A method for setting badger control target numbers for the purpose of the West Gloucestershire and West Somerset pilots

A report to Defra and Natural England

[redacted]

Contents

1. INTRODUCTION

The Government's policy is to allow controlled culling and vaccination of badgers in areas of high incidence of bovine TB in cattle in a carefully regulated way for the purpose of controlling the spread of the disease, in accordance with the requirements set out in Defra's Guidance to Natural England: "Licences to kill or take badgers for the purpose of preventing the spread of bovine TB under section 10(2)(a) of the Protection of Badgers Act 1992" (Defra publication PB13692).

In the first year of culling, a minimum number of badgers must be removed during an intensive cull which must be carried out throughout the land to which there is access, over a period of not more than six consecutive weeks. This minimum number should be set at a level that in Natural England's judgement should reduce the estimated badger population of the application area by at least 70% (para 10(c)(i) and (ii) of the Guidance).

Natural England should aim to ensure that culling will "not be detrimental to the survival of the population concerned" within the meaning of Article 9 of the Bern Convention on the Conservation of European Wildlife and Natural Habitats. For that purpose Natural England should determine appropriate area-specific licence conditions, and set a maximum number of badgers to be removed from the licence area (para 12 of the Guidance).

Hence, there is a requirement to gain an estimate of the population size that is sufficiently accurate and precise for setting a minimum and maximum number of badgers that meets these requirements. The estimate of population size must be based on information that is available to Natural England during the application process, and must relate to the whole culling area, whether culling will take place on a landholding or not. With the possible exception of a full census, any population estimate will have some degree of uncertainty; there will be an interval around the population estimate within which the true population is likely to lie: higher uncertainty leads to a wider interval around the estimated population.

This poses potential challenges for licensing culls (Donnelly and Woodroffe 2012 Nature 485: doi:10.1038/485582a) because targets for removal need to meet both the policy requirement of being at least 70% of the true population and the Article 9 requirement. In this document we have assumed that removal of no more than 95% of the initial population will meet this latter requirement (see section 2). In addition there is limited information about badger population size or factors contributing to population size, and it is not practical to directly count badgers prior to each cull. Hence, uncertainty about populations cannot be easily reduced and nor must it be ignored. The approach proposed in this document is to use information available to Natural England prior to a cull to set a target that acknowledges the uncertainty that is associated with the estimated population size and that is designed to be robust to this uncertainty with a specified degree of confidence subject to reasonable assumptions. While other methods could be applied, this is a pragmatic approach that is consistent with having to assess populations across a wide area, ensuring that the available information is used to the best effect.

In this approach, which applies to the pilots, population estimates are based on the number of active setts observed in a sample of land parcels within each area and the average number of badgers per active sett observed in two West Country areas over a number of years. The population estimate, and its estimated uncertainty are used to provide a cull target that is, with pre-determined confidence, at least 70% of the true population and, with a pre-determined confidence, no more than 95% of the true population. This is subject to the assumption that badger populations in the two West Country areas within which population density was assessed, are representative of the cull areas with respect to the average number of badgers per sett. The size of the pre-determined confidence intervals reflects the level of certainty we can have in estimates derived in this way and under these assumptions. The size of the confidence intervals requires careful consideration to ensure both policy objectives are met sufficiently robustly.

For the purpose of implementing the strategy described in this document, the location of active badger setts on a sample of participating land within the proposed cull area has been mapped by trained surveyors and the analysis of these results submitted to Natural England. All setts were photographed and an independent assessor looked for evidence in each photograph that suggested that the sett had been mis-identified. The misidentification rate was used to reduce the sett count and correct the estimate of uncertainty. Independent repeat surveys were carried out over approximately 15km^2 of the land already surveyed so that the overlap and mis-match between the two surveys were measured and used to estimate the error around estimates of sett abundance. Consequently, the total number of active setts and hence the average active sett density were estimated. Active sett density will now be multiplied by the total area of the pilot area to estimate the total number of active setts, which will be used to calculate the number of badgers within the cull area by multiplying by the expected number of badgers per active sett. This assumes that non-participating land and participating land are equivalent with respect to the density of active setts and badgers.

Explanation of how the average number of badgers per sett was derived is presented in appendix 1. Appendix 2 details the relationship between survey effort and uncertainty, such that confidence in each of the policy objectives (removal of at least 70% of the population, and prevention of a local extinction) can be estimated. A description of how uncertainty about population size can be estimated is given in appendices 3 and 4. A worked example is provided in appendix 5 of how cull targets can be set using sampled active sett location data.

