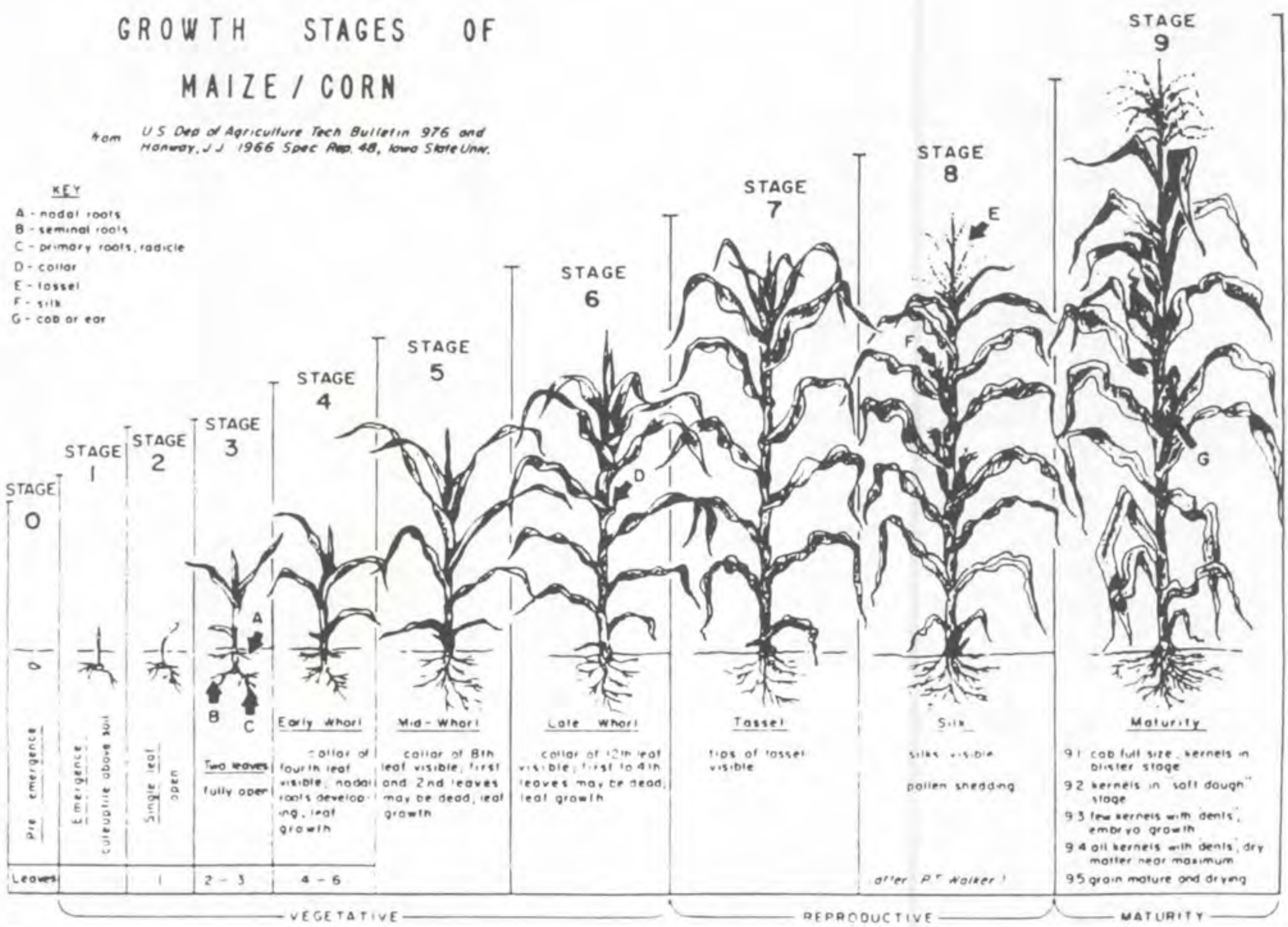


Fig. 5.1: Example of a growth stage key

Useful descriptions of growth stages of field crops (wheat, barley, oats, rice, maize, tobacco, cotton, sunflower, potato, broad bean, soybean, etc.) have been published in the crop loss manual of the FAO (Chiarappa, 1973). These development scales are more useful to extensionists than the more elaborate scales of growth stages published for research purposes. There are also some very useful growth stages for deciduous fruit trees (and vine) worked out, which are used in relation to pest outbreak.

Example: Apple scab (*Venturia inaequalis*) is not expected to develop before the sprouting has reached the stage designated as C3 (fig. 5.2). It is only **at** or **after** this stage that scab attack is imminent if and when weather conditions favour the release of ascospores from overwintering perithecia.

5.2.2 Measurement of crop growth

For some purposes (more detailed analysis) a quantitative assessment of the host population may be needed, e.g. for studying epidemics where an integral component of host growth characteristics is defoliation. Many fungal diseases (mildews, leaf spots, rusts) can infect the host from the seedling stage to the mature plant; in these cases, the amount of host tissue present is not constant and may affect the interpretation of the data obtained in the field. The increase of infected hosts indirectly reflects the increase of the population of pathogens (especially obligate parasites). The growth of a pest population is often measured by counting the number of the insect pest per plant unit, e.g. a leaf, a tiller, a plant, etc.

The following **variables** can be used for measuring host plants and crops **quantitatively**:

- **crop density** :
  - soil coverage in %,
  - number of plants (tillers, stems) per unit area (usually expressed per $1m^2$ soil,
- **plant height and structure**:
  - plant height in cm (m),
  - number of leaves per stem,
  - height of leaves, etc.

- **biomass (leaves, flowers, fruits, roots)**:
  - leaf area index (LAI): amount of leaf area (one side only) per unit of soil area $[L^2 \cdot 1/L]$.
  - leaf area density (LAD): is the LAI differentiated to plant height,
  - flowers and fruits are usually counted or their volume estimated per unit soil area,
  - roots length and volume can also be measured per unit soil area.

Crop density and leaf area index are the variables mostly used; the importance of measuring roots or other variables depends on the host plant/pest-system under consideration.
The information gained from the leaf area index is mainly of analytical importance in relation to pest increase, for instance when evaluating the effect of pesticide application.