# ACCURACY OF HEIGHTS FROM ORDNANCE SURVEY MAPS

## Summary

This document sets out to answer a number of questions:

- What is the accuracy of spot heights at summits and at cols?
- Does the accuracy vary with the map scale?
- Are heights on older Landrangers more accurate?
- How useful are levelled heights on old 6-inch maps?
- How accurate are col heights estimated by interpolation between contours?
- How common are gross errors?
- How accurate are OS trig heights?

The dataset for the analysis comprised 447 summits and 317 cols surveyed by differential GPS instruments between August 2008 and December 2014 for which spot heights or interpolated col heights were available, and 21 benchmarks in the OS legacy database.

It is anticipated that the results will be useful to compilers of hill data and to list authors and peak baggers wanting to identify potentially misclassified hills.

## Key results

1) Spot heights at summits have an accuracy of about ±3.3m at all map scales, in agreement with the published figure.
2) Levelled heights from old 6-inch maps can be a few metres below the summit but are unlikely to be more than a few tenths of a metre higher.
3) The accuracy of old and new 1:50k maps is similar, but where heights have changed the old height has a tendency to be too low.
4) Spot heights at cols have an expected error range of –3.0 to +4.2m (height tends to be overestimated).
5) Interpolated col heights in the DoBIH have an expected error range of –5.2 to +6.6m.
6) About 0.5% of summit heights on current maps are extreme values that are unlikely to arise from normal measurement error.
7) Errors of up to 0.31m were found in 21 flush bracket heights.

## 1 Background

### 1.1 Measures of accuracy

In metrology, the following concepts are used to define the properties of a measurement system (other criteria such as traceability and consistency over time are not relevant to this study):

*Accuracy* is the closeness of the measurements to the true value

*Precision* is the repeatability of the measurements, i.e. the tendency of the measurements to agree among themselves

*Bias* is the tendency of repeated measurements to converge to a value different from the true value.

The statistics used by Ordnance Survey to evaluate these properties are, respectively, root mean square error, standard error and mean error.[1] This document also reports the mean absolute error as an alternative to rms error, and confidence intervals to the OS standard. These statistics are defined below.

[1] J B Harley, *Ordnance Survey Maps a descriptive manual*, HMSO, 1975. Chapter 11 ‘The accuracy of Ordnance Survey maps’. The same statistics are used for position and height.

The measurement error, abbreviated to error, is the difference between the measured and true values. In this study, the error is *map height – surveyed height*. Measurement error has two components, random error and systematic error. These relate directly to the concepts of precision and bias.

Root mean square error (rms error, rmse)

The square root of the average squared error, viz.

$$r = \sqrt{\frac{\sum x^2}{n}}$$

where $x_1$, $x_2$ ..., $x_n$ are the errors of $n$ heights.

OS uses the rms error to evaluate the overall accuracy of a survey.

Due to the squaring of the individual errors, the rms error gives greater weight to large errors. An alternative statistic is the mean absolute error which is the average of the errors without regard to sign, viz. $\frac{\sum |x|}{n}$. Because it weights the errors equally, it is less affected by outliers than the rms error.

Mean error (systematic error)

The average of the errors taking sign into account, viz.

$$b = \bar{x} = \frac{\sum x}{n}$$

OS calls this the systematic error. The observed value of *b* will be the average of a finite number of measurements and is subject to sampling variation. We can perform a statistical test to determine whether its true value differs from zero at a specified level of confidence. The result of this test is given in the Tables. If *b* differs significantly from zero we would conclude that the measurements are biased, i.e. on average too high or too low.

Standard error (standard deviation)

In OS usage, this is similar to the rms error except that it measures the dispersion about the mean error. Hence it estimates the random component of the error.

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

The OS usage is somewhat idiosyncratic because 'standard error' is not normally used to describe the variation in individual values. Current statistical terminology uses *standard deviation* for this purpose. For theoretical reasons it is usual to replace *n* by *n*–1 in the denominator when estimating the standard deviation from a sample. This report follows the OS convention except for the analysis of benchmark heights, where the sample size is much smaller.

The equation relating the rms error to the random and systematic components is $r^2 = \sigma^2 + b^2$.

Maximum error

Scientific literature that quotes the accuracy of a measurement as ±*y* is inconsistent in the definition of *y*. Current OS literature mostly gives the rms error. The probability that the true value of *x* lies within that range is not very high – only 68% if the systematic error is zero and the errors follow a normal distribution. One OS publication mentions the 68% figure but most literature doesn't. Manufacturers of surveying instruments who quote the precision of a measurement as (say) 1.0m usually mean the standard deviation. The interpretation is similar except that it relates to variation about the mean.

OS defines the *maximum expected error* as three times the standard error plus the systematic error. This would be recognised by statisticians as a confidence interval, and is the most useful way of communicating accuracy to laypeople. If the errors come from a normal distribution we can be "99.7% confident" that the true value lies within this range.[2] The OS standard is used in this report and in all G&J Surveys' reports.[3]

Other authorities may have different standards. The US Mapping Agency uses 90% confidence. For a normal distribution of errors, this equates to 1.65 times the standard deviation. OS quotes a 95% confidence interval (common in scientific work) for the accuracy of passive stations; for normal errors this is twice the standard deviation.

