

# THE BEHAVIOUR OF PESTICIDE RESIDUES IN FRUITS AND VEGETABLES: EVALUATION BY DECLINE CURVES

HELMUT FREHSE AND HANS-FRIEDRICH WALTER

## 1. INTRODUCTION

Along with residue studies on harvested crops, reliable data on the behaviour of pesticide residues prior to harvest are an essential prerequisite for the registration of pesticides. Pertinent results are obtained from field or greenhouse experiments, which are carried out under controlled conditions, observing the principles of Good Agricultural Practice and the fundamental criteria for accurate sampling and analysis («supervised trials»). Samples are taken after the (last) application of the pesticide (at «day 0») and in certain intervals thereafter, mostly up to and including the day of regular harvest.

Evaluating the decline of residues as a function of time is an interesting subject and provides an important key to understanding residue behaviour. This paper is aimed at a particular aspect of this problem area, specifically, *how to describe* the decline of residues on treated crops.

It may be feasible to agree on nomenclature first. Pesticide residues diminish measurably with time, although the speed of diminution may differ from compound to compound, as it may vary according to environmental conditions. Physical loss, chemical degradation, and dilution due to growth of the treated plant are the main factors contributing to the diminution. Various terms are used to describe this process, including breakdown, decay, decline, dégradation, or dissipation. In some instances, these terms are used interchangeably or as synonyms, but in other instances, they should preferably be used to describe a particular process, e.g. degradation (in a chemical sense). Although the latter term is often used in a general sense, we prefer the term *decline* for describing the *overall* phenomenon of residues decreasing with time after application of a pesticide.

A problem arises when the results from decline studies are to be evaluated. Simple tabulation of the data is not too instructive. Plotting the measured values on graph paper and drawing a curve through the points «by hand» is also not very satisfactory. The best solution

is to determine a *functional relationship* between residues and time, which will yield a mathematically defined decline curve. Moreover, errors unavoidably inherent in the individual analytical measurements will thus become «smoothed out» and valuable information can be derived from the residue data, e.g.

- the mean («estimated») residue level at a particular point of time within the period of observation,
- the time taken for the residue to drop below a given level or to reach a given percentage reduction (e.g. 50 % or 90 %),
- the confidence interval of the decline curve,
- prediction intervals for the residue at individual points of time or for the entire curve, indicating the levels of residues which may be expected, with a given degree of probability, from future experiments under comparable conditions.

Such information can be essential for establishing pre-harvest intervals and/or maximum residue limits (MRLs). Also, through use of this information, comparisons of residue levels observed under different conditions (e.g. with a different formulation, a different application technique, or on another crop or under different climatic conditions) may become more reliable than those made by just «comparing figures». The question, here, is how to find expressions for the functional relationship. The purpose of this paper is to present several approaches to this problem, which we developed in the past (8 - 11).

## 2. FORMAL APPROACHES

### 2.1. Scale transformations for linear regression of residues on time

Mathematical characterization of the decline of residues and determination of the statistical parameters which describe this process are easiest to achieve when the relationship between decline and time can be expressed as a *linear regression*. This involves fitting a straight line to the measured values in such a way that the sum of the squared deviations of the measured values from the line is minimized (method of least squares). One is more or less compelled to revert to an approach of this kind if residue measurements from only a few points of time are available, and «simple» solutions are desirable. In many practical situations, particularly with rapidly developing crops, it will often not be feasible to take samples at many more than four or five sampling dates. This imposes certain restrictions on the statistical evaluation of the analytical results.

A direct linear decline of residues with time hardly ever occurs. Consequently, time and/or residue values must be transformed to yield a linear relationship. The most common and best known transformation is to interpret residue decline as a «pseudo-first order» reaction. In this case, linearization is achieved by plotting the logarithms of the residues ( $\log R$  or  $\ln R$ ) versus time  $t$ . The father of this concept is the late F.A. GUNTHER of Riverside, California; he first described the decline of residues by this type of reaction kinetics some 40 years ago (3).

In our experience, this interpretation is clearly applicable in about a third to a half of all cases. However, in practice, when plotted in a semilogarithmic coordinate system, it is not unusual for early data points to «sag» and thereby deviate from the pattern of a 1st-order reaction. In other words: compared to a «true» decrease according to the 1st-order model, the residues decline at a faster rate shortly after application of the pesticide than during the subsequent course of the study. In spite of this, some data of this type can still be modelled and verified statistically by a 1st-order regression.

*Figure 1* gives an example of such a relationship. The straight line in *a* still represents a statistically «correct» 1st-order regression line (on a 95 % confidence level). In *b*, the straight line was back-transformed into the original system (linear time and residue axes). Obviously, the curve does not very accurately represent the measurements. A much better fit can be achieved with another type of function (in this case a RF 1st order as described below), represented by the second curve in *b*. Even the model of a «kinked» decline curve in *a*, formed by a sequence of two 1st-order straight lines having different slopes, can be replaced in this way by a curve exhibiting a continuous transition from an initially fast to a progressively slower rate of decline. It should always be kept in mind that «forcing» a regression straight line of 1st order through a series of data pairs, though still legitimate statistically, may lead to falsified statements!

In other cases, the 1st-order model fails altogether. To resolve this problem, we started, in the late 70's, making a search for simple *linear* relationships which would permit decline curves of this kind to be described mathematically. Our attempts were governed by some preconditions which should be fulfilled:

- correspondence between model and «real life» conditions, i.e. no use of a time axis without a zero point (e.g. by using  $\log t$  or  $1/t$ )
- monotonously decreasing function
- avoidance of functions with more than two constants
- avoidance of functions only solvable by iterative methods.

Under these provisos, we investigated scale transformations employing  $\sqrt{t}$  for the time axis (abscissa) and/or  $1/R$  or  $1/\sqrt{R}$  for the residue concentration axis (ordinate) (*Table 1*).

The designation of these model functions as 1st, 1.5th, etc. order reactions simply follows the conventional classification of rate equations for chemical reactions. It does not imply any reaction kinetic interpretation. The functions having a  $\sqrt{t}$  scale are termed «root functions» (RF). For any given set of data, the coefficients *a* and *b* of the straight line will take different values in each model. In models B, C, E, and F the regression straight line has a positive slope. Solving each of these regression equations for *R*, one obtains the equations given in *Table 2* for the decline curves in the *original* system. An example for a decline curve according to a RF 1st order is shown in *Figure 1 b*, as mentioned above.

Linear regression can easily be calculated even with pocket computers, so we will not discuss how to derive the regression line from the measured values. However, three different confidence intervals CI (*Table 3*) are of particular interest in this context. For their computation, see APPENDIX 1.

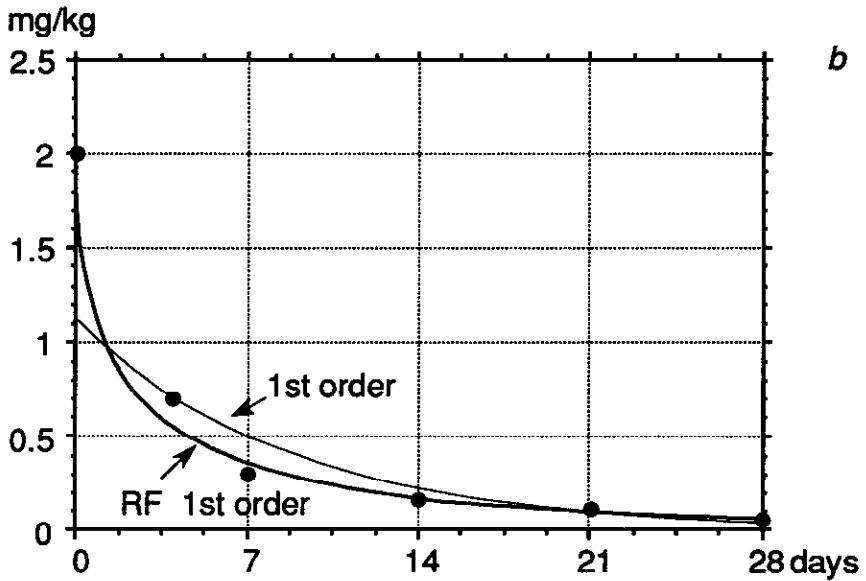
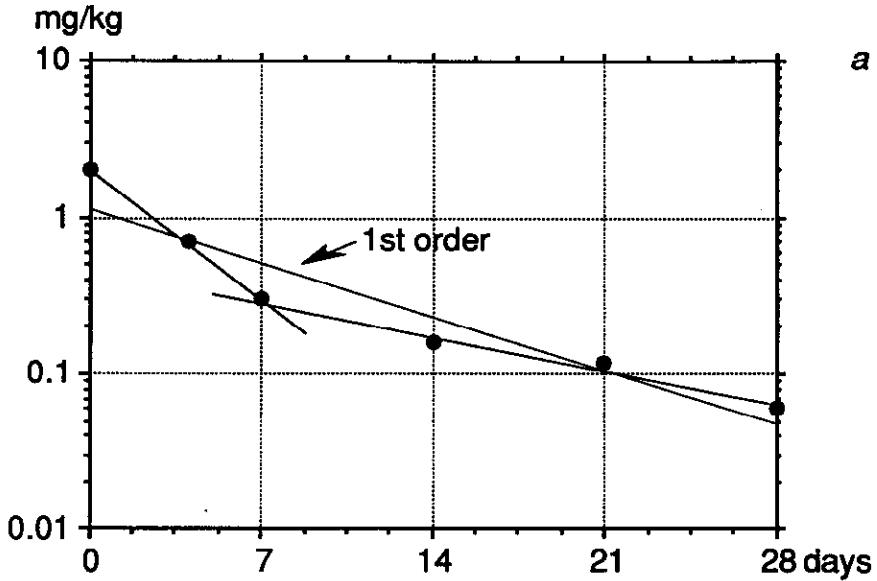


FIGURE 1

Example of an improved curve fit.

The 1st-order straight line in *a* is back-transformed in the original system in *b*.

The "kinked" decline (*a*) can be substituted by a RF 1st-order curve (*b*).

TABLE 1  
Coordinate transformation in formal models A - F

	Reaction Type		Coordinate Transformation	
			Abscissa [X]	Ordinate [Y=f(R)]
A	1 st	order	None	log R
B	1.5 th	order	None	1/√R
C	2 nd	order	None	1/R
D	RF 1 st	order	√t	log R
E	RF 1.5 th	order	√t	1/√R
F	RF 2 nd	order	√t	1/R

RF = "Root function"

The function equation in the transformed system is always  $Y = a + bX$ ,  
where a = intercept on the ordinate at X = 0  
b = slope.

TABLE 2  
Formulae for the decline curves in the original system

Function (Order)	R =	Function (Order)	R =
1st	$10^{a+bt}$	RF 1st	$10^{a+b\sqrt{t}}$
1.5th	$\frac{1}{(a+bt)^2}$	RF 1.5th	$\frac{1}{(a+b\sqrt{t})^2}$
2nd	$\frac{1}{a+bt}$	RF 2nd	$\frac{1}{a+b\sqrt{t}}$

TABLE 3  
Confidence intervals of interest in the evaluation of decline studies

CI <sub>(SL)</sub>	for the entire straight line
CI <sub>(E)</sub>	for the <i>estimated</i> value of Y [E(Y)] at the point X = x
CI <sub>(P)</sub>	for a future (" <i>predicted</i> ") observation Y at the point X = x (Prediction Interval)

The boundaries of these intervals are represented by hyperbolic curves, which envelope the straight line. They are valid for the range of the measured value pairs  $x_1, y_1 \dots x_n, y_n$ . From the upper bounds of the intervals, it is possible to estimate the residues which may possibly be expected under «unfavourable» conditions.

For illustration, *Figure 2* shows a decline curve depicting the decrease of residues of parathion on red currants. In *a*, the 1st-order regression straight line is shown, with its confidence interval  $CI_{(SL)}$  and the prediction interval  $CI_{(P)}$  at day 10. In *b*, the curve and the intervals were back-transformed into the original system.

*For routine assessment of residue data from decline studies, we use a computer program which evaluates the data using all six formal models (Table 1), and then transforms the results back to the original system. The model selected for final evaluation of the study is usually the one that gives the smallest residual sum of squares (SSR) in the back-transformed system after calculation in the transformed system. As required and alternatively, the model to be used for the evaluation can be freely chosen. Initially, only those regression models are accepted which can be confirmed on a 95 % confidence level (calculated in the transformed system). If the calculation of the regression can not be confirmed on this level, it can be repeated on a 90 % level. If it also can not be confirmed on the 90 % level, the individual measurement points are graphically connected to each other in the print-out, without establishing a functional relationship. Preferably, a minimum of 4 to 5 residue/time data pairs should be available for the evaluation. Findings such as «n.d.» (not detectable or not determinable), if occurring within the decline series, are inserted, for calculation, as a concrete figure, e.g. as half the value of the respective limit. If such findings occur in succession at the end of the series, the first of these results is included in the calculation in the same manner, and all subsequent ones are ignored (but will remain in the print-out of the data).*

*To obtain a measure of the quality of the fit, a modified coefficient of determination,  $r^2$ , is computed, and the value of  $r$  is tested with the aid of a test quantity  $D$ ; see APPENDIX 2. If the test quantity differs significantly from zero, the correlation is considered confirmed. Except for the 1st-order and the RF 1st-order model, the functions have a singularity if there are different signs (+ or -) for the constants  $a$  and  $b$ , which will result in a residue value of infinity for a particular decline time. In such cases, the respective function can not be used for the evaluation.*

*The computer print-out contains the main statistical model parameters, the SSR for all six models (if applicable), and the percentage SSR ratio of the best to the second-best model determined. It also provides the graphical representation of the residue decline curve and its confidence interval, back-transformed into the original system, along with the calculated residue values  $R_{mod}$  for the individual values of  $t$  (i.e. the values on the back-transformed curve) yielded by the best (or selected) model. If wanted, the mean residue (cf. Section 2.3) and the prediction interval  $CI_{(P)}$  for selected points of time can also be shown. An example of such a print-out, relating to the curve in *Figure 2*, is given in *Figure 3*.*

We investigated the applicability of the six models first with more than 400 studies of the decline of residues on plants. About one third of these could best be described by the 1st-order model and one third by the RF 1st-order model. This was a surprising observation, since to our

