University of Liverpool
BSc Zoology

BIOL630: Honours Project

Analysis of data on dispersal in southern damselflies (*Coenagrion mercuriale*)

by Stuart David James McHattie

Supervisor: Dr David J. Thompson

2002 / 03

Abstract

The southern damselfly *Coenagrion mercuriale* is a rare species within the UK, of particular interest to conservationist groups. It is know that the damselfly is both a poor coloniser and the range over which it will travel is known to be very short. A good understanding of both how good a coloniser this species is and what distance it will travel to achieve this will help to induce more productive conservation of this endangered species; helping to recolonise nearby areas.

The main objective of this study was to find correlations between size of area and movement of individuals at ten sites distributed around the Hampshire area in the South of England. Other correlations that were hoped to show relevance were that between population density and movement of individuals; correlations with rate of movement and comparisons between males and females with respect to distances travelled and rates of movement. It is also intended to bring light upon the issue of how far *C. mercuriale* will travel around its habitat to discover further habitats and breeding grounds.

The results show that there are definitely correlations between population density and distances travelled (activity) but there are no significant differences between the studied sites with respect to rates of travel by *C. mercuriale*. It was also shown that there the sex of the individual has a significant effect on the activity of the insect but not on the rate of travel. Also, it was noted that individuals crossed the road between Upper and Lower Crockford with great hesitancy and on no occasion was an individual seen travelling further than just over one and a half kilometres.

With this knowledge, and the prospect of further investigation, it is possible to ensure that the conservation of *C. mercuriale* can be maximised and habitats of suitable nature only attempted to be conserved if they are within suitable range of existing habitats.

Acknowledgements

Table of Contents

Introduction

1.1 The story of *Coenagrion mercuriale*

Coenagrion mercuriale is a protected species of damselfly, both within Britain and throughout the world. The species has been listed under Annex II of the EU Habitats Directive and Appendix II of the Bern Convention.[i] Because of this, it has become a model organism for European conservation biology and it is very important that more is known about its conservation to maximise the chances of success.[ii] Within several European countries, Britain included, legislative measures have been taken to protect this species from harm. This produces a select few areas within the UK in which this species of damselfly can be found, and Hampshire is the home to some of the most important of all. Very few populations of *C. mercuriale* exist outside these Sites of Special Scientific Interest (SSSI) or Special Areas for Conservation (SAC).[iii]

The qualities of the habitat *C. mercuriale* will colonise in are well known, due to the analyses made by Buchwald (1989); small running waters near the spring or within the influences of ground water.[iv] But, however, it seems that finding the perfect habitat is the least of the worries of conservationists trying to protect this species and establish new populations. *C. mercuriale* appears to be a terrible disperser and coloniser, often only travelling a foot or two after disturbance, with males being generally confined to the stream in which they live and breed.[v] Although males can and have been found up to 0.5 km. or more from the nearest site after emergence and initial dispersal,[vi] this alone is not enough for colonisation to take place because the females are absent from this area and oviposition doesn't occur.

Many studies have been performed, over the last two decades in particular, to look into the dispersal behaviours of damselflies, but most have been disappointing, usually showing more or less no movement beyond the very close proximity of the stream in question.[vii] On two occasions only, in all of dispersal studies documented within the Journal of the British Dragonfly Society, have there been sightings of a Southern Damselfly travelling more than 1 km. from the first point of sighting.[viii]

Distinguishing the Southern Damselfly from other similar species can be an awkward venture, but the males are of a very bright distinctive nature compared to the females. In a key to distinguish the female of different species of damselfly[ix], it was possible to identify some characteristics specific to *C. mercuriale*: two black stripes on the side of the thorax (as opposed to one) with small eye-spots and a coloured bar between. The males have "mercury marks" on their back, as seen in figure 1 and of which it is possible to get 4 different types. These were recorded during the observation of *C. mercuriale* in this study.
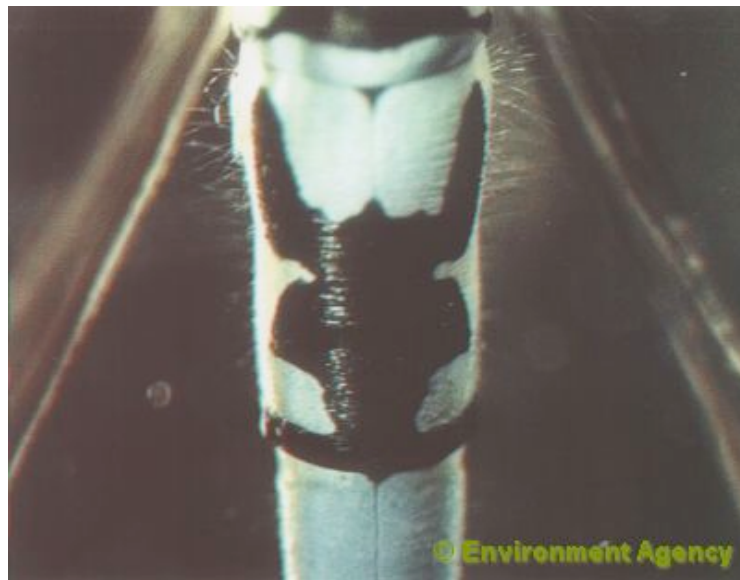
Figure 1 - The mercury mark of a male Southern Damselfly
(*Coenagrion mercuriale*)

It seems apparent in all populations of *C. mercuriale* that the males are predominant, but this is known to be because the males emerge earlier in the season than females and spend more time each day at the water, arriving before the females and being last to leave. Beyond this, males live for later into the season than the females.[x]

1.2 Objectives

The main objectives of this study were two-fold and as follows:

1) To find and assess the nature of correlations between the variables:

- Size of Area
- Population Density
- Sex of Insect

and the variables:

- Distance moved throughout the 34 day period (activity)
- Average rate of travel throughout the 34 days period (speed of movement)

2) To find out if any movement between sites of observation were made and calculate a figure which can be used as a guideline for maximum roaming distance in *C. mercuriale*.

1.3 The data

The data collated in this study was collected in Hampshire in the South of England over the summer of 2002 by Dr. D.J. Thompson and a number of university staff and students. It was collected initially in hand-written form and later converted to Excel as a spreadsheet by a team of transcribers. The final spreadsheet contained 17,846 lines of capture information for a 34 day period including the following columns of data:

- Site Name
- Date of Capture
- Recorder(s)' Initials
- ID Code of Insect
- Sex of Insect
- GPS1 (East-West) Coordinate
- GPS2 (North-South) Coordinate
- Day Number
- Mark or Recapture
- Time of Capture
- Morph of Female
- Mercury Mark of Male; and
- Comments

1.4 The sites

There were, in total, 10 sites within the study area. Table 1 summarises these areas and shows a few vital statistics about each of them.

Name of Site	Approximate Size of Site (m^2)	Total male observations	Total female observations	Total number of unique damselfly
Bagshot	34,000	136	27	34
Deepmoor	555,000	2225	285	496
Greenmoor	520,000	287	52	81
Hatchet	55,000	270	74	64
Roundhill	210,000	3094	530	869
Upper Crockford	770,000	1545	206	375
Lower Crockford	555,000	3765	396	1064
Upper Peaked Hill	1,240,000	4113	838	1091
Frogmoor	███████	1	0	1
Horsebush	███████	2	0	2

Table 1 - Overview of the Sites Observed

Methods

2.1 Overview

The data supplied was in a very raw format. None of the errors in either recording the data or entering it into a computer had been removed. With 17,846 instances of capture or recapture having been recorded with pencil and paper and then entered into a computer, it was of no surprise that some of the information had been corrupted in a very much "Chinese whispers" fashion. The data had also been entered in the order in which it was recorded and so when a damselfly was recaptured ten days after the original marking, this recapture may be 5,000 lines away from the original marking in the data. Therefore some method of rejoining this data together again was of great importance to its analysis.