<u>Outliers</u>

The usefulness of the statistics described above depend on the errors arising from normal statistical variation. Measurement systems may also produce anomalous results from equipment malfunction, mistakes in the way the measurement was conducted, or transcription errors.[4]

Results that are distant from other observations are known as outliers. Exclusion of outliers is a controversial topic. It is desirable to seek an explanation for an outlier before rejecting it. If this is not possible, a statistical test is sometimes applied. If the value is sufficiently unlikely to have come from the same population as the other observations, we might conclude that something has gone wrong. Although the statistical test is objective, the decision to exclude the result is subjective. It helps in the present work that the sample sizes are large, so the underlying distributions are reasonably well characterised.

## 1.2 Published statistics

Harley[5] gives figures for air heights, ground heights and trig points and some information on contours. These pertain to maps published before 1975.

Air heights: ±3.3m for 1:10k maps

Levelled heights: generally ±0.3m; some heights surveyed by telescopic alidade measurements are ±1m

Trig points: no general figure; ±2ft where measured by theodolite, better if spirit levelled.

Contours: 1:10k scale: ±5m for 10m vertical spacing and ±3m for 5m vertical spacing. 1:50k maps: better than ±9m.[6]

## 1.3 Data

### 1.3.1 Surveyed heights

Surveyed heights were supplied by Alan Dawson and G&J Surveys, both using Leica survey-grade GPS

---

[2] As the distribution may not be normal, I prefer to say "over 99% confident". Whatever, the interpretation is "practically certain".

[3] The 99.7% figure, or any alternative probability, assumes σ is known. In practice it will have been estimated from a sample. The effect of σ being an estimate, and therefore itself subject to uncertainty, is to decrease the confidence level. For most of the sample sizes in this report the difference is negligible. For small sample sizes one could increase the factor of 3 to maintain the desired confidence level. It would only be materially different for Harvey maps, where we have 39 spot heights.

[4] An example of a transcription error is provided by the Wainwright summit of Great Yarlside, hill 2575. The circular trig station was given as 1936ft on the original OS 6-inch map. This was misprinted as 1986ft in the 1920 revision. The incorrect height was transferred to both the 1-inch and 2½-inch maps, leading Wainwright to choose that location.

[5] J B Harley, *Ordnance Survey Maps a descriptive manual*, HMSO, 1975. Chapter 11 'The accuracy of Ordnance Survey maps'.

[6] Sources vary. Harley states that current (1973) OS policy is for the standard error not to exceed one quarter of the vertical interval, i.e. 2.5m for contours at 10m spacing and 1.25m for contours at 5m spacing. OS Land-Form PROFILE (a 1:10k digital height product) states rmse = 1.8m for 10m contours and 1.0m for 5m contours. The latter figure is consistent with Harley's experimental data for 5m contours on a provisional 1:10560 map in the Tiverton area, which gave an average rmse of 1.0m. OS Land-Form PANORAMA (an OS OpenData 1:50k product) states "rmse typically better than 3m"; its contours are taken from Landranger maps produced from aerial photography flown in the 1970s.

instruments. They comprise all surveys undertaken by these workers up to the end of 2014. The difference between the map height and the surveyed height was taken as the map error.

For some cols the accuracy of the surveyed height may be ±0.5m or more, particularly where levelling equipment was not employed, but this has a negligible effect on the analysis. If the standard error of the map height is *r* and the standard error of the surveyed height is *s*, the standard error of the difference in heights is $\sqrt{(r^2+s^2)}$. Harley's ±3.3m for the accuracy of an air height is confirmed by this work, so if the survey error is ±0.5m, its contribution to the total error is only 2%.

### 1.3.2 Map heights

Summit spot heights at 1:50k, 1:25k and 1:10k scales, and col heights at 1:25k and 1:10k, were obtained from Geograph mapping, which is derived from OS OpenSpace datasets[7]. Geograph does not say the scale at maximum zoom is 1:10k, but the level of detail is similar. For simplicity I have called the largest Geograph scale "1:10k". I also obtained 117 summit heights from the OS OpenSpace vector map (obtainable with the OS OpenSpace viewer or on the Hill Bagging website by zooming the main map). The vector map results are not reported here as the data are incomplete and all the heights appear on the Geograph 1:10k map[8]. 39 summit heights were obtained from Harvey maps.

Most spot heights at 1:25k and 1:10k in this work are air survey heights derived from metric mapping. The exceptions are a few spot heights on roads at cols which were determined by spirit levelling.[9] 1:50k heights come from a variety of sources. All the heights on the original Landrangers are metric conversions from the 1:63360 Seventh Series. These in turn are a mixture of levelled heights originating from 1:10560 and 1:2500 surveys in the 19$^{th}$ and early 20$^{th}$ centuries, and air heights from photogrammetry. The old heights are gradually being replaced by heights from metric mapping, but the process is far from complete. OS is also committed to eliminating height discrepancies across scales.

Most "old 1:50k" spot heights in this work were taken from Landrangers published before 1990. Note that some heights may have changed more than once.