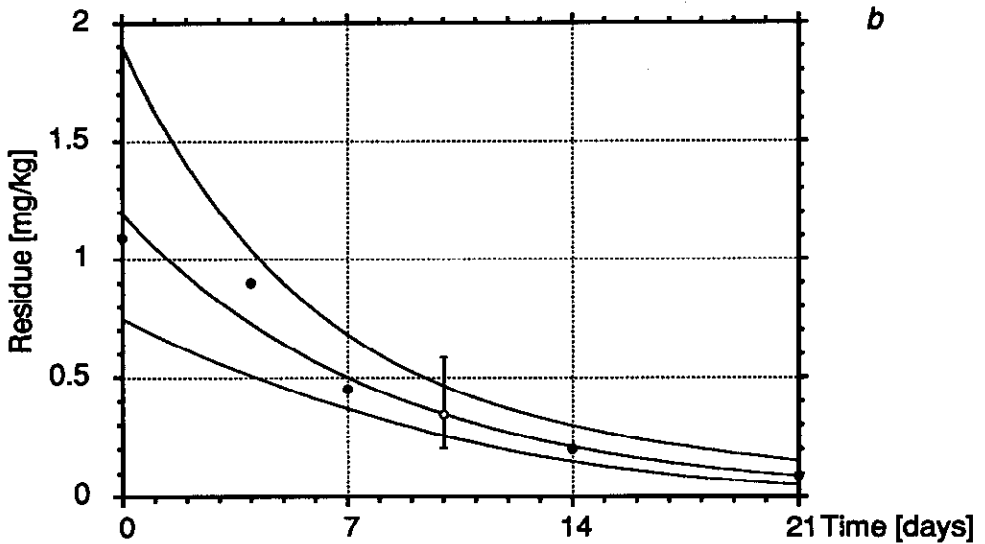
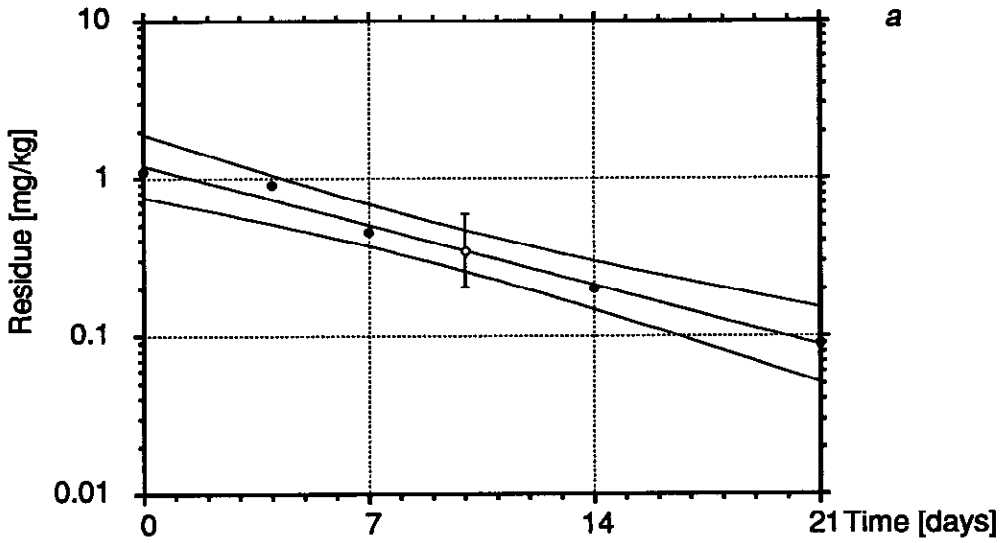


FIGURE 2

Decline of an insecticide on red currants.

*a*: 1st-order straight line with confidence interval.

*b*: Decline curve as in *a*, back-transformed in the original system.

The prediction interval for day 10 is shown for both curves.

Exp. No.: 000000  
 Crop/Soil: Red currants Sample material: Berries  
 Product: E 605  
 Active ingredient: Parathion

Function: 1st order

<i>t</i> (days)	<i>R</i> (mg/kg)	<i>R</i> <sub>mod</sub> (mg/kg)
0	1.09	1.192
4	0.90	0.725
7	0.45	0.500
14	0.20	0.209
21	0.09	0.088

SSR: 0.0435  
 Slope b: -0.0539 CI from -0.0664 to -0.0415  
 Intercept a: 0.0763 CI from -0.0717 to 0.2242  
 Coefficient of determination\*: 0.9426 Significance: 0.0926 at 95 %  
 \* = modified r squared

T/2: 5.58 d CI: 4.29 - 6.87 T/10: 18.54 d CI: 14.25 - 22.83  
*R*<sub>mod</sub> (10 d): 0.344 mg/kg CI: 0.203 - 0.583

SSR Ratio 1st order / RF 1st order: 828.35 %

SSR 1st order: 0.04349 1.5th order: 0.57468 2nd order: -----  
 SSR RF 1st order: 0.40372 1.5th order: 6.97883 2nd order: -----

FIGURE 3

Typical print-out from a computer evaluation of a decline study (slightly modified).  
 The pertinent curves are shown in Figure 2.



knowledge a function of the latter type has not hitherto been described in the literature. The RF 1.5th order accounted for about 10 % of all cases. The other three functions were only seldom applicable. As a rule, the function providing the best fit can not be adequately substituted by one of the other functions. Only in about 25 % of the studies was the model giving the second-best fit virtually identical (SSR < 20 % higher) with that giving the best fit. In such cases, model A was mostly followed by model B, models D and E were mutually close to each other, and model F was followed by model E (cf. Table 1 for designations). In a later evaluation, covering 139 studies, the 1st-order model (A) yielded the best fit in 55 % of the cases, while the RF 1st order (D) gave the best fit in 20 % of the cases, followed by the 1.5th order model (B) with 9 % (cf. Section 4.2.3). The models can also be used for describing residue dissipation in soil. They are not applicable to time spans in which no decrease in residues takes place, e.g. during a «lag phase».

## 2.2. Improving the power of existing models with realistic assumptions about analytical errors

After having discussed simple possibilities for describing decline behaviour and deriving corresponding confidence intervals, let us now take a look at the errors inherent in analytical results in order to obtain improved linear models.

### 2.2.1. «How do you know your results are right?»

With a given analytical method, residues can usually be measured over a wide range of concentrations. It is a well-established fact, however, that in any given case, the variability of measurements, expressed as *repeatability* or *reproducibility*, can be quite substantial (repeatability being the term for the «within (*intra*) laboratory» variability of results associated with a single operator, while reproducibility is the term for the «between (*inter*) laboratory» variability of results associated with different operators in different laboratories). Experience has also shown that in many instances, variability will *increase* with *decreasing* concentration. The smaller the concentration becomes, the greater may be the scatter of the results obtained.

Against this background, the provocative question quoted above was raised in 1981 by William HORWITZ, formerly with the FDA in Washington (4). In his paper, as in a preceding one (6), he published the famous «trumpet» shown in Figure 4. It was based on the results of *interlaboratory* collaborative studies conducted under the auspices of the AOAC over the - then - past 100 years, covering a broad variety of analytical techniques and substrates. This curve is obtained («in an oversimplified fashion, to be sure») when the mean coefficient of variation, *CV* (i.e. the *relative* standard deviation, expressed as per cent of the measured value), expressed as powers of 2, is plotted against the concentration, expressed as powers of 10.

What does this mean? It means that the *CV* increases by one power of 2 for every two powers of 10 by which the concentration (*G*) falls, as is shown in Table 4. HORWITZ expressed the content (*G*) of the analyte in a matrix as fraction of the «pure material» (e.g.  $10^{-1} = 10\%$ ). We may as well express the concentration, *R*, of the analyte in the matrix in mg/kg, as was done in this Table.

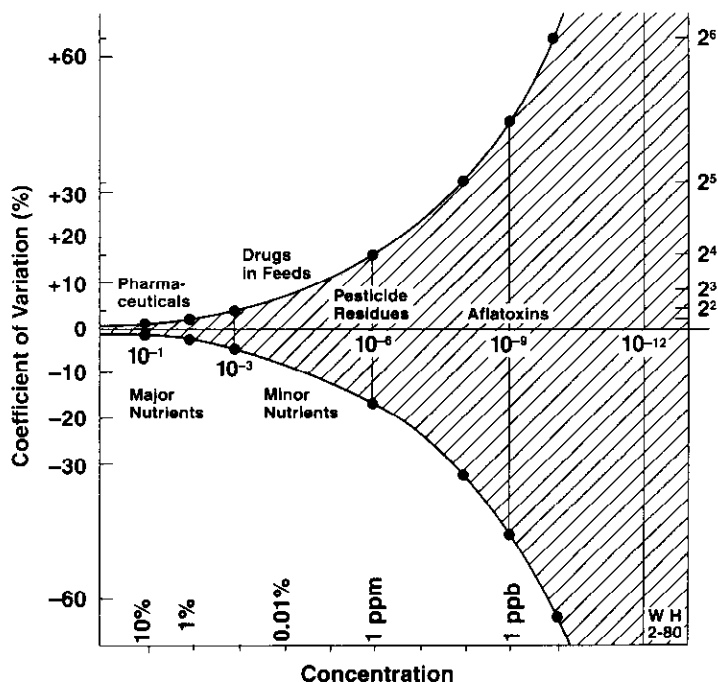


FIGURE 4  
Interlaboratory coefficient of variation as a function of concentration.

Such a relationship can also be plotted as a straight line in a double-logarithmic representation. *Figure 5* gives two examples of such a plot. The lower line represents the figures given by HORWITZ (Table 4). In the case represented by the upper straight line it is assumed that the  $CV = \pm 100\%$  if  $R = 0.01$  mg/kg, and that the  $CV$  decreases by a factor  $K = 0.5$  per increasing order of magnitude of  $R$ , yielding  $CV = 50\%$  if  $R = 0.1$  mg/kg, and so on. These latitudes (designated by the letter  $L$  from here on) were suggested by the Federal Health Office in Germany in 1974 (1) for application in the analysis of organochlorine insecticide residues in food of animal origin. Two examples of the *ranges* of these latitudes are illustrated for  $R = 0.01$  and  $R = 0.1$  mg/kg, respectively.

The mathematical expression for such a straight line in a double-log-system has the general form set out in Table 5, formulae (I a) and (I b). The *absolute* amounts of the latitudes (according to formulae (II a) and (II b) in Table 5) can similarly be represented by *ascending* straight lines (for derivation, see APPENDIX 3).

For the relationships illustrated in *Figure 5*, the general equation (I a) in Table 5 takes the forms as given in the Figure (for derivation of the constants for the «Horwitz curve», see APPENDIX 4).

TABLE 4  
**Coefficients of variation (CV) in relation to  
the concentrations measured (according to HORWITZ)**

Residue Content		CV (%)
$G^a$	$R^b$	
$10^{-2}$	$10^4$	$\pm 2^2 = 4$
$10^{-4}$	$10^2$	$\pm 2^3 = 8$
$10^{-6}$	$10^0$	$\pm 2^4 = 16$
$10^{-7}$	$10^{-1}$	23 <sup>c</sup>
$10^{-8}$	$10^{-2}$	$\pm 2^5 = 32$
$10^{-9}$	$10^{-3}$	46 <sup>c</sup>
$10^{-10}$	$10^{-4}$	$\pm 2^6 = 64$

<sup>a</sup> expressed as fraction of the matrix  
<sup>b</sup> expressed as parts per million, e.g. mg/kg  
<sup>c</sup> calculated from formula in Figure 5 and rounded  
Note that here  $R = G \cdot 10^6$

TABLE 5  
**Formulae for the calculation of latitudes in double-logarithmic representation**

$\log(L)$	=	$\log(L_0) + \log(K) \cdot \log(R)$	(I a)
with			
$L$	=	"latitude" = coefficient of variation (%) of the content determination = relative standard deviation, expressed as percentage of $R$	
$L_0$	=	value of $L$ at $R = 10^0 = 1$ mg/kg (%)	
$R$	=	content of residue in the sample (mg/kg)	
$K$	=	factor by which $L$ decreases as the order of magnitude of $R$ increases (see also Table 8)	
or, by taking antilogarithms,			
$L$	=	$L_0 \cdot R^{\log(K)}$	(I b)
The absolute standard deviation can accordingly be expressed as			
$\log(L_{abs})$	=	$\log(L_{0,abs}) + \log(10K) \cdot \log(R)$	(II a)
or, by taking antilogarithms,			
$L_{abs}$	=	$L_{0,abs} \cdot R^{1+\log(K)}$	(II b)
(cf. APPENDIX 3)			

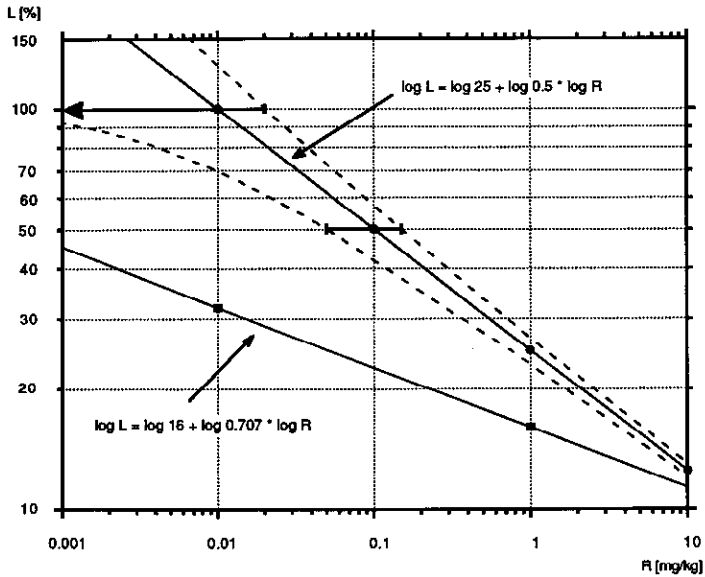


FIGURE 5  
**Latitude straight lines in a double-logarithmic representation.**  
 For explanation, see text.

For these two examples, the *absolute* amounts of the latitudes can be calculated from  $\log L_{abs} = \log 0.25 + \log 5 \cdot \log R$  (for the figures stated by the Fed. Health Office, Germany)

and

$$\log L_{abs} = \log 0.16 + \log (10/\sqrt{2}) \cdot \log R \quad (\text{for the figures given by HORWITZ})$$

or, by taking antilogarithms,

$$L_{abs} = 0.25 \cdot R^{0.6990}$$

and

$$L_{abs} = 0.16 \cdot R^{0.8495},$$

respectively.

