R&D REPORT NO. 50

Statistical Modelling of Lethality Distributions in Canned Foods

October 1997

Campden & Chorleywood
Food Research Association

R&D Report No. 50

Project Number No. 15560

Statistical Modelling of Lethality Distributions in Canned Foods

J E Hall

October 1997

© CCFRA - Campden & Chorleywood Food Research Association

Campden & Chorleywood
Food Research Association

Director-General
Prof. C. Dennis
BSc, PhD, FIFST

Chipping Campden
Gloucestershire
GL55 6LD UK
Tel: +44 (0) 1386 842000
Fax: +44 (0) 1386 842100

Information emanating from this Research Association is given after the exercise of all reasonable care and skill in its compilation, preparation and issue, but is provided without liability in its application and use.

The information contained in this publication must not be reproduced without permission from the Director-General of the Association.

SUMMARY

Any variation in initial temperature, retort temperature and most markedly heating factor can lead to insufficient sterilisation of canned foods. Statistical methods are presented here as an alternative to worst case scenario approaches for prediction of sterilisation values. If variation is known to exist, sufficient sterilisation can be ensured by extending processing times. This may, however, contribute to a greater loss of product quality. To avoid this it is preferable to minimise variability, particularly in heating factor. Whether or not this is possible, statistically valid predictions can be made for the percentage of products which will be sterilised to an adequate level. Results for a product with a heating factor normally distributed with mean 50 and standard deviation 3.5 minutes, processed in UT cans (73 mm diameter × 115 mm height) show F_0 values which appear to be approximately $\log_{10}$- normally distributed. Assuming this distribution allows the likelihood that the F_0 is less than, for example,

6 minutes to be derived from statistical tables. This approach could be helpful for use in process validation or improvement activities.

ACKNOWLEDGEMENTS

This work is part of a project entitled '*A new approach to heat distribution and heat penetration studies for thermal processing of foods in modern batch-type overpressure rotary retorts*' (AIR2-CT94-1017). We sincerely acknowledge the European Commission for financial support of this project.

CONTENTS

		Page No.
1.	INTRODUCTION	1
2.	MODELLING CRITERIA AND METHODS	1
2.1	Model Equations	1
2.2	Monte Carlo Method	2
2.3	Time Series Method	2
3.	RESULTS AND DISCUSSION	3
4.	CONCLUSIONS	15

1. INTRODUCTION

In-container heat sterilisation is a very important method used in the production of shelf stable foods. This process is used both to destroy micro-organisms, spores and enzymes as well as to cook the product. If the sterilisation time is too short there is a real danger of microbial survival which could lead to an outbreak of food poisoning. At the same time, if the product is over processed then quality could be affected in terms of nutritional and organoleptic properties.

To optimise the thermal processing of foods it is important to be able to measure or predict how the temperature changes throughout the food during the process. Discovering which point inside the food will be the coldest and monitoring the temperature at this point is the best way to ensure that the product is fully sterilised.

The use of mathematical models can be beneficial in this area. By determining the thermophysical properties of a food product, accurate predictions can be made of the temperature evolution within a food during heating, as shown by McKenna and Tucker (1991) as well as Tucker, *et al* (1996). Simulation techniques have the distinct advantage that once the model is in place it is very simple to investigate the effects that variations in parameters have on lethality. In the past, this has tended to be done by selecting a parameter which is thought to be particularly influential and running the simulation with a variety of values within a given range. This is a useful practice, but does not fully uncover the random variation present in any real process.

In this work statistical techniques are used to evaluate the effects of random variations inherent in sterilisation processes. These include the random variation of initial temperature of food products, variation in thermophysical properties and also retort temperature which fluctuates throughout the process. This approach can greatly help in making informed decisions about the likelihood of a product not receiving an adequate heat process.

2. MODELLING CRITERIA AND METHODS

2.1 Model Equations

For many food types, the rate of change of temperature (T) with time (t) is equal to the supply by heat conduction. The temperature within a cylindrical can will vary along its radius (r) and its height (x). The equation for heat conduction governing temperature inside a cylindrical can is given by

$$\frac{\partial T}{\partial t} = \alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r}\right)$$

for α the thermal diffusivity of the product. The thermal diffusivity is product dependent and will determine how quickly the product is heated. The temperature within the product is assumed to be constant at the beginning of heating. Cans are made from a highly conductive material. This means that the temperature at the very edge of the product is approximately equal to the retort temperature.

In the majority of cases, the thermal diffusivity will not be available due to the complexities of a multi-component product. When this is the case, an apparent thermal diffusivity is derived from the product heating rate or f_h value. For a finite cylindrical container with radius R and half-height H, the formula given by Ball and Olson (1957) is

$$\alpha = \frac{0.389}{f_h\left(\frac{1}{R^2} + \frac{1.708}{4H^2}\right)}$$

This formula is derived from the first term approximation to the analytical solution to the heat conduction equation for a finite cylinder.

2.2 Monte Carlo Method

The definition of a Monte Carlo method is a heuristic mathematical technique for evaluation or estimation of intractable problems by probabilistic simulation and sampling. This essentially means a simulation which uses random numbers. The product initial temperature and f_h value are picked from normal distributions at the start of the simulation and then kept constant for the duration of the process. The values are selected using NAG FORTRAN library routines. The distributions could equally have been generated using methods outlined by Law and Kelton (1991) if these routines had not been available.

2.3 Time Series Method

The definition of Time Series is a sequence of data indexed by time, often comprising uniformly spaced observations (Mendenhall and Sincich, 1989). Here a Time Series simulation is used for the retort temperature, T_R. This is done by splitting the retort temperature into a general trend, a cyclic variation caused by a temperature controller device, and a residual component (Varga, 1995). The retort temperature is

$$T_R = \text{General Trend} + \text{Cyclic Variation} + \text{Residual}$$

For this work it has been assumed that there is no variation caused by a temperature controller device, i.e. there is no cyclic term. The residual component is chosen in such a way that it is not completely random. This is because if the residual at time t (R_t) is positive then it is more likely that the residual at time t+1 (R_{t+1}) will also be positive rather than negative. In other words the residual is unlikely to jump from being large and positive to being large and negative, it is far more likely to go from being large and positive to a larger or smaller positive value. This does not mean that if the residual is large and positive at one point in time it will be for all points in time, merely that there will be no large jumps. For the residual, the definition used was

$$R_{t+1} = \phi R_t + \varepsilon\left(0, \sigma^2\right)$$

Here ε is a normally distributed random error component. ϕ is taken in the range $0 \le \phi \le 1$ and is a measure of how dependent the retort temperature at a particular time is on the retort temperature at the previous time step. The value of ϕ used in a simulation will depend on the size of time step used. For a small time step size it is expected that the retort temperature will be very dependent on that at the previous time step. For a longer time step the relationship between retort temperatures at consecutive time steps may appear more random.

In this work, the trend for the retort temperature will involve a come up time and a holding period. The temperature will thus comprise of two straight lines, an example of which is illustrated in Figure 1, where no random variation occurs. During the come up time the temperature will rise linearly from an initial value to the holding temperature. The temperature will then remain constant until the end of the process.

It is not possible to find an analytical solution to the above set of equations, so instead a numerical solution procedure has been used. Stochastic equations of this kind have been investigated by Nicolai and De Baerdemaeker (1992) using a finite element technique. As an alternative to this, a finite difference scheme was used. The governing equations have been solved numerically using an Alternating Direction Implicit (ADI) finite difference scheme (Fletcher, 1991).

3. RESULTS AND DISCUSSION

In this section, several examples are given which isolate the effect of random variation in each of the various parameters. Also in this section, ways of reducing the likelihood of not achieving a required sterilisation value whilst minimising the chance of over processing are investigated.