Another point to be taken into consideration was that even damselfly that had not received any recaptures were within the data and, of course, without recaptures it is impossible to relate this data to movement and rate of travel around the surrounding area. Therefore all instances of no recapture had to be removed.

Even after all the data had been collated and irrelevant data removed, there was still the problem of placing the data in a useable format, leaving the window for automation by Excel as wide as possible to make use of the computing power available for the task. By increasing automation, the required results could be achieved much sooner than if the work was done by hand.

2.2 Sorting the Data

The data was initially in date of collection order, with all the recaptures of the individuals intermingled with each other. To separate these recaptures and place them in the order of their capture, the data sorting features of Excel were used. By specifying that the data should be sorted by ID code of insect and then by Date, Excel put the recaptures together using the ID codes, and then it separated the dates of the captures within individuals. The spreadsheet then began with the first individual caught with all its recaptures, followed by the next individual and its recaptures and

so on and so forth. This left the data in a format which would allow the formulae to separate the individuals out as they calculated the final figures and results needed.

2.3   Cleaning the Data

This stage of the process involves cleaning up the mess created naturally by misreading, mistyping and misunderstanding of the data in the process of transferring the raw data from pencil format into computer data within Excel. The tools within Excel proved very useful at this stage, however they were not all that was needed to correct this information and by using macros, a form of programming language for Excel, new tools were able to be created that were more specific to the task in hand.

2.3.1   Find and Replace

This standard feature of Excel, found on the Edit menu is able to parse a large number of cells and identify those which meet specified conditions. Upon finding cells meeting these conditions, it is able to replace sections of the cell to meet new conditions. For example, it can find all cells containing the symbol for Pounds Sterling (£) and replace them with the symbol for U.S. Dollars ($) with one click of a button. This was used in two instances of correcting the data, which were as follows:

2.3.1.1   Correcting the Time

In order for Excel to recognise a time, it must be entered using a colon between the hours and the minutes. As the data was entered by a number of people, some of the data was entered correctly with the colon, but however some of the data was entered using a decimal point instead, and this caused Excel to interpret this figure as a number instead of a time. By using find and replace, it was possible to get excel to find these entries and change them back to a colon. After this correction, the entire column was interpreted by Excel as times.

2.3.1.2   Correcting the Date

In order for Excel to recognise a date, it must be entered using slashes between the day, month and year. This was quite complicated because the spreadsheet shows the date in the format of "DD-MMM" meaning that 02/07/02 would look like this "02-Jul", but however, behind the scenes, Excel is still thinking of the date as 02/07/02. Because of the way the spreadsheet looks to the user, some people had entered the

dates as they saw it, "02-Jul", and this caused Excel to interpret the date as text. To correct this, find and replace was used to convert the dash into a slash on these cases. Excel then interpreted the new layout as a date and after this processing; the date column consisted entirely of dates.

2.3.2 Format Changing

In order for formulae to work with numbers, they must be in the correct format so that Excel can interpret them correctly. Within the data, some of the coordinates went over the 100,000 metre boundary on the map, and therefore began with the digit zero; as all coordinates are made up of 5 digits. Excel would not normally display these preceding zeros on a number and so when someone had entered coordinates beginning with zero, and because the coordinate is 5 digits long, Excel has made them into zip codes. Formulae within Excel cannot work with zip codes in the same way as numbers and so these had to be changed.   By highlighting the entire of the two coordinate columns and choosing the number format, these zip codes were converted back to numbers. The only main drawback was that now the coordinate was not displayed as 5 digits, but this was overcome by then choosing a custom number format and using the custom arrangement of "00000" which would tell Excel to place any leading zeros before the number if it was shorter than 5 digits.

2.3.3 Paste Special

Even after telling Excel to change the coordinates back to numbers, some people had entered the coordinates as text, in an attempt to solve the problem of the missing preceding zeros. This worked perfectly to display the zeros, but left the data useless when the formulae were working on the coordinates. By using Paste Special, it is possible to multiply large numbers of cells by the same figure. Therefore, by entering the number "1" in a spare column and using the standard Copy feature within Excel, the columns of coordinates could be selected and then Paste Special used to multiply them all by 1. This left the cells already containing a number unaffected, but attempted to multiply the text entries by 1. This attempt at multiplying text by a number caused Excel to convert the text to a number to complete the process. After this process, none of the coordinates were anything but numbers and they were ready for the formulae to use them.

2.3.4 Macros

Macros allowed for much more specific changes to the data. Being a programming language means that it can be made to do anything you need it to, within the boundaries of mathematics and logic. The coordinates were all now converted to numbers, and those beginning with zeros were displayed as such, but in order for the formulae to interpret these coordinates (all less than 10000) as being above coordinates of 95000 to 99999 on the map they had to be interpreted as a larger figure. For this reason, a macro was created which would check every cell in the coordinate range; and for each it checked whether the coordinate was a figure lower than 10,000. When it found cells meeting these requirements, it then added 100,000 to them, leading to a figure that reads the same as before, but with a "1" on the beginning; therefore increasing the figures sufficiently to make them appear farther north on the map. I.e. 00096 was converted to 100096 which is now higher than 99999 (the highest coordinate before moving onto 00000). This left the coordinates ready to be interpreted using formula such as Pythagoras' theorem. The code used to do this can be found in Appendix A section 6.1

2.3.5 "If" Statements

An "If" statement in Excel is a form of Boolean expression in which the result of the argument is either true or false and the contents of the cell containing the "If" statement varies with the outcome of the argument. This means that it can make comparisons between data on a spreadsheet and leave a cell reading either one way or another. This was used, in this instance, to find data for deletion or correction. It was used three times in total:

2.3.5.1 Deleting Non-Recaptures

By creating a new column on the spreadsheet an "If" statement was used in all the cells in that column to compare the lines of data. When Excel found a line of data containing just the marking of an individual and no recaptures, it put red writing saying "DEL" in the column adjacent to this line of data. After Excel had finished doing this, it was possible to filter this column for any instance of the word "DEL" and then to delete every line by selecting them and removing the lines from the spreadsheet. After this was complete, the only data left was that of recaptured individuals.

2.3.5.2 Deleting Recaptures within the Same Day

Again, by using an "If" statement in the same way as above, but to identify instances where an individual was seen twice in the same day, it was possible to delete one of the captures; as the formula shown later, which calculates rate of travel, would be dividing the distance by the days taken to travel the distance: it was undesirable to have errors where Excel would be dividing by zero. After these deletions of data 11,436 lines remained.

