the effect of scaling on the smoothing parameter choice

Compare the expression

$$E_p(f) := p\sum_i (y_i - f(x_i))^2 + (1-p)\int (f^{(m)}(x))^2 \mathrm{d}x$$

to be minimized in smoothing with the corresponding expression in terms of $r\tilde{x} := x$. With $\tilde{f}(\tilde{x}) := f(r\tilde{x})$, we have $\tilde{f}^{(m)}(\tilde{x}) = r^m f^{(m)}(r\tilde{x})$ and $\mathrm{d}\tilde{x} = \mathrm{d}x/r$. Hence

$$\begin{aligned}\tilde{E}_q(\tilde{f}) &:= q\sum_i (y_i - \tilde{f}(\tilde{x}_i))^2 + (1-q)\int (\tilde{f}^{(m)}(\tilde{x}))^2 \mathrm{d}\tilde{x} \\ &= q\sum_i (y_i - f(x_i))^2 + (1-q)\int r^{2m}(f^{(m)}(x))^2 \mathrm{d}x/r.\end{aligned}$$

Thus, to obtain the same minimizer, we must have

$$p/(1-p) = q/((1-q)r^{2m-1})$$

or

$$p = q/(q + (1-q)r^{2m-1}).$$

As a particular example, John Bennett changed from degrees to kilometers, going from the interval $[-76\,..\,-73]$ to the interval $[0\,..\,625]$. Taking $\tilde{x}$ to be in degrees and x in kilometers (and ignoring the irrelevant translation), this gives $r = 625/3$. Further, when doing the smoothing in degrees, he used $q = .999$. Thus, when doing the smoothing in kilometers (and given that he does cubic smoothing, i.e., $m = 2$), he should be using

$$p = .999/(.999 + (.001)(625/3)^3) = .0001105$$

instead of .999.

In the y-direction, he goes from a 3-degree range to $[0\,..\,777]$, hence should now use

$$p = .999/(.999 + (.001)(777/3)^3) = .0000575$$

These are seemingly dramatic changes and show the trickiness of dealing with a choice of the smoothing parameter cold, i.e., without reference to the units used.

As a check, recall an earlier calculation which indicated that, for $x_i = a + ih$, all i, the interesting range for p is near $1/(1 + h^3/6)$. Since Bennett has 85 data points, this makes $h = 3/84$, indicating an interesting p near .999992 when working in degrees, while, for the kilometer work in x, $h = 625/84$, hence an interesting p near .014 and, in y, $h = 777/84$, hence the interesting p should be near .0075.