The model can be used to predict the temperature at any point within the cylindrical container. It is the coldest point within the container which is of interest and this is assumed to be the centre point, as is usual during conduction heating. In fact, it is not the actual temperature that is of interest, but the sterilisation value (F) at this point, as described by Lopez (1987). This value is equivalent to the number of minutes required to destroy a specified number of spores at a reference temperature for a given z value. The z value denotes the temperature change required to effect a tenfold change in time to achieve the same lethal effect. The reference temperature (T_{ref}) is generally taken to be 121.1°C and a z value of 10°C is used for the heat resistant spores of *Clostridium botulinum;* these are the values used here to give F_0.

The lethal rate (LR) can then be calculated for each minute as

$$LR = 10^{(T-T_{ref})/z},$$

the F value for the whole process is then the sum of the lethal rates effective in each minute of the process. Alternatively the F value can be calculated as

$$F = \int_0^t 10^{(T(y)-T_{ref})/z}\,dy,$$

from the evaluated time/temperature history.

In this first example, no random variation is allowed for; this is the control that all cases will be compared with. The retort profile used is shown in Figure 1, along with the centre temperature of the product. The retort temperature is split into an initial come up period of five minutes, followed by a holding period of eighty five minutes. The product has an initial temperature of 20°C and an f_h of 50 minutes. The dimensions chosen for the cylinder are those of a UT can (73 mm diameter × 115 mm height). Using the process outlined, the resultant F_0 value is 6.6 minutes.

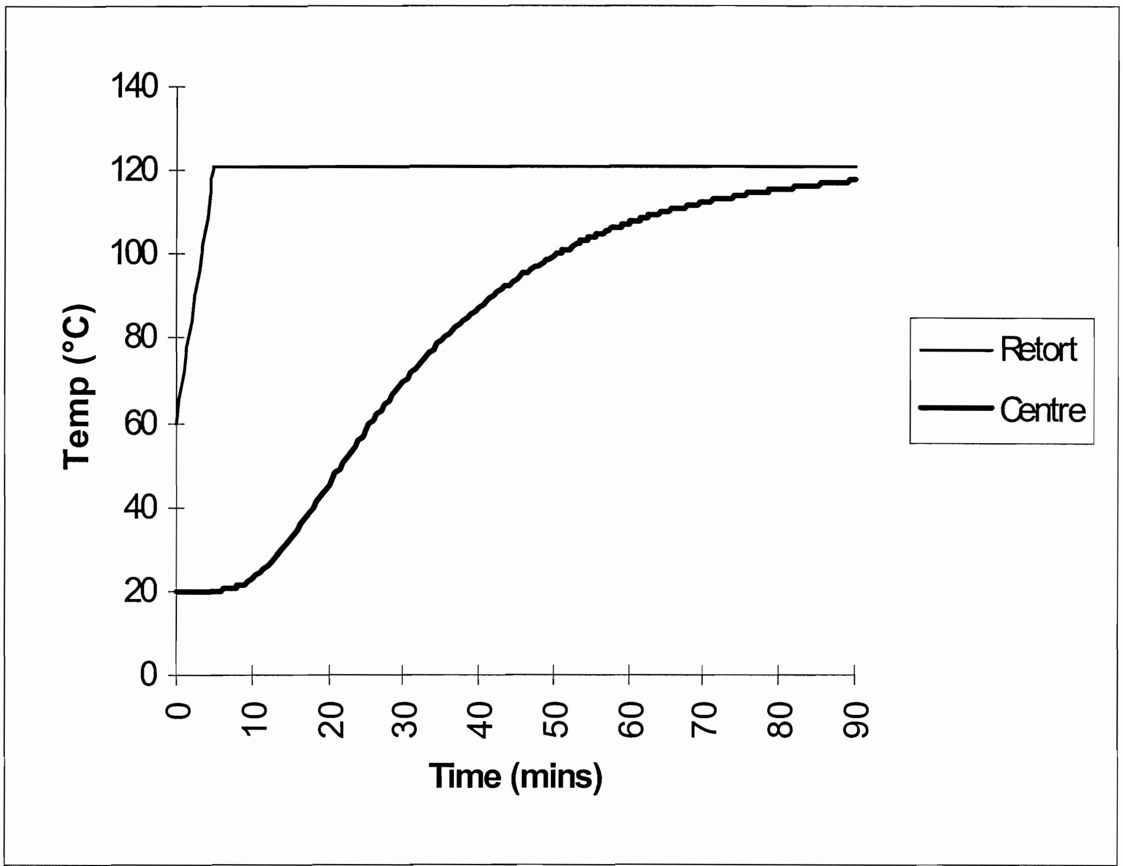

Figure 1: Time/temperature profile for retort and product centre.

Varying each of the parameters in turn, the F_0 values found can be compared with that in the control example. Firstly, the only random variation considered is in the f_h value. f_h has been randomly picked from a normal distribution with mean 50 minutes and standard deviation 3.5 minutes. The 100 f_h values selected from this distribution actually have mean 50.5 minutes and standard deviation 3.5 minutes. The frequency distribution of the f_h values is shown in Figure 2. The corresponding F_0 values have mean 6.6 and standard deviation 2.2 minutes and the frequency distribution of F_0 values is shown in Figure 3.

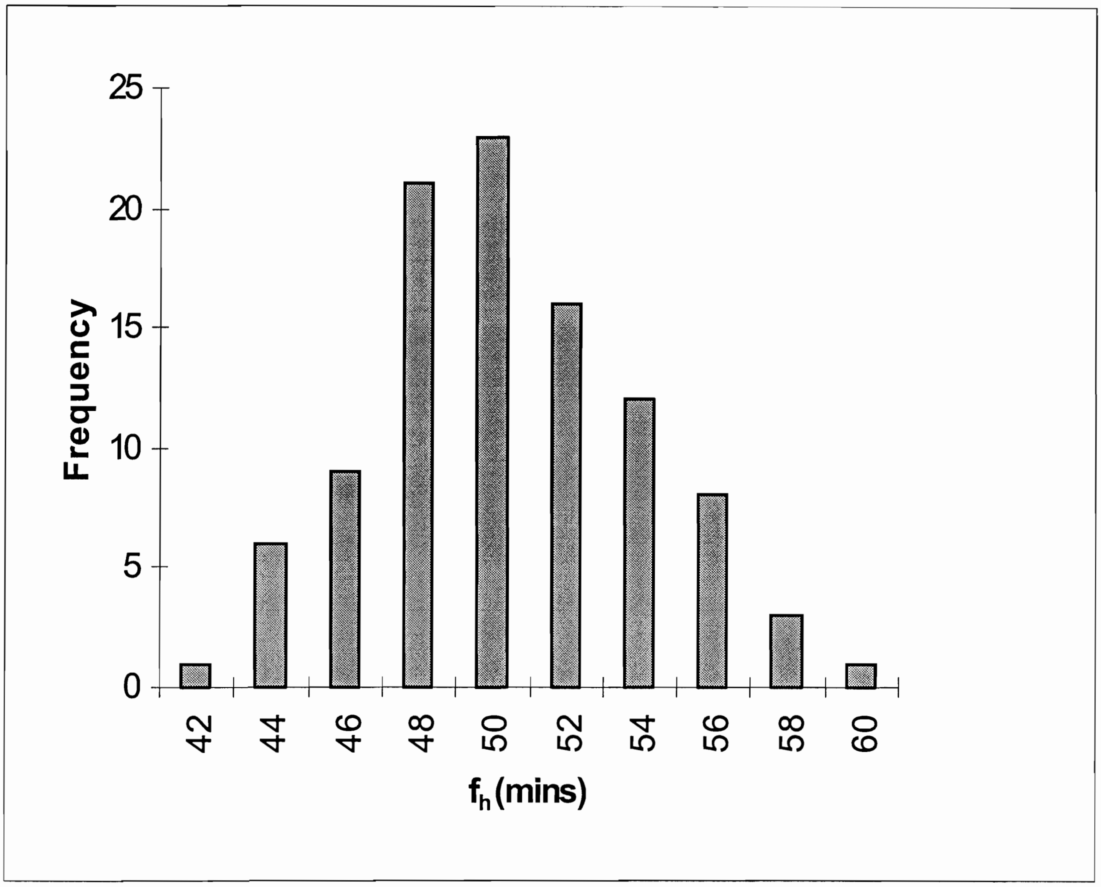

Figure 2: Frequency distribution showing a sample of f_h values, randomly picked from a normal distribution with mean 50 minutes and standard deviation 3.5 minutes. The sample mean and standard deviation are 50.5 and 3.5 minutes.