2.3.5.3 Changing the Status of the First Capture

As the column showing the Mark or Recapture of an individual was entered by hand, there was potential for error which could be shown in the form of Marks being listed as Recaptures or Recaptures being listed as Marks. This would present problems later when the formulae would rely on this information to be correct in order to distinguish between groups of individuals' data. The formulae would need to know which lines were the markings of the individual so as not to compare that line of data with the one above, which would show the final recapture of the individual before. To find these errors, "If" statements were used in the same way as before but this time to find lines of data where the ID code of the individual on two lines was different, but they both displayed Recapture as their status, which is impossible once the data is in date order and each individuals' data was grouped together. This would allow manual editing of these identified lines of data to make sure the Marking of an individual was correctly identified by the Mark or Recapture column.

2.4 Filtering the Data

Once all the sorting and cleaning of the data was complete, the data needed filtering to divide the information off into separate worksheets where they could be worked on in smaller quantities. Excel has built in filtering features to pick out specific lines of data meeting conditions specified. There are two ways to filter: by using auto-filter which gives a list of values from a column and lets you specify the particular value you want to pick out; or by using the advanced filters which could be made much more specific and allow the choosing of data which isn't yet in the column. Auto-filters were sufficient for the tasks in hand. By separating the data, the work load on the computer could be reduced to avoid crashes when running macros that automated parts of the process.

2.4.1 Separating the Sites

Moving the sites onto different worksheets seemed to be the most obvious and easy way to split the data up into manageable pieces. This ruled out the analyses of movement between sites, but these were performed separately as the data involved was much smaller in quantity. Separating the sites was done by creating new worksheets within Excel and then using the filter feature to filter the site column, one site at a time, copying the data selected and pasting it into the relevant worksheet.

2.4.2 Separating the Sexes

Going through each of the newly created worksheets and filtering the sex column for females meant that the female data could be cut from the site's data and placed below the males. This way, the female data could be assessed separately to the males and compared for differences at the analysis stage.

2.4.3 Find Errors in the Coordinates

The coordinates of each site were to be restricted to the area of observation around that site only, but however some of the coordinates turned out to be outside these ranges. For each site, the coordinates were filtered to show those which were outside the specified range. For most of these it was obvious what they were supposed to read, but others were less obvious. Mistakes were made by either mistyping the coordinates on the keyboard or by misreading the coordinates that were written down (e.g. reading a badly drawn 6 to be a 0). After tweaking these coordinates to fit the ranges expected, the data was as clean and prepared as it could be for the analysis of the data.

2.5 Analysing the Data

After the cleaning, sorting and filtering of the data was complete, the data was ready to be analysed. This was done in several easy stages which together combined into a very powerful conversion of the data into something which would be useable and would practically answer the questions originally posed.

2.5.1 Distance between Recaptures

Analysing the distance between recaptures was simply the use of Pythagoras' Theorem on the GPS coordinates of two lines of data. The result of this went into a

new column at the end of the data. The formula was only necessary on lines in which the Mark or Recapture status was "R" so that it would compare recapture data with that of the line before. This way, the formula would never calculate the distance between the final recapture of one individual and the marking point of the next. To make this possible, an "If" statement was used which checked the status of the Mark or Recapture column before inserting the formula. An example of the type of formula used is as follows:

$$=\text{IF } (\textcolor{red}{K2=\text{"R"}}, \textcolor{blue}{\text{SQRT}((I2-I1)\text{^}2+(H2-H1)\text{^}2)}, \textcolor{green}{\text{""}})$$

The part in red is the Boolean function. If the Boolean result is true (i.e. the line of data is a recapture), the cell displays the blue part which is the Pythagoras' Theorem formula for the distance travelled. If the Boolean result is false (i.e. the line of data is a marking), the cell displays the green part (i.e. nothing).

2.5.2 Time between Recaptures

To analyse the time gap between recaptures, an almost identical formula was used as to the distance between recaptures. The only main difference was that the main part of the formula was simply a subtraction of the days between both lines of data to show the time period between the recaptures. Again, it was important to make sure the formula only showed a result on the lines containing data of a recapture, so another "If" statement was used. The formula used was as follows:

$$=\text{IF } (\textcolor{red}{K2=\text{"R"}}, \textcolor{blue}{D2-D1}, \textcolor{green}{\text{""}})$$

Once again, the red part is the Boolean function and if that part is true the blue part is expressed, or if it is false the green part is expressed.

2.5.3 Rate of Travel

The rate of travel was analysed with yet another "If" statement and a simple division which divided the distance travelled in that recapture by the time between observations. The formula looked similar to the following:

$$=\text{IF } (\textcolor{red}{K2=\text{"R"}}, \textcolor{blue}{P2/Q2}, \textcolor{green}{\text{""}})$$

Once again, the red part is the Boolean function and if it is true the blue part is expressed and if it is false the green part is expressed.

2.5.4 Total Distance Travelled and Average Rate of Travel

The total distance travelled and the average rate of travel proved to be more tricky to calculate. The reason for this is that, being a total and an average, there is only one formula to be displayed per individual and each individual had a different number of recaptures. This meant that not only was each formula placed an irregular distance from the previous formula (meaning that copy and paste couldn't be used to quickly fill them in) but also the formula in each case would be covering a different range of data (which definitely ruled out copy and paste). To overcome this problem, the only answer was to use another macro, which could be customised to the job in hand and made more intelligent; enough to interpret the data in much the same way as a human who would enter the formulae by hand.

The macro effectively worked out which part of the data was males and which part was females. It then goes down the list looking at the column with "distance between recaptures" in it. It then looked for the gaps in that column, which is the changeover between one individual and the next. It works out how many pieces of data are present in each individuals' dataset by counting as it finds the next gap. Once it has done this it generates a formula appropriate to each individual and places it in the next available column to calculate total distance travelled. It also generates a second formula which works out the average rate of travel. It then starts looking again and continues to do so until the entire set of data is covered. It is an adaptive piece of code as well, which means that it will adapt itself to suit the datasets of each worksheet. This proved useful to avoid having to recode the macro for each worksheet. The code for this macro can be found in Appendix A section 6.2.

2.5.5 Placing the Data into Ranges

The biggest challenge of the project, in terms of technical complexity, was sorting the data from each dataset into ranges. This involved finding the frequency of individuals found travelling a total distance of between 0 and 25 metres and placing that number into a table to show this; then finding the frequency travelling between 25 and 50 metres and doing the same, all the way up to the frequency of those travelling more than 400 metres. That is as well as keeping the males and female separate and also doing the same type of analysis for average rate of travel. To do this by hand would be laborious and inaccurate, besides taking many hours. However, without the use of macros, this would be the only choice.

By using a very long and complicated macro, the computer would do this scan for itself and for every result it found on the spreadsheet it categorised it within the given ranges and then displayed the results at the end. The macro begins by finding out how many males and how many females there are on the spreadsheet to keep them separate. It then finds each result, categorises it, then moves on, storing all the frequencies in memory. Once every result has been looked at, it places the frequencies in a small table at the top right of the spreadsheet. This, for the computer, takes a maximum of about 3 minutes, which would take a human in the region of 3 days solid labour. It took about 20 minutes to write the code and another half hour to perfect it so that it would work flawlessly in all the spreadsheets which it needed to work within. It is self adaptive, in the same way that the code from the previous section was. The code for this macro can be found in Appendix A section 6.3.

2.5.6 Collating the Data

Having completed the generation of formulae, all the data was ready to be displayed graphically, but to make it easier to do this, the data was moved onto a final worksheet called "Overview". This sheet has all the data of frequencies for both males and females; for both total distances travelled and average rates of travel; and for all the sites. Also included on this sheet are details of movement between sites. This was found by hand when all the data is sorted but not separated into different sites. Finally, at the top of the sheet is a summary of the quantities of data found on each site, the maximum and minimum GPS coordinates of the site and the approximate area of the site.

