watersurface.py User Guide

User Guide for Depth Offset Correction and Diagnostic Output

Purpose. This script estimates the near-surface depth offset from a histogram of raw depth values, subtracts the offset, clips negative values to 0 m, and optionally applies temperature-based correction when justified by the data. It also exports diagnostic figures that allow the correction to be checked visually.

1. Overview

- The script is intended for biologging datasets that contain at least time and depth columns. It is particularly useful for shallow-diving animals, because even a small offset around the surface can affect dive detection and the interpretation of shallow events.
- The recommended workflow is simple: run the script with the default offset-estimation settings, inspect the summary figure, and keep the result if the selected offset matches the near-surface peak and the corrected time series aligns better with the expected surface baseline.

2. Expected Input

Required columns: time, depth

Optional column: temperature

3. Input Parameters

Prompt	Default	Use in most cases	Adjust when...
Input CSV filename	—	Enter the target CSV file. The .csv extension may be omitted.	Only if the file is stored under a different name or path.
Lower/upper bound for peak search in depth0	-0.5 / 0.5 m	Keep the default when the surface peak is expected near 0 m.	Shift or widen the range if the raw near-surface peak is clearly outside the default search window.
Lower/upper bound for histogram display	-2 / 2 m	Use the default for routine inspection.	Adjust when the preview is too zoomed in or too wide to show the relevant peak clearly.
Number of bins for peak detection	100	Use the default for most files.	Increase for finer resolution if the peak is broad; decrease if the histogram is too noisy.
Histogram smoothing sigma	2.0	Use the default for stable peak detection.	Increase if many small local peaks appear; decrease if the main peak becomes overly smoothed.

Estimate the offset automatically?	y	Recommended for standard processing.	Choose n only when visual inspection suggests that the automatic peak is incorrect or ambiguous.
Apply temperature correction?	n	Test only when temperature-dependent drift is suspected.	Use y if shallow corrected values still appear to change systematically with temperature.
Upper depth limit for temperature correction	0.3 m	Good for near-surface screening.	Increase slightly if too few points are available; decrease if shallow dives contaminate the near-surface band.
Correlation threshold |r|	0.6	A conservative threshold for applying temperature correction.	Lower only if you intentionally want a more permissive correction.
Minimum temperature range	0.5 °C	Keep the default.	Change only for unusual deployments with small but biologically meaningful temperature variation.
Minimum number of points	2000	Provides a stable basis for regression.	Lower only for short datasets, recognizing that reliability decreases.
Suffix for output CSV filename	_offset_corrected	Keep the default in most cases.	Change when you need multiple correction versions for the same file.

4. Example Run

```
python watersurface.py
Input CSV filename (extension .csv may be omitted): 9B12280_ORI400-D3GT_23_20110829_standardize

--- Offset estimation (histogram) ---
Lower bound for peak search in depth0 [m] (e.g. -0.5) (Enter = -0.5):
Upper bound for peak search in depth0 [m] (e.g. 0.5) (Enter = 0.5):
Lower bound for histogram display [m] (e.g. -2) (Enter = -2.0):
Upper bound for histogram display [m] (e.g. 2) (Enter = 2.0):
Number of bins for peak detection (e.g. 100) (Enter = 100):
Histogram smoothing sigma for peak detection (in bin units, e.g. 2.0) (Enter = 2.0):
Estimate the offset automatically from the histogram? [y/n] (Enter = y):
[Info] Saved histogram preview for offset estimation: 9B12280_ORI400-
D3GT_23_20110829_standardize_offset_hist_preview.png
[Info] Estimated offset = 0.180000 m (subtracted from depth0)

--- Temperature correction (optional) ---
Would you also like to try temperature correction? (applied after offset correction + zero clipping)
[y/n] (Enter = n): y
Upper depth limit [m] for temperature correction (e.g. 0.3) (Enter = 0.3):
Correlation threshold |r| required to apply temperature correction (e.g. 0.6) (Enter = 0.6):
Minimum temperature range [°C] (e.g. 0.5) (Enter = 0.5):
Minimum number of points required for temperature correction (e.g. 2000) (Enter = 2000):
[Info] Correlation between temperature and depth_offset: r=0.4157, temp_range=4.300 °C, n=171965
[Info] Temperature correction not applied because |r| < 0.6 (offset correction + zero clipping only).
Suffix for output CSV filename (e.g. _offset_corrected) (Enter = _offset_corrected):
[Info] Wrote corrected CSV: 9B12280_ORI400-D3GT_23_20110829_standardize_offset_corrected.csv
[Info] offset=0.180000 m, temp_correction_applied=False
[Info] Saved overlaid histogram: 9B12280_ORI400-
D3GT_23_20110829_standardize_depth_hist_before_after.png
[Info] Saved overlaid time-series plot: 9B12280_ORI400-
D3GT_23_20110829_standardize_depth_timeseries_before_after.png
[Info] Saved summary figure: 9B12280_ORI400-
D3GT_23_20110829_standardize_watersurface_offset_summary.png
```

5. Output Files

- `*_offset_corrected.csv` Main corrected data file.
- `*_depth_hist_before_after.png` Histogram comparison before and after correction.
- `*_depth_timeseries_before_after.png` Time-series comparison before and after correction.
- `*_watersurface_offset_summary.png` Combined diagnostic figure with the key checks on a single page.