As f_h values are randomly picked from a normal distribution, they themselves will be approximately, but not exactly, normal as can be seen in Figure 2. Any deviation from normality influences the shape of the resultant F_0 distribution. However, by taking a large sample, the distribution of values is a good approximation to a normal distribution.

The corresponding F_0 distribution shown in Figure 3 does not have the distinctive shape of a normal distribution as it is skewed to the right. This is because with such an amount of variation in f_h, in order for the F_0 distribution to be normal it would need to be possible for F_0 to take a negative value, which is not possible. This means that although the distribution looks approximately normal to the right of the mean, this is not the case to the left of the mean.

It seems likely that if thermal properties include normal random variation, then centre temperatures will also be normally distributed and likewise F_0 values. Figure 3 shows this not to be the case for F_0 values. However, if the centre temperatures are indeed normally distributed, it seems probable that the F_0 distribution can be transformed into one which is normal. The definition of F_0 value led to the investigation as to whether $\log_{10} F_0$ is normally distributed. The corresponding frequencies are plotted in Figure 4. The distribution does seem to have the characteristic normal shape.

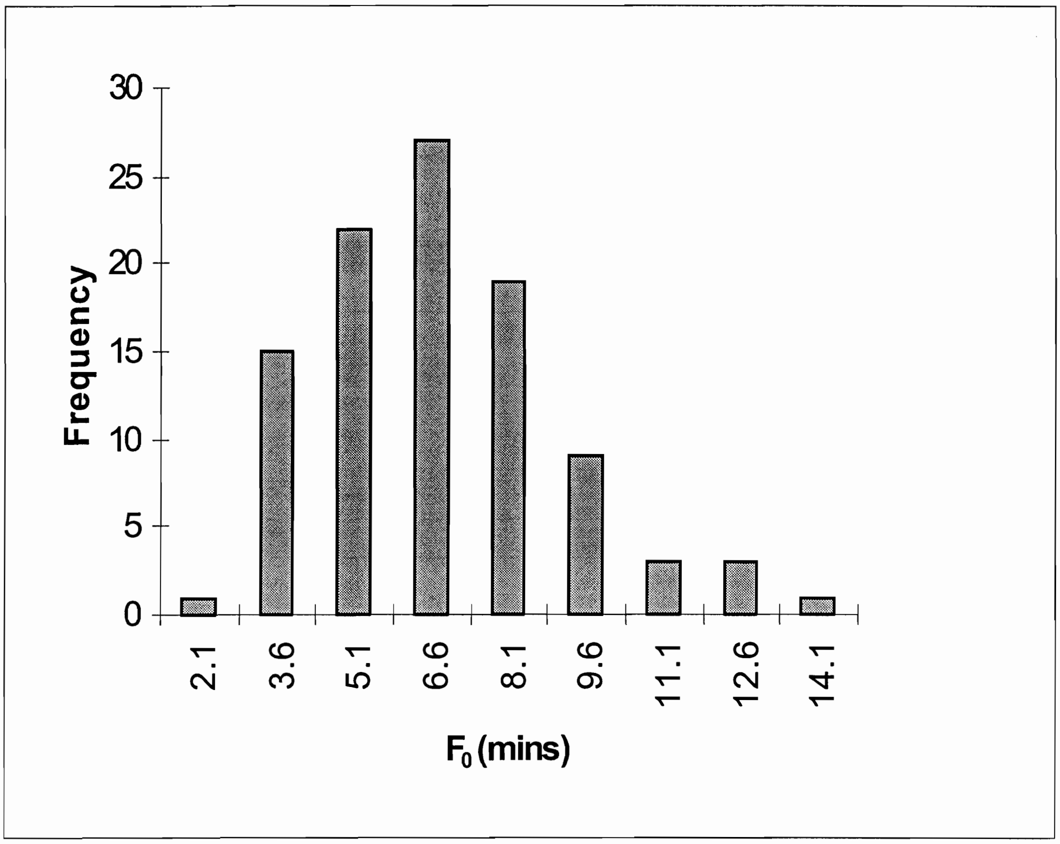

Figure 3: The effect of variation in f_h value on F_0. The f_h values used were sampled from a normal distribution with mean 50 and standard deviation 3.5 minutes. The frequency distribution shows F_0 values with mean and standard deviation 6.6 and 2.2 minutes.

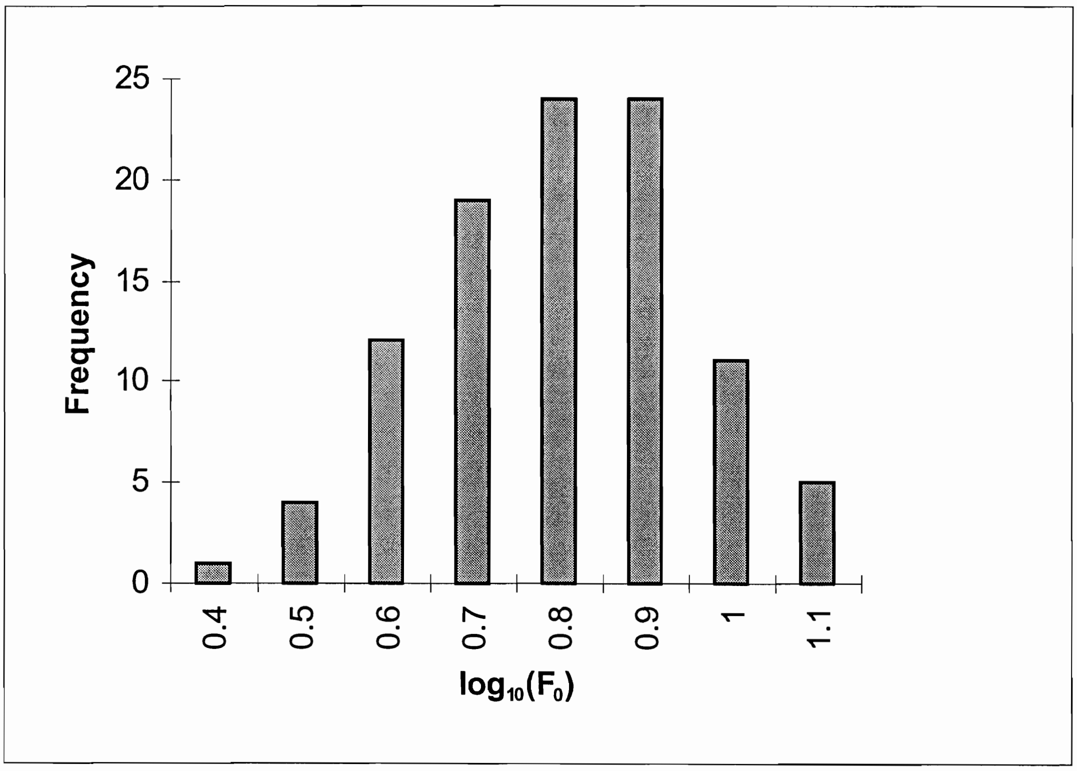

Figure 4: Variation in f_h value gives rise to variation in F_0. By taking the $\log_{10}$ of F_0 values a distribution with normal characteristics is found.