2.5.7 Final Preparation and Graphical Display

In order to produce graphical displays which would be fair comparisons of the frequencies between sites and sexes, the values for frequency had to be converted into percentages. A percentage of 100% in the 0 to 25 metres range for total distance travelled would mean that all the flies from that site and of that sex travelled no more than 25 metres.

The final stage of analysis was to make graphs of the data found. These would demonstrate either the activity or rate of travel of flies from a particular sex and/or

site and compare it to those of the other sex or sites. The graphs and tables produced are as shown in the results section.

Results

3.1 Size of Area and Population Density

In order to compare the outcome of the data in relation to size of area and population density, it is important that these values are first evaluated so that it can be seen which sites are the smallest or least dense population and which are the largest or most dense population. For this reason, these values are tabulated and then placed into a rank order with 1 being the most dense population or largest area and 8 being the least dense population or smallest area. It is important to realise that the approximate size of the site is calculated by placing a rectangle that encloses all the data points (observations) of any particular watercourse and using the area of that rectangle as the size of the area. Also, the density of the population is not necessarily the true density of the population as it is inevitable that not all the individuals of the population were observed, but because the size of the area was calculated from the position of the observations, and the density is calculated from this figure and the number of observations made, it should give a relatively accurate estimate of the true population density. Another important note to make is that Horsebush and Frogmoor are both missing from this data because they presented only 2 and 1 observations respectively which were useful for observing inter-site movement, but not for looking at movement over a period of time.

3.1.1 Direct Comparison of Figures

The first stage is to see what values are present in the data:

Site	Approximate Size of Site (m^2)	Relative Density of Population (observed individuals per 1,000m^2)
Bagshot	34,000	1.00
Deepmoor	555,000	0.89
Greenmoor	520,000	0.16
Hatchet	55,000	1.16
Roundhill	210,000	4.14
Upper Crockford	770,000	0.49
Lower Crockford	555,000	1.92
Upper Peaked Hill	1,240,000	0.88

Table 2 - The Size and Population Density at each Site

3.1.2 Placing the Sites into Rank Order

In order to relate the size of area and population densities to any correlations, they must be placed into order so that any arising correlations will either suit the rank in forward or reverse order:

Rank	Name of Site	Size of Area (m^2)
1	Upper Peaked Hill	1,240,000
2	Upper Crockford	770,000
3	Lower Crockford	555,000
3	Deepmoor	555,000
5	Greenmoor	520,000
6	Roundhill	210,000
7	Hatchet	55,000
8	Bagshot	34,000

Rank	Name of Site	Population Density (x per 1,000m^2)
1	Roundhill	4.14
2	Lower Crockford	1.92
3	Hatchet	1.16
4	Bagshot	1.00
5	Deepmoor	0.89
6	Upper Peaked Hill	0.88
7	Upper Crockford	0.49
8	Greenmoor	0.16

Table 3 - Placing the Sites into Rank Order by Size of Area and Population Density

Note that there are two entries at the third rank for size of area, because the areas were of comparable size.

3.2 Comparing Total Distance Travelled between Sites

Using the data collated on the overview worksheet, the percentages that were calculated for the combined males and females with respect to total distance travelled were plotted:
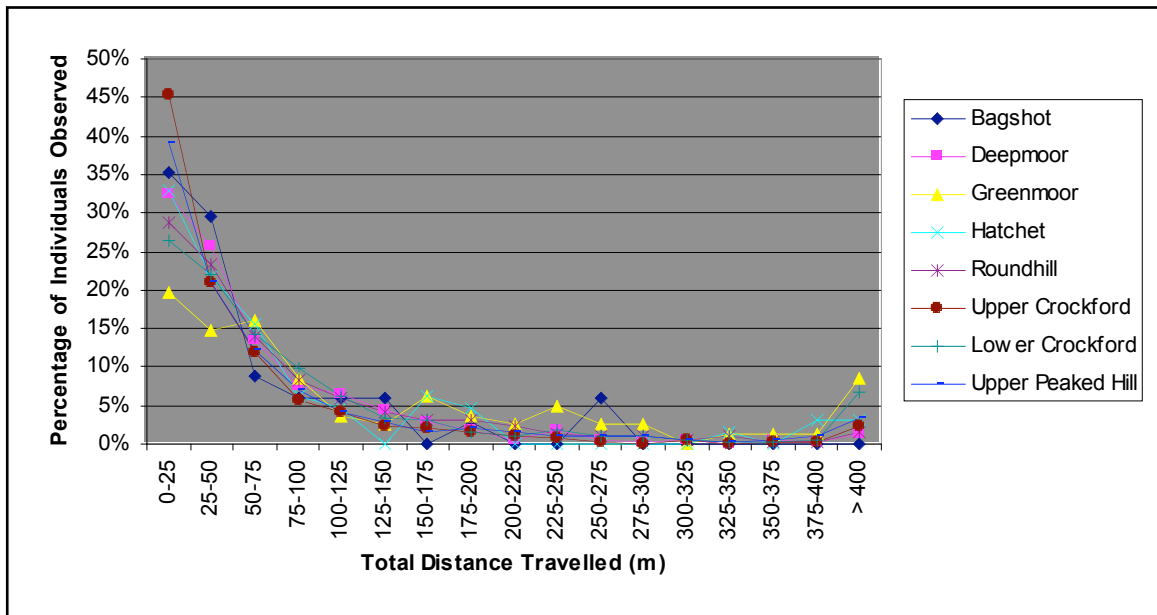
Figure 1 - A Comparison of Total Distance Travelled against Percentage of Individuals Observed

In order to compare this to Size of Area and Population Density, a table was constructed which looks at the order in which the sites appear on figure 2:

Order of Sites at First Data Point on Figure 2 (uppermost first)	Rank Order for Size of Area	Rank Order for Population Density
Upper Crockford	2	7
Upper Peaked Hill	1	6
Bagshot	8	4
Hatchet	7	3
Deepmoor	3	5
Roundhill	6	1
Lower Crockford	3	2
Greenmoor	5	8

Table 4 - Comparison of Total Distance Travelled against Size of Area and Population Density