2. METHODOLOGY - Deriving estimates of pre-cull populations.

2.1 The outline process.

1) Survey a random sample of at least 50km^2 of land parcels (appendix 2) for active setts.
2) Re-survey at least 15km^2 of land parcels for active badger setts using different surveyors. To enhance the value of information from the repeat survey, 1-km squares should be selected within which at least 1 active sett has been found during the first survey.
3) Multiply the average number of setts per km^2 (appendix 1) by the total area (in km^2) of the pilot area and multiply the result by the average number of badgers per sett (5.4 see appendix 1). Calculate the uncertainty (appendixes 3-4). This is the estimate of the population before culling.
4) Set cull targets (section 3 and see appendix 5 for a worked example).

2.2 Calculating pre-cull population size.

The population *N* is estimated by the product of the number of active setts in an area inferred from a survey of a number of one-kilometre squares selected at random from within that area and the expected mean number of badgers per active sett. The efficiency with which setts have been found is estimated via a second survey of a number of the surveyed squares. The efficiency is used to adjust the estimated population (N). Hence

$$N = \frac{A_C \times S \times B}{A_S \times E} \qquad \text{Equation 1}$$

Where A_C is the cull area, A_S is the area surveyed for setts, S is the number of setts found in the survey, E is the estimated efficiency with which setts are found in the survey (appendix 4) and B is the expected mean number of badgers per active sett (appendix 1).

The expected mean number of badgers per active sett is estimated from the observed number of badgers per active sett across two study sites: a long-term study in the west-country, and the “Badger Vaccine Study” area[1]. These currently represent the best-available data on this quantity, which is a limited sample from one region of southwest England and is assumed to be representative of cull areas.

The relative standard uncertainty *R* associated with number of badgers *N* is estimated by the between-year variation (across ‘summers and autumns’) of the mean number of badgers per active sett (RSD=R_T); the between-area variation in the mean number of badgers per active sett (RSD=R_S) (Appendix 1); the relative standard error associated with the estimated number of setts in the cull area (R_H) (Appendix 3); and the relative standard uncertainty associated with the efficiency with which setts are found during a survey (R_E) (Appendix 4).

$$R = \sqrt{R_T^2 + R_S^2 + R_H^2 + R_E^2} \qquad \text{Equation 2}$$

[1] The Badger Vaccine Study area covered 55km^2 and contained 110 badger setts and 63 social groups. Badgers were monitored in this area from 2006 to 2009. See http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&ProjectID=16320&FromSearch=Y&Publisher=1&SearchText=CB0115&SortString=ProjectCode&SortOrder=Asc&Paging=10#Description

3. Setting cull targets.

The estimate of population size will be used to set a cull target and a limit:

1) **Minimum target:** The lowest number which is, with a given confidence, *at least* 70% of the population. The population is defined here as badgers resident within the outer boundary of the culling area. This includes badgers resident on land holdings that are not available for culling. This number should appear as the **minimum target** on the licence or accompanying agreement (**minimum target** on Fig. 1).

2) **Maximum limit:** A number at least as great as the **minimum target**, but below (with a given level of confidence) the estimated pre-cull population size (**maximum limit** on Fig. 1). In practice, this number might be very close to the **minimum target**. There is no clarity about what constitutes detriment to the survival of the population within the meaning of Article 9 of the Bern Convention, and it is not possible to estimate a minimum viable badger population with any certainty, so defining a maximum target with a high degree of certainty is extremely challenging. The **maximum limit** should be set at a number that does not exceed 95% of the population with a given confidence. If we assume that a population of badgers within an area of 300km^2 in the southwest of England numbers approximately 1300, then removal of 95% of the population would leave approximately 65 individuals. This is more likely to constitute a viable population than is one composed of 13 badgers, which would be left following removal of 99% of the population.

