

*Validation of experimental charge density refinement strategies*Regine Herbst-Irmer¹, Lennard Krause¹, Dietmar Stalke¹¹University Of Goettingen, Goettingen, Germany

E-mail: rherbst@shelx.uni-ac.gwdg.de

Modern experimental charge density investigations require the most accurate data, the highest possible resolution and an optimized refinement strategy. As stated in the XD2006 manual the complexity of the model should be increased in a stepwise manner and the final data-to-parameter ratio should not fall below a value of 10. However, following these guidelines does not guarantee a reasonable model or, more precisely, all refined parameters to be reasonable. Common criteria such as residuals of the linear least squares refinement tend to decrease with increasing number of parameters. The statistical method of cross-validation helps to decide whether the addition of parameters increases model quality or introduces bias. This method is well established in macromolecular crystallography, known as the Rfree concept [1]. A fraction of the measured data is excluded from the refinement process and residuals are calculated against the unused data. In charge density investigations there are two concerns that need to be addressed: Firstly, the changes in R-values are small compared to the precision of the Rfree value. Therefore, k-fold cross-validation is utilized. The data set is divided into k subsets. One set acts as validation set, while the remaining sets determine the model. The refinement is repeated k times with each of the subsets serving once as Rfree set [2]. Hence, every single reflection is used for validation. Secondly, the omission of particular reflections may introduce severe bias to the model. This can be monitored by an inspection of the parameter distribution of the k refinements, which can additionally be used for further error diagnostics, e.g. to detect erroneously defined parameters or incorrectly determined reflections. Visualisation tools show the variation of the parameters. The different refinements also provide rough estimates for the standard deviation of topological parameters. This approach is automated for charge density refinement procedures using the XD2006 program suite [3].

[1] Brünger, A. T. (1992). Nature 355, 472-475.

[2] Paul, A., Kubicki, M., Jelsch, C., Durand, P. & Lecomte, C. (2011). Acta Cryst. B67, 365-378.

[3] Krause, L., Niepötter, B., Schürmann, C., Stalke, D. & Herbst-Irmer, R. (2017). IUCrJ submitted.

Keywords: [charge density](#), [cross-validation](#), [multipole model](#)