3.3 Comparing Average Rate of Travel between Sites

Using the data collated on the overview worksheet, the percentages that were calculated for the combined males and females with respect to average rate of travel were plotted:
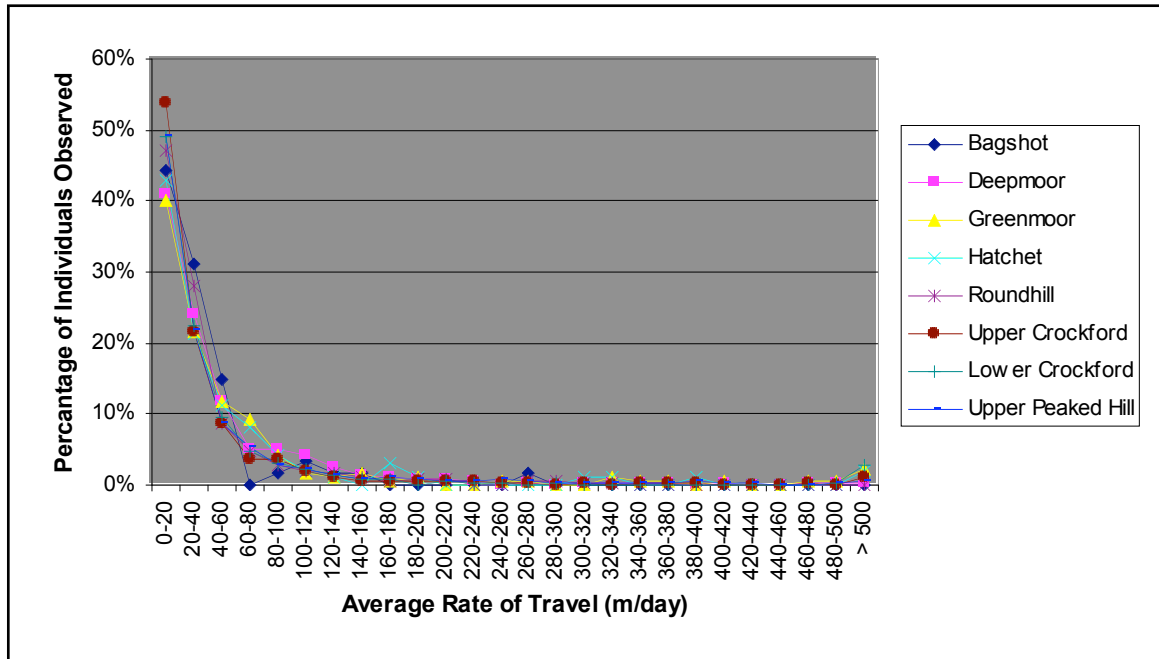

Figure 1 - A Comparison of Average Rate of Travel against Percentage of Individuals Observed

It is obvious from the graph that there is no significant difference between the sites with respect to average rate of travel, therefore it is not necessary to look at the rank orders under these circumstances.

3.4 Comparing Total Distance Travelled between Sexes

By combining the data across all sites for males and then females, a fair comparison between the two sexes can be made. A histogram relating total distance moved in males against total distance moved in females was produced:
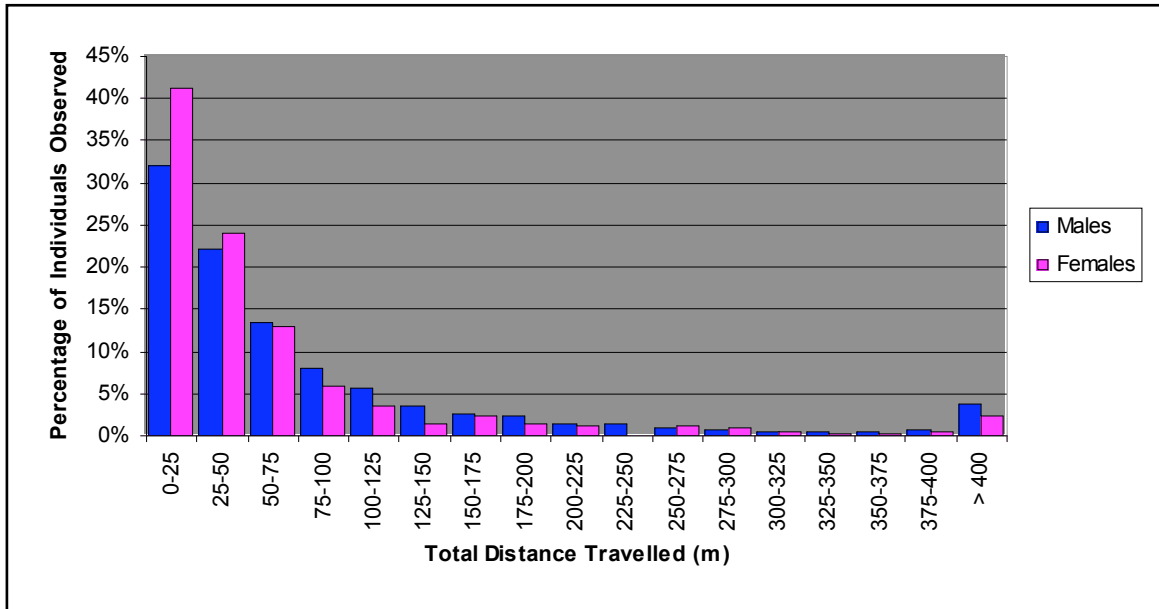

Figure 1 - A Comparison of Difference between Sexes relating to Total Distance Travelled

The plot clearly shows that female bars are higher towards the left hand side of the graph and the male bars only succeed in gaining more height than the females towards the right hand side of the graph, up to around 200 m. distance, after which they are roughly equal.

3.5 Comparing Average Rate of Travel between Sexes

Using combined data once again, it was possible to generate a histogram comparing the average rate of travel in the males to the average rate of travel in the females:
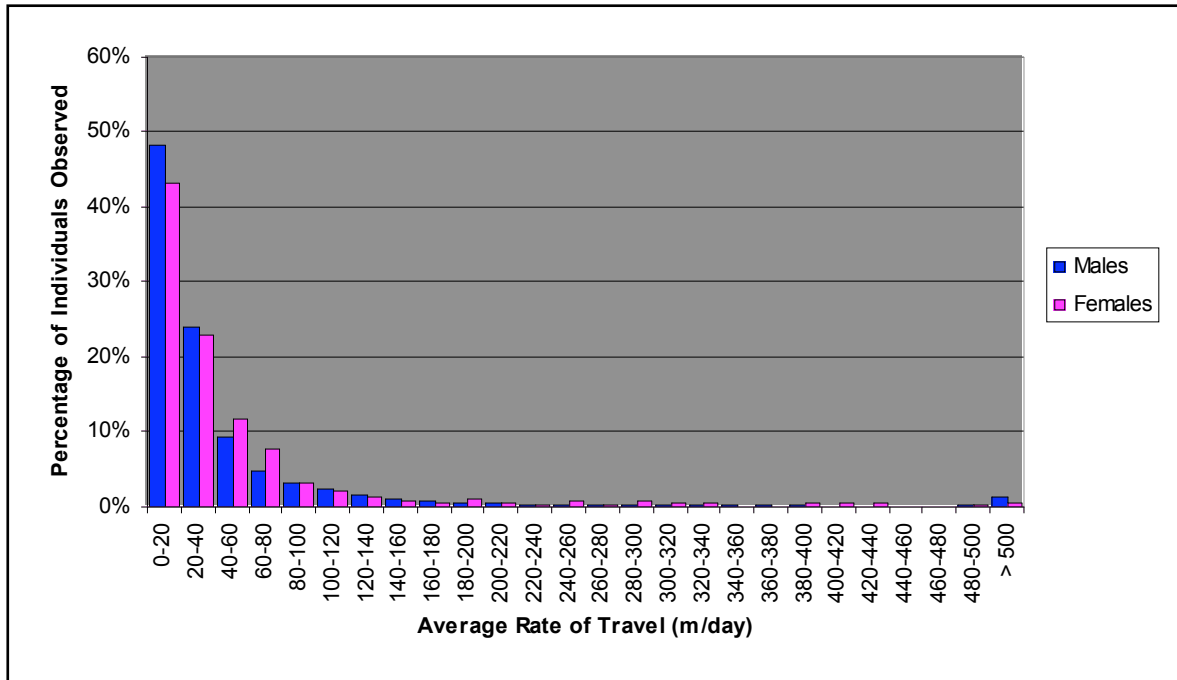

Figure 2 - A Comparison between Sexes in Relation to Average Rate of Travel

The plot is clearly illustrating that the males tend to have higher bars on the left hand side of the graph, whilst the females tend to have higher bars on the right hand side of the graph, up to the point at which both have bars of comparable height from 80 m/day and upwards.