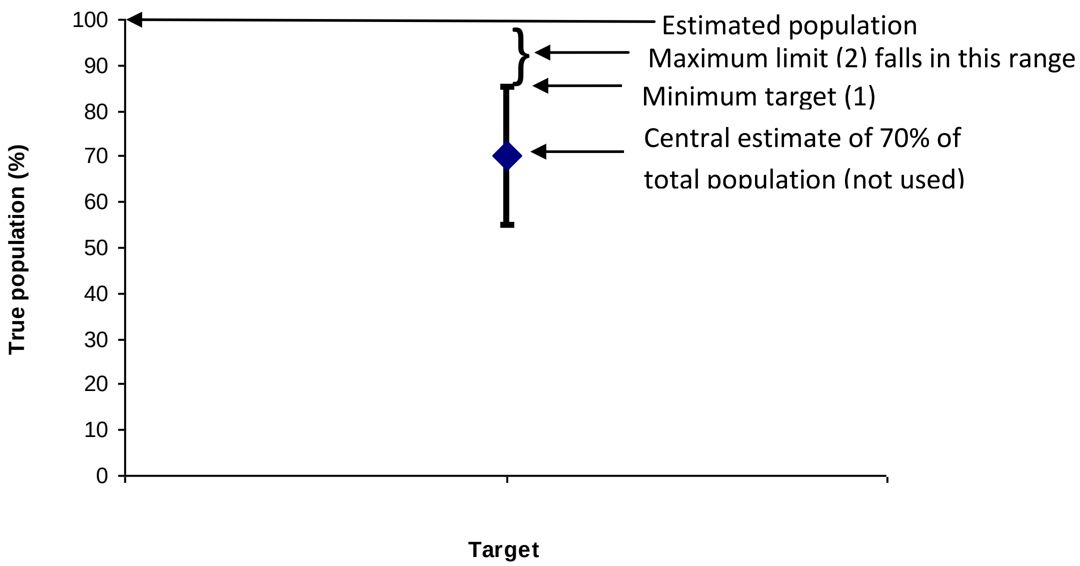

Figure 2. Values used for setting cull targets.

Where uncertainty in N is assumed to be adequately characterised by a log-normal distribution[2] the **minimum target** (m) is given by

$$m = T.exp\left(Ln(N) + \sigma\left(k_m + \frac{\sigma}{2}\right)\right)$$

Equation 3

the **maximum limit** (M) is given by

$$M = L.exp\left(Ln(N) - \sigma\left(k_M - \frac{\sigma}{2}\right)\right)$$

Equation 4

Where N is the estimated population (Equation 1), T is the proportion of badgers that must be removed; k_m is the coverage factor for the confidence interval used to calculate m; L is an upper limit for the proportion of badgers that may be removed; k_M is the coverage factor for the confidence interval used to calculate M; k_V is the coverage factor for the confidence interval used to calculate V and σ is a parameter given by

$$\sigma = \sqrt{Ln(R^2 + 1)}.$$

Equation 5

Values of k_m and k_M are given by the quantile of the standard normal distribution. Table 1 provides some examples.

Table 1. Coverage factors

Confidence	Coverage factor (k)
60.0%	0.253
62.5%	0.319
65.0%	0.385
67.5%	0.454
70.0%	0.524
72.5%	0.598
75.0%	0.674
77.5%	0.755
80.0%	0.842

[2] N is estimated by the product of a number of quantities each having a normal or right-skewed dispersion. Under the central limit theorem sums of quantities are approximately normally distributed and products of quantities are approximately log-normally distributed.

Appendix 1. Estimating the average number of badgers per sett

In order to produce targets based on the number of active setts an estimate of the expected number of badgers per active sett is required. Estimates of the number of badgers per active sett or social group are rare in the literature. More commonly indexes of population size are given, such as a number trapped, the Minimum Number Alive, or measures of activity. Hence some analyses were undertaken specifically for the task of estimating the number of badgers per active sett for setting targets.

A mark-recapture analysis of BVS observations and recent observations at a further southwest study site (SWS) using the Popan implementation of the Jolly-Seber model produced the following estimates of badgers per active sett in summer and autumn (Table 2):

Table 2. Mean number of badgers per active sett

Year	Location	Badgers per active sett
2003	SWS	4.77
2004	SWS	6.38
2005	SWS	6.48
2006	SWS	5.07
2007	SWS	4.76
2006 (autumn)	BVS	4.79
2006 (summer)	BVS	5.69
2008	BVS	4.97

Hence the mean number of badgers per active sett across approximately 80 badger social groups in two West-Country locations over 6 recent years was 5.4 badgers per active sett. The relative standard deviation between years and locations associated with the estimated mean number of badgers per active sett was 13.5%.

Appendix 2. The relation between sampling effort, expected uncertainty and confidence

If uncertainty is described by a log-normal distribution then the approximate relation between confidence, maximum limits, minimum targets, and uncertainty is given by:

$$R \approx \sqrt{exp\left(\left(\frac{Ln(M)-Ln(m)}{k_M+k_m}\right)^2\right)-1} \qquad \text{Equation 6}$$

Where R is the relative standard uncertainty associated with the estimate of the population, *m* is the minimum target for the proportion that must be removed, M is the maximum limit for the proportion that can be removed, k_m is the value of the inverse standard normal distribution for the required confidence that the minimum target is met, k_M is the value of the inverse standard normal distribution for the required confidence that the maximum limit is not exceeded.