To investigate this further, expected frequencies can be compared with observed frequencies, as shown in Table 1. The expected frequencies are found from statistical tables using the sample mean and standard deviation. This distribution fits very well, and can now be used to assess the likelihood of not achieving a target F_0. For example, if we would like to know the likelihood that F_0 is less than 6 minutes, this can be found from statistical tables to be approximately 43.6%. Similarly there is a 1.6% likelihood of obtaining an F_0 value below 3 minutes. By looking once again at Figure 3 it can be seen that these figures are born out by the sample, so this distribution seems a very good approximation.

F_0 (mins)	$\log_{10}F_0$	Observed Frequency	Expected Frequency
up to 2.8	up to 0.45	1	1
2.8-3.5	0.45-0.55	4	4
3.5-4.5	0.55-0.65	12	11
4.5-5.6	0.65-0.75	19	21
5.6-7.1	0.75-0.85	24	26
7.1-8.9	0.85-0.95	24	21
8.9-11.2	0.95-1.05	11	11
11.2-14.1	1.05-1.15	5	4
14.1-17.8	1.15-1.25	0	1

Table 1: Comparison of expected and observed frequencies for F_0 values resulting from f_h variation. Expected frequencies are calculated by assuming a $\log_{10}$-normal distribution.

As speculated already, the skewness in the F_0 distribution seemed inevitable due to the proximity of the mean value to zero, as well as the amount of variation. This led to the investigation of what happens if the process time is extended so that the mean F_0 value will be higher. Lengthening the holding time by 10 minutes, the frequency distribution of F_0 values is given in Figure 6 for the f_h distribution shown in Figure 5. Figure 6 demonstrates that the distribution of F_0 values is indeed less skewed, being far closer in shape to a normal distribution. However, by looking also at the distribution of $\log_{10}F_0$ values in Figure 7 it can once again be seen that the distribution is close to the classic bell shape, although this time the distribution is slightly skewed to the left.

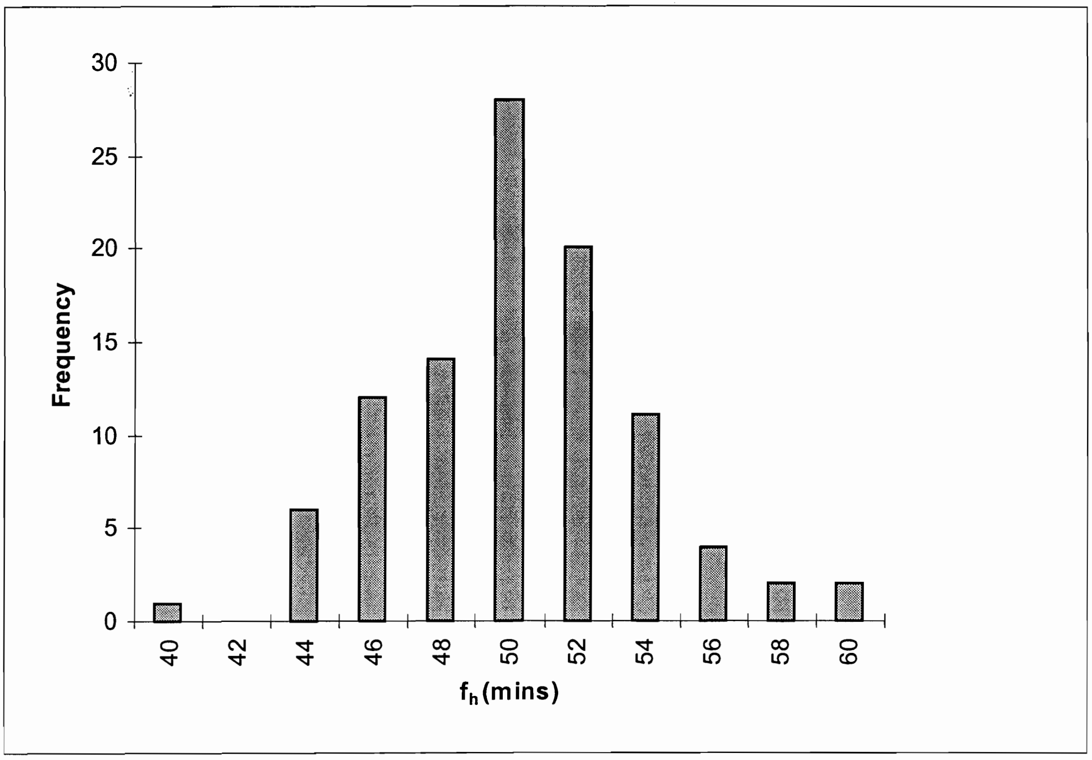

Figure 5: Frequency distribution of f_h values. These are sampled from a normal distribution with mean 50 minutes and standard deviation 3.5 minutes. This sample has mean 50.2 and standard deviation 3.6 minutes.

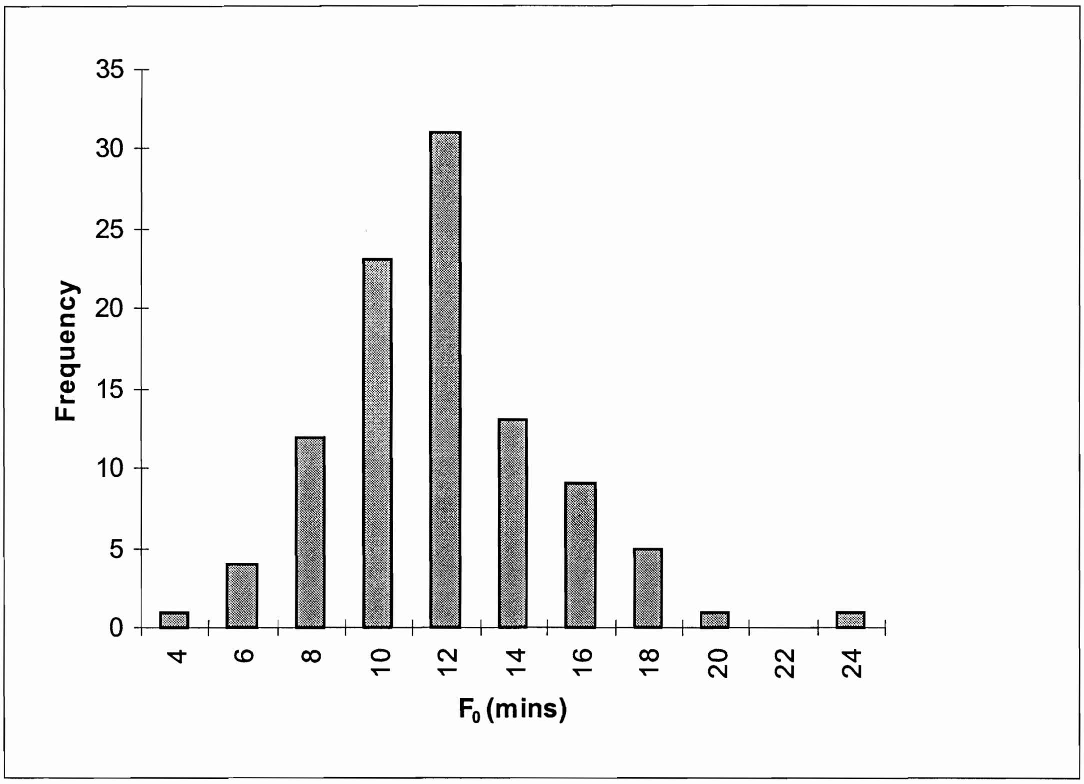

Figure 6: By lengthening the process time the mean F_0 value can be raised. For a mean F_0 which is not close to zero, the distribution is close to normal. F_0 values with mean 12.0 and standard deviation 3.3 minutes result from f_h variation.

