11 Mathematical modelling

Mathematical modelling is an important activity in many occupations and is a valuable skill sought by employers. For students, mathematical modelling provides an opportunity to progress from theoretical learning towards a realistic simulation of the activities of professional users of mathematics. Modelling, through its handling of ill-defined real world systems, allows exciting opportunities to develop the wider numeracy skills of problem solving, reflective discussion and team working.

Mathematical models extend across a broad spectrum from **geometrical** models, through **system** models to **numerical** models.

Geometric models provide realistic representations, often in three dimensions, of machinery, buildings or other structures. These models are generally produced with computer aided design software. They are intended to resolve design problems, and then communicate the final designs to the teams who will carry out the manufacturing or construction.

Geometric modelling may be extended to investigate the performance of the designed object. Examples might include

- the simulation of ventilation and heat flow through an office building,
- simulation of water flow around the piers of a bridge,
- analysis of temperatures within a working internal combustion engine.

Equations from appropriate branches of physics are combined with the three dimensional representations of the objects to produce the simulation results.

System models move away from accurate geometrical representations, and instead use diagrammatic displays of real world systems. Computer animations are often produced. Examples might include

- a simulation of traffic flow through a road junction,
- the simulation of the operation of a factory production line,
- the simulation of fluid flows during processing in an oil refinery.

The objective of this type of model is to investigate whether the proposed system will perform as required, and to identify any necessary modifications to system.

Numerical models focus on producing data, which might be output in the form of tables, graphs or maps. Examples are very wide ranging, from economic models in business planning, through engineering models for the operation of electronic circuits or simulating the flight paths of space vehicles, to environmental models such as weather forecasting or the prediction of sea level changes. Numerical models are valuable for exploring the possible outcomes of events, to allow adequate contingency planning. For example, health officials might wish to estimate the likely numbers of patients who will fall ill during an epidemic so that treatment facilities can be prepared.

Theatre design

Our first modelling example is a system model, produced by a student of computing and performing arts to investigate stage design for theatre productions. The model is based on the stage layout of Theatr Clwyd in North Wales.

The model was created using the programming language ***Processing 2*** (Greenberg, 2007). This is a graphics language which provides functions for the three-dimensional geometry operations discussed earlier in Chapter 7: Shape and space. Of particular use in constructing the theatre model are the commands:

beginShape(POINTS)
  vertex(X1, Y1, Z1)
  vertex(X2, Y2, Z2)   . . . . . . etc.
endShape()

which allow surface patches to be created in three-dimensional space using the X, Y, Z coordinate system. This allows the theatre stage, wings and seating to be created.

texture(image name) which allows a picture image to be added to a surface patch, for example to display a scenery backdrop.

rotateX(angle), **rotateY**(angle), **rotateZ**(angle) which allow a three dimensional model made up from surface patches to be rotated around one of the coordinate axes.

translate(X, Y, Z) allow the three dimensional model to be moved in one or more of the coordinate directions.

scale(ratio) allow the three dimensional model to be enlarged or reduced in size.

These three functions together allow the user to move, zoom or rotate the image of the theatre stage so that it can be viewed from different positions.

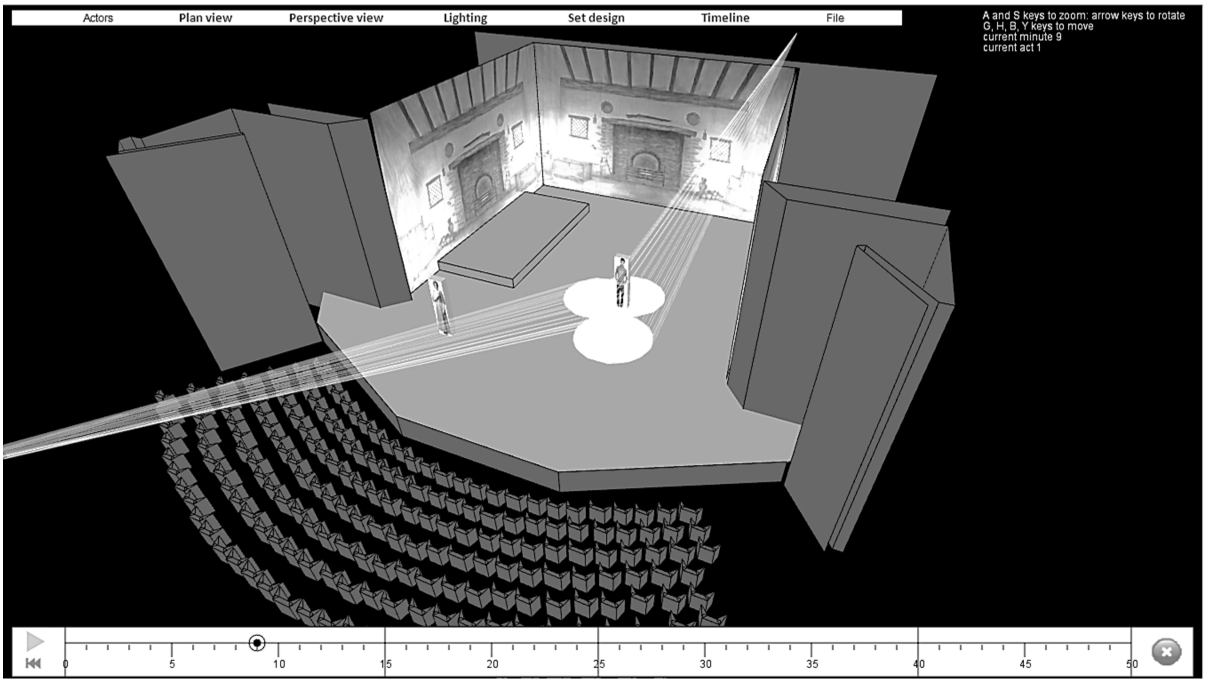


Figure 332: Computer model of the Theatr Clwyd stage

A principal component of the model is a ***timeline***, included at the bottom of the screen. A moving point indicates the current minute during the simulation of a performance. When setting up a model, the user specifies the length of each act, which is then indicated by a marker on the time line. During the simulation, the proscenium curtains will be closed at the end of each act and the stage layout and scenery may be changed.

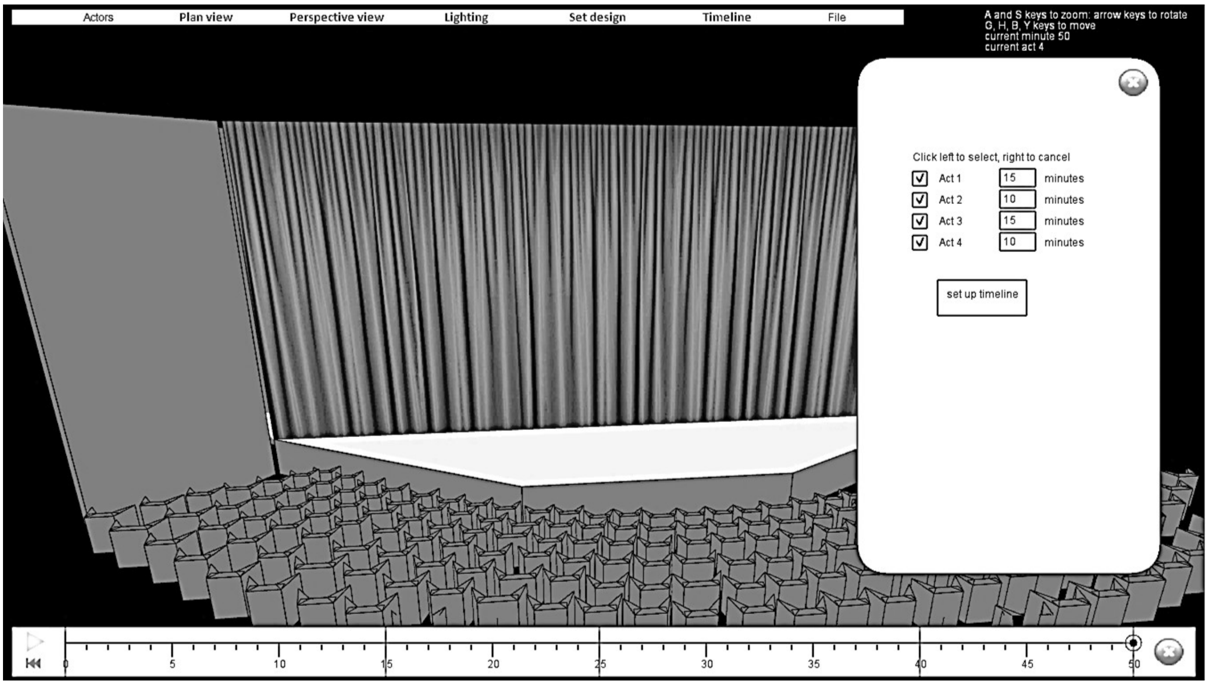


Figure 333: Setting up the performance timeline

In addition to a three dimensional view, the model provides a plan of the stage. This is used to enter the positions of scenery backdrops and any raised platforms required during an act.

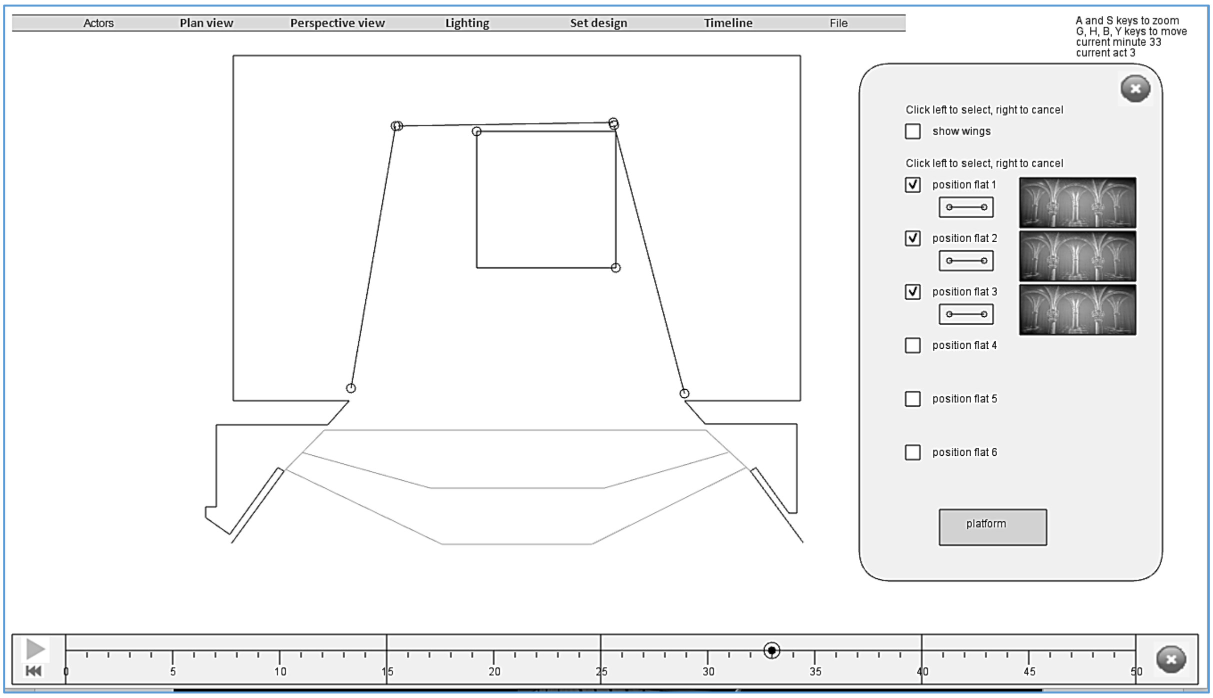


Figure 334: Adding scenery backdrops and stage platforms

Lighting can be specified, using a plan of the lighting gantries of the Theatre Clwyd stage. Directions of light beams can be adjusted by means of the computer mouse, and times that lighting is switched on and off can be linked to the time line.

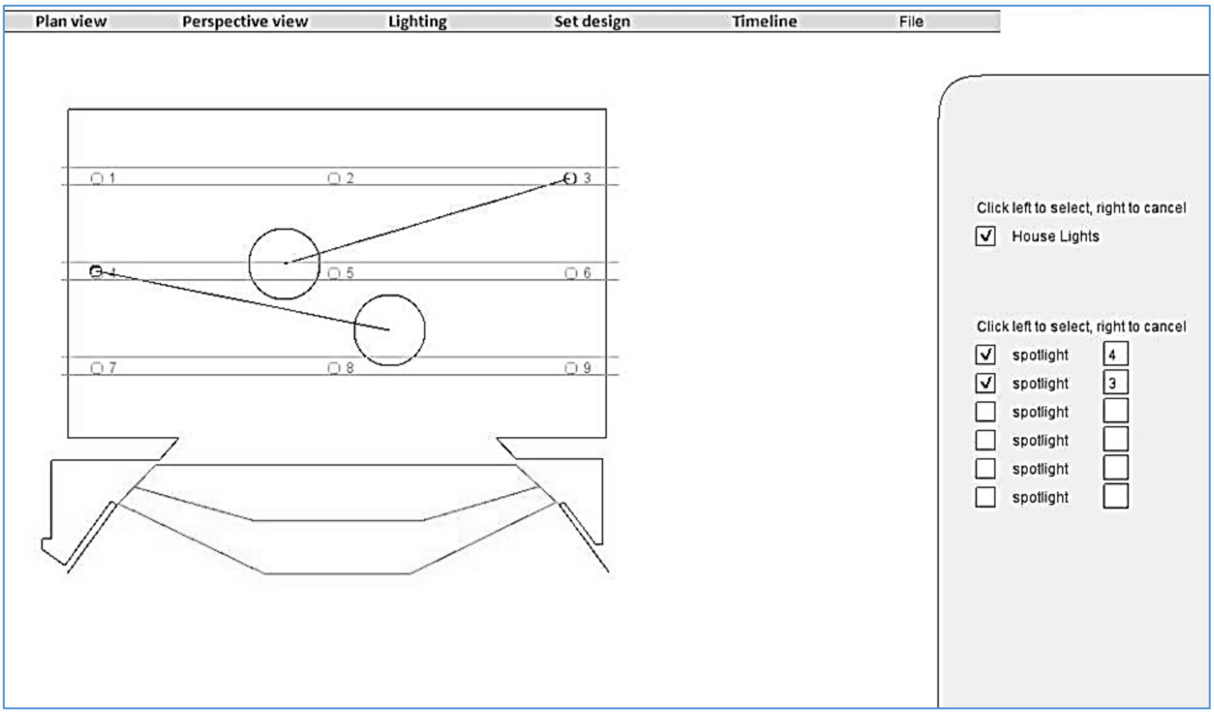


Figure 335: Setting the stage lighting

An important purpose of the model is to plan the movements of the actors during a performance. The times that actors make entrances, their positions as they move about the stage, and the times of their exits, can all be synchronised with the time line.