Heights from old 6-inch maps were converted to the Newlyn datum using the facility on the OS website. Not all the hills were researched, but the 173 heights obtained are sufficient for a robust analysis.

Interpolated heights for 154 cols lacking a spot height were taken from the value in the DoBIH[10] before the survey result was entered. Most of these heights were checked in the 2012-13 data review.

Benchmark heights were taken from the OS legacy database available on the OS website and Trigpointing UK. They were all levelled between 1948 and 1970.

### 1.3.3 Location

Spot heights on summits were excluded if their location was clearly not the same as the surveyed location. This is not easy to ascertain on 1:50k maps because the location of heights transferred from old maps can be in error by up to 80m; see Appendix 1.

Spot heights on cols were allowed irrespective of distance from the surveyed point provided the location was plausible from the map and did not identify a different col. Thus the "map error" of a col includes the error due to OS having chosen an incorrect location (clearly a more difficult task than locating a summit).

[7] In February 2016 OS substituted less detailed maps for Geograph at the two highest zoom levels. The new maps show far fewer spot heights than previously, and the largest scale has lost its contours. The current work was completed before the change.
[8] I have only found one height on the OS Openspace vector map that does not appear on Geograph. This is a 375m spot on hill 5273 Mynydd y Grug.
[9] 1:25k maps show air heights in orange and ground heights in black. However some 1:25k maps in eastern Scotland, e.g. in the Cairngorms, make no distinction and show all heights in black.
[10] The Database of British and Irish Hills: www.hills-database.co.uk and www.hill-bagging.co.uk.

# 2 Results

## 2.1 Summits

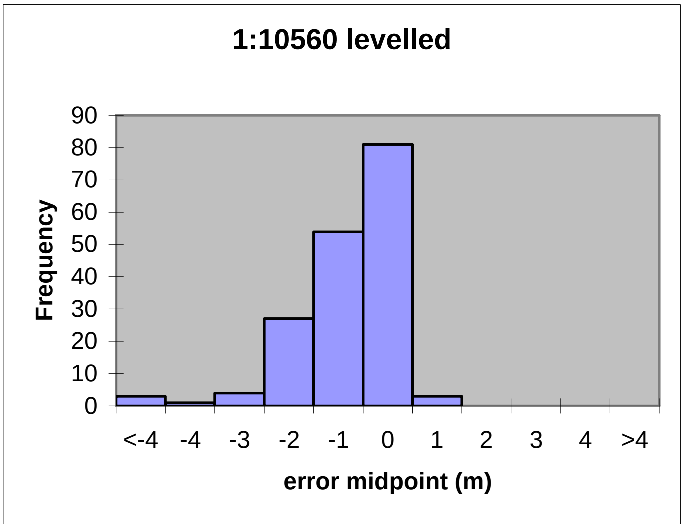
**Table 1. Accuracy statistics for summits**

| | | 1970s-80s Landranger | Geograph | | | historic 1:10560 |
|---|---|---|---|---|---|---|
| | Harvey | old 1:50k | 1:50k | 1:25k | 1:10k | levelled |
| rms error | 1.2 | 1.2 | 1.1 | 1.1 | 1.1 | 1.3 |
| | | | | | | |
| mean error | -0.1 | -0.2 | 0.1 | 0.2 | 0.2 | -0.8 |
| p-value for zero mean error | 0.5103 | 0.0085 | 0.1058 | 0.0001 | 0.0006 | 0.0000 |
| | | | | | | |
| standard error | 1.2 | 1.2 | 1.1 | 1.1 | 1.1 | 1.0 |
| | | | | | | |
| mean absolute error | 1.0 | 0.9 | 0.8 | 0.9 | 0.8 | 0.9 |
| max absolute error in sample | 2.2 | 4.2 | 4.9 | 3.7 | 3.7 | 6.6 |
| | | | | | | |
| maximum expected error - lower | -3.7 | -3.8 | -3.2 | -3.1 | -3.0 | -5* |
| - upper | 3.4 | 3.4 | 3.4 | 3.5 | 3.4 | 1* |
| | | | | | | |
| sample size | 39 | 199 | 328 | 378 | 359 | 172 |

* The OS formula gives –3.8 to +2.2 but is inappropriate due to the highly skewed error distribution

Results are given in Table 1. Statistical outliers (see below) were omitted from the calculations. There is no significant difference between the standard errors at the three Geograph scales, or between old and current 1:50k maps and Harvey maps. The maximum expected errors on current maps are in good agreement with Harley's figure of ±3.3m for air heights.

Spot heights at 1:25k and 1:10k have a small positive bias: on average they are 0.2m too high. Although small, the bias is statistically highly significant. There is no significant bias in online 1:50k maps. However old Landrangers have a significant negative bias of –0.2m. This grows to –0.8m for old levelled heights. The bias is evident from a histogram of the errors (Figure 1).