These results seem to indicate that for a relatively low F_0, particularly with a reasonable amount of variation, by taking $\log_{10}$ of F_0, a distribution which is approximately normal can be found. As the mean value of F_0 increases, the F_0 distribution becomes increasingly less skewed and closer itself to a normal distribution. Consequently, the $\log_{10} F_0$ distribution becomes less normal as its skewness increases. This seems to indicate that, depending on the range of F_0 values, either the normal or the $\log_{10}$ -normal distribution will give the best approximation.

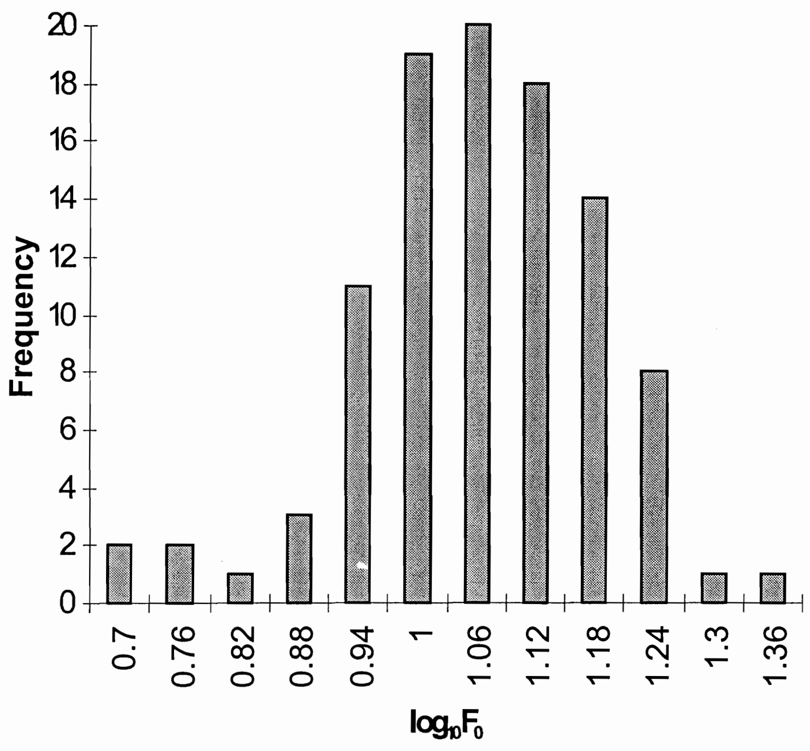

Figure 7: When F_0 values are approximately normal, taking the $\log_{10}$ of F_0 gives a skewed distribution.

The next step was to investigate the effect of a randomly varying initial temperature. In this case 100 values of initial temperature have been picked from a normal distribution with mean 20°C and standard deviation 2.5°C. The sample of size 100 has mean 20.3°C and standard deviation 2.3°C and is illustrated as a frequency distribution in Figure 8. This range of initial temperatures leads to the range of F_0 values displayed in the frequency distribution of Figure 9. The F_0 values have mean 6.7 minutes and standard deviation 0.2 minutes.

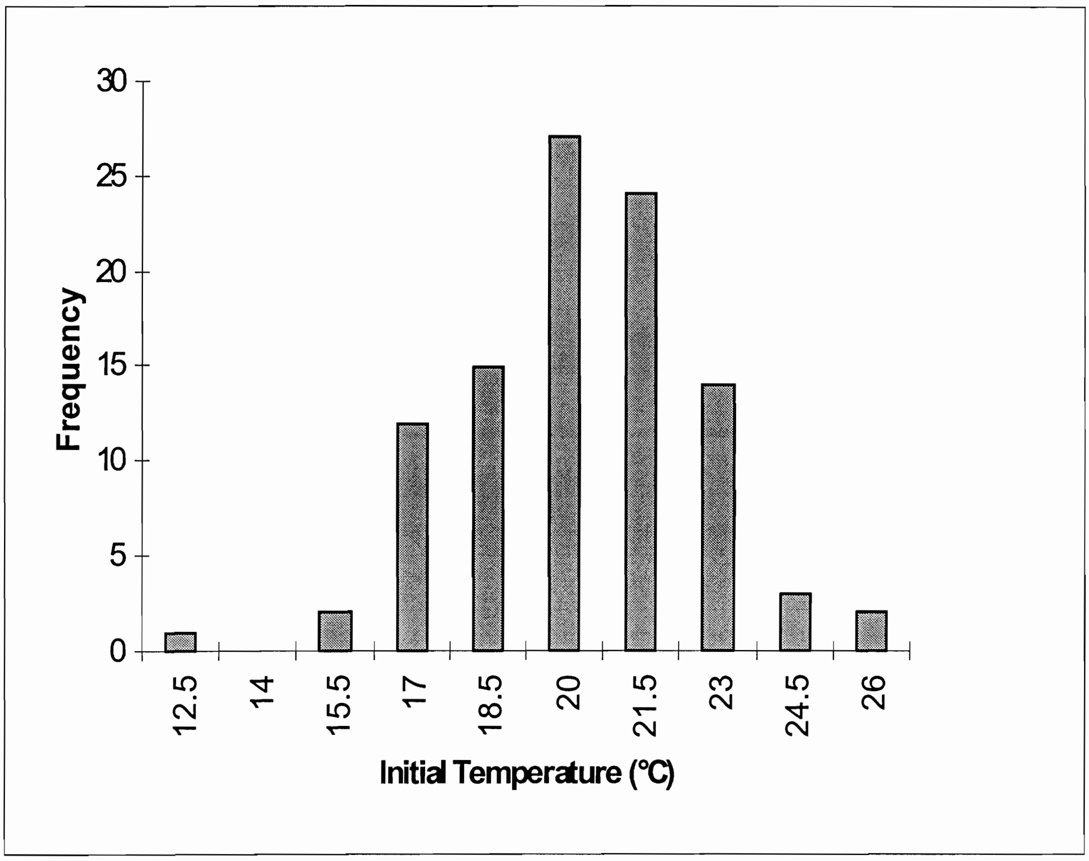

Figure 8: Variation in initial temperature is simulated by sampling from a normal distribution with mean and standard deviation 20°C and 2.5°C. This sample has mean 20.3°C and standard deviation 2.3°C.