3.6 Using ANOVA for Confirmation

By using a one-way ANOVA on the frequencies of individuals fitting each range of Total Distance Travelled at each Site, it is possible to see whether there is statistical evidence showing that the sites vary from one another with respect to Total Distance Travelled. The outcome of such a test gives a P-value of 0.002, which shows a highly significant difference between the data sets: in only 1 in 500 sets of data showing this outcome, there will be no significant difference between the data sets. This means that any trends found are likely to be of significance as the different sites do have differences in the Total Distance Travelled of individuals.

By doing another one-way ANOVA on the frequencies of individuals fitting each range of Average Rate of Travel at each Site, it is possible to see whether there is statistical evidence showing that the sites vary from one another with respect to Average Rate of Travel. The outcome of such a test gives a P-value of 0.093, which shows there is no significant difference between the data sets. This means that any correlations that are drawn from this data are likely to be there by chance alone because the variations between each site are also likely to be produced by uncontrollable factors which have caused disturbances in the data giving the impression of differences between the sites.

3.7 Using t-tests for Confirmation

By doing a t-test on the combined frequencies of all the sites for the Total Distance Travelled, assessing the data sets of the two sexes, it is possible to see if there is a significant difference between males and females. This will indicate whether the sexes behave in different ways to one another or if the differences in behaviour are due to circumstances out of the control of the study. The P-value for this test was 0.030 which shows that there is a significant difference between the two sexes with respect to Total Distance Travelled. Only 3% of datasets showing this result would have differences due to chance alone. This means that any differences seen between the sexes are in fact significant.

By doing the same t-test on the datasets showing the Average Rate of Travel of both sexes, there is a P-value of 0.087 which shows that there is no significant difference

between the sexes with respect to Average Rate of Travel. This shows that all differences seen between the two sexes on a graph of Average Rate of Travel will be most likely due to fluctuations out of control of the study.

3.8 Inter-site Movement

There were, in total, 42 individuals of the 4077 observed that passed between sites. Of these, 6 were female and 36 were male. In the following table, the frequencies of individuals moving between sites are shown and there is an indication of the shortest distance observed in this movement and the longest distance observed:

Between	and	Frequency	Shortest Distance (m)	Longest Distance (m)
Lower Crockford	Upper Peaked Hill	23	77.9	927.7
Deepmoor	Upper Crockford	9	28.4	630.2
Bagshot	Upper Peaked Hill	5	684.6	1105.2
Upper Crockford	Greenmoor	2	1046.0	1097.3
Bagshot	Lower Crockford	1	1165.1	1165.1
Greenmoor	Frogmoor	1	1525.2	1525.2
Upper Crockford	Lower Crockford	1	968.6	968.6

Table 5 - Frequencies of Inter-site Movement and the Distances Travelled

Discussion

4.1 Recapping the Objectives

In order to make the most use of the results of this study it is very important to answer the questions originally posed. For this reason, it is a good idea to recap exactly what the study intended to achieve. The objectives of the study are two-fold and as follows:

1) To find and assess the nature of correlations between the variables:

- Size of Area
- Population Density
- Sex of Insect

and the variables:

- Distance moved throughout the 34 day period (activity)
- Average rate of travel throughout the 34 days period (speed of movement)

2) To find out if any movement between sites of observation were made and calculate a figure which can be used as a guideline for maximum roaming distance in *C. mercuriale*.

4.2 Comparing the Variables

In order to answer the first part of the objectives entirely, it is possible to make every cross between the variables so that every aspect has been fulfilled. This means that six crosses are made, and are as follows:

- Size of Area versus Total Distance Travelled
- Population Density versus Total Distance Travelled
- Sex of Insect versus Total Distance Travelled
- Size of Area versus Average Rate of Travel
- Population Density versus Average Rate of Travel
- Sex of Insect versus Average Rate of Travel

4.2.1 Size of Area versus Total Distance Travelled

Figure 2, from section 3.2, leads to the production of table 4, which shows how Total Distance Travelled relates to the Size of Area. Figure 2 shows that there is a trend of very little movement in *C. mercuriale* during the 34 days period, with most travelling

less than 150 metres or less throughout; and the vast majority travelling less than 25 metres. There is an obvious difference between the sites, with some showing nearly 50% of the observed damselflies moving less than 25 metres, and others having less than 20% of the observed damselflies moving the same distance and greater numbers of them moving longer distances. This is proven when the ANOVA test (section 3.6) performed on the frequencies of Total Distance Travelled shows that the differences between the sites are not due to chance alone. To compare the size of the area to the movement of the individuals, the rank of the sites in relation to size of area were compared to the first plot of figure 2; that is, the percentage of individuals travelling less than 25 metres. The results of this are shown in table 4.

The left column of table 4 shows the sites in order as they appear on the graph. If there is a correlation between the size of area and the total distance travelled the numbers indicating the rank of each site will read in either forward or reverse order in the middle column of table 4. The order presented is 2 - 1 - 8 - 7 - 3 - 6 - 3 - 5. There are two sites in 3rd place because they show almost identical areas. However, it is obvious that there is no correlation as the numbers are almost in a random order. This proves that *C. mercuriale* are not more or less active due to the size of their habitat.

4.2.2 Population Density versus Total Distance Travelled

Population density is almost the same as size of area, but with one key difference, it takes into account the number of damselfly living in the area. A small area with very low numbers of damselfly will have very low population density compared to a very large population on a large area. One of the richest sites was Roundhill, which had nearly four times the population density of any other site, despite being one of the most remote and medium to small sized areas. When looking at table 4 again, this time concentrating on the left and right hand columns, it can be seen that the order of the ranks of population density are as follows: 7 - 6 - 4 - 3 - 5 - 1 - 2 - 8. With the exception of Greenmoor coming last in that order, the numbers are almost in reverse order. This shows that there is a relatively strong correlation between total distance travelled and the population density. It appears that *C. mercuriale* living in lower population densities tend to be less active than those of higher population densities. This could be due to the damselflies interacting with each other more and causing disturbances. Although it was noted by D.K. Jenkins that *C. mercuriale* will only

move a few feet at a time after disturbance,[xi] being disturbed repeatedly by high chances of meeting another individual could lead to movements of long distances.

4.2.3 Sex of Insect versus Total Distance Travelled

Figure 4, in section 3.4, shows that there is a distinct difference between males and females in relation to total distance travelled. This is proven by the t-test performed on this data (section 3.7) which shows there is a significant difference between the two sexes with relation to total distance travelled. Looking at the graph, it is possible to see that females were seen to be less active than males with the peaks of the bars being higher on the left side of the graph, and the males having higher peaks from about 50 up to about 200 metres. After this point, the bars are of similar height for both males and females. This would give the impression that the females do less movement than males in general, but it would seem this is an unlikely conclusion unless the females come to the stream to breed and wait for the males to search for them, which is a distinct possibility. It is more likely, however, that the lower activity of the females is a by-product of their lower abundance in the population. As described in the introduction and in papers such as those by D.K. Jenkins[xii] and David J. Thompson and Bethan Purse[xiii], the females are seen to be less abundant in mark/recapture exercises than males by around four times. This is due to the males being first to hatch in the season and the last to survive at the end of it, as well as being the first on and last off the water during each day. By observing a fewer number of females during this study, each individual female will have received a fewer number of recaptures than the average number in males. This would lead the results to show that females are less active than males, but it would be impossible to say whether this is the only influence that causes the trends from figure 4 or if there is an underlying reason for females to appear less active.