The end points of the latitude ranges (which we might call  $R_{min}$  and  $R_{max}$ ) can be calculated from either  $R$  or  $L$  by the formulae given in Table 6 (2). They can be connected to yield upper and lower «latitude curves», which envelope the straight line. These curves, as shown in Figure 5 for the upper straight line, illustrate the continuous changes of the latitudes over the concentration range of interest. They can be utilized to establish meaningful progressions in setting MRLs in conjunction with analytical feasibility, and they should be taken into account when having to decide whether an MRL has been exceeded. However, these curves must *not* be confused with confidence intervals.

TABLE 6  
Formulae for calculating  $R_{min}$  and  $R_{max}$  of latitudes

$$R_{\pm} = R \pm \frac{R^{1+\log(K)} \cdot L_0}{100}$$

$$R_{\pm} = \frac{100 \pm L}{100} \cdot \left( \frac{L}{L_0} \right)^{1/\log(K)}$$

### 2.2.2. Weighted linear regression

With known (or estimated) values for the factor  $K$ , we can now extend the decline models thus far discussed, by fitting the model constants in such a way that the variation of the scatter of the analytical values with the residue content in the *original* system is accounted for. This will be done by using a linear *weighted* regression (APPENDIX 5).

By applying the error propagation law (APPENDIX 6), we obtain simple terms for the weights  $w$  (Table 7) and can derive confidence intervals for a weighted regression. The terms for  $w$  are interrelated with the amounts of the latitudes. For example, for the functions of 1st order, the relation is given through  $R^{-2 \log(K)} = (L_0/L)^2$  according to (I b) in Table 5. Based on the terms for  $w$  as given in Table 7, the weights can also be expressed as powers of  $R$  (Table 7A).

In linear (unweighted) regression, one of the premises is that the variance of the target (measured) values is *constant* over the entire concentration range. According to a) in Table 8, an unweighted regression ( $w = 1$ ) is obtained if *different* values for  $K$  are used with the individual types of functions. This means that the case of an unweighted regression can not be satisfied by a uniform error law, i.e. by a single factor  $K$ , which would be applicable to all six functions. In order to circumvent this disadvantage, it is necessary to revert to a *weighted* regression. This can be achieved with a predetermined value of  $K$  ( $K < 1$ ). This factor  $K$  then stands for a particular error law which is equally valid for all six types of functions. Table 7 A illustrates this effect.

We can now consider different approaches to calculating weighted decline curves and confidence intervals by using different values for  $K$ , e.g. those given in Table 8. The effect of these weightings is illustrated in Figure 6. Here the decline curve is described by the 1st-order model. The confidence intervals of the line and the prediction interval are shown along with the (back-transformed) decline curve. It can be seen that with the weighted regression (curves b - e), both the confidence and the prediction intervals at the start of the study (when the residues are higher and the latitudes are accordingly smaller) are narrower than with the unweighted regression. This effect eliminates previous disadvantages. If the value chosen for  $K$  is too small (e.g. = 0.1 as in e)), the weighted mean can become displaced so far towards shorter times that the confidence intervals at

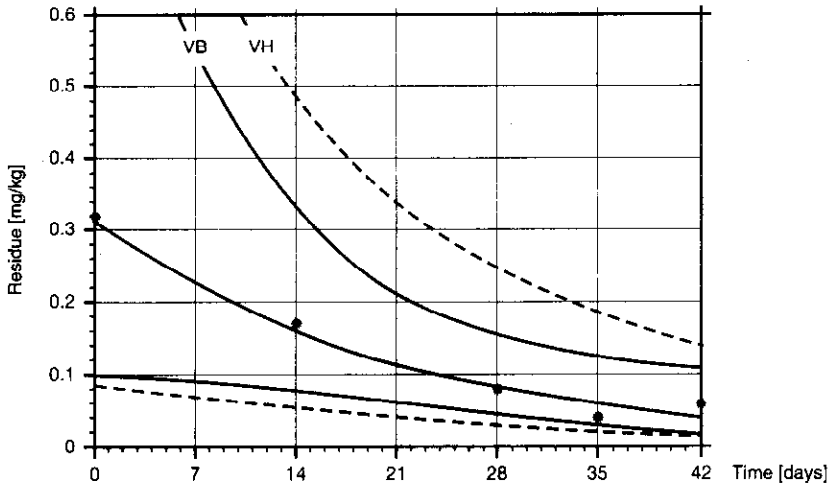
the end of the study period again become distinctly wider. Because of these effects, an approach according to either b) or c) represents an optimal choice. In evaluating his data, every analyst must, of course, make his own decision in attaching weights to his results.

TABLE 7  
Terms for the calculation of weights  $w$

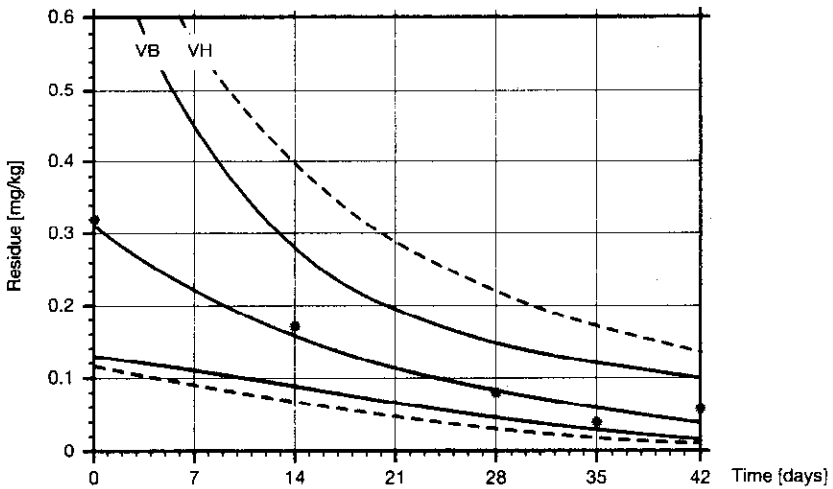
Function (Order)	$w =$
1 st RF 1 st	$R^{-2 \log(K)}$
1.5 th RF 1.5 th	$R^{1-2 \log(K)}$
2 nd RF 2 nd	$R^{2-2 \log(K)}$

TABLE 7A  
Terms for weights  $w$ , expressed as powers of  $R$

Function (Order)	Values of $w$ when $K =$						
	10	$\sqrt{10}$	1	$1/\sqrt{2}$	0,5	$1/\sqrt{10}$	0,1
1 st and RF 1 st	$1/R^2$	$1/R$	1	$R^{\log 2}$	$R^{\log 4}$	$R$	$R^2$
1.5 th and RF 1.5 th	$1/R$	1	$R$	$R^{\log 20}$	$R^{\log 40}$	$R^2$	$R^3$
2 nd and RF 2 nd	1	$R$	$R^2$	$R^{\log 200}$	$R^{\log 400}$	$R^3$	$R^4$



a Unweighted ( $K = 1$ )



b Weighted with  $K = 0.707$  (according to Horwitz)

FIGURE 6  
**Decline of residues of an insecticide in apples.**  
**Curves fitted with 1st-order model with various weights.**  
**VB = Confidence interval, VH = Prediction interval.**

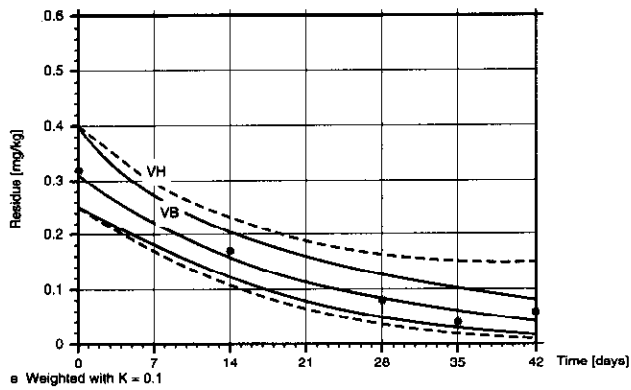
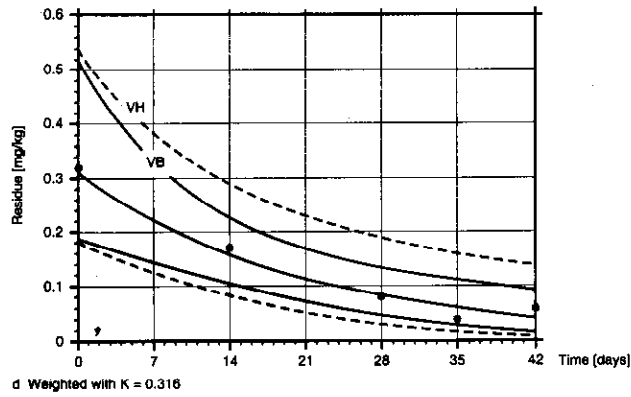
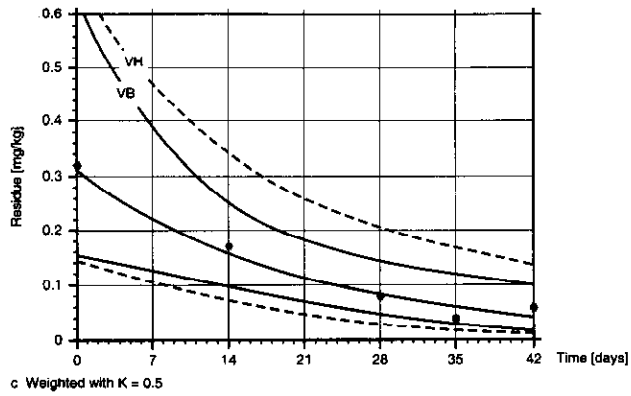


FIGURE 6 (cont.)  
 Decline of residues of an insecticide in apples.  
 Curves fitted with 1st-order model with various weights.  
 VB = Confidence interval, VH = Prediction interval.



TABLE 8  
 Characteristics of various factors  $K$  (see Tables 5 and 7A and Figure 6)

	$K =$	Characteristics
a)	$\cdot$ 10 $\sqrt{10}$ 1	<i>Unweighted</i> <i>(constant variance in the transformed system)</i> in the case of 2nd- and RF 2nd order models in the case of 1.5th and RF 1.5th order models in the case of 1st- and RF 1st-order models <i>In this case, also the coefficient of variation is constant in the <u>original</u> (non-transformed) system, i.e. the <u>relative</u> latitudes are the same over the entire concentration range.</i>
b)	$1/\sqrt{2}$ $\approx 0,7071$	<i>Weighting according to HORWITZ</i> The scatter decreases by a factor of <u>0.5</u> for every <i>two</i> powers of 10 by which the concentration increases (see also APPENDIX 4).
c)	0.5	<i>Weighting as proposed by the German Federal Health Office</i> The scatter decreases by a factor of <u>0.5</u> per <i>one</i> power of 10 by which the concentration increases.
d)	$1/\sqrt{10}$ $\approx 0,3162$	The scatter decreases by a factor of <u>0.1</u> for every <i>two</i> powers of 10 by which the concentration increases.
e)	0.1	<i>Constant variance in the original (non-transformed) system</i> The scatter decreases by a factor of <u>0.1</u> per <i>one</i> power of 10 by which the concentration increases. It follows from Table 5 (II b) that here $L_{abs} = L_{0,abs}$ i.e. the <u>absolute</u> amount of the latitudes is the same as at $R = 1$ mg/kg over the entire concentration range

### 2.3. Decline times

After we have discussed how to describe various decline courses mathematically, we can take the next step and have a look at the time it takes for different degrees of decline to be reached.

In a 1st-order reaction, the relative rate of decline remains constant throughout the entire process and is independent of the initial concentration ( $R_A$ ). Accordingly, the «half-life», or the time that must elapse until the initial concentration of the reacting substance has decreased by one half (yielding  $R_A/2$ ), or until  $R_A/2$  has decreased to  $R_A/4$ , etc., is a constant. For all other functions used here, the rate of decline *decreases* progressively during the course of the reaction. Consequently, the term «half-life» should only be used in conjunction with 1st-order reactions when referring to periods of time required for a 50 % decline of the residue. For all other types of reaction, we need other designations. We have proposed to use the term «T/X times» to determine those periods of time after which the residue concentration has fallen to a fraction  $1/X$  of the *initial* concentration. Thus, T/2 or T/5 would define the time needed for the residue to drop to  $1/2$  or  $1/5$  of  $R_A$ , respectively. The times T/2 and T/10 are often referred to as DT 50 and DT 90. «DT 50» and «DT 90» stand for «decline (decay, degradation, dissipation) time» for 50 % and 90 % loss of residues, respectively.

The decline behaviour of a residue can thus be characterized not only by stating the half-life (in 1st-order reactions) or T/2 (in other types of reactions) but also by stating the longest respective T/X-time which still falls within the period of the study (extrapolations to longer periods should be handled with care!).

The formulae for computing the T/X-times, resulting from the models discussed so far, are given in *Table 9* (right hand column). The T/X times can be calculated simply from the respective coefficients a and b of the regression straight line.

Decline times normally apply to experiments in which the first samples are taken at day 0, i.e. immediately after the (last) application of the pesticide ( $t = 0$ ). In other experiments, the first sampling may be performed at a later date  $t_A$  (*Figure 7*). In these instances, the generally valid formulae given in *Table 9* will apply for computing the T/X-times. For computation, the coefficients a and b, calculated for the curve beginning at  $t_A$ , must be inserted in the formulae. Except for the 1st-order model, the decline times are dependent on  $t_A$ . Meaningful comparisons of decline times from different experiments can, therefore, only be made if there are identical initial points of  $t_A$  or when  $t_A = 0$ . When  $t_A \neq 0$ , stating decline times without at the same time stating  $t_A$  has no informative value (except for 1st-order reactions).

When the different decline times are each placed in relation to the DT 50, it becomes apparent that greatly different rates of decline can be described with these functions. *Table 10* shows for each function the ratio of T/X ( $X > 2$ ) to T/2 for the case  $t_A = 0$ . If  $t_A \neq 0$ , other ratios are obtained for the three *root functions*, each of which depends on the individual figure for  $t_A$ . In this case, for the *other* three functions, the ratios do *not* change. The ratios can be calculated from the formulae in *Table 9* by inserting the value for X (e.g. 3, 4, 5, or 10) and dividing the formula thus obtained by the same formula, but with  $X = 2$  (*Table 10 A*).

Another important question in decline studies is: How long does it take until a given residue level  $Y'$  will be reached? This particular period of time can be calculated, based on the prediction interval, from formula (vii) in APPENDIX 5. (vii) is obtained from the generally valid formula (vi) in APPENDIX 5 by solving this equation for  $X'$ . For computation in case of an unweighted regression, the terms expressing the weights must be set = 1.