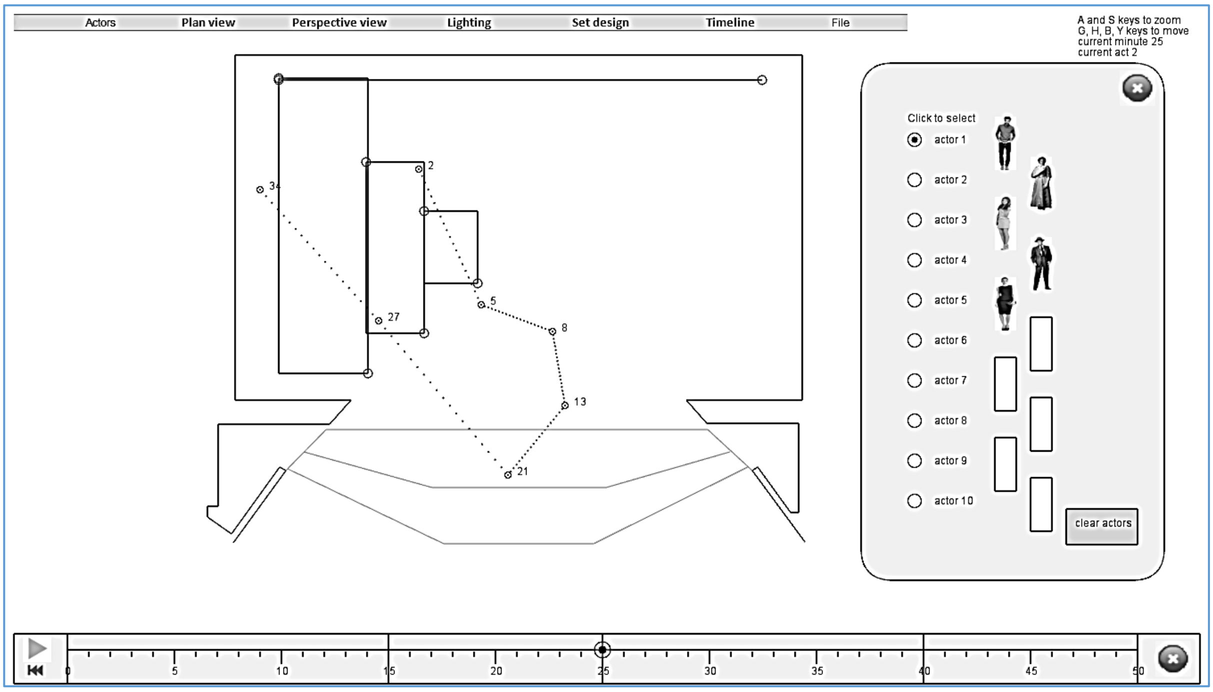


Figure 336: Setting the paths for actors

The model is not intended to be fully geometrically realistic, so actors are represented simply as flat images which can be moved to appropriate positions on stage as the performance is simulated.

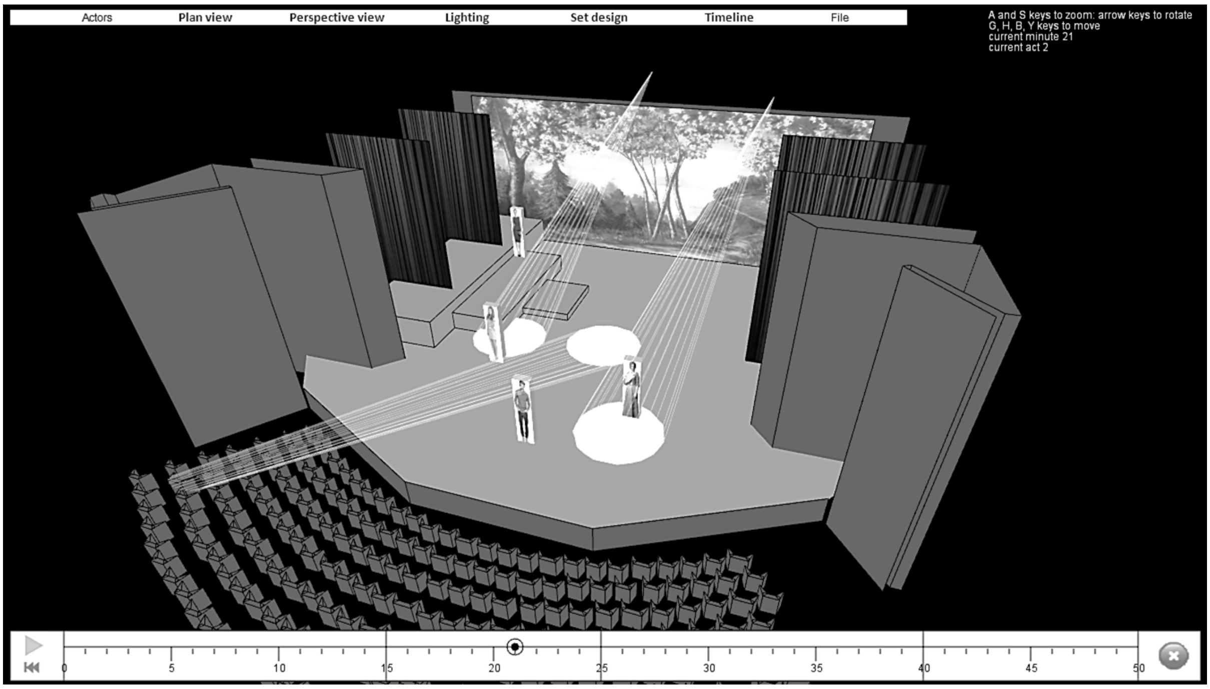


Figure 337: Actors positioned on the stage

An increasing number of graphical programming languages and software systems are now becoming available. These allow the easy construction of three-dimensional animations which can be used in effective and creative ways in a variety of vocational contexts.

In the next sections we present a range of numerical models from different vocational areas. A useful theoretical framework around which to plan and conduct mathematical modelling activities has been provided by Blum and Leiß (Keune and Henning, 2003):

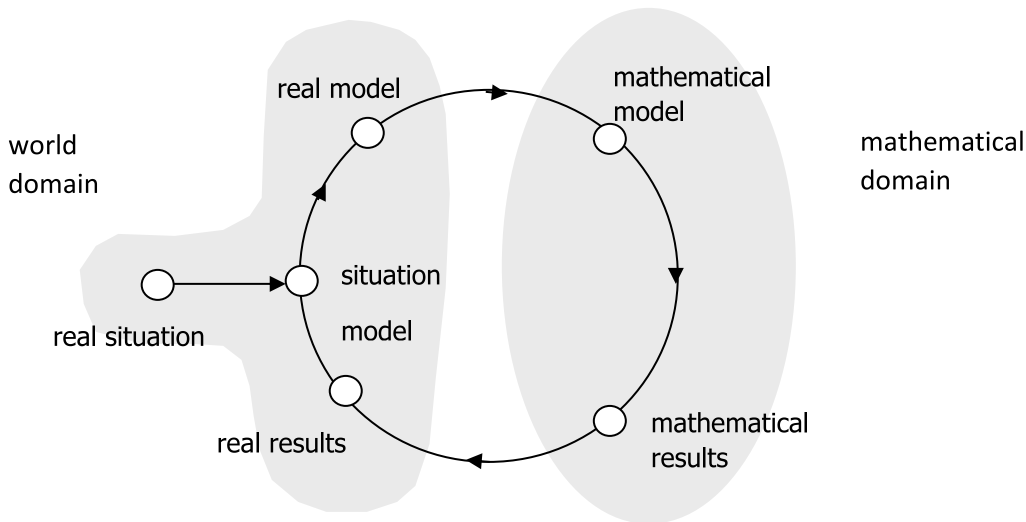


Figure 338: The modelling cycle of Blum and Leiß

Mathematical modelling is seen as a series of connected activities:

- We begin in the world domain with the **real situation**. This will almost certainly be very complex, and it is necessary to pick out only the components which are most important to the model. This creates the **situation model**. For example, we might be interested in modelling the flow of water along a stretch of river. The volume of water entering the section of river, the downstream slope, and the river cross sections at different points are likely to be important. However, in a first approximation we might ignore any additional surface water entering the section of river from fields on the valley sides, or leaving the river by drainage through the river bed. These effects are likely to be small, so will not have a significant effect of the results of the modelling.
- Once the main components of the situation model have been decided, the next step is to identify the ways in which these are related. This leads to the **real model**. At this stage, we specify processes descriptively, with no detailed mathematics involved. For the river flow example, we might decide that an increase in downstream slope is likely to cause faster water flow. A broad, shallow river channel is likely to reduce water flow due to frictional turbulence.
- We are now able to move from the world domain to the mathematical domain. The relationships identified earlier are now formulated mathematically as equations and constructed into a **mathematical model**. It is likely that this will involve computer programming, the use of a spreadsheet or other software package.
- The model is run, producing **mathematical results**. These may be output as a table of data, but will probably be easier to interpret if displayed as graphs or maps as appropriate.
- We now move back from the mathematical domain to the world domain to evaluate the modelling results. This might provide predictions which can be used for decision making.
- Over a period of time, the performance of the model can be assessed against real world events. If the model has failed to provide accurate predictions, it may be necessary to modify the factors which are included in the **situation model,** change the assumptions in the **real model**, or revise the equations in the **mathematical model**. The modelling cycle can be repeated until acceptable results are achieved.

We will now examine several numerical models:

Owls and mice population model

Students were introduced to a scenario in which a nature reserve was being created to conserve a population of Tawny Owls. The task for the group was to use mathematical modelling to advise on the management of the reserve in order to ensure the long term survival of the owl population.

Students worked in small groups to discuss and reach agreement on the factors to be included or excluded from the model. Three factors which might have a major effect on owl populations are:

- Weather and climate. However, these are beyond the control of the nature reserve staff, so these are not useful factors to include in a management plan.
- Human interference with nesting sites. We will assume that the nature reserve staff are already intending to protect the owl population by appropriate control of access by visitors.
- Food supply. The tawny owl population feed principally on mice. The nature reserve staff could influence the size of the mice population by careful ecological management, so this seems an important aspect of the system to model.

It was agreed that modelling would focus on food supply. The first stage is the construction of a **real model** to establish the relationships for the populations of owls and mice.

- We might suppose that the mice population would increase from year to year through breeding in the absence of predators. However, the population would be reduced if the mice are eaten by owls.
- The owl population could increase if there was a plentiful supply of mice, but would be likely to decline if food was limited.

These assumptions can be written as two word equations:

mice next year	=	increased population due to breeding	-	decrease due to owl predation

owls next year	=	fall in population if food supply is short	+	increase if mice available

A mathematical model can now be formulated. The symbols **M** and **T** were introduced to represent the average numbers of mice and tawny owls per hectare. A subscript is used to represent the year, with **n** being the current year, and **n+1** representing next year. Assuming a 10% increase in mouse population if there is no owl predation, the students were able to derive an expression:

$$\underbrace{M_{n+1}}_{\text{mice next year}} = \underbrace{1.1\,M_n}_{\text{increase due to breeding}} - \underbrace{k\,M_nT_n}_{\text{reduction due to predators}}$$

The number of mice eaten by owls was seen to depend on both the numbers of mice and the number of owls in the area.

- If there are a large number of mice, it is more likely that a mouse will be seen and caught.
- If there are a large number of owls searching for food, it is more likely that a mouse will be found.

A parameter **k** was necessary to represent the chances of a mouse being caught when an owl and mouse are both present in the same area. As a first guess, this was given a value of 0.05.

A similar relationship was derived for the number of tawny owls per hectare. We might assume a 20% fall in owl numbers over the year if no mice were available as a food supply. The owls might survive on other poorer food sources, or might migrate away from the nature reserve to a more favourable location.

$$T_{n+1} = \quad 0.8\,T_n \quad + \quad \mathbf{r}\;M_nT_n$$

owls next year

decrease if mice not found

increase if mice available

As previously, we assume that the numbers of mice caught will depend on both the mice and owl populations in the area. A value of 0.005 is suggested for the parameter **r**.

The mathematical problem could then be solved using a spreadsheet. The formulae for mice and owls are interlinked, allowing populations of each species to be calculated if the populations are known from the previous year. It is simply necessary to provide starting populations for year zero.

An example section of spreadsheet showing the formulae is:

	A	B	C	D
1		year	mice	owls
2		0	35	10
3		=B2+1	=1.1*C2-0.05*C2*D2	=0.8*D2+0.005*C2*D2

$$M_{n+1} = 1.1\,M_n - \; 0.05\,M_nT_n \qquad\qquad T_{n+1} = 0.8\,T_n + 0.005\,M_nT_n$$

Figure 339: Spreadsheet for the owls and mice population model

The formulae can then be copied downwards for the required number of years. A mathematical function where each successive value is calculated from the previous result is known as a **recurrence relation**.

When the calculations are completed, we return to the world domain of the modelling cycle to interpret the results in terms of the management of the nature reserve.

For starting numbers of 35 mice and 10 tawny owls per hectare in year zero, it is discovered that the populations are unstable.

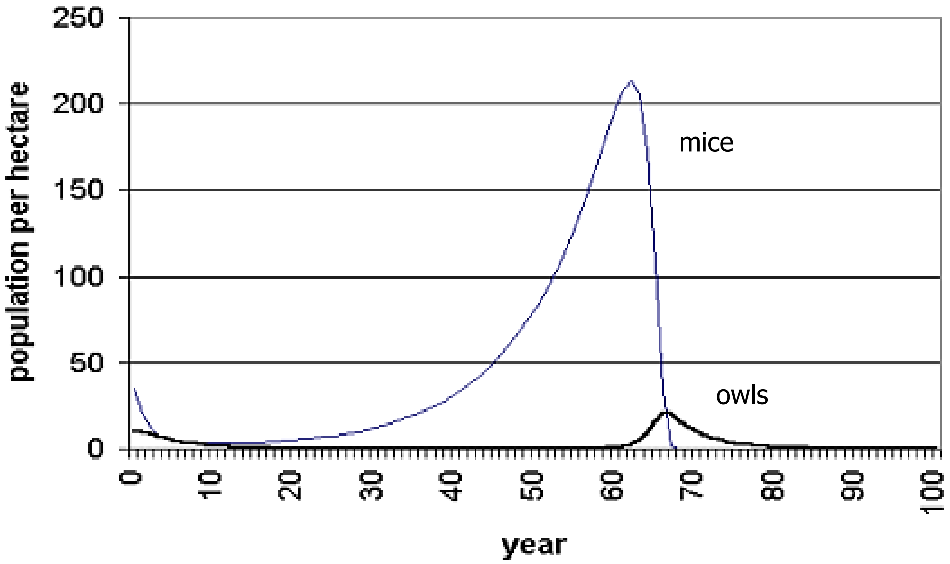


Figure 340: Unstable case of the owls and mice population model

After an initial decline in both mice and owl numbers, the mouse population shows an exponential increase due to a low predation rate. Owls return to the area in increasing numbers, producing a massive increase in predation and collapse in the mouse population. With their principal food source lost, the owls leave and do not return.