4.2.4 Average Rate of Travel

Figure 3, in section 3.3, shows that there is little to no difference between the sites in relation to average rates of travel. This is supported by the ANOVA test performed on this data (section 3.6) which showed that there is no significant difference between the sites in relation to average rate of travel. Looking at figure 5, it would seem apparent that there is a relationship between sex of insect and average rate of travel,

but upon doing a t-test (section 3.7) to test this theory it becomes obvious that there is no difference between males and female with relation to average rate of travel.

4.3 Looking at Inter-site Movement and Long Range Dispersal

Only 42 individuals of the 4077 observed passed between sites. This demonstrates how much of a rare occurrence this is. It has been known for a long time that movement by *C. mercuriale* is very limited and that after initial emergence and dispersal, males are largely confined to their individual stream.[xiv] Of the 42 individuals, 6 were female, which is as expected with roughly 25% of all populations being made up of females. It is unlikely therefore that movement between sites is beyond the boundaries of female movement.

Table 5, in section 3.8, shows that sites of closer proximity tend to get more movement between them, which is as would be expected. Reports made by D.K. Jenkins[xv] and David J. Thompson and Bethan Purse[xvi] describe that no individuals of *C. mercuriale* have really been seen moving more than just over 1 kilometre, but during this study one individual was seen travelling more than 1 and a half kilometres. The reason for this is not apparent, but it has been noted by E.D.V. Prendergast[xvii] and many other authors that weather can have a very strong effect on the density of *C. mercuriale* on as short a period of time as making big daily changes. This could mean that there is an increased population density on some days which could in turn lead to increased movement and cause drifting from site to site. Another way to look at this explanation is that perhaps the strength of the wind picked up on particular days and caused involuntary movement between sites by blowing the damselfly around.

It is suggested that the dispersal of *C. mercuriale* is made mostly immediately after emergence and after becoming a full adult, the damselflies spend the rest of their lives within a small area around the stream[xviii], but it is almost impossible to observe this dispersal because the cuticle of the insect is not developed enough to mark without damaging immediately after emergence, and whilst the insect may be retained whilst the cuticle hardens; by the time it has hardened sufficiently to allow marking of the insect, it is likely that the dispersal part of the life cycle has been passed and release of the insect will no longer present the same levels of dispersal as if the insect was left undisturbed immediately after dispersal.[xix]

One result of particular interest is the movement of only one individual across the road which creates the boundary between Upper and Lower Crockford. This road is a small barrier to cross, yet the movement seems limited. A study by Margaret Feneley[xx] was based on the dispersal behaviour of *C. mercuriale*, with particular interest in this road. It was found that whilst a few individuals were seen crossing this road; it was probable that this was caused by involuntary movement due to strong gusting winds. However, after removal of some of the inhabitable foliage on either side of the road, it was observed that more damselflies were crossing this barrier. It may therefore be that the road is not the blockage causing this lack of mixing between the sites, but in fact the inhabitable areas of land either side. It is also suggested that because some individuals show evidence of being able to travel long distances without inhibition, it is known that it would not be impossible for *C. mercuriale* to travel to new habitats. So by creating "corridors" between sites of current established populations and areas of suitable, but inhabited, habitat, it may be possible to encourage *C. mercuriale* to establish in new areas.

4.4 Further Study Aspects

Further uses of the data from this study are possible, and these include investigation of movement in relation to factors for which information was not available during the time of this project. Several possible relationships are discussed, sometimes in great detail, in the literature, between movement and items such as the weather and the pH of the water. By combining the data from this study with that of studies by the rivers authorities and weather observation organisations, there may be evidence that these factors do or do not have an effect on the movement of *C. mercuriale*.

It was noted by E.D.V. Prendergast[xxi] that from day to day, the presence of C. *mercuriale* was changing in his studies and that it seemed to be closely related to the weather conditions. Knowing that higher densities of population can cause greater amounts of movement, this may directly link weather to movement as well.

In a study by D.K. Jenkins[xxii], it was seen that Upper Crockford stream showed a pH of between 8.02 and 5.94, whilst Lower Crockford stream had a pH between 7.25 and 6.51, which is a much smaller range of values. With enough manpower it would be

possible to survey a large area of the stream all at once to try to relate movement of *C. mercuriale* to the pH of the stream around the area of observation. By keeping all results within one river course, there will not be the complications of introducing sizes of area and population density differences between the different observation sites. Crockford Stream is particularly good for this study as it shows a gradient of pH along the stream; pH rises further downstream.

4.5 Conclusions

In conclusion, *Coenagrion mercuriale* will only travel very short distances throughout its lifetime, most only being seen travelling less than 150 metres total; although some individuals are able to travel a lot more than this.

Higher population densities give rise to greater distances of movement, possibly because the chances of meeting another individual increase and individuals begin to disturb each other and cause movements in this manner.

Statistical tests show that the average rate of travel is neither significantly different between sites nor between sexes.

Females seem to show a less active lifestyle than males and this is proven by statistics, but however this may be a by-product of there being a smaller presence of females in the observations, though this may not account entirely for the differences shown and in fact females could genuinely be less active than males.

C. mercuriale shows a very reluctant behaviour to travel long distances at a time and is therefore a poor disperser, but it may be that barriers to movement include inhabitable landscape and by creating corridors of reduced foliage, more inter-site movement could be seen.

The road between Upper and Lower Crockford is acting as a strong barrier against the movement of *C. mercuriale*, but this may again be attributed to the growth of foliage either side of the road which creates a visual barrier to the damselfly.

References

[i] **Thompson, D.J., Purse B.** (1999) A search for long-distance dispersal in the Southern Damselfly *Coenagrion mercuriale*. *Journal of the British Dragonfly Society* **15(2)**:46-50

[ii] **Martens, A.** (1999,2000) Group oviposition in Coenagrion mercuriale (Charpentier). *Odonatologica* **29(4):**329-332

[iii] **Purse, B.** (2001) *The Ecology and Conservation of the Southern Damselfly (Coenagrion mercuriale)* 19

[iv] **Martens, A.** (1999,2000) Group oviposition in Coenagrion mercuriale (Charpentier). *Odonatologica* **29(4):**329-332

[v] **Jenkins, D.K.** (1986) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site using a modified "Pollark walk". *Journal of the British Dragonfly Society* **2(1):**17-20

[vi] **Jenkins, D.K.** (1998) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site. Part 7. Mark/recapture used to determine the extent of local movement. *Journal of the British Dragonfly Society* **14(1):**1-4

[vii] **Jenkins, D.K.** (2001) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site. Part 8. Short range dispersal. *Journal of the British Dragonfly Society* **17(1):**13-19

[viii] **Thompson, D.J., Purse B.** (1999) A search for long-distance dispersal in the Southern Damselfly *Coenagrion mercuriale*. *Journal of the British Dragonfly Society* **15(2)**:46-50