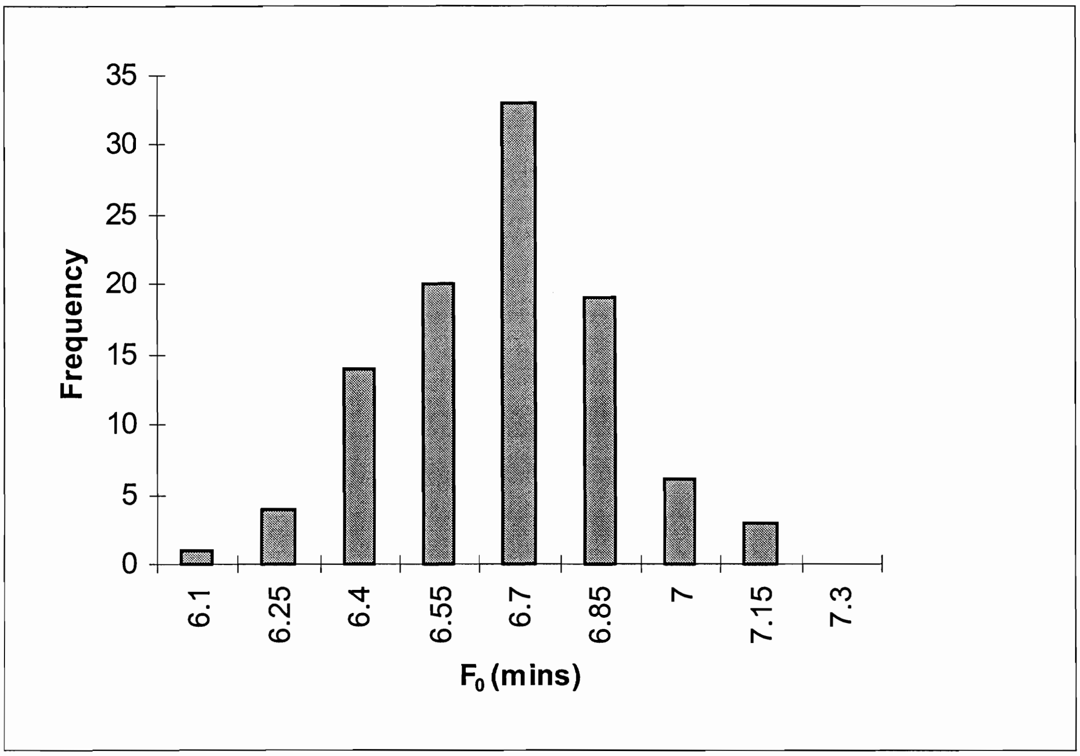

Figure 9: Variation in initial temperature is less influential than f_h variation. A mean F_0 of 6.7 with standard deviation 0.2 minutes results from initial temperatures sampled from a normal distribution with mean 20°C.

Figure 9 demonstrates that the initial temperature has far less effect on F_0 than f_h value did. Due to this, for this amount of variation in f_h, the F_0 distribution is clearly not affected by the proximity of zero to the extent demonstrated in the previous cases.

In this case, the distribution does not look visibly skewed, and looks close to normal. This seems to follow the trend of previous results, as with this small amount of variation the closeness to zero is not a factor. Table 2 shows that the mean initial temperature gives an F_0 value of 6.6 minutes. An initial temperature of the mean minus one standard deviation gives an F_0 of 0.2 minutes below the likely mean F_0 value, so it seems reasonable to take 0.2 minutes as the standard deviation of the F_0 population, assuming the population is approximately normal. The table shows that this value works on all levels except the mean plus one standard deviation where the initial temperature of 22.5°C gives an F_0 of 6.9 minutes rather than the expected 6.8 minutes, although these values are very close.

	Initial Temperature (°C)	**F_0 (mins)**
mean minus two standard deviations	15.0	6.2
mean minus one standard deviation	17.5	6.4
mean	20.0	6.6
mean plus one standard deviation	22.5	6.9
mean plus two standard deviation	25.0	7.1

Table 2: F_0 values corresponding to the mean and one and two standard deviations from the mean initial temperature.

To demonstrate the effect that a varying retort temperature profile can have, $R_{t+1}=0.5R_t + \varepsilon(0,0.25)$ has been used to generate retort temperature profiles, an example of which is shown in Figure 10. Here R_t and R_{t+1} are the residual components of the retort temperature at times t and t+1, which add random variation to the trend. $\phi=0.5$ has been chosen fairly arbitrarily whilst bearing in mind that the time step size relates to 30 seconds. This value of ϕ means that the retort temperature has some reliance on the previous time steps temperature, but not as much as for $\phi=1$. This seems reasonable for the length of time step selected, giving a variation of approximately plus or minus one or two degrees. To reduce the variation, the variance of 0.25 can be reduced in the ε normal distribution. A frequency distribution of F_0 values for 100 different retort profiles generated in this way is illustrated in Figure 11. The F_0 values have mean 6.6 minutes and standard deviation 0.1 minutes.

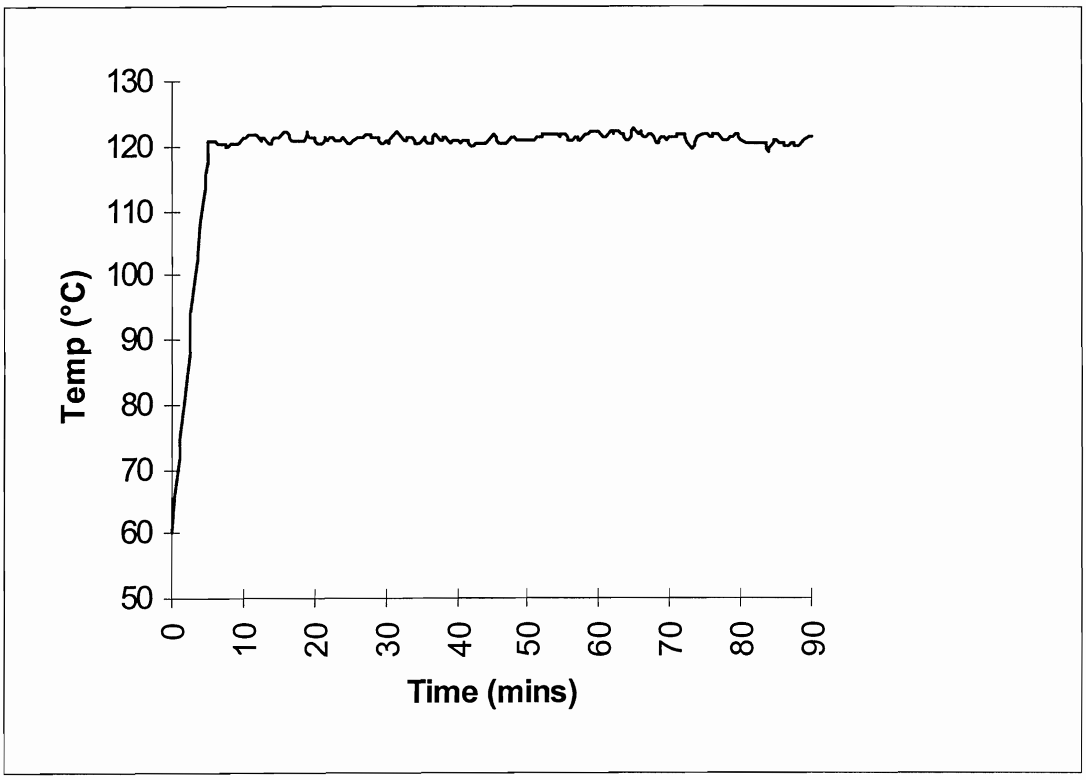

Figure 10: Retort temperature profile including random variation.