In addition to values for the times  $T/X$ , the accuracy of these values is also of interest. This is stated in the form of the appropriate standard deviation  $S_{T/X}$  ( $T/X \pm S_{T/X}$ ). The formulae for calculating  $S_{T/X}$  are given in Table 11. For their derivation, see APPENDIX 7.

TABLE 9  
Formulae for calculating decline times  $T/X$

Function (Order)	General formulation	Formulation for $t_A = 0$
1 st	$\frac{\log X}{-b}$	$\frac{\log X}{-b}$
1.5 th	$\frac{(a+bt_A)}{b} (\sqrt{X} - 1)$	$\frac{a}{b} (\sqrt{X} - 1)$
2 nd	$\frac{(a+bt_A)}{b} (X - 1)$	$\frac{a}{b} (X - 1)$
RF 1 st	$\left[ \frac{\log X}{(-b)} \right]^2 + 2\sqrt{t_A} \cdot \left[ \frac{\log X}{(-b)} \right]$	$\left( \frac{\log X}{-b} \right)^2$
RF 1.5 th	$\left[ \frac{(a+b\sqrt{t_A})}{b} \cdot (\sqrt{X} - 1) + \sqrt{t_A} \right]^2 - t_A$	$\left( \frac{a}{b} (\sqrt{X} - 1) \right)^2$
RF 2nd	$\left[ \frac{(a+b\sqrt{t_A})}{b} \cdot (X - 1) + \sqrt{t_A} \right]^2 - t_A$	$\left( \frac{a}{b} (X - 1) \right)^2$

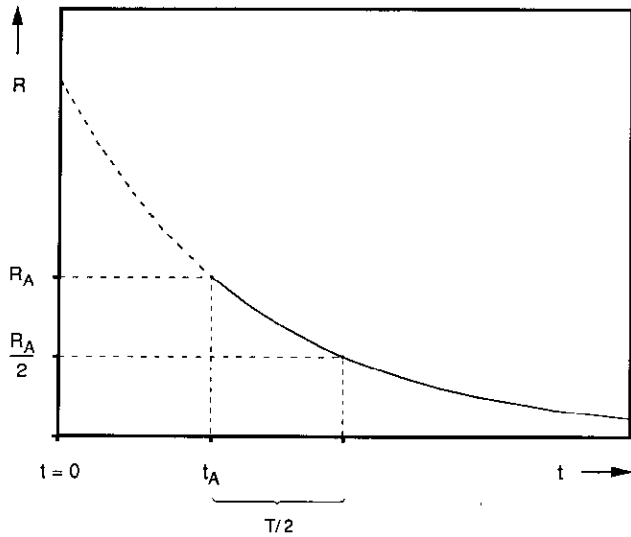


FIGURE 7

**Schematic representation of decline times for the case  $t_A \neq 0$ .**  
 **$t = 0$ : point of time of the last application of the pesticide.**  
 **$t_A$  = initial point of time of the residue study.**

TABLE 10  
**Ratios of decline times  $T/X$  to  $T/2$  for  $t_A = 0$**

Decimals are rounded. – No decimals: integers.

Function (Order)	Ratio of $T/X$ to $T/2$ when $X =$			
	3	4	5	10
1 st	1.6	2	2.3	3.3
1.5 th	1.8	2.4	3.0	5.2
2 nd	2	3	4	9
RF 1 st	2.5	4	5.4	11.0
RF 1.5 th	3.1	5.8	8.9	27.3
RF 2 nd	4	9	16	81

TABLE 10A  
General formulae for calculating the ratios T/X: T/2

Function (Order)	$\frac{T/X}{T/2}$
1 st	$\frac{\log X}{\log 2}$
1.5 th	$\frac{\sqrt{X}-1}{\sqrt{2}-1}$
2 nd	$(X-1)$
RF 1 st	$\frac{(\log X)^2 - 2b\sqrt{t_A} \cdot \log X}{(\log 2)^2 - 2b\sqrt{t_A} \cdot \log 2}$
RF 1.5 th	$\frac{[(a+b\sqrt{t_A})(\sqrt{X}-1)+b\sqrt{t_A}]^2 - b^2 t_A}{[(a+b\sqrt{t_A})(\sqrt{2}-1)+b\sqrt{t_A}]^2 - b^2 t_A}$
RF 2 nd	$\frac{[(a+b\sqrt{t_A})(X-1)+b\sqrt{t_A}]^2 - b^2 t_A}{[(a+b\sqrt{t_A})+b\sqrt{t_A}]^2 - b^2 t_A}$

TABLE 11  
Standard deviations  $S_{T/X}$  of the decline times T/X

Function (Order)	$S_{T/X}$
1 st	$\frac{(T/X) \cdot S_b}{(-b)}$
F 1 st	$\frac{2 \cdot (T/X) \cdot S_b}{(-b)}$
1.5 th and 2 nd	$(T/X) \cdot u$
RF 1.5 th and RF 2 nd	$2 \cdot (T/X) \cdot u$

$S_b$  = root of the variance  $\text{Var}(b)$

$$u = \frac{\text{Var}(a)}{a^2} + \frac{\text{Var}(b)}{b^2} - 2 \frac{\text{Cov}(a,b)}{ab}$$

(see APPENDIX 7)

### 3. MODELS WITHOUT LINEARIZING TRANSFORMATIONS (NON-LINEAR REGRESSION MODELS)

On closer examination, the use of linearizing transformations before fitting the model constants is a complication that is not necessarily desirable. On one hand, there is the advantage of being able to use known linear regression procedures. On the other hand, there are several disadvantages. For example, the distribution of the analytical data, which forms the basis of the confidence intervals, is subject in the original system to constraints which must be laboriously incorporated into the transformed system, e.g. as in the calculation of weights. Also, the optimal model can not be identified until all models have been transformed back to the original system. Furthermore, the need for linearization restricts the choice of the possible models.

Let us therefore examine how errors associated with the model parameters can be established and how confidence intervals can be defined in the case of *non-linear* dependency on the model parameters. To achieve this goal, the same principle and procedure can be employed as in the analogous problems of multiple linear regression.

Some principles of multiple linear regression and the computation leading to

- regression coefficients
- an estimate of the accuracy of the model parameters
- confidence intervals

with the aid of a matrix notation are described in APPENDIX 8.

Dispensing with linearization then gives a generally applicable procedure that can also be used for other non-linearizable models, such as the two-zone model presented in Section 4.3. However, once linearizing transformations have been abandoned, there are no longer any simple formulae available for calculating confidence intervals. Moreover, the model constants must be derived by direct fitting procedures.

As an example, the residue data used for Figure 6 were subjected to these computations. For the sake of simplicity, the model constants were taken from the calculations of the linear regression (1st-order model). The resulting confidence and prediction intervals are shown in Figure 8. They must be compared with the curves for the scatter parameter  $K = 0.1$  (Figure 6 e), since in both cases, the calculation was based on the assumption that the variance in the original system is constant (cf. Table 8). Excellent agreement is found between the directly fitted curves of Figure 8 and the curves of Figure 6.

In principle, an extension of the non-linear regression presented here, which includes a weighting to account for analytical error is possible, but has not been attempted.

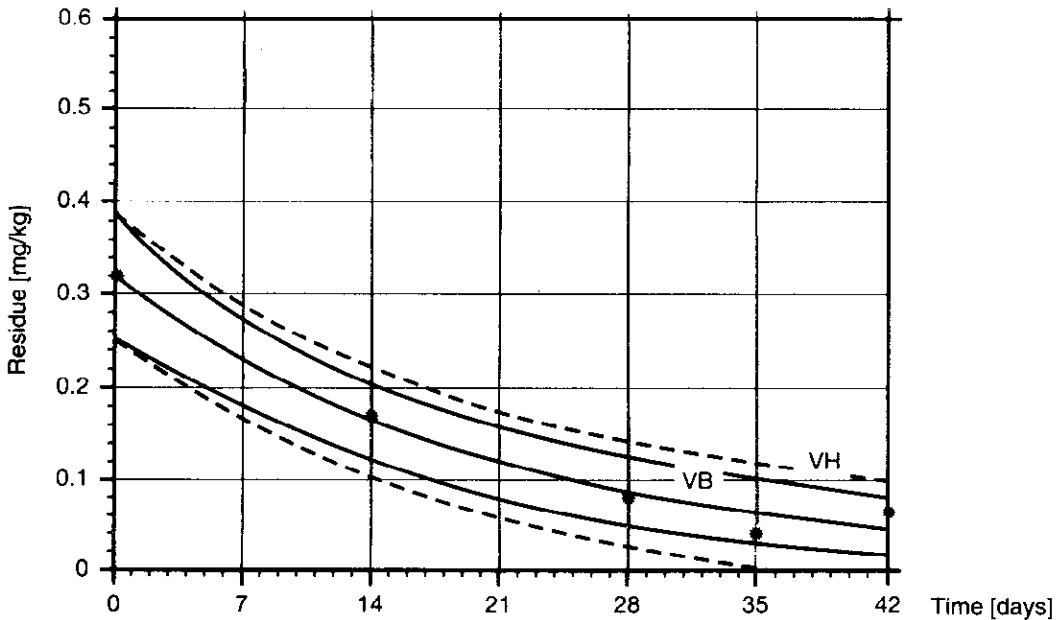


FIGURE 8  
**Residue data and labelling as in Figure 6.**  
 Curve fitted by direct calculation without linearization.

## 4. PHYSICALLY BASED MODELS

### 4.1. Introduction

The models discussed so far have proved their worth in providing a purely formal mathematical *description* of the decline of residues. Of course it would be desirable to be able to formulate the decline behaviour in terms of *physically* based models. We undertook several attempts in this direction, whereby three aspects were considered:

- the «dilution» of residues by the increase in plant mass during the growing stage
- the combination of the increase in plant mass and the simultaneous biological/chemical degradation of the residues
- the distribution and degradation of residues as a two-zone model.

## 4.2. Models taking account of plant growth

### 4.2.1. Evaluation of growth data

We evaluated published data on the increase in plant mass during growth of fruits and leafy vegetables (8). In order to establish growth curves (expressed as increase in weight, volume, or diameter with time), the data were normalized. This was done by expressing the growth stages relative to the final state, e.g. by  $D/D_{\max}$  (where  $D_{\max}$  = diameter when ripe). Similarly, the time scale was normalized, expressing any point of time during growth as a fraction of the time  $T_{\max}$  taken to reach maturity.

The growth data were evaluated with the aid of direct curve fitting procedures. The results showed that the growth profile can be reproduced by functions of the type shown in Table 12 (I) and (II). For the crops included in our investigation, the coefficients of equations (I) and (II) were as given in Table 13.

Figure 9 shows a typical growth curve obtained for apples. The vertical line at  $T/T_{\max} = 0.55$  indicates  $T_p$ . The interval between  $T_p$  and  $T_{\max}$  is the pre-harvest interval, which is generally of interest in decline studies, while  $T_p$  corresponds to «Day 0» of the residue decline curve.

In many cases, the growth during the pre-harvest interval, or else during the time  $T_p/T_{\max} \approx 0.5 - 1$ , can be approximated by the linear function (II) in Table 12 even if the overall growth

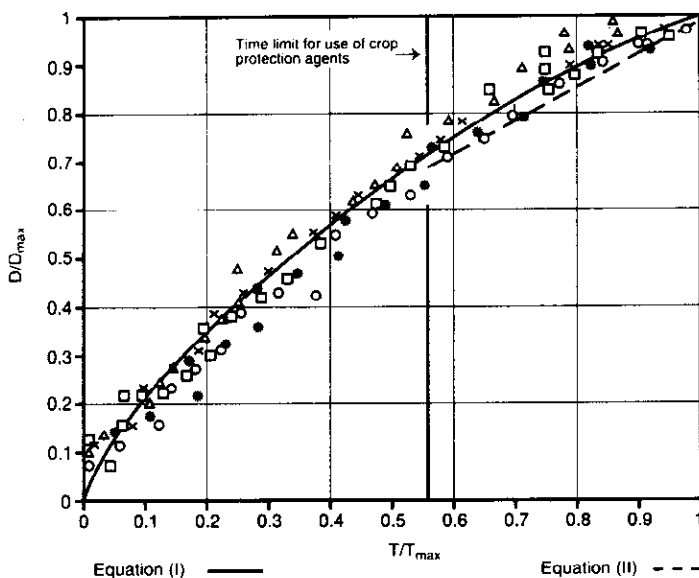


FIGURE 9  
**Growth profile of apples (5 varieties).**  
 Based on data in Kolbe (1979), Höfchen Research Station.



TABLE 12  
Plant growth functions

$$Y/Y_{\max} = b (T/T_{\max} - c)^d + a \quad (\text{I})$$

$$Y/Y_{\max} = A + B (T/T_{\max}) \quad (\text{II})$$

Parameter Y = growth stage expressed as diameter

Parameter T = time scale (e.g. in days) expressing the "life time" of the fruit until maturity ( $T_{\max}$ )

If t = time after the application of the pesticide,

$$T/T_{\max} = t/T_{\max} + T_p/T_{\max} \quad (\text{III})$$

$T_p$  = point of time up to which plant protection measures were maintained

TABLE 13  
Coefficients in the growth functions (Table 12)

Crop (Mean $T_{\max}$ , days)	$T_p/T_{\max}$ *	Equation I				Equation II	
		b	c	d	a	A	B
Apples (105)	0.55	1.09	0	0.8	0	0.3	0.7
Nectarines (68)	0.5	1.6	0.5	2	0.6	0.2	0.8
Peaches (104)	0.5	1.23	0.5	1.3	0.5	0	1
Pears (126)	0.7	1	0	1	0	0	1
Cucumbers (48)	0.4	1	0	0.25	0	0.8	0.2
Head lettuce (60)	**	1	0	4	0	-	-
Tomatoe plants (40)	**	1.3	0	1	-0.3	-0.25	1.25
White cabbage (102)	**	1	0	1.2	0	-0.15	1.15

\* Relative time scale (see Table 12); values refer to the average of the decline studies used for the evaluation

\*\* No data or not relevant

is not already linear, as indicated by the broken line in Figure 9. The *entire* growth curve, if it is not linear, can be described by equation (I). We will not enter here into more details about these evaluations (8).

#### 4.2.2. Fitting residue data to a 1st-order degradation law

The interesting facet of the growth curves is that apparently they can be used to describe the *overall* decline of residues as well as both their *dilution* by plant growth and their decrease due to true *degradation*. In exploratory calculations, this could be shown by combining the growth function with the assumption of *1st-order* degradation kinetics (APPENDIX 9). Taking apples again as a representative crop, the decline curves for a fungicide shown in Figure 10 a were obtained in this way. The unbroken line shows the overall residue decline, the two other curves show how decline would proceed if it were due solely to the increase in plant mass or solely to the degradation of the chemical.