[ix] **Welstead, A.R., Welstead N.I.** (1984) A key to identify females of three species of *Coenagriidae*. *Journal of the British Dragonfly Society* **1(3):**43-44

[x] **Jenkins, D.K.** (1995) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest. Part 6. Mark/recapture programme. *Journal of the British Dragonfly Society* **11(1):**10-14

[xi] **Jenkins, D.K.** (1986) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site using a modified "Pollark walk". *Journal of the British Dragonfly Society* **2(1):**17-20

[xii] **Jenkins, D.K.** (1995) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest. Part 6. Mark/recapture programme. *Journal of the British Dragonfly Society* **11(1):**10-14

[xiii] **Thompson, D.J., Purse B.** (1999) A search for long-distance dispersal in the Southern Damselfly *Coenagrion mercuriale*. *Journal of the British Dragonfly Society* **15(2)**:46-50

[xiv] **Jenkins, D.K.** (1986) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site using a modified "Pollark walk". *Journal of the British Dragonfly Society* **2(1):**17-20

[xv] **Jenkins, D.K.** (1998) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site. Part 7. Mark/recapture used to determine the extent of local movement. *Journal of the British Dragonfly Society* **14(1):**1-4

[xvi] **Thompson, D.J., Purse B.** (1999) A search for long-distance dispersal in the Southern Damselfly *Coenagrion mercuriale*. *Journal of the British Dragonfly Society* **15(2)**:46-50

[xvii] **Prendergast, E.D.V.** (1996) The Southern Damselfly *Coenagrion mercuriale* (Charpentier) on the Ministry of Defence. *Journal of the British Dragonfly Society* **12(1):**2-10

[xviii] **Jenkins, D.K.** (1986) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site using a modified "Pollark walk". *Journal of the British Dragonfly Society* **2(1):**17-20

[xix] **Thompson, D.J., Purse B.** (1999) A search for long-distance dispersal in the Southern Damselfly *Coenagrion mercuriale*. *Journal of the British Dragonfly Society* **15(2)**:46-50

[xx] **Feneley, M.** (2001) *Dispersal ability of the southern damselfly Coenagrion Mercuriale (Charpentier)* p. 20

[xxi] **Prendergast, E.D.V.** (1996) The Southern Damselfly *Coenagrion mercuriale* (Charpentier) on the Ministry of Defence. *Journal of the British Dragonfly Society* **12(1):**2-10

[xxii] **Jenkins, D.K.** (1994) A population study of *Coenagrion mercuriale* (Charpentier) at a New Forest site. Part 5. Temperature and pH *Journal of the British Dragonfly Society* **10(1):**1-5

Unreferenced Material

http://www.dragonflysoc.org.uk - The British Dragonfly Society Webpage

http://www.english-nature.org.uk - The English Nature website (Life in UK Rivers Project)

d'Aguilar, J., L Dommanget J., Préchac R. (1986) *A Field Guide to the Dragonflies of Britain , Europe and North Africa* pp. 195-196

Begon, M. (1979) *Investigating Animal Abundance*

Fowler, J., Cohen L., Jarvis P. (1998) *Practical Statistics for Field Biology: Second Edition*

Appendix A

6.1 Macro Code for Adding 100,000 to Coordinates

```
Sub Add1()
  For x = 2 To 5000
    Range("I" & x).Select
    If ActiveCell.FormulaR1C1 < 10000 Then
      ActiveCell.Formula = ActiveCell.Formula + 100000
    End If
  Next
End Sub
```

6.2 Macro Code for Entering Formulae for Total Distance Travelled and Average Rate of Travel

```
Sub Totals()
  Range("B2").Select
  males = ActiveCell.Value
  Range("B4").Select
  females = ActiveCell.Value
  For cycle = 3 To (males + females + 3)
    If cycle = males + 3 Then GoTo getout
    Range("p" & cycle).Select
    If ActiveCell.Value = "" Then
      For found = cycle - 1 To 1 Step -1
        Range("p" & found).Select
        If ActiveCell.Value = "" Then
          Range("s" & (cycle - 1)).Select
          ActiveCell.Formula = "=SUM(p" & (found + 1) & ":p"
            & (cycle - 1) & ")"
          Range("t" & (cycle - 1)).Select
          ActiveCell.Formula = "=average(r" & (found + 1) &
            ":r" & (cycle - 1) & ")"
          GoTo getout
        End If
      Next
getout:
    End If
  Next
End Sub
```

6.3 Macro Code for Splitting the Data into Ranges

```
Sub Ranges()
Dim dist(50), travrate(50)
    'get range of data
    Range("B2").Select
    males = ActiveCell.Value
    Range("B4").Select
    females = ActiveCell.Value

    'flush variables
    For x = 1 To 50
        dist(x) = 0
        travrate(x) = 0
    Next

    'males
    For cycle = 2 To (males + 1)
        Range("s" & cycle).Select
        If ActiveCell.Value <> "" Then
            contents = ActiveCell.Value
            For y = 2 To 17
                If contents >= (y - 2) * 25 And contents < (y - 2) * 25
                    + 25 Then dist(y) = dist(y) + 1
            Next
            If contents > 400 Then dist(18) = dist(18) + 1
        End If

        Range("p" & cycle).Select
        If ActiveCell.Value <> "" Then
            contents = ActiveCell.Value
            For y = 2 To 26
                If contents >= (y - 2) * 20 And contents < (y - 2) * 20
                    + 20 Then travrate(y) = travrate(y) + 1
            Next
            If contents > 500 Then travrate(27) = travrate(27) + 1
        End If
    Next

    'enter data
    For d = 2 To 18
        Range("w" & d).Select
        ActiveCell.Value = dist(d)
        If dist(d) = 0 Then
            Range("x" & d).Select
            Selection.ClearContents
        End If
```

```
Next
For t = 2 To 27
    Range("ac" & t).Select
    ActiveCell.Value = travrate(t)
    If travrate(t) = 0 Then
        Range("ad" & t).Select
        Selection.ClearContents
    End If
Next

'flush variables
For x = 1 To 50
    dist(x) = 0
    travrate(x) = 0
Next

'females
For cycle = (males + 2) To (males + females + 2)
    Range("s" & cycle).Select
    If ActiveCell.Value <> "" Then
        contents = ActiveCell.Value
        For y = 2 To 17
            If contents >= (y - 2) * 25 And contents < (y - 2) * 25
                + 25 Then dist(y) = dist(y) + 1
        Next
        If contents > 400 Then dist(18) = dist(18) + 1
    End If

    Range("p" & cycle).Select
    If ActiveCell.Value <> "" Then
        contents = ActiveCell.Value
        For y = 2 To 26
            If contents >= (y - 2) * 20 And contents < (y - 2) * 20
                + 20 Then travrate(y) = travrate(y) + 1
        Next
        If contents > 500 Then travrate(27) = travrate(27) + 1
    End If
Next

'enterdata
For d = 2 To 18
    Range("y" & d).Select
    ActiveCell.Value = dist(d)
    If dist(d) = 0 Then
        Range("z" & d).Select
        Selection.ClearContents
    End If
```

```vba
    Next
    For t = 2 To 27
        Range("ae" & t).Select
        ActiveCell.Value = travrate(t)
        If travrate(t) = 0 Then
            Range("af" & t).Select
            Selection.ClearContents
        End If
    Next
End Sub
```