By careful adjustment of the numbers of owls and mice in year zero, the students were able to find conditions which allowed the survival of the owl population.

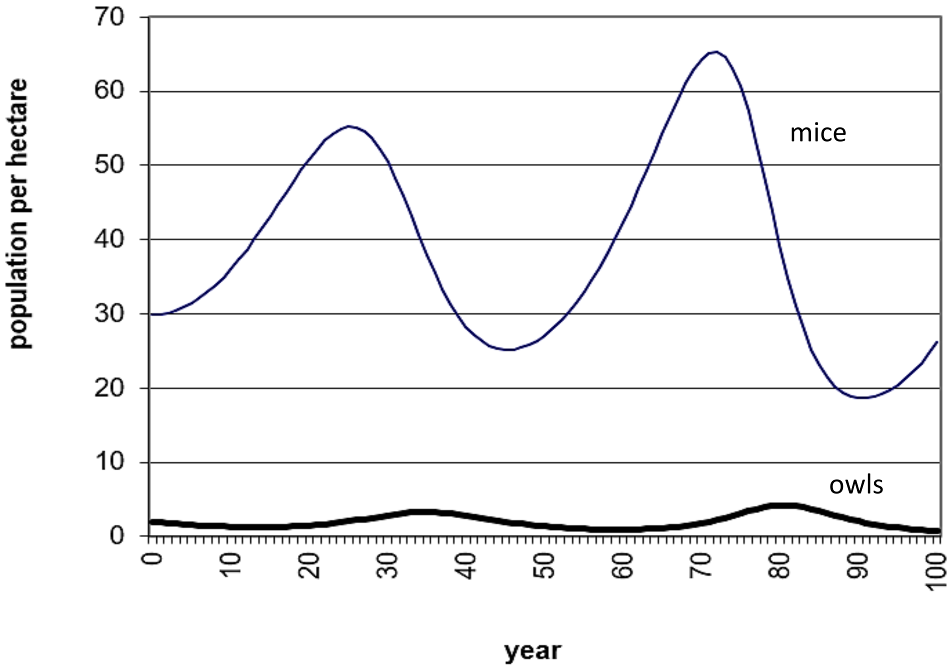


Figure 341: Owls and mice population model indicating survival of both species

It is interesting to plot the populations of owls and mice against one another, producing an outwards expanding spiral from year to year.

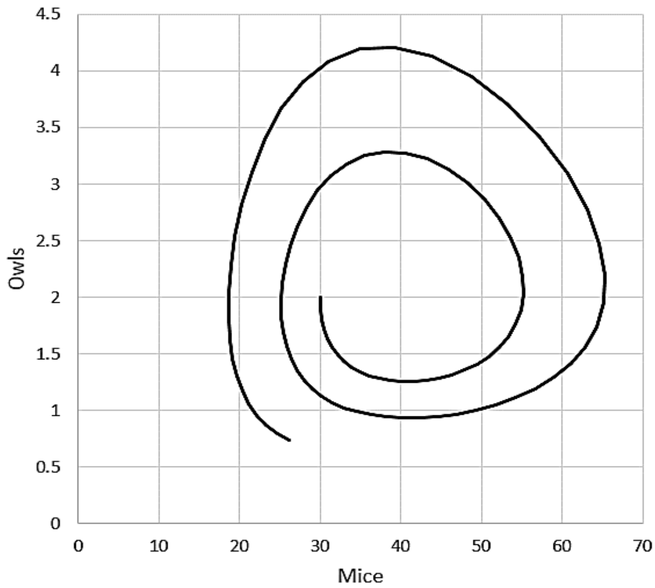


Figure 342: Owl population plotted against mice population

Some instability still exists, with population numbers of owls and mice fluctuating from year to year. However, the model predicts the survival of both species for at least 100 years.

The students were able to formulate advice to the managers of the nature reserve as to the stable ranges of owl and mouse population which the nature reserve could support. If necessary, the numbers of mice could be increased by providing suitable vegetation cover, or the numbers of owls could be increased by providing nest boxes.

A final stage of the modelling cycle would be to monitor the performance of the model over several years, to determine whether predictions of the owl and mouse populations were correct. If necessary, the values of the parameters **k** and **r** can be changed to improve future predictions.

Spread of an illness

Our next numerical model simulates an epidemic of a non-fatal illness such as influenza. The modelling project was undertaken by computing students, who were provided with a published article to explain the set of recurrence relations which might be used (Keeling, 2001). In this simple epidemic model, the population is treated as three groups:

- **Susceptible**: those who can catch the illness
- **Infected**: those who have the illness, and could infect others
- **Recovered**: those who cannot catch the illness again and are no longer infectious to others

We again begin by formulating the relationships between the groups of the population:

- **Susceptible** people catch the illness. They then move into the **infected** group, and the number of people who remain susceptible is reduced.
- The number in the **infected** group increases when susceptible people catch the illness. However, the number of infected persons will fall as patients get better and move to the **recovered** group.
- The number of people in the **recovered** group gradually increases as more and more patients recover from the illness.

These relationships can be written as a set of word equations. In a time period, which might represent a single day during the epidemic:

decrease in the number of persons susceptible	=	number of persons who become infected

change in the number of persons infected	=	number of persons who are newly infected	–	number of persons who newly recover

increase in the number of persons recovered	=	number of persons who have newly recovered

A mathematical model can now be formulated. The symbols **S**, **I** and **R** were introduced to represent the fractions of the population who are ***susceptible***, ***infected*** or ***recovered***. A subscript is used to represent the time period, with **n** being the current period, and **n-1** representing the previous period.

$$\underbrace{S_n}_{\text{new number susceptible}} = \underbrace{S_{n-1}}_{\text{previously susceptible}} - \underbrace{b \; S_{n-1} I_{n-1} \Delta t}_{\text{number newly infected}}$$

Using a similar argument to the predator-prey model, we assume that the number of people becoming infected will depend on:

- the number of people currently susceptible and able to catch the disease, $\mathbf{S_{n-1}}$
- the number of people currently infected and able to transmit the disease, $\mathbf{I_{n-1}}$
- the length of the time period, $\mathbf{\Delta t}$

It is also necessary to introduce a parameter, which we have called **b**, which represents the infectiousness of the disease.

$$I_n \;=\; I_{n-1} \;+\; (\, b\; S_{n-1} I_{n-1} \;-\; k\; I_{n-1}\,)\,\Delta t$$

new number infected | previously infected | newly infected | newly recovered

When formulating an equation for the new number of persons infected, we add the persons who have newly joined from the susceptible population, but we remove the persons who have now recovered. A parameter **k** has been introduced to represent the proportion of infected persons who recover during the time interval.

$$R_n \;=\; R_{n-1} \;+\; k\; I_{n-1}\,\Delta t$$

new number recovered | previously recovered | number newly recovered

The final equation calculates the new number recovered by adding the number of persons who have moved from the infected group during the current time interval.

Before carrying out the modelling calculations, we can add one further refinement to the equation to determine the number of newly infected persons. The parameter **b** representing the infectiousness of the disease can be split into two components:

$$b = \beta\,.\,\kappa$$

where

β indicates how often a contact between an infected and susceptible person leads to infection

κ indicates how many contacts occur during a time period between each infectious person and susceptible individuals in the population

A spreadsheet can be set up to represent the model:

	A	B	C	D
1	day	susceptible	infected	recovered
2	1	1000	1	0
3	=A2+1	=B2-(0.08*10*C2*(B2/1000))	=C2+(0.08*10*C2*(B2/1000))-C2/8	=D2+C2/8

$$S_n = S_{n-1} - (0.08 * 10)\, I_{n-1} * (\, S_{n-1} / 1000\,)$$

$$R_n = R_{n-1} + I_{n-1} / 8$$

$$I_n = I_{n-1} - (0.08 * 10)\, I_{n-1} * (\, S_{n-1} / 1000\,) - I_{n-1} / 8$$

The numerical values of parameters used in the model are:

infection rate **β** = 0.08, **κ** = 10
recovery time **k** = 8 days
initial susceptible population = 1000, initially infected = 1, initially recovered = 0

We can now run the model for a simulated period of 60 days. During this time, it is predicted that just over half of the population would be ill with the disease. This represents a high infection rate.

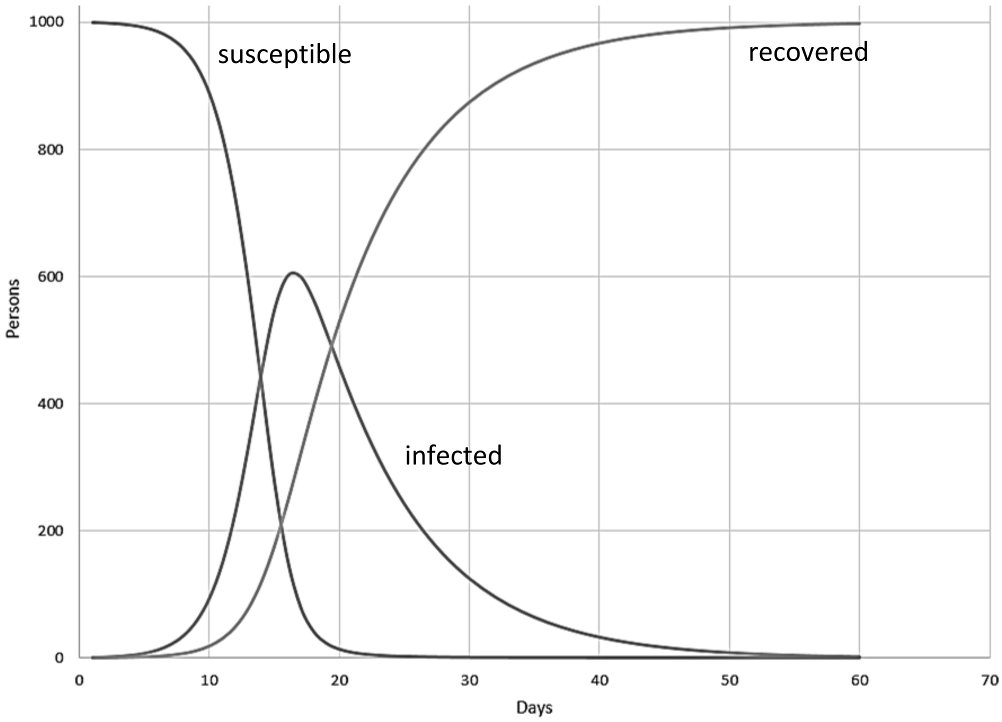


Figure 343: Epidemic model

It is interesting to investigate the effects of changing the model parameters. If the number of contacts, **κ,** between infectious and susceptible individuals is reduced from 10 to 5, the effect is to delay the spread of the epidemic. The overall number of persons catching the illness is reduced to about half of the original number. Early quarantining of infectious individuals is therefore a very effective control measure.

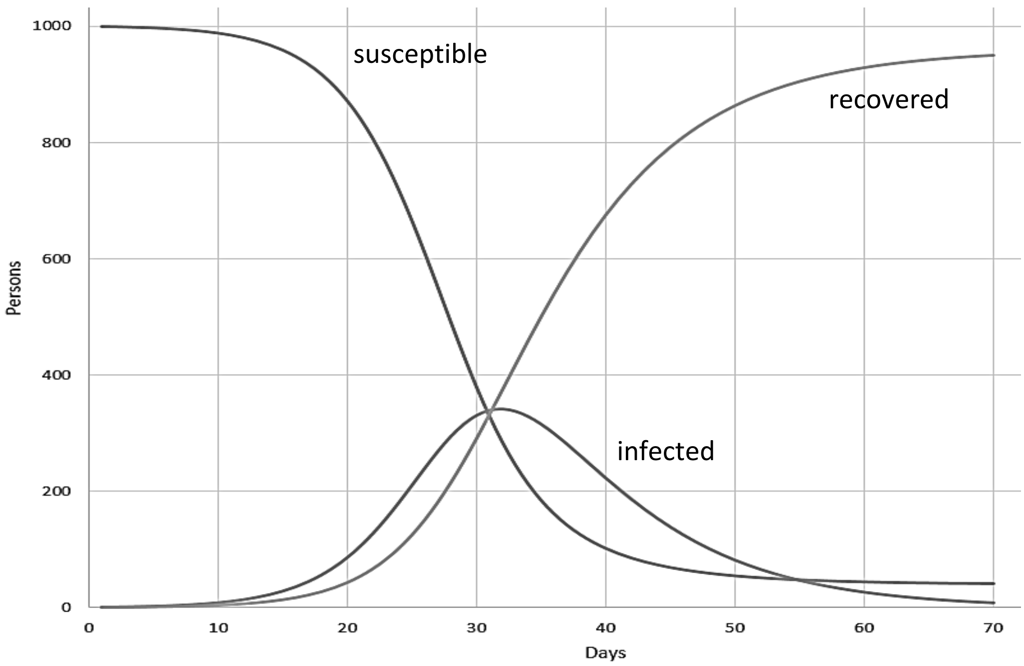


Figure 344: Epidemic model with reduced infectious contact rate

From runs of the model, it is seen that the peak of the epidemic always occurs at the time when equal numbers of susceptible and recovered persons are present in the population:

> Up to this point, the epidemic is increasing. When an infected person makes contact with another individual in the population, there is a greater than 50% chance that they will be susceptible.
>
> Beyond this point, the chances of an infected person coming into contact with a susceptible person begins to fall. There is a greater than 50% chance that they will come into contact with a person who has recovered and is now immune.

The total number of individuals which a person directly infects in the population during his or her entire infective period is known as the **epidemiological parameter R_0**. This is a measure of the infectiousness of a disease. Typical values are:

AIDS	2 to 5
Smallpox	3 to 5
Measles	16 to 18
Malaria	over 100

This gives an indication as to why malaria is perhaps the most serious health risk in many developing countries.

In the next two examples, we examine the relationship between **macroscopic** and **microscopic** approaches to modelling:

Traffic flow model

Traffic flow modelling is important for the planning of safe and efficient road networks, and for managing road congestion problems. Results from modelling can provide information on the most suitable layout of traffic lanes and the most appropriate speed limits to set.