These two curves were respectively obtained by setting  $k = 0$ , and by considering only the exponential function in formula (iii) of APPENDIX 9 (see also APPENDIX 10), with  $T_p = 93$  days,  $T_{max} = 135$  days,  $k = 0.02594$ . The *degradation* half life was  $\ln 2/k = 26.7$  days, as opposed to a  $T/2$  of 13.9 days for the *overall* decline. The time for the overall decline can in each case be obtained by multiplying the two separate residue values, e.g. 0.41 (dilution)  $\times$  0.34 (degradation) = 0.14 mg/kg on day 42 in Figure 10 a.

Apart from a slight deviation in the early phases, the curve for the overall decline compares quite well with that obtained with the formal model RF 1st order (Figure 10 b). Assuming that reliable data on the increase in plant mass are available and that the course of the decline curve can be adequately described by equation (iii) in APPENDIX 9, an interesting opportunity arises for considering the two effects separately. It appears worthwhile to investigate this approach in more detail.

#### 4.2.3. Fitting residue data to a degradation law of arbitrary order

The approach described under 4.2.2 was based on the simplifying assumption that biologically and/or chemically induced degradation of residues will obey 1st-order kinetics. The next question is: What do we get if we assume kinetics of an arbitrary order  $n$  in the *physical* model? The theoretical background for these considerations is described in APPENDIX 10.

For comparison, we evaluated residue data from 139 decline studies by two approaches:

- using the *formal* model of individually optimal order according to Section 2.1 (unweighted)
- using the *physical* model with each of the reaction orders  $n = 1, 1.5$  and  $2$ ; in each case the best fit (method of least squares) was chosen for comparison.

In the calculations using the *formal* approach, the best fit was obtained with a 1st-order model in 55 % of these studies, followed by the RF 1st order (20 %) and the function 1.5th order (9 %). In the calculations using the *physical* approach, 53 % of the studies were optimally described when  $n = 1$ , 36 % when  $n = 2$ , and 11 % when  $n = 1.5$ . In nearly half of the studies,

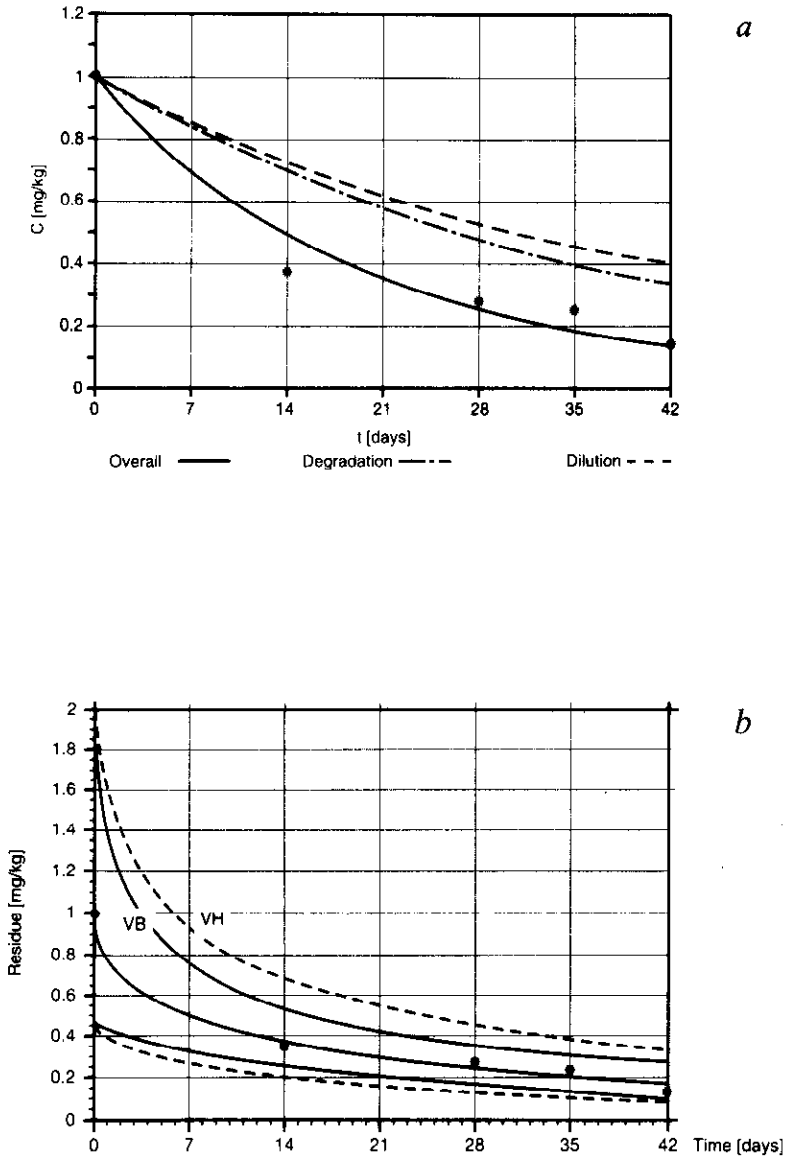


FIGURE 10

**Decline of residues of a fungicide in apples.**

**a:** Representation according to a 1st-order degradation law (unbroken curve).  
 The two other curves show how decline would proceed if it were due solely to the increase in plant mass or solely to the degradation of the residue .

**b:** Curve fitted with RF 1st-order formal model, unweighted.

VB = Confidence interval, VH = Prediction interval.

the physical model using  $n = 1$  surpassed the formal 1st-order model, and in 13 % of the studies the physical model using  $n = 2$  yielded a better fit than the RF 1st order. Overall, only in 20 % of the studies the physical model gave a poorer fit than the formal model. The improvements were distinct, but not spectacular in most instances, as exemplified by *Figure 11 a* (we shall revert to *Figure 11 b* later). As set out in APPENDIX 10, and in analogy to 4.2.2, the separate effects of dilution and degradation on the decline can be determined also with these models.

In view of the satisfactory fit achievable with this physical model, we might even dare to claim that it gives support to the merely formal, «non-physical» approach. However, there also appears to be a disadvantage associated with the physically based model. It can *not* give a better fit to experimental data in which the concentration/time curve shows strong «sagging» shortly after the application of the pesticide. All comparisons clearly showed that in most cases in which the best fit was attained with the physical model, and not with the optimal formal model, the reaction order in the *formal* approach was *not a root function*. In comparison to the formal approach, the «sagging phenomenon» could be better described by the two-zone model presented in the following Section.

### 4.3. The Two-zone Model

This model is based on the following assumption:

A plant part, e.g. a fruit, can be divided into two zones,

- the volume near the surface,
- the internal volume enclosed by it.

When a pesticide is applied to the exterior of the crop, its active ingredient will gradually penetrate to the interior. Degradation of the compound at the surface and in the outer zone obeys different laws than does degradation in the inner zone. Obviously, degradation at the surface is influenced by light and by biochemical reactions and may be accompanied by losses due to evaporation and washoff. In the *interior*, degradation is entirely due to biochemical processes. Shortly after application of the pesticide, degradation in the outer zone will predominate and moreover will proceed more rapidly than in the inner zone.

Schematically, this model can be visualized as shown in *Figure 12*. The mass transport of the compound between the two zones occurs by diffusion and thus is linearly related to the concentration of the compound in the zones, and to the interface between them. For simplicity, it is assumed that this interface is proportional to the volumes of the zones and that back-transport from the interior to the exterior is negligible. The volumes are known functions of time, as was discussed in Section 4.2.1.

These considerations can be formulated mathematically (for derivation, see APPENDIX 11) as set out in *Table 14*. Based on the general formulation (i), equation (ii) describes the degradation of compounds in crops according to the two-zone model. The function  $V_{\text{tot}}(t)$  describes the growth profile of the crop.

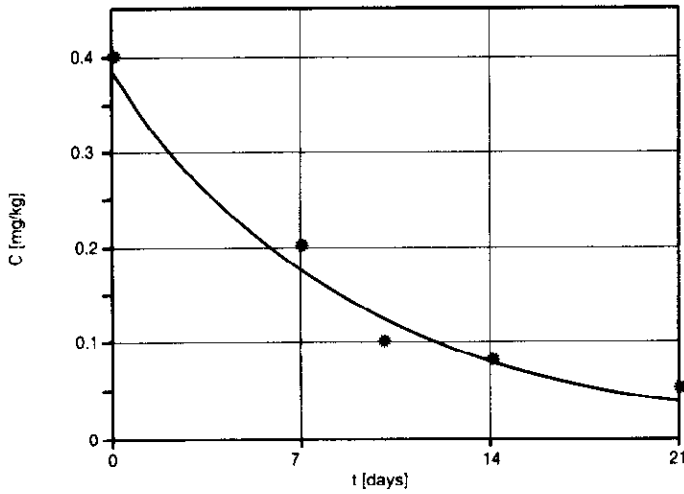
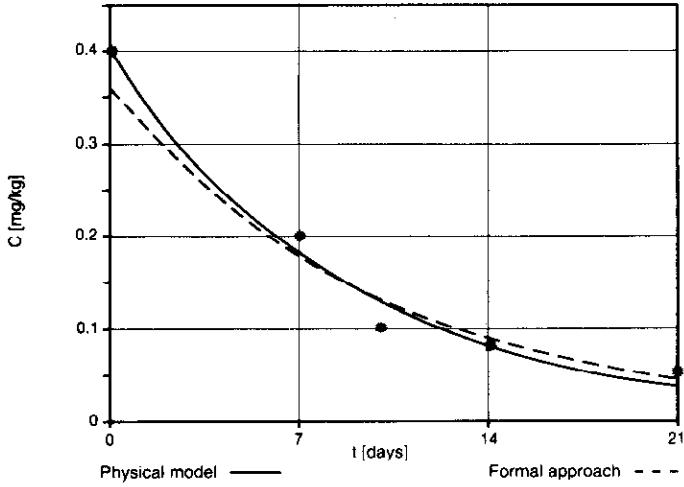


FIGURE 11

**Decline of residues of a fungicide in apples.**

**a: Curves fitted with the 1st-order formal model and the 1st-order physical model.**

**b: Curve fitted with the Two-zone model.**

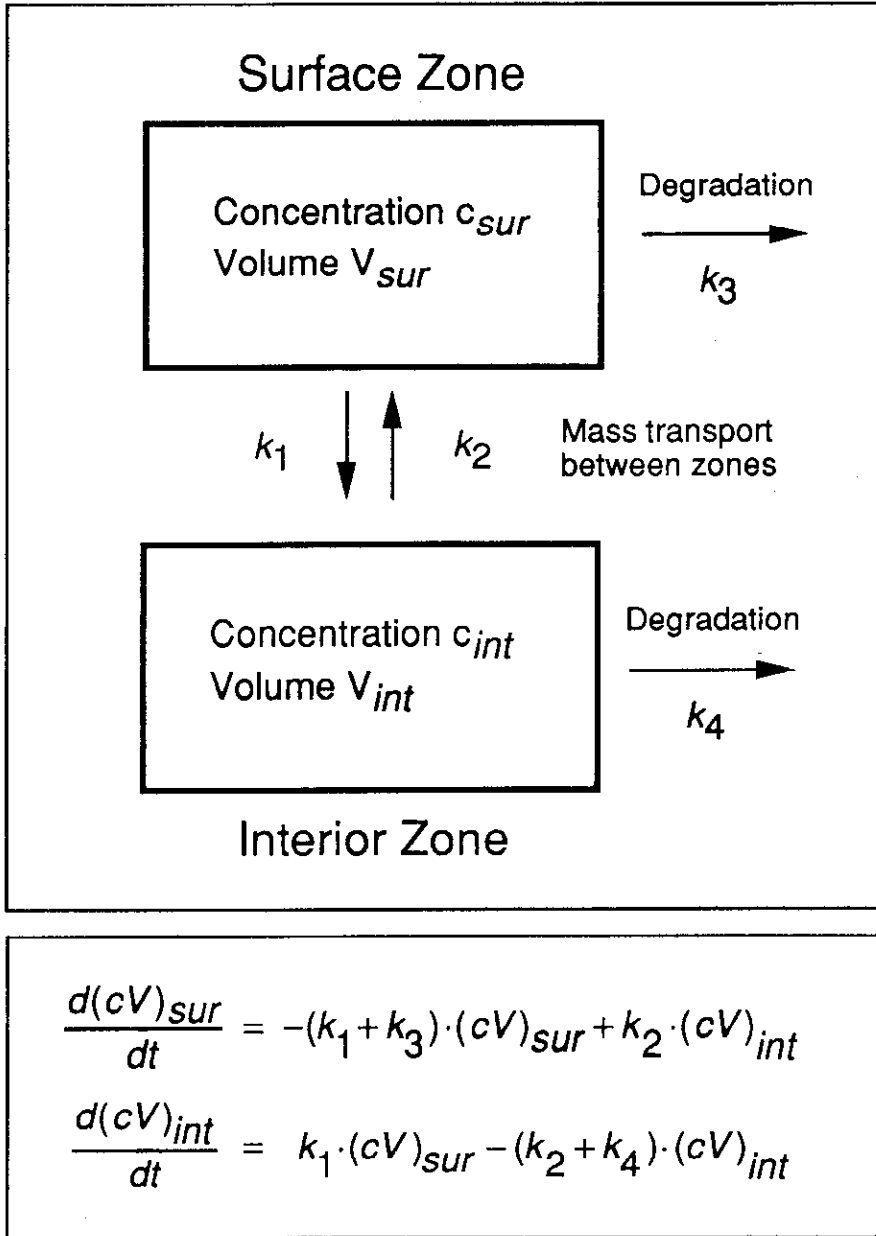


FIGURE 12  
**Schematic representation of the Two-Zone Model.**  
 For explanation of  $k_1$ - $k_4$ , see APPENDIX 11.

The growth profile can be described by making use of the coefficients given in Table 13 and formula (III) in Table 12. In addition, it must be taken into account that the volume of a spherical fruit is proportional to the 3rd power of the diameter  $D$  ( $V = \pi/6 \times D^3$ ). Therefore we can express the total fruit volume as a function of time, with a suitable proportionality constant,  $K$ , by the following terms for  $V_{\text{tot}}(t)$ :

$$\begin{aligned} V_{\text{tot}}(t) \text{ for: } & \text{Apples} & = & K \{0.3 + 0.7 (t/105 + 0.55)\}^3 \\ & \text{Peaches} & = & K \{0.5 + t/104\}^3 \\ & \text{Pears} & = & K \{0.7 + t/126\}^3 \\ & \text{Nectarines} & = & K \{0.2 + 0.8 (t/68 + 0.5)\}^3. \end{aligned}$$

If no function to describe the growth profile of a given crop is available, the residue values can be fitted by a function that ignores growth. This can be done on the assumption that an adequate fit will be obtained with the aid of the four coefficients,  $P_1 - P_4$ , which are obtained when the denominator in (ii) Table 14 is ignored.

