t-distribution with degrees of freedom $\mu \in (15, 25)$ is a good approximation of the empirical distribution of $\Delta^{\text{ppm}}$ for masses > 2000.

![Figure 6: qqplot - of $\Delta^{\text{ppm}} = m_M - c_1 \cdot m_N - c_0$ versus the t-distribution with 19 degrees of freedom for four mass ranges $m \in (500 - 530), m \in (1000 - 1110), m \in (2000 - 2200)$ and $m \in (3400 - 3700)$.](image)

Sensitivity analysis

The input parameters to the model of the peptide mass cluster centres included:

- $f_i$ – frequencies of the amino acids.
- Cleavage specificity of the protease $R_C$
- $|P|$ - Protein length
- $p_c$ - Cleavage probability

To examine how the output of the model is influenced by these factors we varied the protein length $|P|$ in steps of 100 from 300 to 800 amino acids per protein. We determined the amino acid frequencies $f_i$ for 9 sequence databases (cf. Methods) and used them as inputs to the model. Furthermore, six cleavage