When temporary road works are taking place on a motorway, contra-flow systems are often put in place with traffic constrained to a single lane with no opportunity for overtaking. We will consider how this situation could be modelled mathematically.

Figure 345:

Traffic contra-flow

One approach to the modelling problem is to take a **macroscopic** view. The traffic flow is considered as a fluid, using an analogy with water flowing through a length of pipe. A section of the pipe with reduced diameter represents the contra-flow restriction. It is then possible to use fluid flow equations to model the traffic flow. The numbers of vehicles able to pass through the contra-flow system and re-join the normal motorway can be predicted for different arrival rates and speed limits. Traffic modelling is made more complicated because drivers sometimes behave unusually. Vehicles may accelerate, reduce speed or brake at moments when this would not be expected from the surrounding traffic conditions. Unpredictability can be added to a fluid model by introducing randomness or turbulence into the flow.

Whilst fluid flow models give interesting results, and can be justified theoretically, it can be difficult to relate the model to events taking place on an actual road network. An alternative approach is to take a **microscopic** view in which the movement of each individual vehicle is modelled. The vehicle movements can then be combined to determine overall traffic flow parameters. We will develop a microscopic model for the contra-flow traffic system.

A simple four-step model for the simulation of traffic flow has been proposed by Vandaele et al. (2000):

Step 1: acceleration

A maximum velocity $\mathbf{v}_{\mathbf{max}}$ is set for the road. All cars that have not already reached the maximal velocity will accelerate.

Step 2: safety distance

We determine the length of empty road ahead of each vehicle. If a vehicle has $\boldsymbol{d}$ empty units of road ahead and its velocity (in distance units in each time interval) is larger then $\boldsymbol{d}$, then the velocity is reduced to $\boldsymbol{d}$ so as to match the speed of the vehicle in front.

Step 3: randomization

Drivers may behave unpredictably and reduce speed unexpectedly. With a probability $\boldsymbol{p}$, a vehicle reduces its velocity by one unit.

Step 4: driving

The new velocity $\mathbf{v}_\mathbf{n}$ for each vehicle $\boldsymbol{n}$ is projected forward to find the new position of the vehicle for the start of the next time step.

A microscopic traffic flow model can be run manually, using paper and pencil, but it is an interesting exercise for students to set up the model on a computer. Either a spreadsheet or programming language can be used. In this example, a Java program has been produced with a Windows interface (figure 346). The length of the contra-flow system has been divided into 100 sections, with each represented by a numbered column in the tables. Details of the motion of any vehicle currently at each location will be shown.

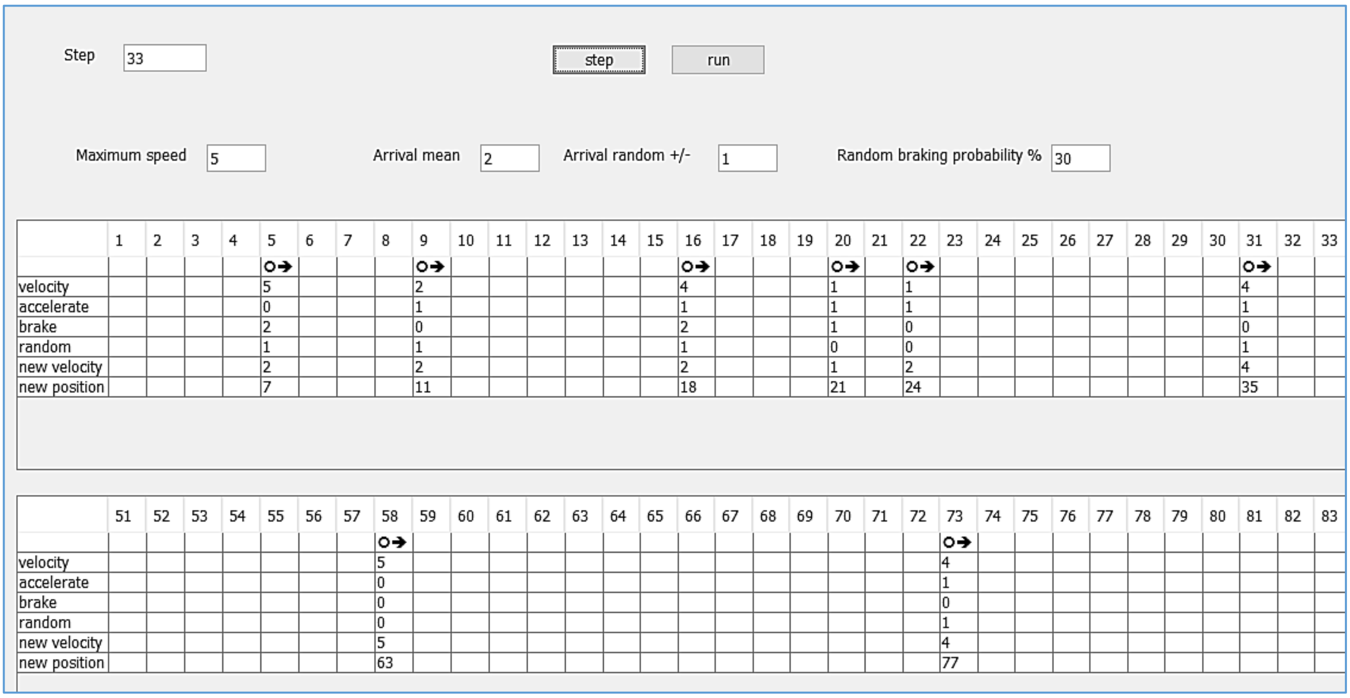


Figure 346: Microscopic traffic flow model

Variables for each vehicle are:

velocity
The velocity of the vehicle at the start of the current time step. For the purpose of this model, only qualitative results were needed and velocity is specified in arbitrary units. A further development would be to introduce accurate road speeds.

accelerate
The maximum permitted speed is specified when running the model. Any vehicles travelling at less than the maximum speed are given an acceleration of 1 unit. For example: the vehicle at location 5 is already travelling at the maximum speed, so is given zero acceleration. However, the vehicle at location 9 is travelling below the maximum speed so is given an acceleration.

brake
The number of empty road units ahead of each vehicle is checked, and braking is applied if the current velocity exceeds this value. Braking is also applied if the vehicle would be exceeding the maximum permitted speed, as might be the case when vehicles arrive at the start of the contra-flow.

random
Some vehicles randomly brake by one unit, according to a probability which is chosen before running the model. Whether a particular vehicle will brake is decided by generating a random number between 0 and 99, then checking whether this is less than the probability chosen.

new velocity
The velocity for each vehicle during the next time step is found by adding the acceleration to the previous velocity, then removing any braking.

new position
The new road position can be calculated by adding the number of velocity units to the current position. For example, consider the vehicle currently at position 5:

- the initial velocity is 5 units
- there is no acceleration, as the vehicle is already travelling at the maximum permitted speed
- there is a required braking of 2 units because of the spacing before the vehicle ahead
- there is also random braking of 1 unit
- deducting braking from the initial velocity gives a new velocity of 2 units
- the new velocity is added to current road location 5 to give a new road location 7 for the next time step.

Positions of other vehicles will be advanced in a similar way.

The arrival pattern for vehicles can be adjusted using two parameters:

arrival mean sets the number of time intervals between each arriving vehicle at the start of the contra-flow.

arrival random will randomly adjust the arrival rate within the specified range of time intervals.

It is therefore possible to simulate a steady flow of vehicles arriving, or to allow vehicles to arrive randomly in bunches.

Each vehicle leaving the system is counted, and its time for passing through the contra-flow can be recorded. This allows detailed analysis of the traffic flows produced under different simulated conditions. Several interesting conclusions can be drawn:

Speed

The overall volume of traffic able to pass through the contra-flow section can be affected by the specified speed limit:

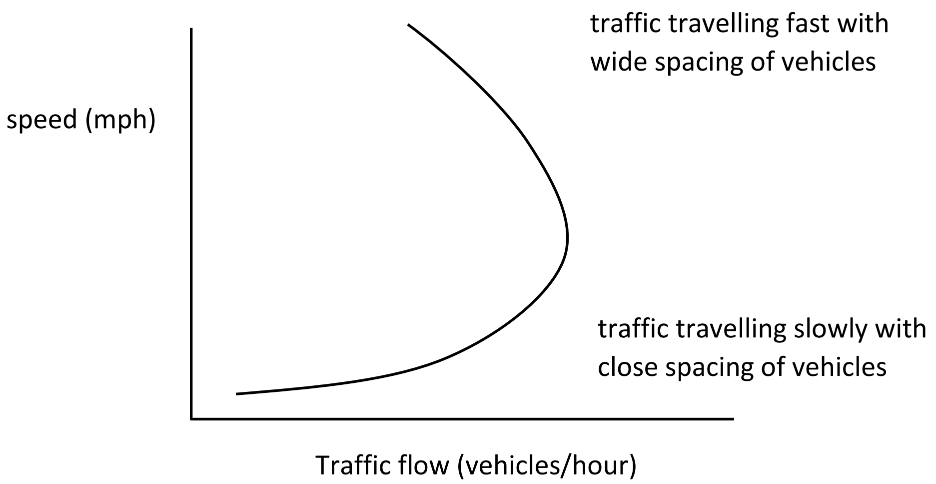


Figure 347: Effect of the contra-flow mean speed on traffic flow

If the speed limit is low, the total traffic flow will be low. It is more surprising, however, that a high speed can also lead to a reduced total traffic flow. Drivers naturally maintain a larger separation distance when travelling in fast traffic. Where traffic is constrained to a single lane, there is a high probability that drivers will sense that they are travelling too fast and brake. This braking might be more severe than is really necessary, and can cause a wave effect through the following traffic, disturbing the smooth flow. It is found that maximum traffic flow capacity can be achieved at an intermediate speed.

Bunching of vehicles

Runs of the model can be carried out for different patterns of vehicle arrivals. If arrivals are evenly spaced, this tends to produce a smooth transmission of traffic through the contra-flow. However, the arrival of fast moving bunches of traffic can produce severe braking and disruption to the smooth flow, leading to a reduction in overall traffic flow capacity.

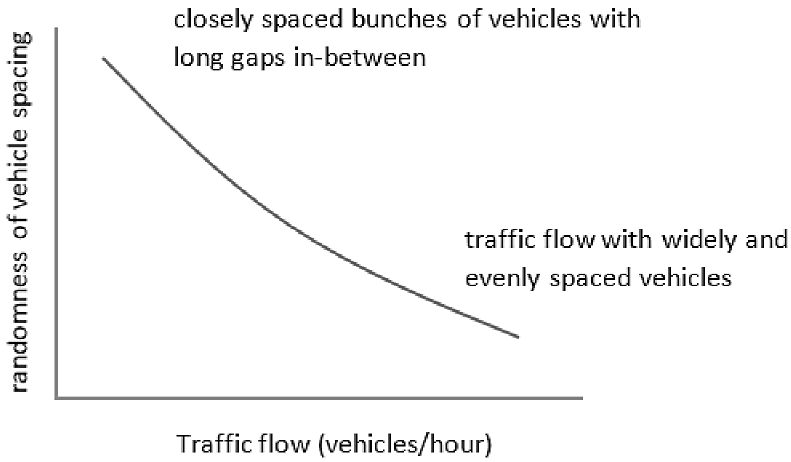


Figure 348: Effect of grouping of arriving vehicles on traffic flow

If measures can be found for producing a more even flow of traffic in the approaches to the contra-flow section, this is likely to reduce congestion and delays.

The links between the microscopic and macroscopic models become apparent. Higher traffic flows can be achieved if the flow is smooth and turbulence is avoided.

Hillslope water flow

Another use of a microscopic approach is in the modelling of water flows down hillslopes in response to rainfall. This is an important topic in geography for explaining processes such as flooding and the distribution of natural vegetation.

Rather than model a hillslope as a single object, we can represent the slope as a series of **stores**. Each store is able to receive inputs of water from rainfall, and downslope flows of surface water and through-flow in the soil (figure 349).

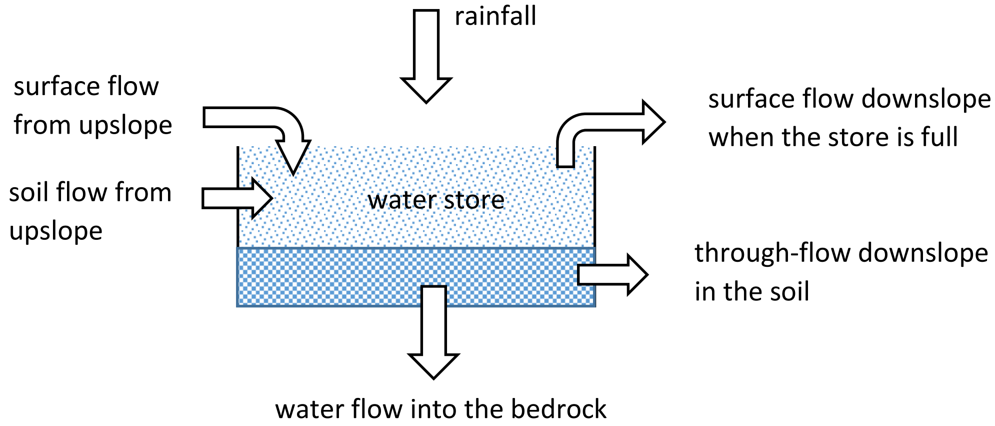


Figure 349: A hillslope water store

The amount of water that can be held in a store depends of the porosity of the soil. The store can lose water, both through downslope flow through the soil, and into the bedrock. If the store becomes full, water can overflow at the ground surface and run down the hill slope.