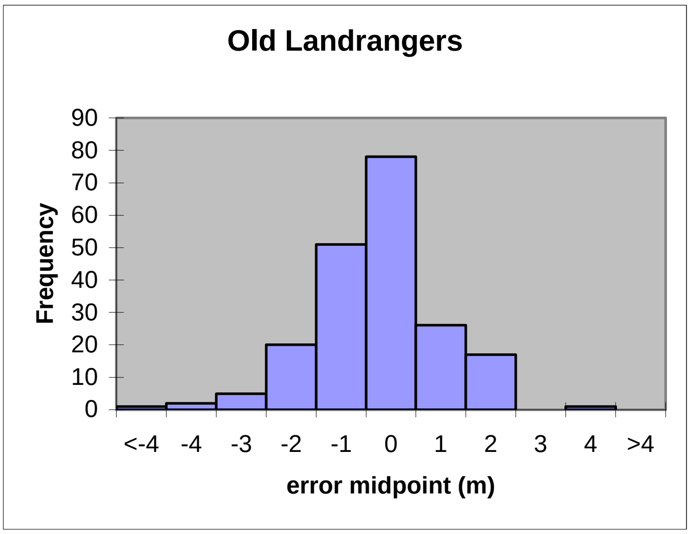




**Figure 1. Histograms of errors in heights on old maps (outliers included)**

The distribution of errors in levelled heights is revealing. The overall accuracy is about the same as for air heights but the distribution is skewed to the left (map height too low). Of 173 hills, 81 have an error between –0.5 and +0.5m, 89 hills have a negative error exceeding –0.5m, and only 3 hills have a positive error exceeding +0.5m (+0.59, +0.69 and +0.75m). The knowledge that a levelled height may be significantly lower than the true height but is unlikely to be higher suggests how these heights are best used[11].

There is a ready explanation for this finding. Most of the old heights are trigonometric stations that have a block buried below the surface or a rivet driven into rock. The surveyor would have chosen a convenient location rather than necessarily the highest ground, as with the later Hotine pillars which are frequently not at the highest point.

The reason for the positive bias in air heights is less obvious, but might be due to the top of some cairns being spotted. It is not usually possible for a cartographer to distinguish between a cairn and a rock on the photogrammetry plate. Only a minority of cairns are marked on OS maps. The bias at 1:25k grows to +0.8m for the 32 hills showing a different height at 1:50k or 1:10k. This finding is discussed further in section 3.1.

The analysis was repeated for the 64 hills that show a height difference between old and new 1:50k maps. There is no significant difference in standard error but the old 1:50k has a bias of –0.7m. For most of these hills the old 1:50k height corresponds to an old levelled height, so the explanation is similar.

### 2.1.1 Outliers

All the map scales have one or more extreme values. The frequency distributions suggest they are more likely to be due to faulty data than to a heavy tailed distribution. The 10 largest negative and 10 largest positive errors are given below. Those highlighted are statistically significant. Details of the outliers are given in Table 2.

old Landrangers
-15.3 -4.2 -3.6 -3.4 -2.8 -2.7 -2.6 -2.5 -2.5 –2.5 . . . . . . . . 1.8 1.9 1.9 2.1 2.1 2.2 2.2 2.2 2.4 3.6

1:50k
-20.9 -15.3 -10.6 -8.9 -4.9 -3.6 -3.3 -3.1 -2.9 -2.8 . . . . . . 2.3 2.3 2.4 2.5 2.6 2.6 2.9 3.1 3.2 17.6

1:25k
-10.6 -3.6 -3.3 -3.2 -2.5 -2.2 -2.2 -2.1 -1.9 -1.9 . . . . . . . . 2.6 2.8 2.9 2.9 3.1 3.1 3.2 3.2 3.7 17.6

1:10k
-10.6 -3.6 -3.3 -3.2 -2.2 -2.2 -2.2 -2.1 -2.0 -1.9 . . . . . . . . 2.6 2.8 2.9 2.9 3.1 3.2 3.2 3.7 6.7 17.6

levelled
-14.3 -6.6 -4.9 -4.4 -3.4 -2.9 -2.8 -2.7 -2.5 -2.5 . . . . . . . . . 0.1 0.1 0.2 0.2 0.2 0.2 0.2 0.6 0.7 0.8

The +17.6m and –10.6m errors are common to three different scales. It is a matter of conjecture as to whether the heights were independently obtained from the same photogrammetry or were copied. For hill 2950 the 1980s LR50 and LR56 have only an 840m contour. For hill 3389 the 1983 LR124 has only a 520m contour, like the modern map. In all three Landrangers the contours are from the metric survey. For hill 902, the highest contour at 1:50k would be 890m if we can assume that most of the 860m contour is missing. There is no height on 6-inch maps, but there is a 2950ft (899m) contour on the 1-inch Seventh Series map.

One of the hills in Table 2 (Craig yr Hafod) was surveyed after repeated measurements on Garmin GPS instruments suggested the map was incorrect, hence the dataset may contain more outliers than would be expected from a random sample of hills. On current evidence, around 0.5% of spot heights on current maps are spurious, giving values well outside the expected error ranges in Table 1.