Using data on active sett locations collected by the time of writing table 3 shows the number of nominally 1-km^2 areas required to be surveyed in the West Gloucestershire pilot area to deliver a population estimate with sufficiently low relative standard uncertainty (Equation 2) to meet different confidence requirements (Equation 6). Table 4 shows the same outputs for the West Somerset pilot area. Tables 3 and 4 have been derived assuming that at least 15 one-kilometre squares are surveyed in the second survey, and that no more than 6 setts are found by only one surveyor (R_E<0.054). Hence if 49-51 one-kilometre squares are surveyed in the main survey in West Gloucestershire, the uncertainty that is associated with population estimates is expected to be sufficient to provide:

1) 80% confidence that the minimum target is at least 70% of the population and 65% confidence that the maximum limit is no more than 95% of the population,

2) 75% confidence that the minimum target is at least 70% of the population and 70% confidence that the maximum limit is no more than 95% of the population,

3) 70% confidence that the minimum target is at least 70% of the population and 75% confidence that the maximum limit is no more than 95% of the population,

4) 65% confidence that the minimum target is at least 70% of the population and 80% confidence that the maximum limit is no more than 95% of the population,

Table 3. Survey effort (number of 1-km squares) necessary to deliver a range of confidence intervals (assuming a second confirmatory survey of 15 x 1-km squares) in West Gloucestershire.

		Confidence that target is at least 70% of the population						
		50%	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**
Confidence that the target is no more than 95% of the population	**50%**	<10	<10	<10	<10	<10	15	24
	55%	<10	<10	<10	<10	14	21	30
	60%	<10	<10	<10	13	20	28	41
	65%	<10	<10	13	19	27	38	51
	70%	<10	14	20	27	37	49	67
	75%	15	21	28	38	49	65	89
	80%	24	30	41	51	67	89	120

If 50 one-kilometre squares are surveyed in the main survey in West Somerset, the uncertainty that is associated with population estimates is expected to be sufficient to provide:

1) 80% confidence that the minimum target is at least 70% of the population and 75% confidence that the maximum limit is no more than 95% of the population,

2) 75% confidence that the minimum target is at least 70% of the population and 80% confidence that the maximum limit is no more than 95% of the population,

Table 4. Survey effort (number of 1-km squares) necessary to deliver a range of confidence intervals (assuming a second confirmatory survey of 15 x 1-km squares) in West Somerset.

		Confidence that target is at least 70% of the population						
		50%	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**
Confidence that the target is no more than 95% of the population	**50%**	<10	<10	<10	<10	<10	<10	<10
	55%	<10	<10	<10	<10	<10	<10	15
	60%	<10	<10	<10	<10	<10	14	20
	65%	<10	<10	<10	<10	13	19	27
	70%	<10	<10	<10	13	18	26	36
	75%	<10	<10	14	19	26	35	51
	80%	<10	15	20	27	36	51	82

The differences in sampling effort required for given levels of confidence between the two areas are caused by the lower variation in the number of setts between survey squares in Somerset in comparison with Gloucestershire (Tables 5-7).

The traditional confidence limit used by scientists is 95%; it covers the majority of outcomes that might be expected, whilst excluding the 5% most extreme ones. However, in wildlife management different confidence intervals are generally considered to be acceptable (Vaske, 2002 Human Dimensions of Wildlife 7: 287-300), mainly because measurements of wild populations can be highly uncertain, yet decisions on action nevertheless need to be made. For example, when setting roe deer cull targets according to population estimates derived from dung counts, best-practice guidance is to use an 80% confidence limit (Ratcliffe & Mayle 1992 Roe deer biology and management. Forestry Commission Bulletin 105). Therefore, confidence limits that are lower than 95% may be suitable for the purpose of setting badger cull targets. In all likelihood, confidence intervals greater than 50% will mean that delivery of the minimum target will result in removal of slightly more than 70% of the estimated population, more than 50% of the time.