A microscopic model is made up from a series of interconnected stores:

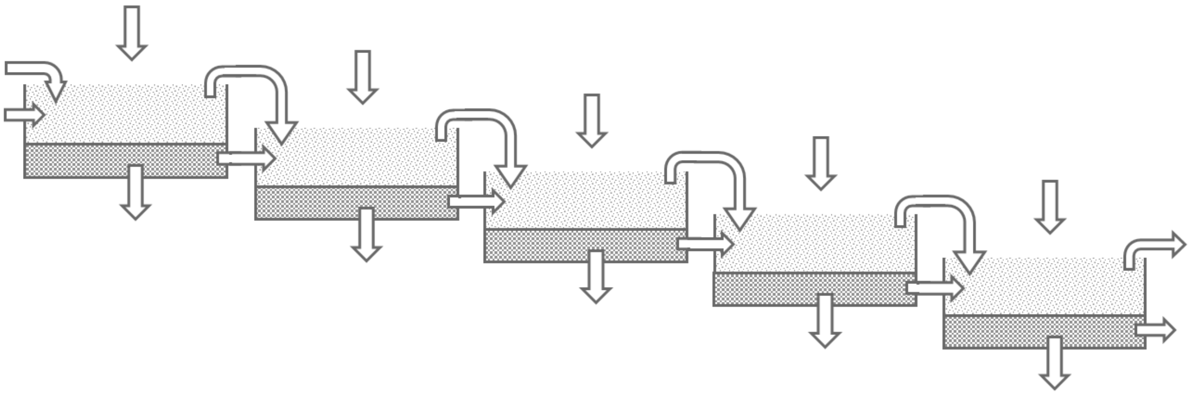


Figure 350: A hillslope hydrological model

To operate the model, we make a series of assumptions:

- Rainwater and surface water can only enter a store if it is not already full. Once the store becomes full, any additional water will run-off across the land surface to the next store down the slope.
- The rate at which water can enter a store is limited by water movement through the pores of the soil. If water arrives at too fast a rate, the excess will run-off along the ground surface.
- The rate at which water flows into the bedrock is proportional to the amount of water in the store. If the store is full, there will be a greater downwards water pressure, leading to a faster loss of water.
- In a similar way, downslope through flow in the soil will depend on the amount of water in the store. If the store is full, water will flow out through the soil more

rapidly. Downslope through flow will also depend on the slope angle, with faster flow through the soil on a steep slope.

The hillslope model can be constructed using a spreadsheet, or a computer program can be produced. By developing and experimenting with the model, geography students can gain a greater understanding of hydrological processes. Different flood mechanisms can be simulated:

- **Saturation excess overland flow**, when a long period of rainfall has filled the soil water stores and any further rainfall now runs off along the ground surface. This situation can occur after several days of moderate steady rain.
- **Infiltration excess overland flow**, where extremely intense storm rainfall reaches the ground surface faster than it can infiltrate into the soil. The excess water runs off down the slope. This situation can occur during thunder storms.

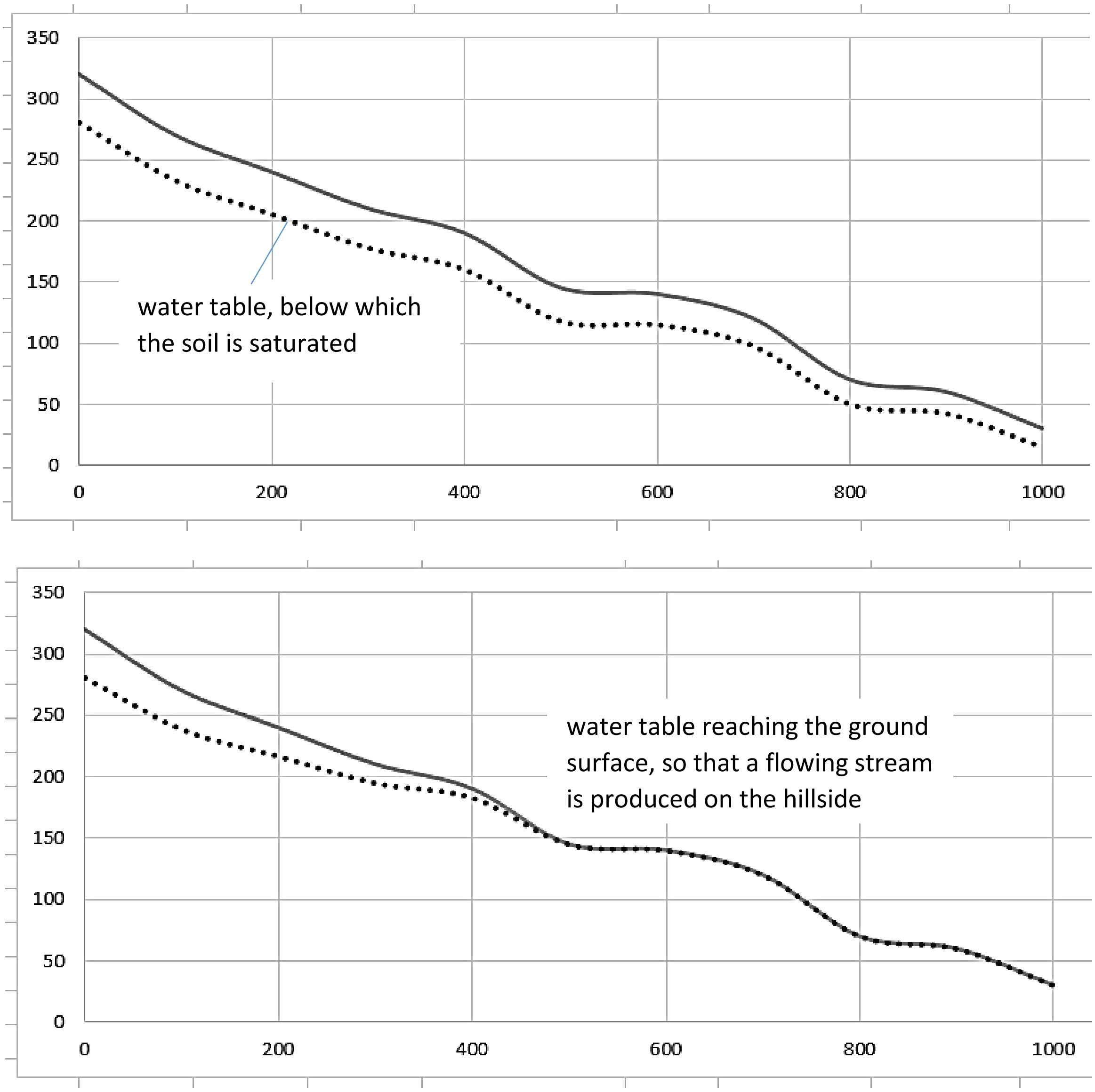


Figure 351: The hillslope hydrological model, showing the water table rising to the surface after a period of simulated rainfall, producing saturation excess overland flow

Scree slope angle

Mathematical modelling often has the objective of identifying patterns in data and producing an algebraic formula to describe some real world situation. The formula can then be used to make future calculations.

Real data usually shows some randomness. A common method of obtaining a formula from data is to plot data points on a scatter graph, then find the best fit of curve through the points. The curve chosen is often a straight line with an equation:

$$y = mx + c$$

where m is the gradient of the straight line and c is the intercept on the vertical axis. As an example, we might look at a project carried out by geography students to investigate the slope angles of scree.

Scree was formed at the end of the Ice Age, at a time when temperatures were oscillating around freezing point. During the day, melting of snow and ice produced water which entered cracks in rock surfaces. At night, freezing caused expansion which fractured the rock, allowing fragments to break away and fall under gravity to produce accumulations of scree below cliff faces.

Scree slopes have a constant, or **rectilinear**, slope. The actual angle of the slope, however, varies for different locations. Scree at Bird Rock, Abergynolwyn, has a slope angle of 34°.

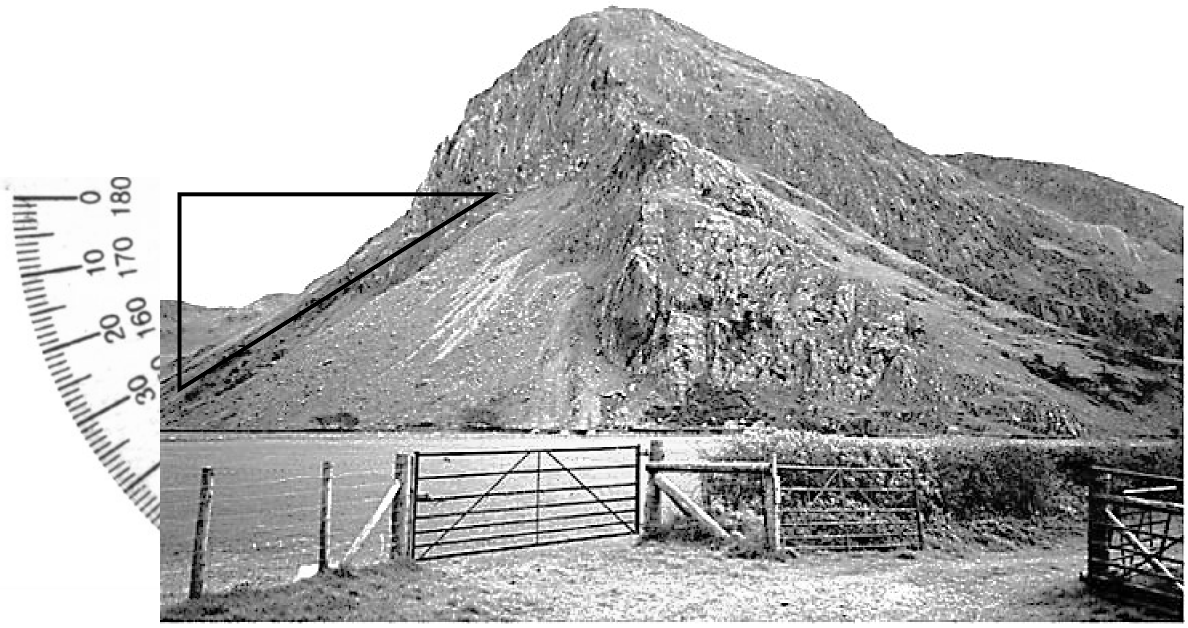


Figure 352: Scree at Bird Rock, Abergynolwyn

The first stage in mathematical modelling is to decide on the significant factors which should be included in the model. It is reasonable to suggest that the most import factor controlling slope angle is the shape of the rock fragments in the scree.

Blocky fragments tend to pile up against one another, whilst flat slabs can slide more easily downhill over one another.

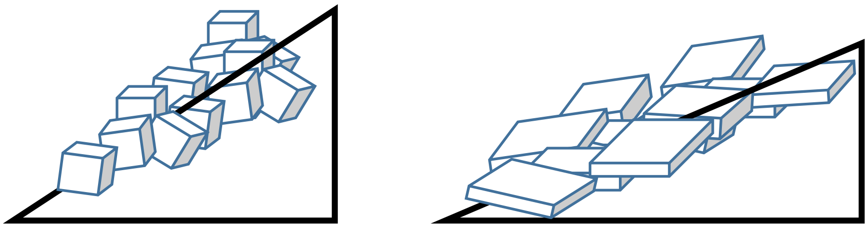


Figure 353: Fragment shape affecting scree slope angle

Rocks which break down to cubic fragments, such as granite, tend to produce steeper scree accumulations. Rocks which break down to flat sheets, such as slate, tend to have gentler scree angles.

Fragment shape can be measured by means of the **sphericity index**. This is defined by the ratio:

$$sphericity = \frac{volume\ of\ the\ rock\ fragment}{volume\ of\ the\ smallest\ sphere\ which\ would\ enclose\ it}$$

In practice, it is difficult to measure the volume of a rock fragment. A quicker method which gives a good approximation of the sphericity index is to measure the maximum, intermediate and shortest dimensions of the rock fragment along three axes at right angles:

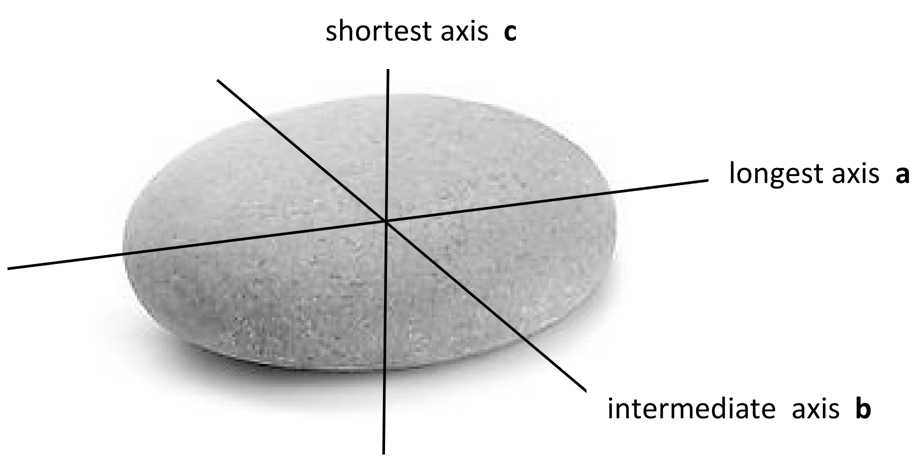


Figure 354: Dimension measurements for a rock fragment

The sphericity index is determined using the formula:

$$sphericity = \sqrt[3]{\frac{bc}{a^2}}$$

A perfectly spherical fragment has a sphericity index of 1.0, with flatter shapes having lower sphericity index values.

A survey was carried out at a number of scree slopes in North Wales which were developed in different types of rock. An average sphericity value was obtained for each site after measuring 50 randomly selected fragments from the scree. Results for the sites are displayed in figure 355. It is seen that there is a reasonably strong linear relationship between slope angle and sphericity index.

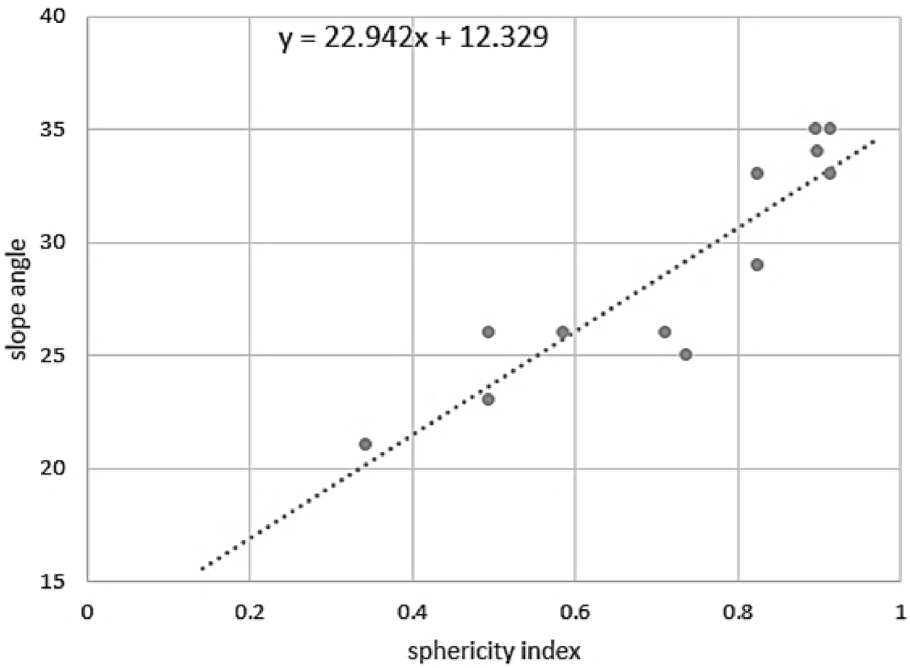


Figure 355: Relationship between slope angle and sphericity index in scree deposits

A straight line has been fitted to the data by the spreadsheet software, using linear regression. The equation of the line gives the relationship:

$$slope\ angle = (22.942\ \times sphericity) + 12.329$$

The final stage of the modelling cycle is to test the model against the real world by making and checking predictions. Using this equation, it is possible to:

- estimate the sphericity for fragments in a scree slope, if the slope angle is known
- estimate the likely slope angle of a scree formed from fragments of a known shape.