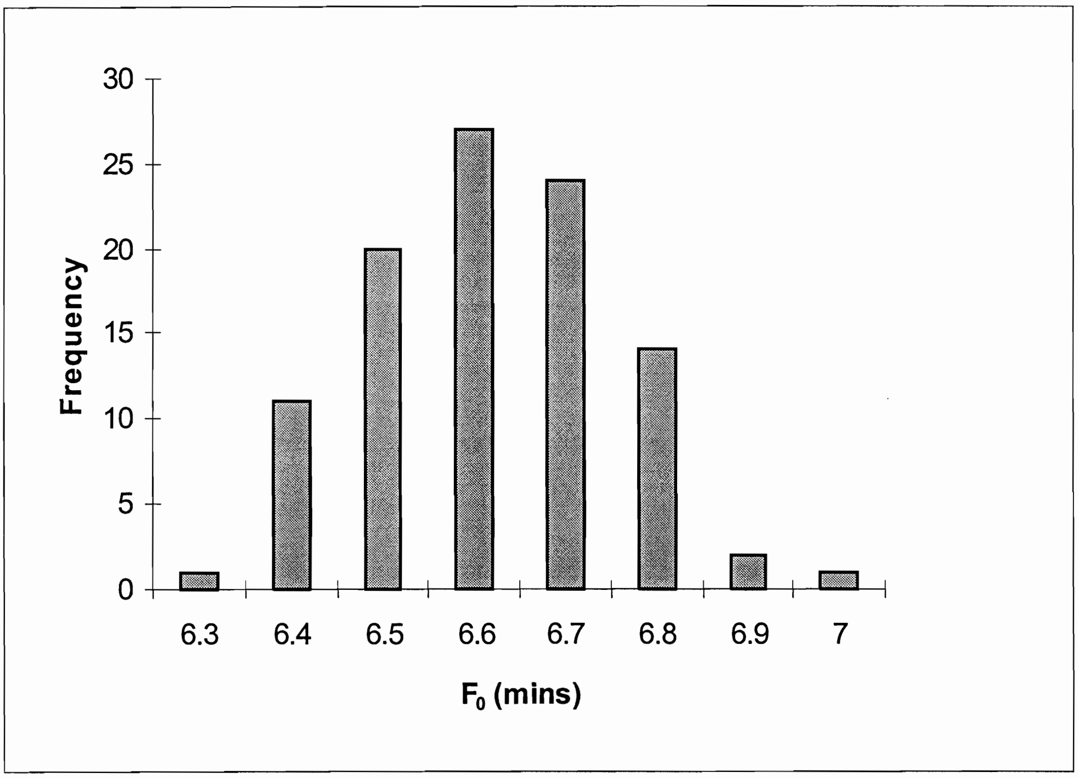

Figure 11: A retort temperature profile which includes random variation results in an F_0 distribution which is close to normal. F_0 values have mean 6.6 and standard deviation 0.1 minutes.

Once again Figure 11 shows that the F_0 values show only a small amount of variability and so appear to be approximately normally distributed. Also the mean F_0 is 6.6 minutes, the F_0 value found in the control example.

Finally, all the previous work is brought together to look at what happens if the f_h value, the initial temperature and the retort temperature are varied simultaneously. Here the initial temperature includes normal random variation with mean 20°C and standard deviation 2.5°C; f_h is normally distributed with mean 50 minutes and standard deviation 1.5 minutes. The frequency distribution of Figure 12 shows f_h values, which have mean and standard deviation 50.3 and 1.4 minutes. The actual mean and standard deviation for initial temperature are 19.8°C and 2.6°C, the values shown on the frequency distribution of Figure 13. The retort temperature is generated in the same way as for the previous example. This gives the range of F_0 values shown in the frequency distribution in Figure 14, with mean 6.5 and standard deviation 0.8.

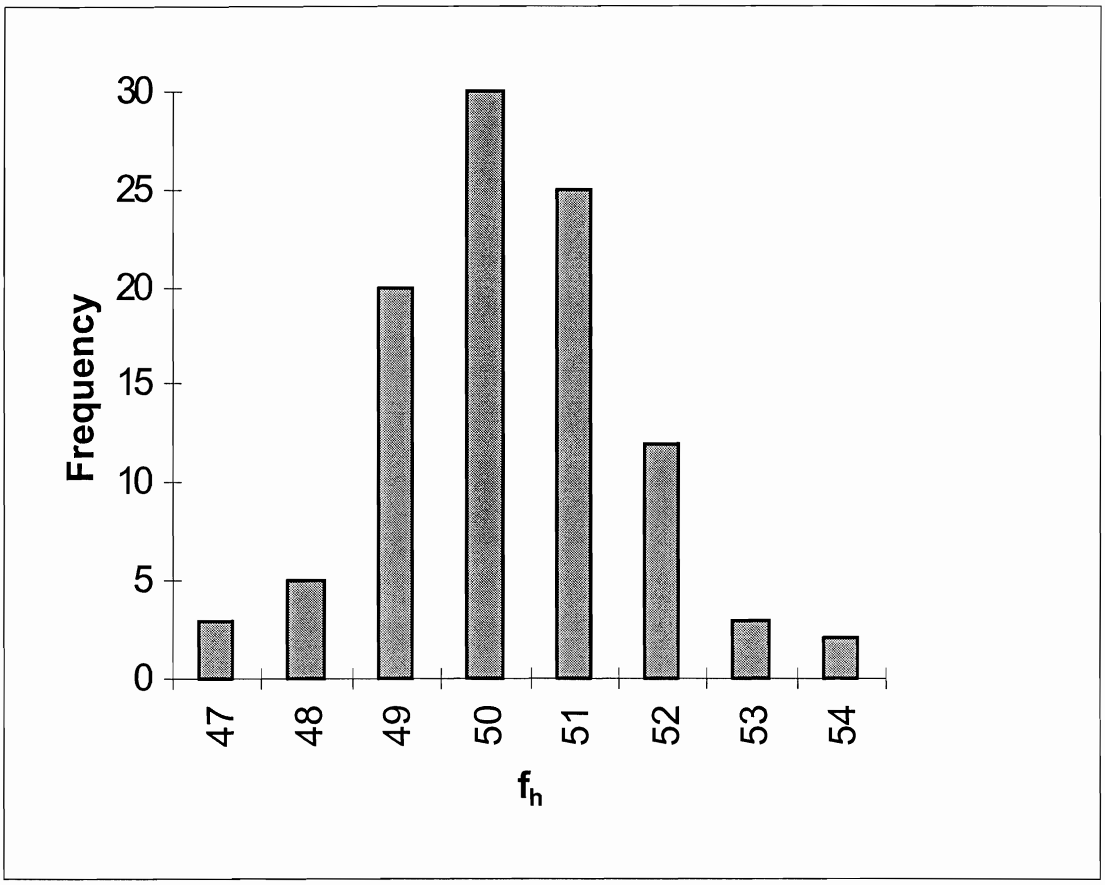

Figure 12: Randomly varying f_h can be used alongside randomly varying initial temperature and retort temperature to investigate the combined effect on F_0. This sample of f_h values has mean 50.3 and standard deviation 1.4 minutes.