[11] Subsequent to this work an old levelled height has been found with an error exceeding +19m. Beinn an Lochain (hill 1419) was for many years classified as a Munro on the strength of a 3021ft spot height on the 1874 1:10560 Sheet CXXXIV, reproduced in the 1924 revision. As with much of the Scottish Highlands, the mountain was not resurveyed after the original 1870 survey until the National Grid survey in 1975, which gave an air height of 901m. The 3021ft spot is missing from the 1900 Sheet CXXXIV.SE, which however retains a spurious 2992ft spot between the 3021ft spot and a 2955ft trig. All 1-inch maps from 1876 onwards give only the 2955ft trig. The summit was surveyed by Alan Dawson as 901.7m.

**Table 2. Outliers excluded from the summary statistics**

| hill no. | hill name | map height | diff GPS | error | scale | comments |
|---|---|---|---|---|---|---|
| 902 | Sgurr na Conbhaire | 881 | 901.9 | -20.9 | 1:50k | 1:10k has 901; 1:25k has 900 contour |
| 2950 | Stac a'Chuirn | 870 | 852.4 | 17.6 | all scales | inside 860 contour; old 6" map has 850.1 |
| 243 | Drochaid Ghlas | 1009 | 1024.3 | -15.3 | 1:50k | no 1010,1020 contours at any scale; rocky ridge |
| | | | | -14.3 | levelled | |
| 3389 | Craig yr Hafod | 523 | 533.6 | -10.6 | all scales | no 530 contour |
| 18914 | Meall an Daimh | 753 | 761.9 | -8.9 | 1:50k | old 6" map has 757.5, in error by -4.4m but not an outlier |
| 340 | Meall na Duibhe | 578 | 571.3 | 6.7 | 1:10k | 1:10k has 578 and 573 close together; online 1:50k & 1:25k have 573; former 578 spot on vector map has been replaced by 573. Transcription error? |
| 1253 | Sgurr a'Fionn Choire | 930 | 936.0 | -6 | Harvey | due to terrain? |

## 2.2 Cols

**Table 3. Accuracy statistics for cols**

| | 1:25k | 1:10k | interpolated |
|---|---|---|---|
| rms error | 1.3 | 1.4 | 2.1 |
| | | | |
| mean error | 0.5 | 0.6 | 0.7 |
| p-value for bias | 0.0000 | 0.0000 | 0.0000 |
| | | | |
| standard error | 1.2 | 1.2 | 2.0 |
| | | | |
| mean absolute error | 1.0 | 1.1 | 1.6 |
| max absolute error in sample | 4.3 | 4.3 | 6.7 |
| | | | |
| maximum expected error - lower | -3.1 | -3.0 | -5.2 |
| - upper | 4.2 | 4.3 | 6.6 |
| | | | |
| sample size | 95 | 156 | 154 |

Results are given in Table 3. Two outliers at the Geograph 1:10k scale were omitted from the calculations.

The errors at 1:25k and 1:10k are greater than for summit heights. This is to be expected because a difference in location was ignored provided the spot height identifies the same col; hence they are not purely measurement errors. There is no significant difference between the errors at the 1:25k and 1:10k scales. The error in interpolated heights is about 60% greater than for spot heights.

There is a significant positive bias for both spot heights and interpolated heights, i.e. cols are on average somewhat lower than expected from the map. This bias gives rise to the asymmetric confidence intervals. Histograms of the errors in the interpolated heights (Figure 2) are somewhat skewed towards positive errors. The histograms of 1:25k and 1:10k are similar. An examination of the individual hills might reveal particular types of terrain that are more prone to systematic and/or random errors in the col height.