The maximum limit presents a different situation, because Natural England needs to be satisfied that culling "will not be detrimental to the survival of the population concerned", within the meaning of article 9 of the Bern Convention on the Conservation of European Wildlife and Natural Habitats. We recommend that the maximum limit should be set at a number that does not exceed 95% of the population with a given degree of certainty (see section 2). However, eradication of a high-density population typically requires extremely high effort during the latter stages (reviewed for feral pig removal by McCann & Garcelon 2008 *J. Wildl. Manage*. **72**: 1287-1295). Indeed, unless three key principles are met (rate of removal exceeds rate of replacement at all population densities; immigration is prevented; all individual capable of reproduction must be removed), efforts to eradicate a population are highly unlikely to succeed (Bomford & O'Brien 1995 *Wildlife Society Bulletin* **23**: 249-255). Therefore, in practice, a lower level of certainty in the maximum limit than the traditionally used 95% confidence interval is nevertheless likely to avoid prejudicing the persistence of badgers within the area following culling. Furthermore, there will be pockets of land inaccessible to culling (potentially up to 30% of the total area) so it is unlikely that all badger could be removed in the 6 week culling period.

Careful consideration is required when deciding levels of confidence that are acceptable. Tables 3 and 4 above offer the opportunity to trade-off confidence intervals for the two policy objectives for varying survey effort for both pilot areas.

Appendix 3. Uncertainty associated with sampling badger setts

Where sett surveys are undertaken on a proportion of the total cull area (<50%), the approximate relative standard uncertainty of the number of estimated setts in the cull area is given by

$$R_H = \frac{t}{\bar{S} \times \sqrt{n}}$$

Where $\bar{S}$ is the average number of setts found in each square, t is the standard deviation of the number of setts between one-kilometre squares and n is the number of squares that are surveyed. Where a higher proportion of squares are surveyed

$$R_H < \frac{t}{\bar{S} \times \sqrt{n}}$$

and tends to zero as the survey approaches a full census of the 300 squares

Approximate expected value of R_H

A survey of randomly selected 1km^2 areas of land within each area in August and September produced the following observations of putative active setts (Tables 5 and 6).

Table 5. Putative active sett survey results of all 1km squares surveyed in Gloucestershire at the time of writing.

Location	Setts		Location	Setts		Location	Setts
1	3		20	3		39	1
2	2		21	0		40	0
3	0		22	0		41	0
4	2		23	0		42	0
5	0		24	1		43	0
6	2		25	4		44	0
7	1		26	1		45	6
8	0		27	2		46	1
9	1		28	2		47	3
10	3		29	0		48	1
11	3		30	3		49	0
12	1		31	0		50	3
13	17		32	0		51	4
14	6		33	0		52	3
15	0		34	0			
16	5		35	0			
17	3		36	0			
18	0		37	0			
19	14		38	0			

Note: These are un-validated preliminary results, uncorrected for variation between locations in area surveyed. Final estimates of active sett density may be based on different values.

Table 6. Putative active sett survey results of all 1km squares surveyed in Somerset.

Location	Setts		Location	Setts		Location	Setts
1	3		20	2		39	3
2	4		21	8		40	0
3	2		22	7		41	3
4	5		23	2		42	1
5	6		24	3		43	0
6	9		25	14		44	0
7	11		26	4		45	0
8	15		27	4		46	0
9	4		28	2		47	0
10	4		29	8		48	0
11	2		30	2		49	0
12	1		31	6		50	0
13	1		32	4		51	0
14	0		33	2		52	0
15	6		34	3		53	0
16	9		35	5		54	0
17	11		36	9		55	0
18	3		37	2			
19	9		38	5			

Note: These are un-validated preliminary results, uncorrected for variation between locations in area surveyed. Final estimates of active sett density may be based on different values.

A negative binomial distribution (a standard distribution used for describing counts that are clustered) was fitted to the observations (parameters, Gloucester mean=2.001, over-dispersion relative standard deviation=1.402; Somerset 3.783, over-dispersion relative standard deviation 1.046). Of the two areas, Somerset has a higher observed average sett density that varies less between locations.

A number of simulated surveys were undertaken by generating a number of 300 one-kilometre sett counts using the fitted negative binomial distributions. The sum of the 300 counts represented a true simulated population and the sum of a sample from the 300 counts (10 to 300) was used to simulate a survey of that population.

The differences between the estimated number of active setts based on the simulated survey and the sum of the 300 counts that represented the true value was used to estimate the expected standard error associated with each survey size.

Relative Standard errors are shown in Table 7.