If the model does not perform adequately, then further data can be added and a new equation calculated.

Linear regression can be a powerful technique for deriving equations from sets of measurements. In this simple example the slope angle was dependent on only one variable, the sphericity index. In more complex models, the entity of interest may be dependent on a number of different variables. In the next example, we will investigate the use of **multiple linear regression** in analysing a more complex problem of this type.

Expedition Journey Time

Students who are training to become outdoor pursuits instructors are required to make reasonably accurate estimates of the time which expeditions will take over mountainous terrain as part of the procedure for safety planning. A mathematical formula known as *Naismith's Rule* can be used for estimating journey time. This determines a time based on walking speed over flat ground, then adds extra time for the amount of ascent during the journey:

> The rule was devised by William W. Naismith, a Scottish mountaineer, in 1892. Allow 1 hour for every 5 kilometres (3.1 mi) forward, plus 1 hour for every 600 metres (2,000 ft) of ascent.

Naismith's Rule has subsequently been modified and improved by the mountaineer Tranter, who introduced corrections to allow extra journey time if the level of fitness of the party was low, if heavy rucksacks were being carried, or if the group was walking in the dark or in bad weather.

Students using Naismith's Rule with Tranter's corrections have found that the time calculations for expeditions in the mountainous area of North Wales are still very inaccurate. This is due to wide variations in the time taken to cross different types of terrain. Walking speed is much slower across moorland or through overgrown forest in comparison to well-constructed footpaths. Scrambling over large, loose rocks on mountainsides or crossing peat bogs with deep pools of soft moss is particularly slow, as great care has to be taken.

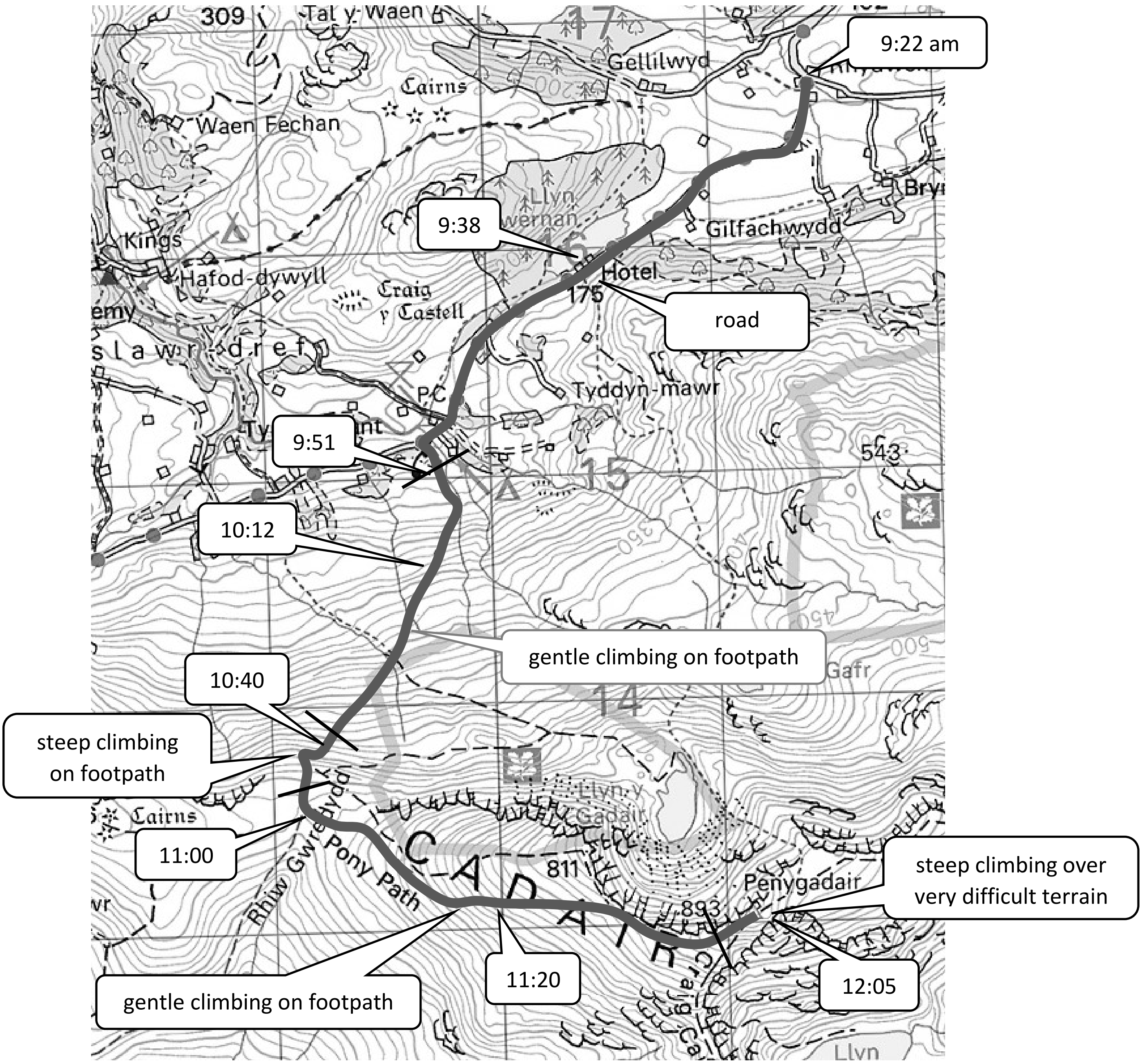
To create a mathematical model for estimating journey times, distances for six terrain factors were identified as significant:

- **total distance** for the journey,
- **gentle climbing**, along a reasonable path where speed is reduced due to the ascent,
- **steep climbing**, along a reasonable path where speed in greatly reduced due to very steep ascent,
- **steep descent**, along a reasonable path with speed reduced due to the care needed,
- **difficult terrain**, such as crossing rough moorland or forest along poorly defined paths, where speed is reduced,
- **very difficult terrain**, such as crossing boulder fields or peat bogs, where speed is greatly reduced.

All distances are measured in kilometres, making use of Ordnance Survey maps, air photographs and other photographic images which will help to illustrate the nature of the terrain which will be crossed. Each section of the journey is only counted within one sub-category, so for example: a steep descent over loose rocks would be counted as very difficult terrain.

In order to calibrate the model, it was necessary to collect data on the journey times and route characteristics for a number of expeditions. For example, timings for an ascent of

Cader Idris are illustrated in figure 356. Characteristics of each section of the journey are recorded:



total time	2 hours 43 minutes	
total distance	6.7 km	
gentle climb	3.3 km	slow
steep climb	0.3 km	very slow
steep descent	-	slow
difficult terrain	-	slow
very difficult terrain	0.2 km	very slow

Figure 356: Journey data for the ascent of Cader Idris

The series of journey data sets were collected, then analysed using **multiple linear regression**. The objective of this technique is to apply a loading factor to each of the variables. This leads to an equation of the form:

$$journey\ time = (A \times total\ distance) + (B\ \times gentle\ climb) + (C \times steep\ climb) + (\ D \times steep\ descent) + (E\ \times diff.terrain) + (F \times v.diff.terrain)$$

The journey time for a new route can then be predicted by entering the numbers of kilometres within each terrain category, then multiplying by the appropriate constant A – F.

Software for carrying out multiple linear regression is available for download from the Internet. The software used by our students is the RegressIt package produced by Duke University and the University of Texas. This is installed as a data analysis extension to the Microsoft Excel spreadsheet. This package is available from: http://regressit.com/

Data is first entered into an Excel worksheet, with variable names provided as column headings. The software requires that the number of letters in each variable name is the same:

TIMEJ	journey time
DISTN	distance total
GENTL	gentle climb
STEEP	steep climb
DESNT	steep descent
DIFFT	difficult terrain
VDIFF	very difficult terrain

Distances in kilometres for eleven expedition journeys were entered:

	A	B	C	D	E	F	G	H
1	Journey	TIMEJ	DISTN	GENTL	STEEP	DESNT	DIFFT	VDIFF
2	1	2.72	6.7	3.3	0.3	0	0	0.2
3	2	2.05	6.7	0	0	0.3	0	0.2
4	3	6	14.4	5	1	1	0.02	0.4
5	4	5.8	12.8	4.8	1.8	3.2	0.8	0
6	5	5.4	11.2	1.6	3.2	1.9	1.4	0.5
7	6	6.9	12.6	2.8	4.4	3.8	1.1	0.3
8	7	6.2	12.1	3.9	2.1	4.5	0	0.1
9	8	5.9	11.4	3.2	2.6	3.9	0.2	0.2
10	9	2.1	6.2	1.3	2.2	2.2	1.2	2.9
11	10	9.2	28.4	6.7	4.2	4.3	2.2	0.3
12	11	3.1	9.2	1.5	3.4	2.9	1.6	1.2

Figure 357: Journey data for a series of expeditions

The software can produce a set of scatter graphs for each of the independent variables in relation to our dependent variable of **journey time**. These are shown in figure 358.

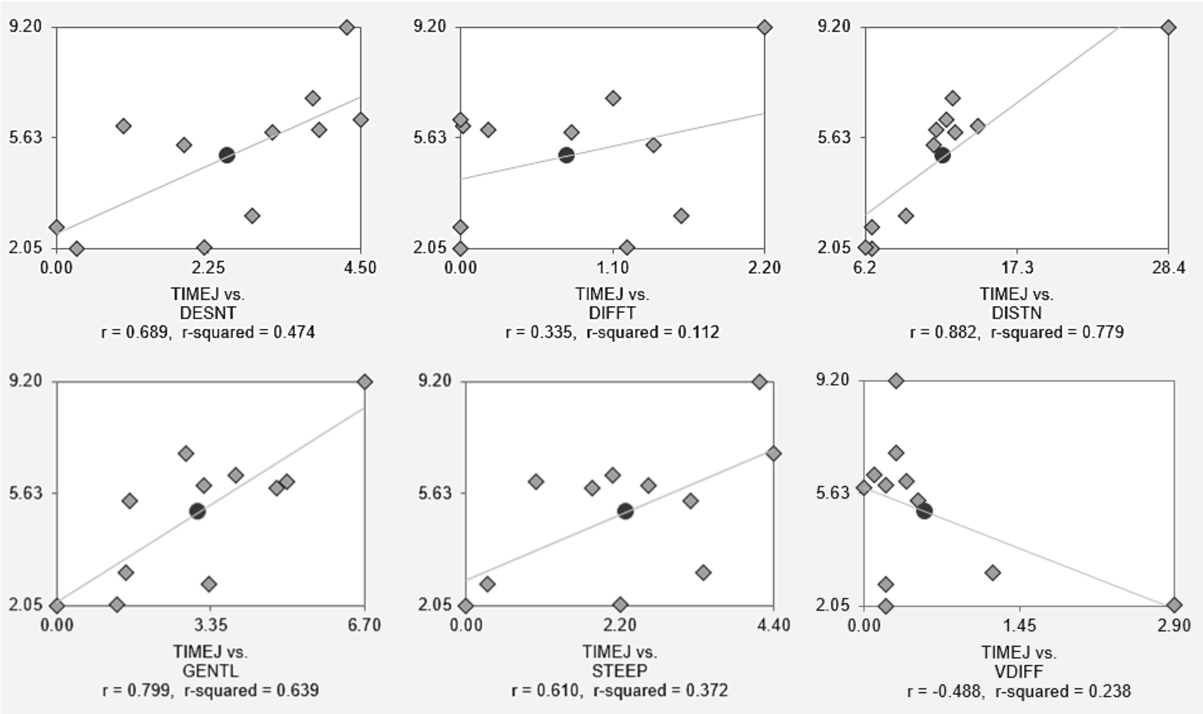


Figure 358: Scatter graphs for journey time against each of the independent variables

Best fit regression lines are shown, along with the correlation coefficient **r**. As we might expect, there is a fairly good correlation between journey time and total distance, and also some correlation against gentle and steep ascent distances. Journey time is likely to be longer for longer routes, and when slower climbing is involved. Correlations with the other variables, steep descent, difficult and very difficult terrain, are less clear. This is because these types of terrain are often not significant components of the expeditions recorded.