The model was verified using about 280 decline studies. With the aid of direct fitting procedures, the fitting parameters  $P_1 - P_4$  and the residual sum of squares, as a measure of the goodness of fit, were calculated. The growth profile was taken into account in 70 % of the studies. In 46 % of the cases, one addend in the model equations contributed less than 5 % to the result of the fit. No correlation with certain active ingredients or crops was detectable.

In about 90 % of the cases, the decline of the residues could be reproduced more accurately with this model than with the formal models described in Section 2. Particularly the «sagging» phenomenon often observed shortly after treatment could be described to a very close approximation by the two-zone model. 98 % of the cases that previously could be described only by one of the root functions were described more successfully by the new approach.

In Figure 11 *b* an example was given of residue data fitted with the two-zone model in comparison to curves fitted with both the formal and the physical model of 1st order. This Figure shows that the curve obtained with the two-zone model compares well with that obtained by the physical model.

In Figure 13, another comparison is presented. In *a*, the curves were fitted with the formal model RF 1.5th order and the 2nd order physical model, respectively (showing that the formal approach gave a better fit). In *b*, the two-zone model was used.

The curve in Figure 13 *b* reveals a peculiarity that we occasionally observed when applying the two-zone model. In these cases, it must be assumed that the residue declines extraordinarily rapidly for a short time immediately after application of the pesticide. For such extremely short intervals, there are usually no experimental data to verify the interpretation made by the model. Formally, however, the fit of the curves to the measured values is superior to those obtained with the two other approaches, as may be seen from the sum of squared deviations (SSR) for Figure 13 in Table 15. For comparison, the SSR is also given for the curves of Figure 11. For exemplification, the Table lists the fitting parameters  $P$  which were obtained in the calculations for Figures 11 *b* and 13 *b*. In these calculations, account was taken of the growth profile for apples.

TABLE 14  
**Formulae for computing residue decline according to the Two-Zone Model**

$$c = \frac{(cV)_{sur,0}}{V_{tot}} \cdot \frac{k_1 e^{-k_4 t} + (k_3 - k_4) e^{-(k_1 + k_3) t}}{k_1 + (k_3 - k_4)} \quad (i)$$

where c = concentration of the compound averaged over the total volume of the plant part  
 t = time (e.g. days) after the (last) application of the pesticide

On merging the constants, we obtain

$$c = \frac{P_1 e^{-P_2 t} + P_3 e^{-P_4 t}}{V_{tot}(t)} \quad (ii)$$

resulting in:

$$P_1 = \frac{(cV)_{sur,0} \cdot k_1}{[k_1 + (k_3 - k_4)]} \quad k_4 = P_2$$

$$P_2 = k_4 \quad k_1 = \frac{1}{(cV)_{sur,0}} \cdot P_1 (P_4 - P_2)$$

$$P_3 = \frac{(cV)_{sur,0} \cdot (k_3 - k_4)}{[k_1 + (k_3 - k_4)]} \quad k_3 = P_4 - k_1$$

$$P_4 = (k_1 + k_3) \quad k_2 = 0$$

TABLE 15  
**Comparison of the goodness of fit of three models, and model parameters in the fitting function, for Figs. 11 b and 13 b**

	Figure 11			Figure 13		
	Formal fit	Physic. Model	Two-Zone Model	Formal fit	Physic. Model	Two-Zone Model
SSR · 10 <sup>4</sup>	29	14	14	4	15	1
P <sub>1</sub>	—	—	0.1300	—	—	0.0291
P <sub>2</sub>	—	—	0.0878	—	—	8.2300
P <sub>3</sub>	—	—	0.0000	—	—	0.0416
P <sub>4</sub>	—	—	—	—	—	0.0131



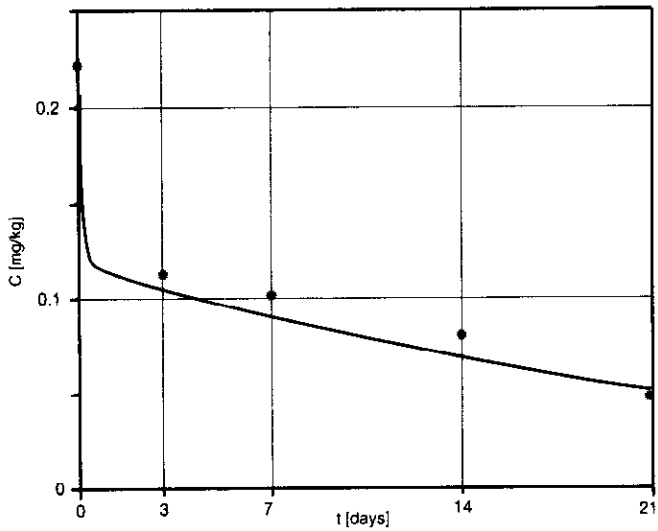
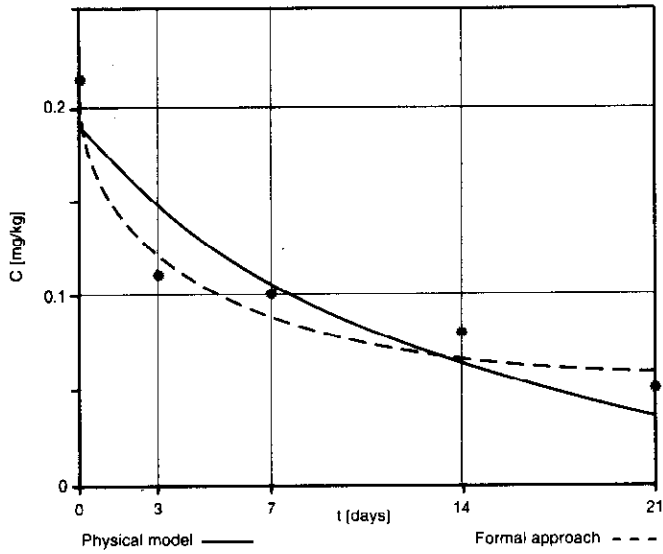


FIGURE 13

**Decline of an acaricide in apples.**

*a:* Curves fitted with the RF 1.5th-order formal model and the 2nd-order physical model.

*b:* Curve fitted with the Two-zone model.

As for the physical models which account for plant growth, we presented the two-zone model hoping that those interested in residue kinetics might feel encouraged to test it for further investigation of its merits or demerits.

## ADDENDUM

*Reference: Section 2.1.*

It should be noted that, because of the different ways of transformation of the residue values,  $R$ , in the formal models ( $R$  being transformed into either  $\log R$  or  $1/\sqrt{R}$  or  $1/R$ ), the resulting confidence intervals are not directly comparable if a given set of data is evaluated by different models.

*Reference: Section 2.2.1.*

If the amount of  $L_0$  is not known beforehand, e.g. in an *unweighted* linear regression, it can be obtained by computing the following terms:

$$\text{in 1st order functions: } L_0 = (10^S - 1) \cdot 100$$

$$\text{in 1.5th order functions: } L_0 = \left[ \left( \frac{1}{1-S} \right)^2 - 1 \right] \cdot 100$$

$$\text{in 2nd order functions: } L_0 = \left[ \left( \frac{1}{1-S} \right) - 1 \right] \cdot 100$$

with  $S$  as defined in APPENDIX 1.

These terms refer to the *upper* limits of the latitude ranges.

## ACKNOWLEDGEMENT

The studies reported here were performed in collaboration with several colleagues of the Bayer AG, Leverkusen, Germany, in their various expert capacities: Gabriele Timme and V. Laska (particularly concerning the formal models), N. Kalkert and J. Lahrs (in the studies reported under 4.2.1 and 4.2.2), H.-G. Frosch (in the studies reported under 4.2.3), and S. Quest, who formulated the Two-zone model.

## LITERATURE CITED

- (1) **Bundesgesundheitsamt (Beck, H.):**  
Untersuchungsmethoden zur Bestimmung der Rückstände von Chlorkohlenwasserstoff-Pestiziden in oder auf Lebensmitteln tierischer Herkunft.  
Bundesgesundheitsblatt **17**, 269 - 276 (1974).
- (2) **Frehse, H. and Timme, G.:**  
Quantitative residue analytical reliability: Beatitude through application of latitude.  
Residue Reviews **73**, 27 - 47 (1980).
- (3) **Gunther, F.A. and Blinn, R.C.:**  
Analysis of Insecticides and Acaricides, p.135 ff.  
Interscience Publishers, New York, London (1955).
- (4) **Horwitz, W.:**  
Analytical measurements: How do you know your results are right?  
in: Bandal, S.K. et al. (Eds.): The Pesticide Chemist and Modern Toxicology.  
ACS Symposium Series, No. 160, 411-438, Washington, D.C. (1981).
- (5) **Horwitz, W.:**  
Evaluation of analytical methods used for regulation.  
J. Assoc. Off. Anal. Chem. **65**, 525-530 (1982).
- (6) **Horwitz, W., Kamps, L.R. and Boyer, K.W.:**  
Quality assurance in the analysis of foods for trace constituents.  
J. Assoc. Off. Anal. Chem. **63**, 1344 - 1354 (1980).
- (7) **Kolbe, W.:**  
Seasonal course of top and small fruit development stages in relation to annual weather and crop protection measures.  
Pflanzenschutz-Nachr. Bayer **32**, 93 - 159 (1979), Engl. Edit.
- (8) **Quest, S., Walter, H.-F., Frehse, H. and Timme, G.:**  
Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. IV. Physically based methods.  
Pflanzenschutz-Nachr. Bayer **46**, 289 - 314 (1993), Engl. Edit.
- (9) **Timme, G. and Frehse, H.:**  
Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. I.  
Pflanzenschutz-Nachr. Bayer **33**, 47 - 60 (1980), Engl. Edit.

- (10) **Timme, G., Frehse, H. and Laska, V.:**  
Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. II.  
Pflanzenschutz-Nachr. Bayer **39**, 187 - 203 (1986), Engl. Edit.
- (11) **Walter, H.-F., Frehse, H. and Timme, G.:**  
Statistical interpretation and graphic representation of the degradational behaviour of pesticide residues. III. Extension of the formal description of degradation.  
Pflanzenschutz-Nachr. Bayer **46**, 265 - 288 (1993), Engl. Edit.
- (12) **Walter, H.-F.:**  
in: TIMME et al. (Ref. 10).

**APPENDIX 1: Confidence intervals for decline curves in linear regression**

First, we require the following terms:

$$\begin{aligned}
 s_{xx} &= \sum (x_i - \bar{x})^2 &= \sum x^2 - (\sum x)^2/n \\
 s_{yy} &= \sum (y_i - \bar{y})^2 &= \sum y^2 - (\sum y)^2/n \\
 s_{xy} &= \sum (x_i - \bar{x}) \cdot (y_i - \bar{y}) &= \sum xy - (\sum x \cdot \sum y)/n &= \sum xy - \bar{y} \sum x \\
 s_{sr} &= \sum (y_i - \hat{y})^2 &= s_{yy} - b \cdot s_{xy} &= s_{yy} - (s_{xy})^2/s_{xx} \\
 &= \text{sum of squared residuals.}
 \end{aligned}$$

With the aid of the residual variance

$$S^2 = \frac{ssr}{n-2}$$

we can calculate the two standard deviations,

$$\begin{aligned}
 s_E &= S \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{s_{xx}}} \\
 &= \text{standard deviation of an } \textit{estimated} \hat{y} \text{ (value on the straight line) at a given point } x
 \end{aligned}$$

$$\begin{aligned}
 s_P &= S \cdot \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{s_{xx}}} \\
 &= \text{standard deviation of a } \textit{predicted} \text{ value } y \text{ at a given point } x.
 \end{aligned}$$

and with them, the confidence intervals

$$CI_{(SL)} = \hat{y} \pm s_E \cdot \sqrt{2} F$$

$$CI_{(E)} = \hat{y} \pm s_E \cdot t$$

$$CI_{(P)} = \hat{y} \pm s_P \cdot t$$

where  $\hat{y}$  = estimated value(s) of y on the regression line  
 $F$  =  $F_{(2, n-2)}$  from tables of the F-distribution, normally at  $\alpha = 5\%$   
 $t$  =  $t_{(n-2)}$  from tables of the t-distribution, normally at  $\alpha = 5\%$  (two-sided).

Here, S is identical with the term  $s_{Rest}$  used in Appendices 5 and 7.

**APPENDIX 2: Test for the goodness of fit**

The coefficient of determination (according to WALTER (12)) is

$$r^2 = 1 - \frac{\sum(R_i - R_{mod})^2}{\sum(R_i - \bar{R})^2}$$

with  $R_i$  = residue values measured  
 $R_{mod}$  = residue values on the (back-transformed) decline curve  
 $\bar{R}$  = mean of the measured  $R_i$ .

This computation is performed in the *back-transformed* system because the selection of the best fit model is also performed in this system.

The test quantity D is

$$D = |r| - \frac{t}{\sqrt{t^2 + (n-2)}}$$

with  $|r|$  = absolute value of the coefficient of correlation,  $r$   
 $t$  = t-value (two-sided), with  $n - 2$  degrees of freedom, from the table of the t-distribution.