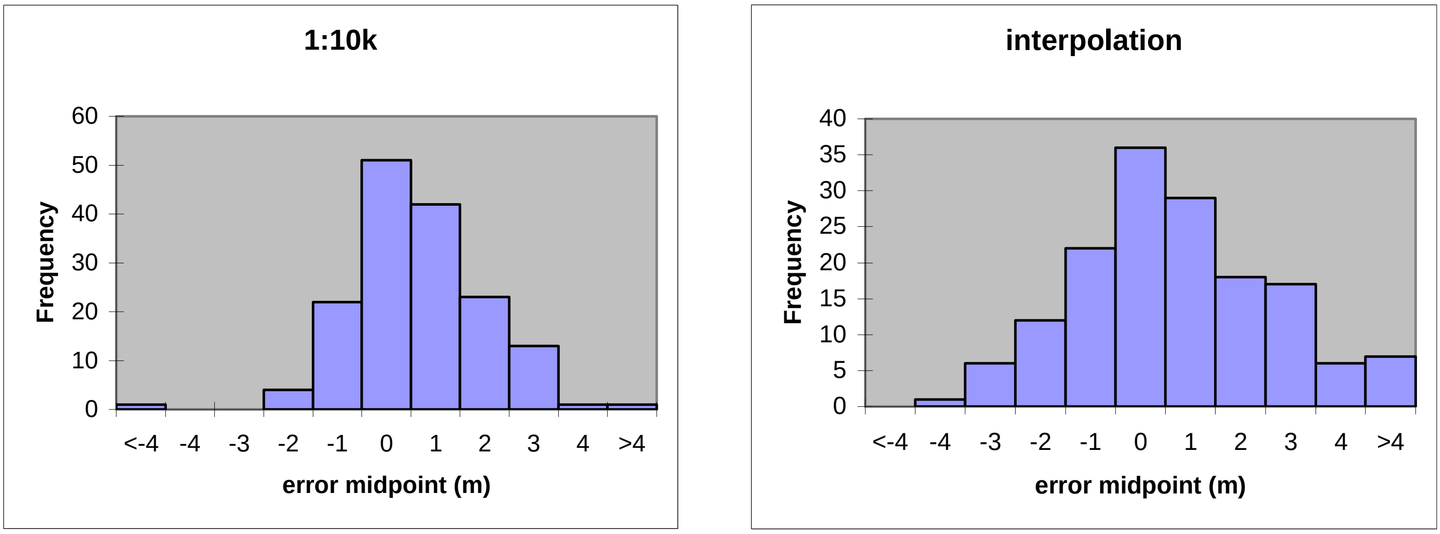



**Figure 2. Histogram of errors in col heights (outliers included in 1:10k)**

The maximum expected error in Tables 1 and 3 pertains to the error in the map height. It may be more helpful to the user of the statistics to give a confidence interval for the true height, by reversing the signs. Thus if the map height is *x* metres, the true height is expected to be within –6.6 and +5.2m of *x* for an interpolated col.

### 2.2.1 Outliers

There are only two statistical outliers, both significant at 99% confidence.

1:25k
-2.0 -1.9 -1.6 -1.3 -1.3 -1.1 -1.1 -1.0 -0.9 -0.9 . . . . . . . 2.1 2.3 2.5 2.7 2.7 3.0 3.1 3.2 3.5 4.3

1:10k
-7.1 -2.0 -1.9 -1.6 -1.6 -1.3 -1.3 -1.2 -1.1 -1.1 . . . . . . . 2.8 2.8 2.9 3.0 3.1 3.2 3.4 3.5 4.3 8.3

interpolated
-4.2 -3.1 -3.1 -3.1 -3.0 -2.6 -2.6 -2.4 -2.2 -2.0 . . . . . . . 3.9 4.0 4.5 4.6 4.6 4.7 4.9 5.6 5.7 6.7

**Table 4. Outliers excluded from the summary statistics**

| hill no. | hill name | map height | diff GPS | error | scale | comments |
|---|---|---|---|---|---|---|
| 244 | Stob Diamh | 873 | 864.7 | 8.3 | 1:10k | between 860 and 870 contours, possibly bump? |
| 2306 | Hedgehope Hill | 560 | 567.1 | -7.1 | 1:10k | Transcription error? 1:25k has 566 |

## 2.3 Benchmarks

21 flush brackets, mostly on trig pillars at hill summits, were measured by G&J Surveys using 1 hour or longer collection times. OS heights were taken from the OS legacy database. 11 benchmarks are order 3, 9 benchmarks are order 2, and 1 benchmark is order 1.

**Table 5. Accuracy statistics for benchmarks**

| | OS database |
|---|---|
| rms error | 0.11 |
| | |
| mean error | -0.02 |
| p-value for bias | 0.4204 |
| | |
| standard deviation | 0.11 |
| | |
| mean absolute error | 0.07 |
| max absolute error in sample | 0.31 |
| | |
| sample size | 21 |

Statistics for the differences in heights are given in Table 5. There are no outliers. For consistency with the summit and col analyses, the error is defined as *OS database height – GPS height*. The standard deviation of the G&J measurements is 0.019m for 60 minutes data collection and the rms error is conservatively estimated as 0.028m[12]. The GPS error contributes only 3% to the standard deviation in Table 5 (see section 1.3.1) and slightly reduces the rms error.

The mean error of –0.02m is not remotely significant. The largest error is –0.314m for Thack Moor, which is an order 2 pillar. The next largest are Earl's Hill +0.20m, Calf Top –0.16m, Moel Tryfan +0.13m, and Beinn Talaidh –0.13m. The order 3 benchmarks are more accurate than the order 2 benchmarks, which despite the small sample size is statistically significant. This result is unexpected, as the accuracy should fall with order.

## 3. Implications for hill data

### 3.1 Conflicting heights

When faced with choosing between different heights, authors of hill lists have generally preferred those at larger scales. OS advised the Nuttalls to prefer 1:10k or 1:25k heights to those on 1:50k maps, which at the time of their research were all conversions of imperial heights. Our results give no reason to prefer larger scales.

Theoretically the best statistical estimator of height would be a weighted average of independent determinations (with equal weights in our case). Most often the spread is 1m. The average, rounded to the nearest metre, then equates to choosing the height that occurs twice. The principle assumes the locations are the same, which may be difficult to verify on 1:50k maps because positional errors of spots transferred from imperial maps can be considerable (Appendix 1). It also breaks down if the heights are not independently obtained, e.g. if one was copied. This is probably the case with many hills and will become increasingly likely in the future as the OS is committed to resolving discrepancies across scales.

The “2 versus 1 rule” was used by the editors of the Database of British and Irish Hills in data reviews covering many thousands of hills as it is very simple to apply, and in the event of the heights not being independent there was at the time no reason to prefer one height to the other. There are only 28 such hills in this dataset, but the height occurring twice is significantly *less* accurate: it has a bias of +0.9m and a higher standard error. For 16 hills the 1:50k height is the odd one out, and for 12 hills it is the 1:10k height. There is no significant difference in accuracy between unique 1:50k heights and unique 1:10k heights.

