Appendix C: Derivation of the distributions for cross-validating parameters

Difference between the m estimates, Eq. 5

If the number of years in the parameterization and evaluation periods are set equal, $n_e = n_p = n$, the variance of the m estimates, $s^2_{m,R}$, will be the same if s_p and s_{np} remain constant in the two periods. Using this assumption, we can derive the t-distribution of the t statistic in Eq. 5.

$$\frac{(\hat{m}_{R,p}-\hat{m}_{R,e})}{\sqrt{\frac{df_{slp}\hat{s}^2_{slp,p}+df_{slp}\hat{s}^2_{slp,e}}{2df_{slp}}\left(\frac{1}{n-L}+\frac{1}{n-L}\right)}} = \frac{(\hat{m}_{R,p}-\hat{m}_{R,e})}{\sqrt{\left(\frac{(n-L)s^2_{m,R}}{n-L}+\frac{(n-L)s^2_{m,R}}{n-L}\right)}}\,\frac{1}{\sqrt{\frac{df_{slp}\hat{s}^2_{slp,p}+df_{slp}\hat{s}^2_{slp,e}}{2df_{slp}(n-L)s^2_{m,R}}}}$$

$$= \frac{(\hat{m}_{R,p}-\hat{m}_{R,e})}{\sqrt{\left(\frac{(n-L)s^2_{m,R}}{n-L}+\frac{(n-L)s^2_{m,R}}{n-L}\right)}}\,\frac{1}{\sqrt{\frac{s^2_{slp}}{(n-L)s^2_{m,R}}}}\,\frac{1}{\sqrt{\frac{1}{2df_{slp}}\left(\frac{df_{slp}\hat{s}^2_{slp,p}}{s^2_{slp}}+\frac{df_{slp}\hat{s}^2_{slp,e}}{s^2_{slp}}\right)}}$$

This has the distribution $\frac{1}{\sqrt{g}}\,\mathrm{Normal}(0,1)/\sqrt{c^2_v/n}$ which has the distribution $\frac{1}{\sqrt{g}}t_v$

where $v = 2df_{slp}$ and $g = \frac{s^2_{slp}}{(n-L)s^2_{m,R}} = \frac{s^2_{slp}}{\left(\frac{2}{L(n-L)}s^2_{np}+s^2_p\right)}$.

The degrees of freedom, df_{slp}, are given in Appendix B.

Ratio of the s^2_{slp} estimates

If the number of years in the parameterization and evaluation periods are set equal, $n_e = n_p = n$, and if we assume that the non-process and process error variance remain constant within the two periods, the expected value of $\hat{s}^2_{slp,p}$ and $\hat{s}^2_{slp,e}$ are the same and the degrees of freedom for the χ^2 distributions that approximate $\hat{s}^2_{slp,p}$ and $\hat{s}^2_{slp,p}$ are the same. Thus,

$$\frac{\hat{s}^2_{slp,e}}{\hat{s}^2_{slp,p}} \sim \frac{s^2_{slp}c^2_{df_{slp}}/df_{slp}}{s^2_{slp}c^2_{df_{slp}}/df_{slp}} \sim \frac{c^2_{df_{slp}}}{c^2_{df_{slp}}} \sim F(df_{slp}, df_{slp})$$