

Filtering protocol for temp files, with intermittent temp failures and or intermittent unrealistic temperature spikes

T.McIntyre

Mammal Research Institute, University of Pretoria

In Microsoft Excel

- 1) Ensure that the number of temperature profiles add up to a factor of 12 (of the alternative number of temp values reported per profile). If it does not, identify the profile(s) where the number is not 12 (usually less than 12) and remove. This can be done manually, or by writing a little formula that will flag values if there are not 12 successive lines with exactly the same date and time data.
- 2) Flag temp values where any value is more than 40% different (+ or -) from the previous value, *within the same profile*. The excel formula would look something like this:

$$= \text{IF}(\text{OR}((t2 > (0.4 * t1) + t1), (t2 < (t1 - (0.4 * t1))))$$
$$* \text{AND} (\text{time2} = \text{time1}), \text{"NOT"}, \text{"OK"})$$

Where $t2$ = the second temp value in a profile

$t1$ = the first temp value in a profile

- 3) Now flag the *profiles* that contain identified values of "NOT" previously. Such a formula would need to be run over sets of 12 temp points, and would look something like this:

$$= \text{IF}(\text{OR } t12 = \text{"NOT"}, t11 = \text{"NOT"} \dots, t1 = \text{"NOT"}), \text{"NOT"}, \text{"OK"})$$

Where $t12:t1$ = the temp values of points 1 through to 12 of each profile

- 4) Sort the dataset based on temp values and remove values that are extreme [in most cases on the Marion dataset I removed values $> 15^\circ\text{C}$; and $< -3^\circ\text{C}$, but this is obviously related to the oceanographic area that the track encompassed].
- 5) Sort the dataset based on the flag values created in point 3, and remove all values flagged as "NOT".