---

[12] The standard deviation was estimated from repeated measurements at a fixed location over a 2 week period and is more than double the figure reported by the instrument on individual surveys (the latter figures are given in G&J Surveys’ reports). There is a systematic difference of –0.02m between replicate measurements on the Daresbury FBM and the OS value for both the Leica 530 and GS15 instruments, i.e. they give results 2cm lower than the OS figure. This is not necessarily the instruments’ bias because the standard error of OS passive stations is stated to be 0.033m.

There are no hills among the 28 where the 1:25k height is unique. Possibly the 1:25k height is being copied to other scales. There are a further 6 hills where the height is missing at one scale and differs between the other two scales, giving a total of 34 hills with height conflicts. Analysis of the 34 hills at each scale shows no significant difference in their standard errors, but the 1:25k heights have a significant bias of +0.8m. A possible explanation is that the cartographers on this series are sometimes spotting the tops of cairns.[13] The 1:50k and 1:10k heights are unbiased. Hence a better rule for 2 versus 1 situations might be "prefer the height that does not appear at 1:25k". However it might be unwise to draw firm conclusions from this relatively small sample.

For the six hills offering a choice between two singly occurring heights (1 versus 1) there is no significant difference in accuracy between the higher and lower height, but the sample is much too small to draw robust conclusions. The DoBIH generally chooses the higher on the grounds that it might be closer to the summit. When the higher spot occurs at 1:50k it could well originate from an old levelled height and our results indicate that an old height should be preferred if higher. Hill 368 Mullach Coire nan Nead has spot heights of 921m at both 1:50k and 1:25k, and 922m at 1:10k. The old 6-inch map gives 3025ft which converts to 922.0m[14]. The finding that levelled heights may be lower but are seldom higher than the true height by more than 0.5m led the DoBIH to take 922m.

When summit and col are close together we would expect the errors from the photogrammetry to be correlated, and the evidence points that way. It may therefore be advantageous to take the summit and col heights from the same map when calculating drop. In the past this has occasionally given a conflict with the rules above. Hill 535 Cnap a'Chleirich is 1172m at 1:50k and 1:25k, and 1174m at 1:10k. However only the 1:10k map spots the col. One could choose 1172 or 1173m for the summit and 1174m for calculating the drop and accept the discrepancy between the col height and the drop, or perhaps lower the col height. The 1174m spot is c.15m NE of the 1172m spot and the summit located by a walker's GPS is a further 5m NE. The summit feature is a rock tor, which raises the possibility that only the 1:10k map has spotted the summit. Hence 1174m was adopted for the DoBIH. Preferring the height that doesn't appear on the 1:25k map would have prevented a conflict in the first place, but it might give conflicts on other hills.

### 3.2 Use of benchmarks

Some summits in the DoBIH were surveyed by levelling to a trig flush bracket, most often by Abney Level when the highest ground is close to the pillar. The accuracy of OS benchmarks justifies giving a decimal height when the levelling is accurate to 0.1m, and the DoBIH does so. For very marginal hills such as Thack Moor and Calf Top the accuracy of the benchmark is inadequate. The error in the OS height for the Thack Moor trig is 0.3m and from Harley's figure we can expect some trigs to exceed this.

### 3.3 Interpolating cols

Many hills lack a spot height at the col. British list authors have become more sophisticated in their use of maps and are now much more likely to interpolate the height by visualising the topography, rather than take the midpoint of the contours enclosing the col or the lowest contour in the hill-hill direction (the latter method, which gives a biased estimate of drop, is prevalent in the US). The DoBIH interpolates col heights in this way. It takes some skill to create a mental model of the 3-dimensional surface, but the editors have improved with practice. The statistics in Table 3 indicate that on average, the error is 60% higher than when a spot height is available. If the contribution of contour error is removed, the standard error of the human interpolation is 1m. This is a creditable result and about half the error from taking the midpoint of the contours. There is a tendency for both spot heights and interpolations to overestimate the col height.

Irregularities in the terrain will decrease the accuracy of the interpolation. This is particularly likely in rocky ridges. Digital elevation models suffer from the same limitation. Garmin Topo map will interpolate heights when the underlying vector data is not too sparse. An analysis of 23 Topo heights versus surveyed heights,

[13] OS suggested this explanation for the +3.7m error on hill 1025 Beinn Dearg Mor.
[14] 922m also appears on the Geograph 1:250000 map.

excluding two steep-sided cols for which the granularity of the data would be insufficient, gave an rms error of 2.0m, a mean error of +1.0m, and a standard error of 1.8m. These values are no better than the statistics for human interpolation in Table 3.