**APPENDIX 3: Derivation of the absolute standard deviation in expressing analytical latitudes**

The formula for the calculation of the relative standard deviation  $L$  (Table 5) can be modified to represent the *absolute* standard deviation  $L_{abs}$ . By inserting the terms

$$L_{abs} = \frac{L \cdot R}{100} \quad \text{and} \quad L_{0,abs} = \frac{L_0}{100}$$

in formula (1 a) in Table 5 we obtain

$$\begin{aligned} \log(L_{abs}) &= \log(L_{0,abs}) + \log R (\log K + 1) \\ &= \log(L_{0,abs}) + \log R (\log K + \log 10) \\ &= \log(L_{0,abs}) + \log R (\log 10 K) \end{aligned}$$

or, by taking antilogarithms,

$$L_{abs} = L_{0,abs} \cdot R^{\log(10 K)} = L_{0,abs} \cdot R^{1 + \log(K)}$$

<b>APPENDIX 4: Conversion of HORWITZ's formula</b>
--

HORWITZ (5) expressed the relationship between the  $CV$  of the measured values and the percentage content  $G$  of the analyte in a sample by the (here slightly modified) form

$$CV(\%) = 2^{1-0.5 \log G}$$

But  $\log G = \log(R \cdot 10^{-6})$ ,

and so the HORWITZ formula becomes

$$\begin{aligned} CV(R) &= 2^{1-0.5[-6 + \log(R)]} = 2^{4-0.5 \log(R)} \\ &= 16 \cdot \left(\frac{1}{\sqrt{2}}\right)^{\log(R)} \end{aligned}$$

Since  $a \log(b) = b \log(a)$ , we may interchange  $\left(\frac{1}{\sqrt{2}}\right)$  and  $R$  in this equation, giving

$$CV(R) = 16 \cdot R^{\log\left(\frac{1}{\sqrt{2}}\right)}$$

A comparison with formula (I b), Table 5, and the formula in Figure 5 gives, as stated ( $CV = L$ ):

$$L_0 = 16 \%, \quad K = \left(\frac{1}{\sqrt{2}}\right) \approx 0.7071.$$

<b>APPENDIX 5: Linear weighted regression</b>
---

In *unweighted* linear regression, the variance of the target (residue) values is regarded as constant. By using a linear weighted regression it is possible to formulate a non-constant variance so that for every value of  $Y_i$  the variance (Var) is given explicitly by

$$Var(Y_i) = \frac{\sigma^2}{w_i} \quad (i = 1, 2, \dots, n). \quad (i)$$

Here the weights  $w$  are known positive numbers which will be defined in such a way that they correspond to a certain value  $K$  (Table 5). If the magnitude of  $\sigma^2$  is unknown, the residual variance,  $s_{Res}^2$ , is used as the best estimate.

#### Calculation of weights

If we regard the formulae given in Table 5 as expressing the relative and absolute standard deviation of the analytical error in the determination of  $R$ , we can use (II a) to calculate the variance. We obtain

$$Var(R) = L_{abs}^2 = R^2 \cdot L_{0,abs}^2 \cdot R^{2 \log(K)}. \quad (ii)$$

According to the error propagation law (see APPENDIX 6) the variance of the transformed analytical values  $Y = f(R)$  (see Table 1) is directly related to  $Var(R)$  as follows:

$$Var(Y = f(R)) \approx \left[ \frac{df}{dR} \right]^2 Var(R). \quad (iii)$$

Thus in conjunction with (i) above, the weights may be calculated as

$$w = \frac{\sigma^2}{Var(Y)} \approx \frac{\sigma^2}{\left[ \frac{df}{dR} \right]^2 Var(R)}. \quad (iv)$$

The function  $f$  differs according to the model employed, being

- $f = \log(R)$  for the 1st- and RF 1st-order models
- $f = 1/\sqrt{R}$  for the 1.5th- and RF 1.5th-order models
- $f = 1/R$  for the 2nd- and RF 2nd-order models.

The appropriate forms of  $f$  must be employed in calculating the weights  $w$  for models A – F (Table 1).

In these calculations, any constant factor occurring in the right hand side of formula (iv) above can be omitted, because it would cancel out (e.g. in formulae (v) and (vi) below). This fact was employed in the calculation of the weights, whereby formula (iv) resulted in the expressions listed in Table 7.

*Calculation of confidence intervals*

To fit the model constants  $a$  and  $b$  (according to Table 1), the method of least squares is employed in such a way that now a minimum is sought for the *weighted* sum

$$S = \sum_{i=1}^n w_i (Y_i - a - bX_i)^2.$$

We then obtain for  $a$  and  $b$  the expressions

$$b = \frac{S_{wxy}}{S_{wxx}} \quad a = \bar{Y}_w - b \cdot \bar{X}_w$$

where  $n$  = number of value pairs (Y; X)

$$\bar{X}_w = \text{weighted mean of all } X \text{ - values} = \frac{\sum wX}{\sum w}$$

$$\bar{Y}_w = \text{weighted mean of all } Y \text{ - values} = \frac{\sum wY}{\sum w}$$

$$S_{wxx} = \sum wX^2 - \sum w(\bar{X}_w)^2 \quad \text{and} \quad S_{wxy} = \sum wXY - \sum w\bar{X}_w\bar{Y}_w.$$

As residual variance, whose expectation is proportional to  $\sigma^2$ , we obtain



$$s_{Rest}^2 = \frac{1}{n-2} \left[ S_{wyy} - \frac{S_{wxy}^2}{S_{wxx}} \right]$$

$$\text{with } S_{wyy} = \sum wY^2 - \sum w(\bar{Y}_w)^2.$$

For our evaluation, we require the confidence interval for the entire line (CI<sub>(SL)</sub>) and/or the prediction interval CI<sub>(P)</sub>.

The confidence interval of the line is given by

$$Y_{\pm} = \bar{Y}_w + b(X' - \bar{X}_w) \pm k_5 \cdot \Delta s_5 \quad (\text{v})$$

$$\text{with } k_5 = \sqrt{2 F_{95\%}(2; n-2)} \quad \text{and} \quad \Delta s_5 = s_{Rest} \cdot \sqrt{\frac{1}{\sum w} + \frac{(X' - \bar{X}_w)^2}{S_{wxx}}},$$

where  $F_{95\%}(2; n-2)$  is the critical value of the F-distribution for a confidence level of 95 % with 2 and n-2 degrees of freedom, and  $X'$  is the point of the running coordinate.

The corresponding formula for the prediction interval is:

$$Y_{\pm} = \bar{Y}_w + b(X' - \bar{X}_w) \pm k_6 \cdot \Delta s_6 \quad (\text{vi})$$

$$\text{with } k_6 = t_{95\%, f=n-2} \quad \text{and} \quad \Delta s_6 = s_{Rest} \cdot \sqrt{\frac{1}{w'} + \frac{1}{\sum w} + \frac{(X' - \bar{X}_w)^2}{S_{wxx}}}.$$

The factor  $t$  is taken from a table of critical values of the t-distribution for n-2 degrees of freedom and a confidence level of 95 % (two-tailed). The weight  $w'$  at the point of the running coordinate  $X'$  is calculated in accordance with formula (iv).

Formula (vi) can be used to determine the time at which a given residue level  $Y = Y'$  will exist, taking the prediction interval into account. To do this, the equation is solved for  $X'$ :

$$X' = \bar{X}_w + \frac{b(Y' - \bar{Y}_w)}{C} \pm k_7 \cdot \Delta s_7 \quad (\text{vii})$$

$$\text{with } k_7 = t_{95\%, f=n-2} \quad \text{and} \quad \Delta s_7 = \frac{s_{Rest}}{C} \sqrt{\left[ \frac{1}{w'} + \frac{1}{\sum w} \right] C + \frac{(Y' - \bar{Y}_w)^2}{S_{wxx}}}$$

$$\text{where the quantity } C \text{ is defined as } C = b^2 - \frac{(t_{95\%, f=n-2} \cdot s_{Rest})^2}{S_{wxx}}.$$

**APPENDIX 6: Gauss's error propagation law**

With the aid of the error propagation law the error of a quantity  $z$  can be calculated from the errors (if known) of the quantities

$$x_i \quad (i = 1, 2, 3, \dots)$$

if values of  $z$  and  $x$  are linked by a known function  $f$  in accordance with

$$z = f(x_1, x_2, x_3, \dots).$$

The variance of  $z$ ,  $Var(z)$ , is defined by ( $E$  = expectation)

$$Var(z) = E \left[ (z - E(z))^2 \right].$$

For the sum of two values  $z_1$  and  $z_2$  we have

$$Var(z_1 + z_2) = E \left[ (z_1 + z_2 - E(z_1 + z_2))^2 \right].$$

Because

$$E(z_1 + z_2) = E(z_1) + E(z_2),$$

by applying the binominal formula  $(a + b)^2 = a^2 + 2ab + b^2$  we obtain

$$Var(z_1 + z_2) = Var(z_1) + Var(z_2) + 2Cov(z_1, z_2)$$

where  $Cov$  is the covariance defined by

$$Cov(z_1, z_2) = E \left[ (z_1 - E(z_1))(z_2 - E(z_2)) \right].$$

If we replace  $z_1$  by  $k_1 z_1$  and  $z_2$  by  $k_2 z_2$  with two arbitrary constants  $k_1$  and  $k_2$ , then since

$$E(kz) = kE(z) \quad \text{and} \quad Var(kz) = k^2 Var(z),$$

we obtain the somewhat more general equation

$$Var(k_1 z_1 + k_2 z_2) = k_1^2 Var(z_1) + k_2^2 Var(z_2) + 2k_1 k_2 Cov(z_1, z_2). \quad (i)$$

Now the function  $f$  at the "0" position can be expanded into a Taylor series and broken off after the first-order terms:

$$z \approx z_0 + \left( \frac{\partial f}{\partial x_1} \right)_0 \cdot (x_1 - x_{10}) + \left( \frac{\partial f}{\partial x_2} \right)_0 \cdot (x_2 - x_{20}) + \dots$$

Repeated application of equation (i) to this relation gives

$$Var(z) \approx \sum_i \left( \frac{\partial f}{\partial x_i} \right)_0^2 Var(x_i) + 2 \sum_{i < j} \left( \frac{\partial f}{\partial x_i} \right)_0 \left( \frac{\partial f}{\partial x_j} \right)_0 Cov(x_i, x_j). \quad (ii)$$

**Special case I:** Only one variable  $x_1 = R$ , so that  $z = f(R)$

Here the covariances vanish and the partial derivatives become ordinary derivatives, giving

$$Var(z) = f(R) \approx \left[ \frac{df}{dR} \right]^2 Var(R)$$

(see APPENDIX 5).

**Special case II:** Accuracy of  $(T/X)$ , e.g. for the 1st-order model (see Section 2.3):

As in Special case I, there is again just one variable  $x_1 = b$ , and we have

$$z = (T/X) = \frac{\log X}{-b} = f(b)$$

so that, by analogy with Special case I, we obtain

$$\begin{aligned} Var(T/X) &\approx \left[ \frac{\log(X)}{b^2} \right]^2 Var(b) \\ &= \left[ \frac{(T/X)}{-b} \right]^2 Var(b) \end{aligned}$$

(see Tables 9 and 11).

#### APPENDIX 7: Accuracy of times $T/X$

The standard deviations of the decline times are derived with the aid of the error propagation law (APPENDIX 6), using the model parameters  $a$  and  $b$  for the variables  $x_i$ . For the variances of  $a$  and  $b$  it is necessary to use

$$Var(a) = \left[ \frac{1}{\sum w} + \frac{\bar{X}_w^2}{S_{wxx}} \right] \cdot s_{Rest}^2 \quad \text{and}$$

$$Var(b) = \frac{s_{Rest}^2}{S_{wxx}}$$

and to take into account that both parameters are not independent of each other:

$$Cov(a, b) = \frac{\bar{X}_w \cdot s_{Rest}^2}{S_{wxx}}$$

(see APPENDIX 5 for definition of the individual terms).

For the individual models, we then obtain the corresponding expressions for calculating the accuracy of the times  $T/X$ , expressed as the standard deviation  $S_{T/X}$  (Table 11). These expressions are obtained in analogy to Special case II in APPENDIX 6.

**APPENDIX 8: Multiple linear regression**

In multiple linear regression there are k independent variables

$$x_1, x_2, x_3, \dots, x_k$$

to be related to the dependent variable y as follows:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k + e \tag{i}$$

$$y = y_c + e, \tag{ii}$$

where

$$y_c = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k$$

represents the value calculated from the regression function.

If there are n samples available for determination of the regression coefficients b and the residues e, the relationship (i) for the n data sets can also be formulated, very simply, in matrix notation:

$$y = Xb + e \tag{iii}$$

y-values	x-values				
y <sub>1</sub>	x <sub>11</sub>	x <sub>12</sub>	x <sub>13</sub>	...	x <sub>1k</sub>
y <sub>2</sub>	x <sub>21</sub>	x <sub>22</sub>	x <sub>23</sub>	...	x <sub>2k</sub>
⋮	⋮	⋮	⋮	⋮	⋮
y <sub>n</sub>	x <sub>n1</sub>	x <sub>n2</sub>	x <sub>n3</sub>	...	x <sub>nk</sub>

Here y is a (n,1) vector comprising the n sample values of y, b is a (k+1,1) vector comprising the regression coefficients b, and e is a (n,1) vector comprising the residues:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}; \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix}$$

X denotes the following (n,k+1) matrix:

$$X = \begin{pmatrix} x_{10} & x_{11} & x_{12} & x_{13} & \dots & x_{1k} \\ x_{20} & x_{21} & x_{22} & x_{23} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n0} & x_{n1} & x_{n2} & x_{n3} & \dots & x_{nk} \end{pmatrix}$$

To simplify the notation, here the additional variable x<sub>0</sub> is introduced, whose sample values x<sub>i0</sub> are all identical to 1.

Further calculation with the aid of matrix algebra leads to

- values for the model parameters (regression coefficients)
- an estimate of the accuracy of the model parameters
- a definition of confidence intervals.

The regression coefficients  $b$  are determined, as usual, by the method of least squares. The defining equations for the coefficients  $b$  are the so-called normal equation system.

In matrix notation the normal equation system is

$$(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y},$$

where the prime (') indicates the transposed matrix obtained by interchanging the rows and columns.

Determination of the inverse matrix with matrix elements  $C_{ij}$

$$(\mathbf{X}'\mathbf{X})^{-1} = \{C_{ij}\}$$

finally yields for the coefficients  $b$ :

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{iv}$$

Accuracy of the model parameters (variance-covariance matrix)

To access the accuracy of the model parameters, the variances of the regression coefficients are required. These are best obtained from the so-called variance-covariance matrix of the regression coefficients, the diagonal elements of which are the required variances.

The variance-covariance matrix takes the form

$$\text{Var}(\mathbf{b}) = E\{[\mathbf{b} - E(\mathbf{b})][\mathbf{b} - E(\mathbf{b})]'\}, \tag{v}$$

where  $E$  is the expectation. If  $\text{Var}$  denotes the variance and  $\text{Cov}$  denotes covariance, the variance-covariance matrix in its expanded form is

$$\text{Var}(\mathbf{b}) = \begin{pmatrix} \text{Var}(b_0) & \text{Cov}(b_0, b_1) & \text{Cov}(b_0, b_2) & \cdots & \text{Cov}(b_0, b_k) \\ \text{Cov}(b_1, b_0) & \text{Var}(b_1) & \text{Cov}(b_1, b_2) & \cdots & \text{Cov}(b_1, b_k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(b_k, b_0) & \text{Cov}(b_k, b_1) & \text{Cov}(b_k, b_2) & \cdots & \text{Var}(b_k) \end{pmatrix}.$$

Insertion of (iv) in (v) yields

$$\text{Var}(\mathbf{b}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}. \tag{vi}$$

If  $\sigma^2$  is unknown, the residual variance

$$s_{Rest}^2 = \frac{1}{n - k - 1} \sum_{i=1}^n e_i^2$$

is used as the best estimate.