Table 7. Expected relative standard errors associated with estimates of the number of putative active setts in 300 km^2

Number of nominally 1km^2 areas surveyed	Relative standard error of sett number estimate:	
	Gloucester	Somerset
10	0.537	0.340
20	0.378	0.243
30	0.292	0.195
40	0.249	0.162
50	0.210	0.145
60	0.190	0.132
70	0.168	0.121
80	0.150	0.114
90	0.143	0.108
100	0.133	0.093
110	0.119	0.085
120	0.113	0.080
130	0.109	0.078
140	0.102	0.073
150	0.095	0.068
200	0.068	0.047
250	0.041	0.031
300	0.000	0.000

Appendix 4. Uncertainty associated with the efficiency of finding active setts

There may be some probability that active setts are missed by surveyors even though they are present in the surveyed squares. The probability of active setts being missed, and the additional uncertainty associated with this factor can be estimated by employing blind-duplicate surveying of a number of squares.

Assuming that the surveyor and blind-duplicate surveyor are equally efficient then the average probability of a survey missing an active sett can be estimated from

$$p = \frac{A}{A + 2B}$$

Where A is the number of active setts that are found by only one survey (original or blind-duplicate) and B is the number of active setts found by both surveys.

And the efficiency with which an active sett is found is given by

$$E = 1 - p$$

The uncertainty associated with the binomial proportion p depends on the number of active setts observed by one survey and the number observed by both surveys. A slightly conservative estimate of the standard uncertainty associated with the probability p can be made by treating it as being equivalent to the uncertainty that is associated with the binomial proportion A/2B.

Hence for A>0 the relative standard uncertainly R_E associated with E is estimated by

$$R_E = \sqrt{\frac{\frac{A}{2B}\left(1 - \frac{A}{2B}\right)}{2B \times E^2}}$$

For A=0 the relative standard uncertainty R_E associated with E can be estimated by[3]

$$R_E = \frac{0.431}{(2B)^{0.968}}$$

The other error potentially associated with surveying is that a hole in the ground created by some other mechanism is mistakenly identified as an active badger sett. This ‘false positive’ error will be controlled for by photographing and independently scrutinising the identity of reported active setts, seeking evidence to reject the surveyors’ claims.

---

[3] This is an empirical numerical approximation to the modified Jeffrey’s interval described in Brown LD, Cail T, DasGupta A, 2001, Interval estimation for a binomial proportion, Statistical Science, 16(2), 101 – 133.

Appendix 5. Illustrative example: setting cull targets

This example assumes that a target for a cull is required which, with 80% confidence, is at least 70% of the population. A limit that with as high a confidence as is practicable is no more than 95% of the population is also required.

A survey of 50 one-kilometre squares found 80 active badger setts. The relative standard error of this quantity was estimated to be 0.210 (in practice this will be estimated using the method described in Appendix 3). 15 squares were selected at random for a second survey. 27 active setts had been found in these squares in the first survey. The second survey also found 27 setts. However, it missed 3 setts that were found in the first survey and found 3 setts that were not observed in the first survey. Hence (Appendix 4) A=6, B=24

$$p = \frac{6}{6 + 2 \times 24} = 0.111$$

$$E = 1 - 0.111 = 0.889$$

$$R_E = \sqrt{\frac{\frac{6}{48}\left(1 - \frac{6}{48}\right)}{48 \times 0.889^2}} = 0.0537$$

The mean number of badgers per active sett is estimated to be 5.4 (Appendix 1).

Hence (Equation 1, A_C=300, A_S=50 S=80, E=0.892, B=5.4) the population in the 300 km^2 area is estimated to be

$$N = \frac{300 \times 80 \times 5.4}{50 \times 0.889} = 2916$$

Relative standard uncertainties are:

$\sqrt{R_T^2 + R_S^2}$=0.135 (Appendix 1)

R_H=0.210 (estimated from survey data taken from Appendix 3 for this example)

R_E=0.0537

Hence (Equation 2) the relative standard uncertainty associated with the estimated population is given by

$$R = \sqrt{0.135^2 + 0.210^2 + 0.0537^2} = 0.255$$

Hence (Equation 5)

$$\sigma = \sqrt{Ln(0.255^2 + 1)} = 0.251$$

For 80% confidence that the target is at least 70% of the population (k_m= 0.842, Equation 3)

$$m = 0.7exp\left(Ln(2916) + 0.251\left(0.842 + \frac{0.251}{2}\right)\right) = 2603$$

For 64% confidence that the limit is no more than 95% of the population (k_M= 0.385, Equation 4)

$$M = 0.95exp\left(Ln(2916) - 0.251\left(0.385 - \frac{0.251}{2}\right)\right) = 2612$$