## 4. Implications for hill lists

The majority of hill lists use classification criteria based on height, drop, or both. Inevitably there will be hills whose correct classification is not assured because the potential measurement error exceeds the difference from the threshold.[15] Knowledge of the measurement error and elementary statistical theory enable the qualification probability to be calculated for any hill. Such information can be used to draw up a list of candidates for surveying, or by baggers to decide which hills outside the list ought to be climbed, either to insure against future promotions or simply to provide assurance that all hills meeting the criterion have been climbed.[16]

For most purposes I would advocate a shortlist based on the maximum error ranges given in Tables 1 and 3. If they seem conservative, bear in mind that the cumulative probability of at least one marginal hill changing classification soon becomes substantial as the number of hills increases.

The error in drop has contributions from both the summit and col heights. If both summit and col have spot heights, the maximum error in drop is –4 to +6m. If the summit has a spot height and the col is interpolated, the maximum error is –6 to +8m. These ranges do not take into account the possible correlation in errors at summit and col. The effect of correlation would be to reduce the upper and lower bounds by up to 1m. If the summit has been surveyed, only the error in col height is relevant which can be taken from Table 3. As the errors are *map height – surveyed height* the shortlist should subtract the error ranges from the threshold, e.g. for Marilyns the list should include hills with drops of 144–154 or 142–156m according to whether the col is spotted or interpolated, respectively.

Outliers will remain a problem. Non-survey grade GPS units of the type used by walkers have an accuracy of about ±13m for height and, as with maps, can give extreme values.[17] Hence a single measurement is of little diagnostic use. However repeated readings at odds with the mapping would suggest something is amiss. Hill 3389 Craig yr Hafod has a spot height of 523m at all scales. After four walkers had submitted GPS data with heights averaging 15m in excess of this figure, the hill was surveyed in November 2011 and found to be 533.6m.[18]

**Acknowledgments**

I am grateful to Alan Dawson and G&J Surveys for supplying their survey data in a user-friendly format. Thanks are due to Dave Marshall for assisting with the old maps research and extracting interpolated col heights from past releases of the DoBIH, and to those who commented on the draft.

---

Chris Crocker
29 March 2016

.../Appendix 1

---

[15] At present only the Munros are free from such uncertainties, all the marginals having been surveyed at the instigation of The Munro Society.

[16] The approach is described by the author in www.hills-database.co.uk/probMunros.html. It was first published in 1998 in *The Angry Corrie*.

[17] A precision of 12.6m (3 standard deviations) was obtained by Graham Jackson from 100 replicate measurements on a Garmin eTrex over 7 weeks. A precision of 12.2m was obtained by the author from 1099 GPS measurements on trig pillars submitted to the DoBIH by subtracting the OS height (at the top of the pillar) from the GPS height. The five largest errors were +30, +25, -21, -19, –18m. The analysis also found a bias of +1.4m in the GPS height and a heavy tailed distribution of errors (probably due to the variety of instruments and modes of use) so a better interval for the error giving 99% confidence would be –11 to +14m. The positive bias has been noted by many users so is probably endemic, at least to Garmin models.

[18] Ordnance Survey agreed to change the height on its maps but has not done so.

## Appendix 1. Positional errors on 1:50k maps

Spot heights and other features on the original Landranger maps were transferred from 1:63360 (1 inch to the mile) maps. According to Oliver[19], errors in the positions of these features arise from at least two sources:

1) Change in projection. Early surveys used the Cassini projection. This was easy to construct and reasonably accurate for small areas but unsuitable for large scales, which caused problems when the 1-inch maps were produced. Transverse Mercator was adopted for newly drawn maps in 1931, but it was not until the completion of the metric survey in the early 1980s that OS had a complete set of data surveyed to TM. To quote Oliver, "the geodetic basis of the mapping may explain apparent discrepancies in the position of a feature apparently unchanged on the ground ... it follows that if mapping drawn on the Cassini projection is superimposed on that drawn on the Transverse Mercator projection the distortion of angles in the former and of distance in the latter means they are liable to be at odds with each other, even if there is no distortion from other causes".

2) Paper distortion due to shrinkage over time. Apparently the County Series maps were digitally scanned on flatbed scanners without an allowance for paper shrinkage, even though the OS could have made an approximate adjustment using the printed scale. Oliver found that positions of features transferred from 6" maps could be in error by up to 70m.

For pre-digital printed maps, misalignment of the overlays offers another potential source of error. I have seen an old Landranger in which the N-S gridlines were displaced 100m from their correct positions.

I took a sample of 20 hills for which the heights on online 1:50k and 1:25k maps can be assumed to relate to the same feature. The difference between the two scales ranged from 5m to 80m. The mean difference was 30m.

A well-known discrepancy concerns the northern summit of hill 2318 Housedon Hill. The 267m spot on both old and current 1:50k mapping is well inside the wood and coincident with a 260m contour. The 6-inch map shows an 877ft trig (267m) only a few metres east of the forest fence. There is a cairn at the latter point which two independent surveys found to be the highest ground. As there is no spot at 1:25k or 1:10k, the 1:50k spot has almost certainly been transferred from the 1-inch Seventh Series map which shows 877 next to a cairn in the same incorrect location. This 1:50k spot is 50m out of position.

[19] R Oliver, *Ordnance Survey Maps a concise guide for historians*, Charles Close Society for the Study of Ordnance Survey Maps, c/o The Map Library, British Library; Second Edition, 2005.