6. Reading the Summary Figure

Top panel: Offset estimation (peak region)

This panel shows the histogram of raw depth values near the surface together with the smoothed curve used for peak detection. The selected offset should pass through or very close to the dominant near-surface peak. In the present example, the peak is centered around 0.18 m, so subtracting 0.18 m is reasonable.

Middle panel: Depth distribution (full range)

This panel compares the full depth distribution before and after correction. A successful correction shifts the near-surface distribution toward 0 m without distorting the overall depth structure.

Bottom panel: Time series comparison

This panel is often the most intuitive diagnostic. Blue depth0 and orange corrected depth should show the same dive structure, but the corrected series should align better with the expected surface baseline. In this file, the corrected line is clearly shifted upward by approximately 0.18 m while preserving the pattern of deeper events.

7. Interpretation of the Present Example

For this file, subtracting 0.18 m provides an appropriate depth correction. The histogram-based offset is supported by the near-surface peak in the top panel, the shift of the full distribution toward 0 m in the middle panel, and the improved surface alignment in the time-series panel. Temperature correction was tested but not applied, because the temperature–depth relationship ($r = 0.4157$) did not exceed the chosen threshold of 0.6. This is therefore a successful example of offset correction without temperature correction.

8. Practical Notes

- In routine processing, the default offset-estimation parameters are usually sufficient.
- Manual offset selection is most useful when multiple near-surface peaks are present.
- Temperature correction should be treated as optional and conservative; it should be applied only when the relationship is both strong and based on sufficient shallow data.
- The summary figure is the most efficient single diagnostic because it combines the statistical basis of the correction with its effect on the full time series.

Example Output Figure

This figure (Fig. 1) shows an example of the offset estimation and depth correction workflow applied to a Streaked Shearwater dataset.

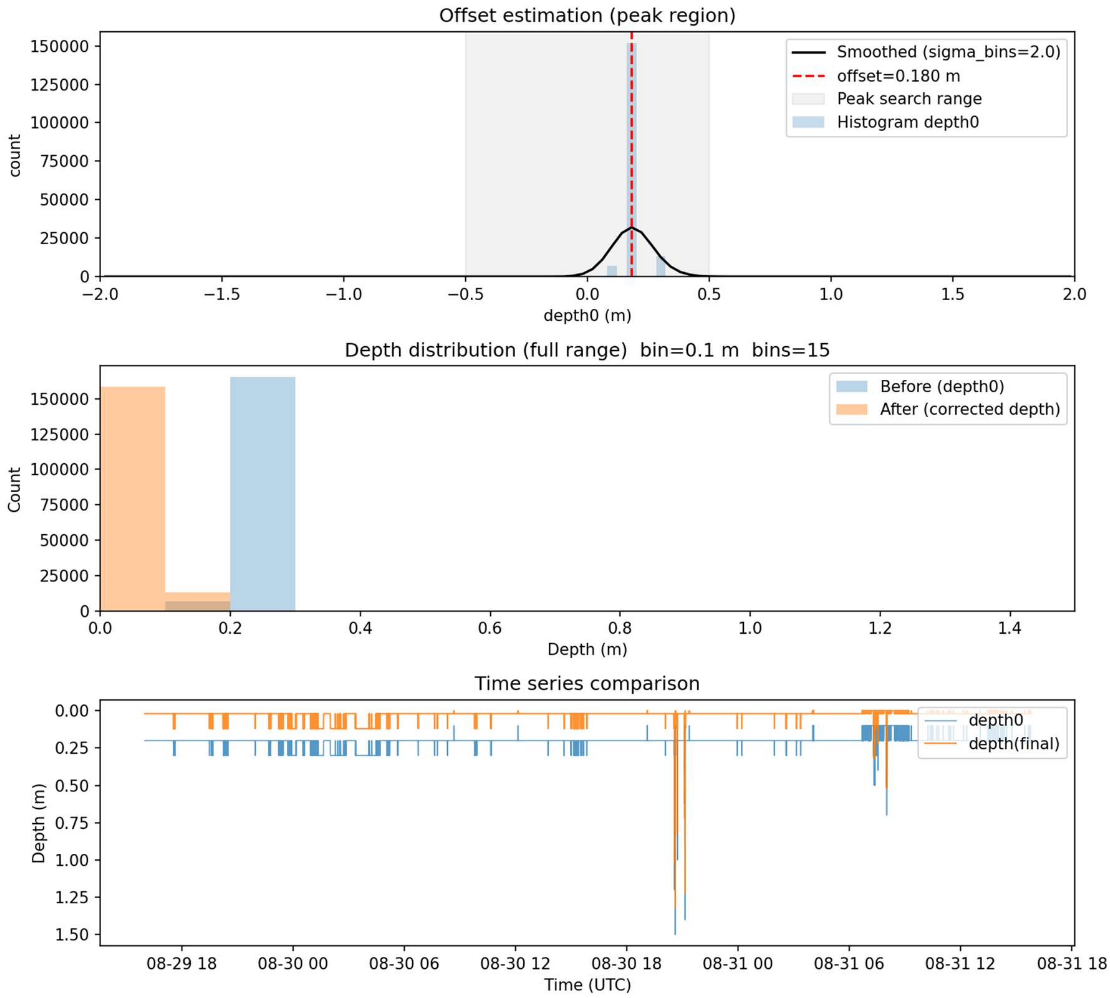

Fig. 1. Summary output from the watersurface.py script. The upper panel shows the offset estimation based on the histogram peak of shallow depth values. The middle panel shows the depth distribution before and after correction. The lower panel shows the time-series comparison between the original depth (depth0) and the corrected depth.