Thus the formula for calculating the variance-covariance matrix is finally

$$\text{Var}(\mathbf{b}) \approx s_{Rest}^2 (\mathbf{X}' \mathbf{X})^{-1} \tag{vii}$$

Definition of confidence intervals

To establish a confidence interval for the true value  $E(y)$  at the  $n$  sample positions, the variance for the difference between  $y_c$  (see equation (ii)) and  $E(y)$  at these sites must be determined.

In matrix notation  $\mathbf{y}_c = \mathbf{Xb}$  and  $E(y) = \mathbf{XE}(\mathbf{b})$ , so that

$$\text{Var}\{\mathbf{y}_c - E(y)\} = \sigma^2 \mathbf{X}(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' = \{V_{ij}\}.$$

Only the diagonal elements  $V_{mm}$  of this matrix are used.

With the matrix elements of the variance-covariance matrix, denoted above by  $C_{ij}$ , we obtain

$$V_{mm} \approx \sum_{i=0}^k \sum_{j=0}^k C_{ij} x_{mi} x_{mj} s_{Rcst}^2.$$

The confidence interval at the  $m^{\text{th}}$  measurement point is then

$$y_c - t_{\alpha/2, f=n-k-1} \cdot \sqrt{V_{mm}} \leq E(y) \leq y_c + t_{\alpha/2, f=n-k-1} \cdot \sqrt{V_{mm}}, \tag{viii}$$

where  $\alpha$  is the significance level and  $t$  is the one-sided critical value of the  $t$  distribution.

If simultaneous confidence intervals (the confidence intervals of the curves) are to be stated, the  $t$  factor in (viii) must be replaced by  $\sqrt{2F \cdot (1+k)}$ , where  $F$  is the critical value of the  $F$  distribution for  $f_1 = k + 1$  and  $f_2 = n - k - 1$  degrees of freedom.

Confidence intervals for non-linear regression

For the case in which the model function that is to be fitted,  $f$ , depends non-linearly on the model constants in accordance with

$$y_i = f(b_0, b_1, \dots, b_k; x_{1i}, x_{2i}, \dots, x_{ki}) + e_i = f(b; x_i) + e_i \quad \text{where } i = 1, 2, 3, \dots, n,$$

the procedure is as follows:

Direct fitting using the method of least squares gives a set of model constants  $b_{j_0}$ . These can be used to linearize the function  $f$  and replace it with the tangential plane:

$$y - y_0 \approx \left( \frac{\partial f}{\partial b_0} \right)_0 \cdot (b_0 - b_{0_0}) + \left( \frac{\partial f}{\partial b_1} \right)_0 \cdot (b_1 - b_{1_0}) + \dots + \left( \frac{\partial f}{\partial b_k} \right)_0 \cdot (b_k - b_{k_0}).$$

In matrix notation we then obtain

$$\mathbf{y} - \mathbf{y}_0 \approx \mathbf{J}(\mathbf{b} - \mathbf{b}_0), \tag{ix}$$

where  $\mathbf{J}$  is the Jacobi matrix, in which the second subscript on the partial derivatives refers to the sampling position:

$$\mathbf{J} = \begin{pmatrix} \left(\frac{\partial f}{\partial b_0}\right)_{0,1} & \left(\frac{\partial f}{\partial b_1}\right)_{0,1} & \cdots & \left(\frac{\partial f}{\partial b_k}\right)_{0,1} \\ \left(\frac{\partial f}{\partial b_0}\right)_{0,2} & \left(\frac{\partial f}{\partial b_1}\right)_{0,2} & \cdots & \left(\frac{\partial f}{\partial b_k}\right)_{0,2} \\ \left(\frac{\partial f}{\partial b_0}\right)_{0,3} & \left(\frac{\partial f}{\partial b_1}\right)_{0,3} & \cdots & \left(\frac{\partial f}{\partial b_k}\right)_{0,3} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\partial f}{\partial b_0}\right)_{0,n} & \left(\frac{\partial f}{\partial b_1}\right)_{0,n} & \cdots & \left(\frac{\partial f}{\partial b_k}\right)_{0,n} \end{pmatrix}$$

### Referring back from non-linear regression to multiple linear regression

Equation (ix) is structured exactly like equation (iii), however with  $y$  and  $b$  being replaced by the differences from the values at the linearization point. But, since

$$\text{Var}(\mathbf{b} - \mathbf{b}_0) = \text{Var}(\mathbf{b}),$$

the calculation of the variances is not hindered by this. The matrix  $\mathbf{X}$  is replaced by the Jacobi matrix  $\mathbf{J}$ , and so all the formulae for the variance-covariance matrix and the confidence intervals that were derived for the multiple linear case can be adopted directly after replacing  $\mathbf{X}$  by  $\mathbf{J}$ .

### Confidence intervals and prediction intervals

The formulae for the confidence interval can be taken directly from (viii), the matrix elements  $C_{ij}$  being determined by the Jacobi matrix.

To obtain the formulae for the prediction interval, all that is necessary is to replace the quantity  $V_{\text{mm}}$  in (viii) by  $V_{\text{mm}} + s_{\text{Rcst}}^2$  as in the case of multiple linear regression.

### APPENDIX 9: Fitting to a degradation law of 1st order

By using the coefficients given in Tables 12 and 13, and with  $V = \pi/6 \cdot D^3$ , apple growth can be described by

$$V/V_{\text{max}} = 1.09^3 (T/T_{\text{max}})^{2.4} \quad (\text{i})$$

Biologically, e.g. enzymatically, induced degradation obeying 1st-order kinetics is described by

$$\frac{d}{dt} Vc = -Vkc \quad (\text{ii})$$

where

- $V$  = fruit volume
- $c$  = residue concentration
- $k$  = reaction constant for degradation
- $t$  =  $T - T_P$  = time from  $T_P$ .

By substituting (i) in (ii) and integration (initial conditions  $t = t_0 = 0, c = c_0$ ) we obtain

$$c(t) = c_0 \frac{\exp(-kt)}{(1+t/T_p)^{2.4}} \quad \text{(iii)}$$

where  $t_0 =$  time at which the first sample for residue measurement is taken  
 $= T_p$   
 $c_0 =$  residue concentration at time  $t_0$ .

**APPENDIX 10: Fitting to a degradation law of arbitrary order**

In analogy to formula (ii) in APPENDIX 9, biologically induced degradation obeying kinetics of *arbitrary* order is described by the more general expression

$$\frac{d}{dt}Vc = -Vkc^n, \quad \text{(i)}$$

where  $n$  is the order of the reaction.

In the case  $k = 0$  (no degradation), equation (i) has the solution

$$\frac{c}{c_0} = \frac{V_0}{V},$$

so the function  $f = \frac{V_0}{V}$  stands for the *dilution* effect, where  $V_0$  is the volume at  $t = 0$ .

Assuming a multiplicative model, we can now introduce a function  $g$  standing for the degradation effect, and where the *combined* effect is given by

$$\frac{c}{c_0} = f(t) \cdot g(k,t) = \left(\frac{V_0}{V}\right) \cdot g. \quad \text{(i a)}$$

According to (i),  $g$  is a solution of the equation

$$\frac{dg}{dt} = -k \cdot \left(\frac{c_0 V_0}{V}\right)^{n-1} \cdot g^n \quad \text{with } g(t=0) = 1.$$

If the growth profile for the period of interest, from the last use of the pesticide to maturity of the crop, is represented by

$$V = V_0 + Bt, \quad \text{(ii)}$$

then solution of the differential equation (i a) with the growth law (ii) gives the following equations for the *degradation* effect alone:



$$g(k, t) = e^{-kt} \quad \text{for } \underline{n = 1} \quad \text{(iii a)}$$

$$= \left[ 1 - \frac{1-n}{2-n} \cdot \frac{k}{c_0^{1-n}} \cdot \frac{V_0}{B} \cdot \left( \left( \frac{V}{V_0} \right)^{2-n} - 1 \right) \right]^{1/(1-n)} \quad \text{for } \underline{1 < n < 2, B \neq 0} \quad \text{(iii b)}$$

$$= \left[ 1 + \frac{V_0}{B} \cdot c_0 \cdot k \cdot \ln \left( \frac{V}{V_0} \right) \right]^{-1} \quad \text{for } \underline{n = 2, B \neq 0} \quad \text{(iii c)}$$

$$= \left[ 1 + (n-1) k c_0^{n-1} t \right]^{1/(1-n)} \quad \text{for } \underline{B = 0, n \neq 1} \quad \text{(iii d)}$$

(no increase in plant mass).

(i a) is then the expression for the combined effect of dilution and degradation.

To fit the solution (i a) to the experimental results, the constants  $V_0$  and  $B$  were first determined from the weights of the individuals ( $m \sim V$ ) by regression analysis. Here it is assumed that the density of the plant material remains constant over time. Next, the constants  $k$  and  $c_0$  at constant reaction order,  $n$ , were determined via a direct fit. This step was repeated for various (assumed) reaction orders. The goodness of fit was evaluated in terms of the sum of squares

$$F = \sum (c_{\text{calculated}} - c_{\text{experimental}})^2.$$

The optimal reaction order is then the one for which the sum of squares is the smallest (method of least squares). In addition, the coefficient of determination (square of the correlation coefficient) was considered:

$$r^2 = 1 - F / \sum (c_{\text{experimental}} - \bar{c}_{\text{experimental}})^2.$$

**APPENDIX 11: Derivation of the formulae for describing degradation according to the Two-Zone Model**

By introducing a mean concentration for the plant part as a whole

$$c = \frac{c_{sur} V_{sur} + c_{int} V_{int}}{V_{sur} + V_{int}} \quad \text{(i)}$$

and the total volume

$$V_{tot} = V_{sur} + V_{int} \quad \text{(ii)}$$

and integrating the two equations given in Figure 12, we obtain the solution

$$c = \frac{(cV)_{sur,0}}{V_{tot}} \cdot \frac{k_1 e^{k_4 t} + (k_3 - k_4) e^{-(k_1 + k_3) t}}{k_1 + (k_3 - k_4)} \quad \text{(iii)}$$

Integration of the system of equations

To integrate the differential equations let us make the substitution

$$y_{sur} = c_{sur}V_{sur} \tag{iv a}$$

$$y_{int} = c_{int}V_{int} \tag{iv b}$$

We then obtain the following linear system with constant coefficients:

$$\frac{dy_{sur}}{dt} = -(k_1 + k_3)y_{sur} + k_2y_{int} \tag{v a}$$

$$\frac{dy_{int}}{dt} = k_1y_{sur} - (k_2 + k_4)y_{int} \tag{v b}$$

For the general solution we assume the usual exponential law for the function y:

$$y_{sur} = A e^{-rt} \tag{vi a}$$

$$y_{int} = B e^{-rt} \tag{vi b}$$

Insertion of (vi a) and (vi b) in (v a) and (v b) leads to a homogeneous system of equations:

$$(r - k_1 - k_3)A + k_2 B = 0 \tag{vii a}$$

$$k_1A + (r - k_2 - k_4) B = 0 \tag{vii b}$$

Equations (vii a) and (vii b) only have non-trivial solutions when the determinant of coefficients is zero:

$$\text{Det} = \begin{vmatrix} r - k_1 - k_3 & k_2 \\ k_1 & r - k_2 - k_4 \end{vmatrix} = 0.$$

This gives the so-called characteristic equation for r, which has the following roots:

$$r_{1,2} = \frac{1}{2}(k_1 + k_2 + k_3 + k_4) \pm \sqrt{\frac{1}{4}[k_1 + k_2 + k_3 + k_4]^2 - k_1k_4 - k_2k_3 - k_3k_4}.$$

In every case, there are two non-negative real roots satisfying the following equations:

$$r_1 + r_2 = k_1 + k_2 + k_3 + k_4$$

$$r_1r_2 = k_1k_4 + k_2k_3 + k_3k_4.$$

The general solution of (v a) and (v b) is thus a sum of two exponential functions:

$$y_{sur} = A_1e^{-r_1t} + A_2e^{-r_2t}$$

$$y_{int} = B_1e^{-r_1t} + B_2e^{-r_2t}.$$

To determine the four integration constants, the initial conditions must be taken into account.

Moreover the  $A_i$  and  $B_i$  are interconnected, e.g. via equation (vii b):

$$y_{sur}(t=0) = c_{sur}(t=0) \cdot V_{sur}(t=0) = c_{sur,0}V_{sur,0}$$

$$y_{sur}(t=0) = 0.$$

$$k_1A_1 + (r_1 - k_2 - k_4) B_1 = 0$$

$$k_1A_2 + (r_2 - k_2 - k_4) B_2 = 0.$$

Hence the general solution is

$$y_{sur} = \frac{c_{sur,0} V_{sur,0}}{r_2 - r_1} \left[ (r_2 - k_1 - k_3) e^{-r_1 t} - (r_1 - k_1 - k_3) e^{-r_2 t} \right]$$

$$y_{int} = \frac{k_1 c_{sur,0} V_{sur,0}}{r_2 - r_1} \left[ e^{-r_1 t} - e^{-r_2 t} \right].$$

From these two functions the mean concentration

$$c = \frac{y_{sur} + y_{int}}{V_{sur} + V_{int}}$$

is now calculated, finally giving

$$c = \frac{c_{sur,0} V_{sur,0}}{(V_{sur} + V_{int})(r_2 - r_1)} \left[ (r_2 - r_1) e^{-r_1 t} - (r_1 - k_3) e^{-r_2 t} \right].$$

Growth is represented only by the bracket containing the total volume  $V_{sur} + V_{int}$ , so only the growth represented by the total volume  $V_{tot}$  is relevant.

For the special case  $k_2 = 0$  (no back transport), the expression for  $c$  becomes as given in formula (iii) above.

#### Explanation of symbols (if not explained elsewhere)

$k_1, k_2$	Exchange coefficients describing mass transport between surface zone and interior zone
$k_3$	Rate constant of degradation reaction in surface zone
$k_4$	Rate constant of degradation reaction in interior zone
$V_{sur,0}$	Volume of surface zone at time $t = 0$
$V_{int,0}$	Volume of interior zone at time $t = 0$