The multiple regression formula is calculated by selecting the dependent variable and independent variables from a list of fields:

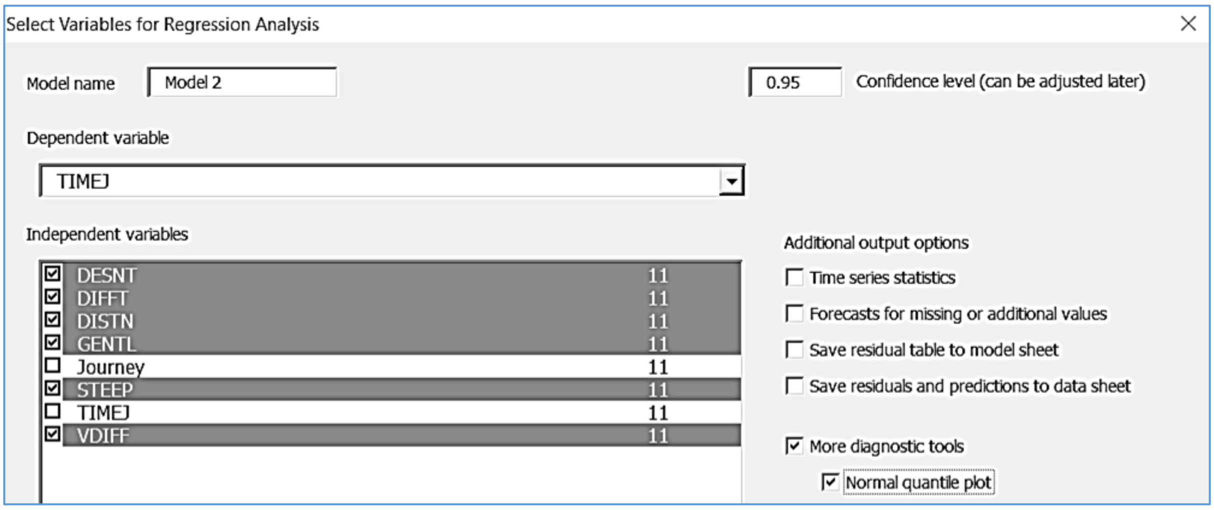


Figure 359: Selecting variables to carry out multiple linear regression

Results output by the software show the coefficients to be used in the regression formula:

Dependent Variable: TIMEJ

Regression Statistics: Model 2 for TIMEJ (6 variables, n=11)

R-Squared	Adj.R-Sqr.	Std.Err.Reg.	Std. Dev.	# Cases	# Missing	t(2.50%,4)	Conf. level
0.964	0.911	0.675	2.263	11	0	2.776	95.0%

Coefficient Estimates: Model 2 for TIMEJ (6 variables, n=11)

Variable	Coefficient	Std.Err.	t-Stat.	P-value	Lower95%	Upper95%	Std. Dev.	Std. Coeff.
Constant	0.782	0.676	1.156	0.312	-1.095	2.659		
DESNT	-0.039	0.265	-0.146	0.891	-0.773	0.696	1.591	-0.027
DIFFT	-1.309	0.706	-1.854	0.137	-3.270	0.651	0.780	-0.452
DISTN	0.208	0.095	2.198	0.093	-0.055	0.471	6.119	0.563
GENTL	0.248	0.231	1.076	0.343	-0.392	0.889	1.947	0.214
STEEP	1.012	0.408	2.478	0.068	-0.122	2.145	1.463	0.654
VDIFF	-0.377	0.368	-1.023	0.364	-1.399	0.645	0.834	-0.139

Figure 360: Output of the regression model

We have obtained the constants which were required to produce a formula for calculation of journey time:

$$\begin{gathered} journey\ time = 0.208\ total\ distance + 0.248\ gentle\ climb + 1.012\ steep\ climb \\ -0.039\ steep\ descent - 1.309\ diff.terrain - 0.377\ v.diff.terrain + 0.782 \end{gathered}$$

A graph is produced to show the extent to which this formula would produce correct results for our eleven known journey times. We see that most are predicted very accurately, with maximum errors of no more than half an hour.

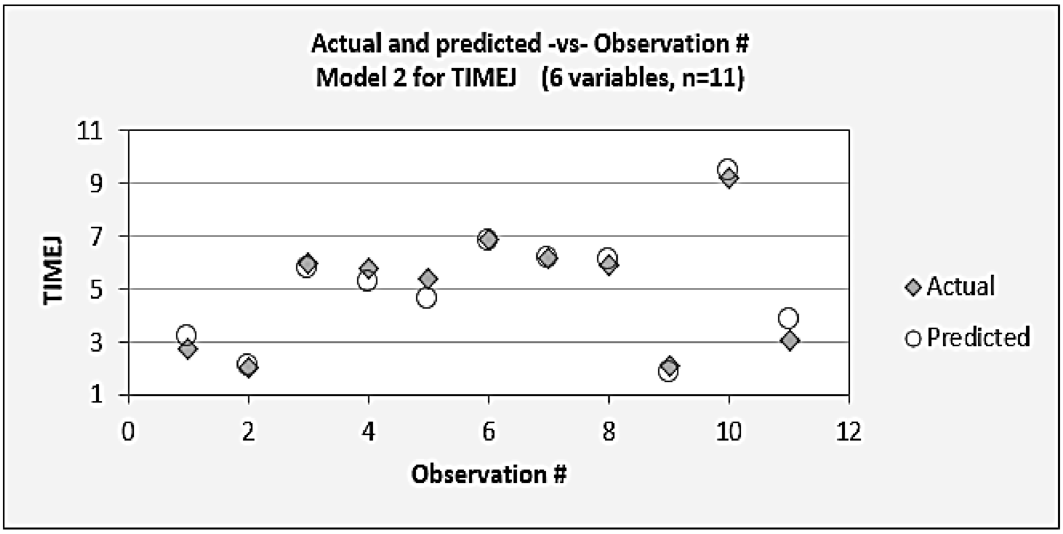


Figure 361: Comparison of measured journey times with predictions using the regression formula

The formula produced during the modelling will probably continue to give predictions of journey times accurate to within half an hour. However, the formula might be further improved if distances are recorded for future expeditions, along with the actual times taken, and this data is added to the regression model.

The model developed so far has included only the terrain factors of gradient and the nature of the ground. We are assuming that other factors will affect all expeditions in a similar way and can be ignored, which is probably an over-simplification. After obtaining a general time prediction, corrections might need to be added for the particular circumstances of this expedition, such as:

- Group factors: fitness of the party, weight carried, and stops planned, e.g. for food or photography,
- Weather factors which might slow the group: hot sun, heavy rainfall, mist, snow or ice.

Expedition leaders must still be aware of the possibility of unforeseen events and have contingency plans in place. Things which could go wrong include: route finding errors, sections of the route closed (e.g. a bridge out of use), or an accident or illness affecting a member of the group.

Diet and nutrition

The final modelling example in this chapter looks at the areas of diet and nutrition. Health professionals make recommendations on the daily intake of nutrients for different groups of people, with variations according to age, gender and occupation. Men tend to need more nutrients than women, with the exception of salt and fibre. Typical daily amounts recommended for a healthy, balanced diet for maintaining rather than losing or gaining weight are:

	Men	Women
Energy (kcal)	2500	2000
Protein (g)	55	45
Carbohydrates (g)	300	230
Sugar (g)	120	90
Fat (g)	95	70
Saturates (g)	30	20
Fibre (g)	24	24
Salt (g)	6	6

In addition, a range of vitamins and minerals are needed for correct body function. Examples of recommended daily intake are:

Vitamin A	900 micrograms
Vitamin C	90 milligrams
Vitamin D	15 micrograms
Vitamin B_6	2 milligrams
Vitamin E	15 milligrams

Nutritional information is becoming increasingly available, both for menu ingredients and for cooked meals in restaurants. We might consider a couple of examples shown on web pages:

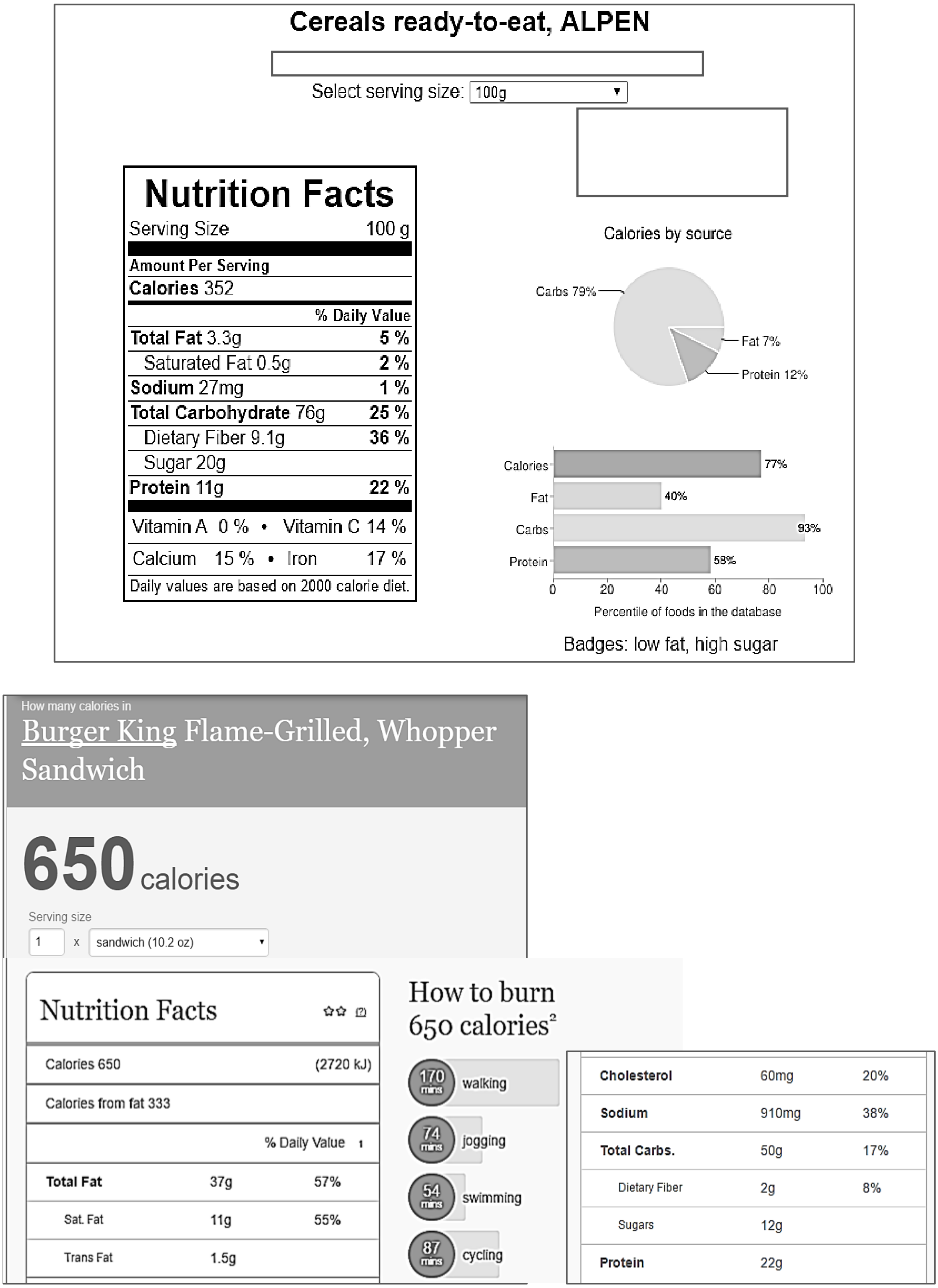


Figure 362: Examples of published nutritional information on menu ingredients and cooked meals

Software applications are available on-line which allow details of a daily or weekly diet to be entered, and will then calculate the nutritional content of the selected meal items. The program can highlight a consumption of particular nutrients, vitamins or minerals which is significantly above or below recommended levels.

An interesting project for computing or health care students is to develop their own dietary analysis software, using either a spreadsheet or programming language. In the example below, a computing student produced a program which allows the selection of gender and age group, then displays daily nutrient requirements. Food items can be selected, and the nutritional content of portions displayed. Food items can be added to a daily list, and percentages of the recommended daily amounts for each nutrient can be calculated.

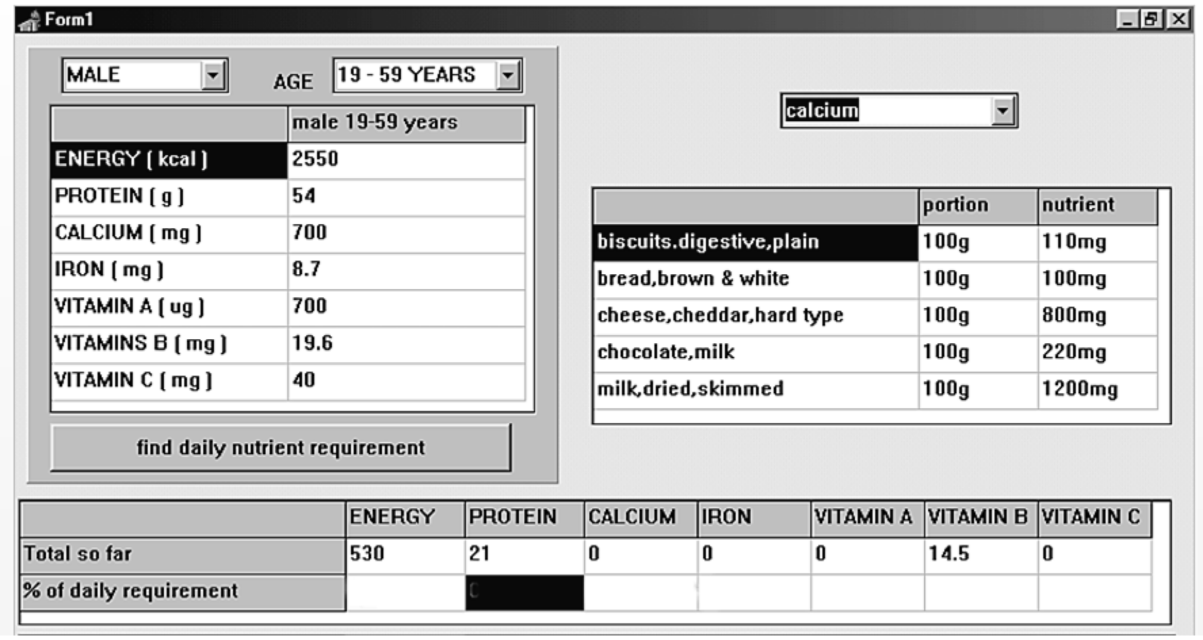


Figure 363: Student project to analyse diet

By experimenting in this way, students can gain a greater knowledge and understanding of dietary planning for groups with differing dietary requirements.

A common problem that health and social care students encounter is the planning of balanced diets at minimum cost from a choice of possible meals. Young families or elderly people may need advice on how to eat well on very limited budgets. Techniques to solve this type of problem are found in a branch of mathematics known as **operational research**.

To demonstrate the principles of this approach, we will consider a simplified scenario using data presented earlier in this section:

A man wishes to consume the recommended daily amounts of:

fat	95 grams
carbohydrate	300 grams

For this very simple example we will assume that only two foods are available:

Burger, containing 37 grams of fat and 50 grams of carbohydrate

Cereal, containing 3.3 grams of fat and 76 grams of carbohydrate

The prices of the two foods are:

Burger: £3.60

Cereal: £2.79 for 750 grams, so that a 100 gram portion costs 37 pence.

What would be the most economical way of achieving the daily recommended intake of fat and carbohydrate from portions of just these two foods?

We can answer this question using **linear algebra**. We begin by defining two variables:

$\mathbf{x_1}$ represents the number of burgers

$\mathbf{x_2}$ represents the number of portions of cereal

We can now set up a series of equations. Considering first the **fat** content. Each burger contributes 37g and each portion of cereal contributes 3.3g. Overall, the diet must contain at least 95g of fat. Using algebra to represent the numbers of burgers or portions of cereal which must be consumed to meet the fat requirement:

$$37x_1 + 3.3x_2 \geq 95$$

We can now consider the **carbohydrate** content. Each burger contains 50g and each portion of cereal contains 76g. Overall, the diet must contain at least 300g of carbohydrate. The numbers of burgers or portions of cereal which must be consumed to meet the carbohydrate requirement are given by:

$$50x_1 + 76x_2 \geq 300$$

We can add two further equations:

$$x_1 \geq 0 \qquad x_2 \geq 0$$

These simply indicate that the numbers of burgers or portions of cereal must be zero or a positive quantity. Negative amounts are not possible.