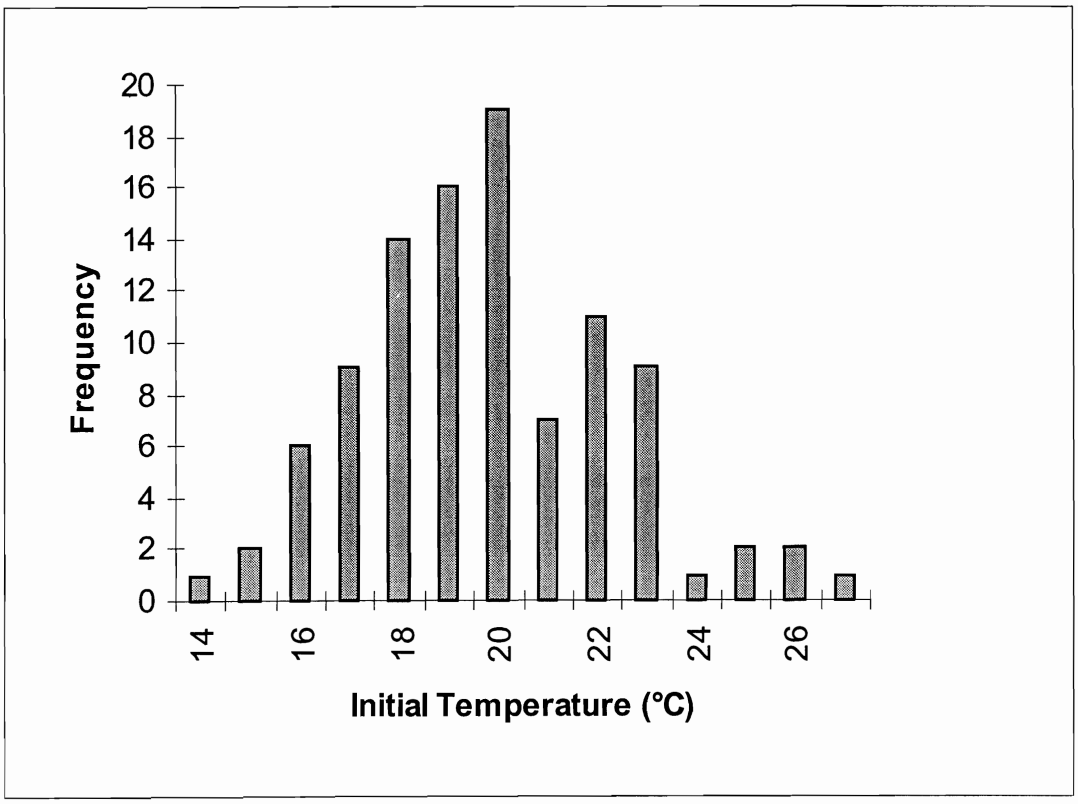

Figure 13: A sample of initial temperature with mean 19.8°C and standard deviation 2.6°C.

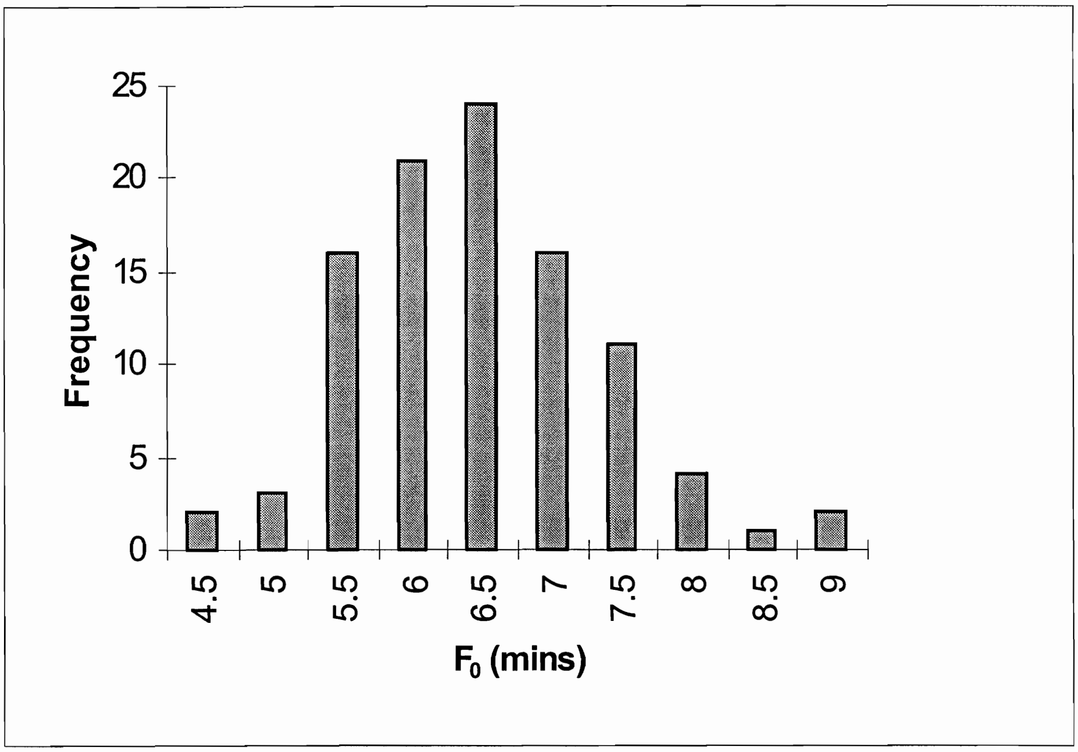

Figure 14: A combination of variation in initial temperature, f_h and retort temperature gives F_0 values which are approximately normal. The F_0 values have mean 6.5 and standard deviation 0.8 minutes.

As could be expected from the previous results, the F_0 distribution appears to be approximately normal, as the variation in all parameters is relatively small. Once again it is speculated that these values could quite adequately be approximated by a normal distribution with mean 6.6 minutes.

The cases that have been looked at here have clearly shown that variation in heating factor, initial temperature and retort temperature all lead to variation in sterilisation values. Heating factor is clearly the most significant factor showing the importance of consistency of thermal properties between products. Variation could be caused by variation in the proportion of product components for example.

4. CONCLUSIONS

This work has shown just how important random variation can be when predicting sterilisation values. Any variation in initial temperature, retort temperature and most markedly heating factor can lead to insufficient sterilisation. It is usual to take a worst case scenario approach to prediction of sterilisation values, and statistical methods such as those presented here are an extremely useful alternative to this. By using this type of technique it has been shown that minimising variation is all important, but also that if this is not possible, predictions can be made as to what percentage of products will be sterilised to an adequate level. It has been shown that if the heating factor is normally distributed with mean 50 and standard deviation 3.5 minutes then the associated F_0 appears to be approximately $\log_{10}$- normally distributed. In this case the likelihood that the F_0 is less than 6 minutes can be found from statistical tables to be approximately 43.6%, whilst an F_0 of less than 3 minutes is expected in 1.6% of cases. This type of information could be very helpful for use in process validaticn.

The results show that to reduce the likelihood of not achieving a required F_0 value either the average heating factor or initial temperature can be increased, the process time extended, or the variation in these parameters reduced. It has been shown that if the mean values are merely increased, the F_0 distribution will simply be shifted and become less skewed, so that by reducing the number of low values the amount of high values is also increased as is the average value, as it is if the process time is simply increased. This may contribute to a greater loss of product quality. If possible, it is better to reduce variability, particularly in heating factor.

REFERENCES

Ball, C. O. and Olson, F. C. W. (1957). Sterilization in food technology. Theory, practice and calculation. Published by McGraw-Hill, New York, USA.

Fletcher, C.A.J. (1991). Computational techniques for fluid dynamics, volume 1. pp 252-255, Published by Springer-Verlag.

Law, A. M. and Kelton, W. D. (1991). Simulation, modelling and analysis. Published by McGraw-Hill, New York, USA.

Lopez, A. (1987). A complete course in canning and related processes, book II - packaging, aseptic processing, ingredients. Published by The Canning Trade Inc., Baltimore, USA.

McKenna, A. B. and Tucker, G. S. (1991). Computer modelling for the control of particulate sterilization under dynamic flow conditions. Food Control, 224-233.

Mendenhall, W. and Sincich, T. (1989). A second course in business statistics: regression analysis. pp 490-545, Published by Collier MacMillan Publishers, London.

Nicolai, B. M. and De Baerdemaeker, J. (1992). Simulation of heat transfer in foods with stochastic initial and boundary conditions. Transactions of the IChemE, 70, 78-82.

Tucker, G. S., Noronha, J. F. and Heydon, C. J. (1996). Experimental validation of mathematical procedures for the evaluation of thermal processes and process deviations during the sterilization of canned foods. Transactions of the IChemE, 74, 140-148.

Varga, S. (1995). Application of statistical variability to in-pack thermal processing, Periodic Progress Report Part II, EU project AIR2-CT94-1017.