

## The Use of Statistics in Biodiversity Indicators

JNCC is a statutory advisor to UK Government and devolved administrations. During my time with the Environmental Department (Defra), they worked on a response indicator, which focused on public involvement in biodiversity. It was important to have an index of volunteer time spent in biodiversity conservation across the leading conservation organisations as this is critical to the successful delivery of many of the department's objectives.

Data were not available for all organisations in all years, so missing values were estimated using extrapolation. These were based on the trend in the data that had been provided.

The points to be extrapolated for the Woodland Trust are not far beyond the known data so linear extrapolation would be an appropriate method; which means extending a tangent line beyond the end of the known data.

If the two data points nearest the point  $x_*$  to be extrapolated are  $(x_{k-1}, y_{k-1})$  and  $(x_k, y_k)$ , linear extrapolation gives the function:

$$y(x_*) = y_{k-1} + \frac{x_* - x_{k-1}}{x_k - x_{k-1}}(y_k - y_{k-1}).$$

(Which is identical to linear interpolation if  $x_{k-1} < x_* < x_k$ ).

For example, we can extrapolate the number of volunteer hours completed at the Woodland Trust in 2001:

$$\begin{aligned} y(2001) &= 113229 + \frac{2001 - 2002}{2003 - 2002} \cdot (101163 - 113229) \\ &= 113229 + 12066 \\ &= 125295 \end{aligned}$$

Woodland Trust	
Year	Number of Volunteer Hours
2000	-
2001	-
2002	113,229
2003	101,163
2004	143,520
2005	272,113
2006	253,405
2007	231,237
2008	224,662
2009	202,658

As some of the figures were extrapolated, and because of the difference in data quality, the number of volunteer hours were converted to an index and the final graph shows the increase relative to 2000 rather than the absolute numbers of volunteer hours.

Year	Total Hours	Index (2000 = 100)
2000	4805716	100
2001	4290962	89
2002	5157637	107
2003	5170069	108
2004	5460046	114
2005	5854246	122
2006	6854238	143
2007	7169219	149
2008	7227861	150
2009	7257816	151

The total for 2000 was assigned a value of 100, and then the values for subsequent years were calculated according to their relative size compared to this base year.

e.g.  $(4290962/4805716) \cdot 100 = 89$

This has allowed the Environmental Department to draw useful conclusions and to assess their success in key objectives:

*“There has been a 51 per cent increase in time spent volunteering between 2000 and 2009.” (JNCC)*

Correlations were also used to investigate whether trends could be estimated from only one or two of the major data suppliers. For example, Spearman's rank correlation coefficient could be used to calculate the correlation between the British Trust for Ornithology and the British Trust for Conservation Volunteers.

First, we must find the value of the term  $d_i^2$ . To do so we use the following steps:

1. Sort the data by the first column ( $X_i$ ). Create a new column  $x_i$  and assign it the ranked values 1,2,3,... $n$ .
2. Next, sort the data by the second column ( $Y_i$ ). Create a fourth column  $y_i$  and similarly assign it the ranked values 1,2,3,... $n$ .
3. Create a fifth column  $d_i$  to hold the differences between the two rank columns ( $x_i$  and  $y_i$ ).
4. Create one final column  $d_i^2$  to hold the value of column  $d_i$  squared.

British Trust for Ornithology	British Trust for Conservation Volunteers
1,338,328	987,500
1,074,150	1,104,600
1,468,875	1,042,600
1,514,358	1,041,500
1,568,059	1,039,200
1,622,738	1,106,400
1,612,169	1,594,032
1,713,055	1,657,702
1,771,958	1,395,132
1,777,173	1,379,124

With  $d_i^2$  found, we can add them to find  $\sum d_i^2 = 4$ . The value of  $n$  is 10. So these values can now be substituted into Spearman's equation:

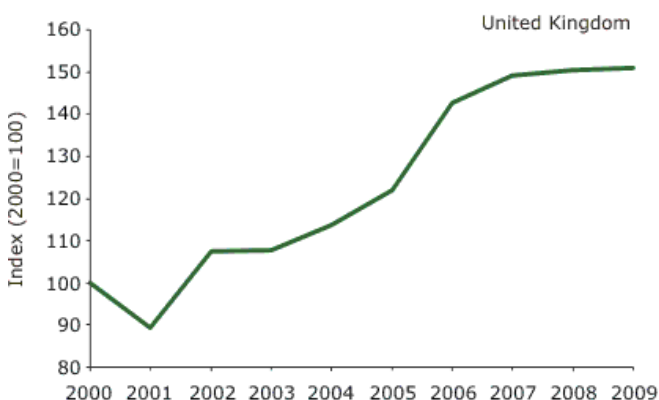
$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

$$\rho = 0.9758$$

BTO (Xi)	BTCV (Yi)	(xi)	(yi)	(di)	(di)^2
1,074,150	1,104,600	1	2	-1	1
1,338,328	987,500	2	1	1	1
1,468,875	1,042,600	3	3	0	0
1,514,358	1,041,500	4	4	0	0
1,568,059	1,039,200	5	5	0	0
1,612,169	1,594,032	6	7	-1	1
1,622,738	1,106,400	7	6	1	1
1,713,055	1,657,702	8	8	0	0
1,771,958	1,395,132	9	9	0	0
1,777,173	1,379,124	10	10	0	0

This high value shows that there is a considerable amount of correlation between volunteering hours at the two major data suppliers, so they could be used to help judge the accuracy of any extrapolations.

These various statistical devices have enabled Defra to complete their conservation volunteering indicator and to be confident in its accuracy; making important policy decisions based on its findings.



**Notes:** 1. Interpolated data have been used by Defra to fill missing years for Woodland Trust (2000 and 2001), Butterfly Conservation (2000 to 2002), and The Wildlife Trusts (2000 to 2004).

2. As data provided RSPB were for financial years as opposed to calendar years, 2008-2009 data were allocated to 2008 and Defra estimates were made for 2009.

**Source:** Bat Conservation Trust, British Trust for Conservation Volunteers, British Trust for Ornithology, Butterfly Conservation, Natural England, Plantlife, Royal Society for the Protection of Birds, The Wildlife Trusts, Woodland Trust.

**Figure 18 (i): Index of volunteer time spent in biodiversity conservation in selected UK conservation organisations, 2000-2009**