We can now make some simple calculations. If all the fat is provided by burgers, with no portions of cereal eaten at all, the number of burgers needed to meet the daily fat requirement would be:

$$\frac{95}{37} = 2.6$$

If all the fat is provided by cereal, with no burgers eaten, the number of portions required would be:

$$\frac{95}{3.3} = 28.8$$

(This is clearly not a reasonable number of portions, and the need for a mixed diet is apparent!)

In a similar way, we can calculate that all the required carbohydrate could be provided by 6 burgers and no cereal, or by 3.9 portions of cereal and no burgers. These figures are shown graphically in figure 364.

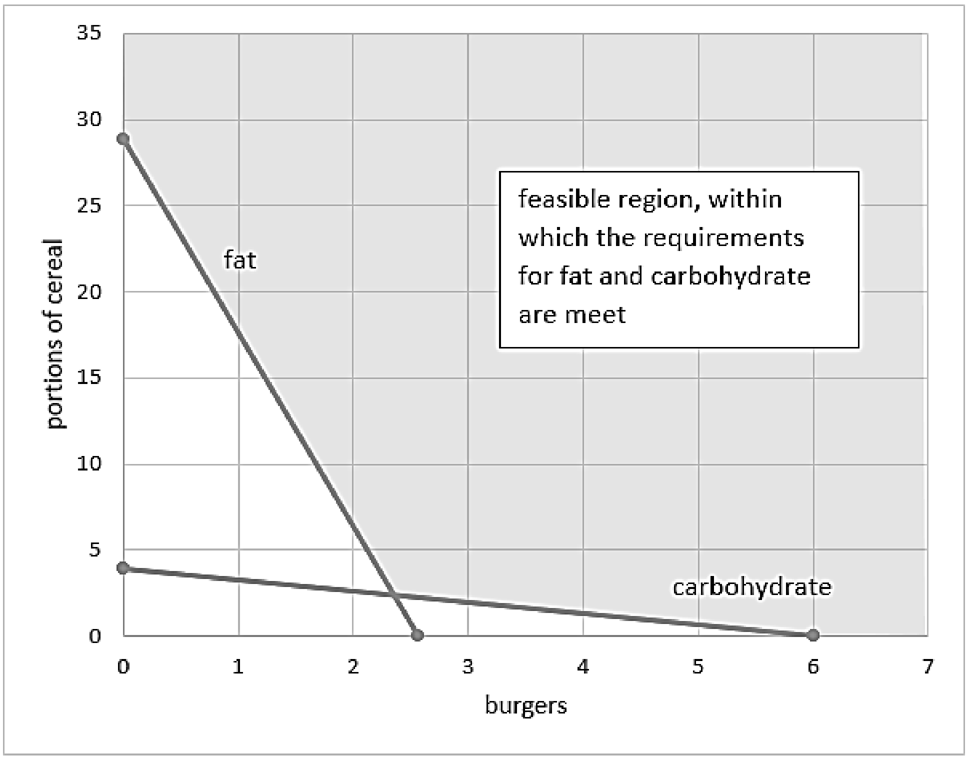


Figure 364: Feasible region which meets fat and carbohydrate requirements

The daily requirements for fat and carbohydrate can be met by any combination of burgers and portions of cereals within the shaded region of the graph, for example: 2 burgers and 10 cereal portions, or 3 burgers and 5 cereal portions. The next problem is to decide which combination of the foods would be cheapest.

We can set up an equation to represent the total cost of the food consumed, based on a price of £3.60 for a burger and 37p for a portion of cereal:

$$total\ cost = 3.60x_1 + 0.37x_2$$

Suppose that we wish to spend £15.00 on food. We could eat entirely burgers and no cereal. The number of burgers which could be purchased (ignoring for the moment the requirement of whole numbers) is:

$$\frac{15.00}{3.60} = 4.2$$

Supposing instead that we choose to spend our £15.00 entirely on cereal, the number of portions purchased would be:

$$\frac{15.00}{0.37} = 40.5$$

A line representing possible numbers of burgers or portions of cereal which could be purchased for £15.00 has been added to the graph in figure 365.

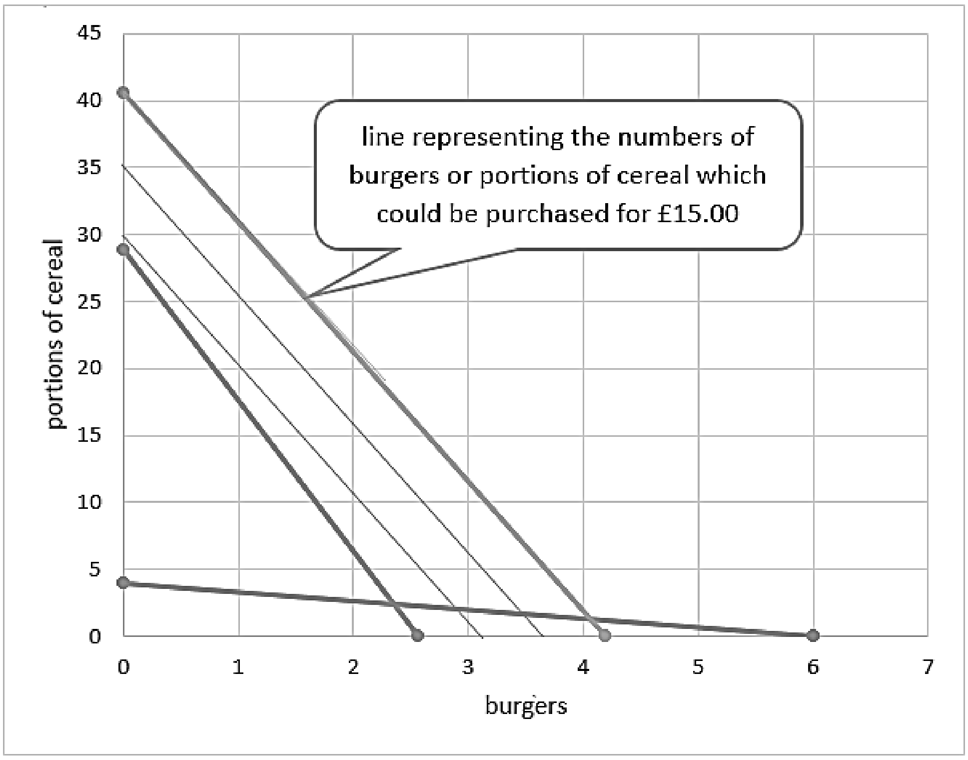


Figure 365: Cost of food items

If, instead of £15.00, we had chosen to only spend £14.00 on food, a new line could be plotted on the graph to show the numbers of burgers or portions of cereal which we could now purchase. This line would be parallel to the original line, but closer to the zero origin of the graph. We could continue to reduce the food bill in this way, producing a series of parallel lines ever closer to the graph origin. However, to meet the dietary requirement for fat and carbohydrate, we must remain within the feasible region of the graph. Lowest cost is achieved at the point where the fat and carbohydrate graph lines intersect. The region of the graph intersection is enlarged in figure 366:

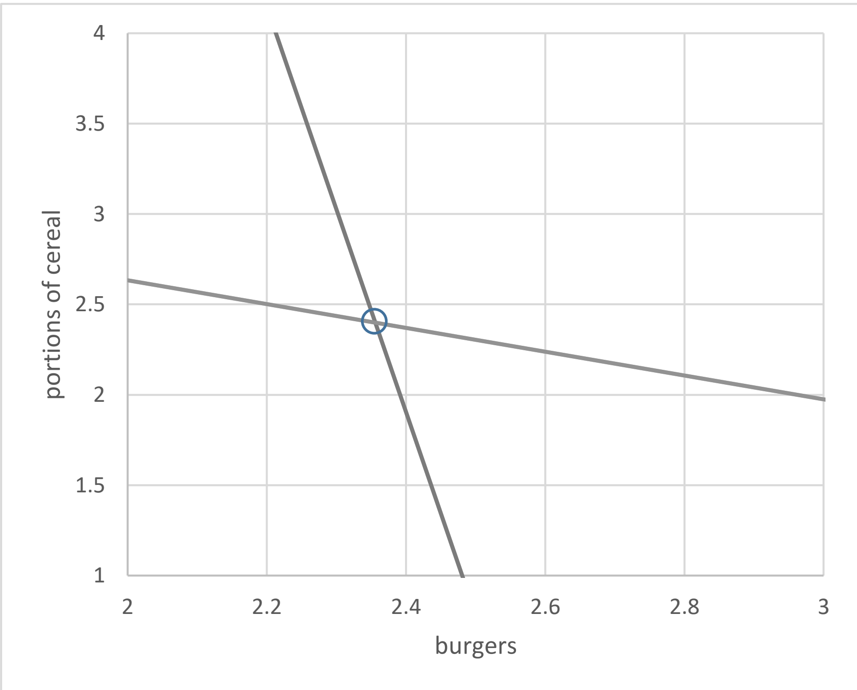


Figure 366: Point where dietary requirements are met at lowest cost

We deduce that the daily requirements for fat and carbohydrate could be met most economically by eating about 2½ burgers and 2½ portions of cereal.

This example has obviously been greatly oversimplified:

- In a real dietary situation, many more nutrients than just fat and carbohydrate would need to be considered,
- A far more varied diet would be available than just burgers and cereal,
- We would need to be sensible about specifying quantities. Whilst it might be possible to measure out 2½ portions of cereal, it would not be possible to order 2½ burgers.

The method we have used can be readily extended to any number of variables. However, it becomes impossible to solve more complex problems graphically. Solutions are usually obtained by means of linear algebra software applications. An example is **Linear Program Solver**, which can be downloaded from:

download.cnet.com/Linear-Program-Solver/3000-2053_4-75401989.html

To demonstrate the use of this application, we can re-run the burger and cereal example. We begin by entering the equations for food cost, fat and carbohydrate totals. Notice that we are asking the program to find a minimum for the cost which will meet the nutritional requirements.

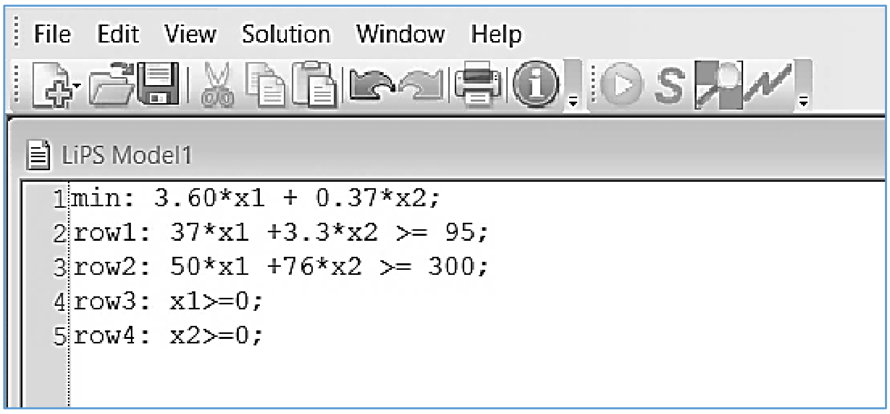


Figure 367: Entering a set of equations into the Linear Program Solver application.

The software then calculates the optimum solution to the problem. We find that the lowest cost is achieved by eating 2.4 burgers and 2.4 portions of cereal, which is consistent with the earlier graphical result. The program has calculated the minimum cost to be £9.36.

The project could be developed by adding further equations to represent other nutrient requirements, such as sugar or fibre. Maximum recommended daily quantities, as well as minimum quantities, could be specified. Additional food items could be introduced by adding further variables, such as x_3, x_4, to each equation.

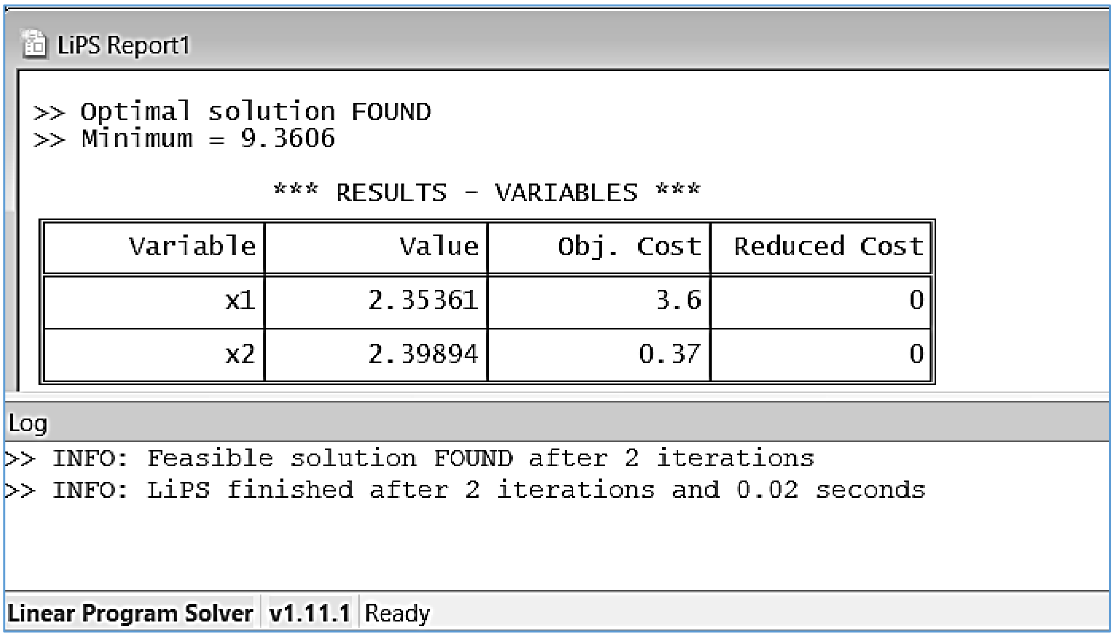


Figure 368: Output of results from the Linear Program Solver application.

In this chapter we have briefly reviewed a range of approaches to mathematical modelling. In each case, there have been opportunities to develop broader numeracy skills, including:

- Development of mathematical techniques in the context of specialist subject knowledge,
- Problem solving, which can be effectively carried out by experimentation,
- Use of a range of technologies, including spreadsheets and specialist mathematical application software,
- Communication, both in developing specialist mathematical terminology, and in being able to present complex mathematical results in a format with is easily understood by the lay person.

We have found that mathematical modelling can be an interesting and motivating activity for students in a wide range of vocational areas.